ON THE MINIMUM DIAMETER OF PLANE INTEGRAL POINT SETS

SASCHA KURZ AND ALFRED WASSERMANN

ABSTRACT. Since ancient times mathematicians consider geometrical objects with integral side lengths. We consider plane integral point sets $\mathcal{P}$, which are sets of n points in the plane with pairwise integral distances where not all the points are collinear.

The largest occurring distance is called its diameter. Naturally the question about the minimum possible diameter $d(2, n)$ of a plane integral point set consisting of n points arises. We give some new exact values and describe state-of-the-art algorithms to obtain them. It turns out that plane integral point sets with minimum diameter consist very likely of subsets with many collinear points. For this special kind of point sets we prove a lower bound for $d(2, n)$ achieving the known upper bound $n^{c_2 \log\log n}$ up to a constant in the exponent.

A famous question of Erdős asks for plane integral point sets with no 3 points on a line and no 4 points on a circle. Here, we talk of point sets in general position and denote the corresponding minimum diameter by $\dot{d}(2, n)$. Recently $\dot{d}(2, 7) = 22\,270$ could be determined via an exhaustive search.

1. INTRODUCTION

In radio astronomy systems of antennas are used. To avoid frequency losses the distance between each pair of antennas have to be an integer multiple of the used wave length. So there is some interest in the construction and properties of m-dimensional integral point sets $\mathcal{P}$, i.e. sets of n points in the Euclidean space $\mathbb{E}^m$ with pairwise integral distances where not all the points are contained in a hyperplane of $\mathbb{E}^m$. For other applications we refer to [3].

In this article we focus on the planar case $m = 2$ and refer to [12, 13, 15] for $m \geq 3$. At most $n - 1$ points are allowed to be collinear. A point set is said to be in semi-general position if no three points are collinear. If additionally no four points are located on a circle we talk of general position. To describe integral point sets $\mathcal{P}$ we denote the largest occurring distance as its diameter $\operatorname{diam}(\mathcal{P})$. From the combinatorial point of view there is a natural interest in the determination of the minimum possible diameter $d(2, n)$ of plane integral point sets. For plane integral point sets in semi-general position and general position we denote the minimum diameter by $\overline{d}(2, n)$ and $\dot{d}(2, n)$, respectively.

Although the study of integral point sets has a long history, see [3] for an overview, not much has been known about the exact values of $d(2, n)$, $\overline{d}(2, n)$, and $\dot{d}(2, n)$, previously. The first exact values are [3, 7]:

$$(d(2,n))_{n=3,\dots,9} = 1, 4, 7, 8, 17, 21, 29,$$
$$\left(\overline{d}(2,n)\right)_{n=3,\dots,9} = 1, 4, 8, 8, 33, 56, 56,$$

and

$$\left(\dot{d}(2,n)\right)_{n=3,\dots,6} = 1, 8, 73, 174.$$

Apart from $d(2, n) \leq \overline{d}(2, n) \leq \dot{d}(2, n)$ the best known bounds are given in [22] and [4], respectively,

$$c_1 n \leq d(2,n) \leq \overline{d}(2,n) \leq n^{c_2 \log\log n}.$$

2000 *Mathematics Subject Classification.* 52C10; 11D99, 53C65.

Key words and phrases. integral distances, diameter, exhaustive search, orderly generation.

For pictures of the corresponding point sets we refer to [3, 10, 12]. It is worth noting that the points of all known integral points sets in semi-general position with minimum diameter lie on a common circle each. We remark that there are constructions known which lead to a dense subset of points of the unit circle with pairwise rational distances, see i.e. [3]. A famous Erdős problem asks for seven points in the plane no three on a line, no four on a circle with pairwise integral distances or more generally for the existence and value of $\dot{d}(2,n)$ (still open). There are some constructions for infinite series of plane integral point sets in general position with 6 points and coprime side lengths [7]. Recently such an example consisting of seven points was found [10].

This is the story of plane integral point sets so far. The paper is arranged as follows:

In Section 2 we give algorithms for the exhaustive generation of plane integral point sets. A variant of orderly generation which combines two integral point sets to obtain a new one instead of extending one point set is described in Subsection 2.1. Clique search is utilized in Subsection 2.2. In Section 3 we analyze properties of the characteristic char$(\mathcal{P})$, this is the squarefree part of $(a+b+c)(a+b-c)(a-b+c)(-a+b+c)$ of a non-degenerated subtriangle with side lengths a, b, and c of a plane integral point set $\mathcal{P}$. Subsections 4.1, 4.2, and 4.3 are devoted to computational results using the algorithms described Section 2:

Theorem 1. *For $n \le 36$ the plane integral point sets in semi-general position with n points and minimum diameter consist of points on a circle with radius $r = \frac{z}{k}\sqrt{k}$ where $k \in \{3, 7, 15\}$ and z is an integer with many prime factors.*

Conjecture 1. *The points of a plane integral point set in semi-general position are situated on a circle.*

Theorem 2.

$$\dot{d}(2,7) = 22\,270.$$

This improves the previous known bound $\dot{d}(2,7) > 320$ [7] and answers Erdős' question positively. The corresponding integral point set was first announced in [10].

Theorem 3. *For $9 \le n \le 122$ the plane integral point sets with n points and minimum diameter $d(2,n)$ consist of a subset of $n-1$ collinear points and one point apart from this line.*

Conjecture 2. *For $n \ge 9$ a plane integral point set with minimum diameter contains a subset of $n-1$ collinear points*

Theorem 3 motivates us to investigate plane integral point sets with many collinear points in Section 5. We give a link between these special plane integral point sets and factorizations of integers. This link enables us to generate them very efficiently and to give the following lower bound.

Theorem 4. *For $\delta > 0$, $\varepsilon > 0$, and a plane integral point set $\mathcal{P}$ consisting of n points with at least n^δ collinear points there exists a $n_0(\varepsilon)$ such that for all $n \ge n_0(\varepsilon)$ we have*

$$\mathit{diam}(\mathcal{P}) \ge n^{\frac{\delta}{4\log 2(1+\varepsilon)}\log\log n}.$$

We end with some remarks on lower bounds for $d(2,n)$ in Section 6 and give a conclusion.

2. Exhaustive search

To determine some further exact values of $d(2,n)$, $\overline{d}(2,n)$, and $\dot{d}(2,n)$ we have applied an exhaustive search. In this Section we describe the used algorithms.

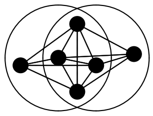

FIGURE 1. Combination of two integral point sets.

2.1. **Orderly generation by combination.** For the construction of plane integral point sets $\mathcal{P}$ in semi-general position at first our used method is to combine two point sets consisting of $n-1$ points having $n-2$ points in common to an integral point set consisting of n points, see Figure 1. We remark that this is similar to the approach of [7].

As an algorithm for the combination of integral point sets we use a variant of orderly generation [2, 16, 17, 21]. The big advantage of orderly generation is that the isomorphism problem can be solved without comparing every constructed pair of discrete objects. There is no need to access a large set of constructed structures during the algorithm, so memory is not the bottleneck any more, which is the case for other types of enumeration algorithms. For more information about the concept of orderly generation we refer to [21] or the more general overview on enumeration algorithms [6].

Since our variant of orderly generation is applicable for the enumeration of universal discrete structures which can be described by an equivalency relation we have to be more general and technical in our description.

At first we have to describe the objects of a discrete structure which we like to construct as an equivalency relation $\simeq$ on an ordered list L. In our case L is the set of distance matrices which correspond to plane integral point sets. To describe $\simeq$ we introduce a mapping $\rho : L \to \mathbb{N}$. Here, we simply map onto the number q of rows or columns of the distance matrix. This number also equals the number of points of the corresponding point set. We denote the total ordering of L by $\prec$. If $\rho(l_1) < \rho(l_2)$ or $\rho(l_1) > \rho(l_2)$ for $l_1, l_2 \in L$ we define $l_1 \prec l_2$ or $l_1 \succ l_2$, respectively. In the remaining case $\rho(l_1) = \rho(l_2)$ we continue $\prec$ by a column-lexicographic ordering on the upper right triangle of l_1 and l_2. By L_q we denote the ordered list containing all elements $l \in L$ with $\rho(l) = q$. If $\rho(l_1) \neq \rho(l_2)$ for $l_1, l_2 \in L$ we define $l_1 \not\simeq l_2$. Otherwise we define $\simeq$ by the natural group action [8] of the symmetric group S_q on the $q = \rho(l_1) = \rho(l_2)$ points of l_1 and l_2.

Because we want to combine only pairs of distance matrices l_1, l_2 consisting of $n-1$ points which have $n-2$ points in common we need a mapping $\downarrow$which deletes the last row and the last column of a distance matrix so that we can rewrite our condition to $\downarrow l_1 = \downarrow l_2$. The combination of two distance matrices l_1 and l_2 itself is done by a mapping $\Gamma_q : L_q \times L_q \to L_{q+1}^{\star}$ where $L_{q+1}^{\star}$ denotes the set of lists of arbitrary length containing elements from L_{q+1}. We define $\Gamma_q(l_1, l_2)$ by an example. Let $l_1 = \begin{pmatrix} 0 & 4 & 4 \\ 4 & 0 & 2 \\ 4 & 2 & 0 \end{pmatrix}$ and $l_2 = \begin{pmatrix} 0 & 4 & 2 \\ 4 & 0 & 4 \\ 2 & 4 & 0 \end{pmatrix}$ be distance matrices with $\downarrow l_1 = \downarrow l_2$ then each distance matrix Δ in $\Gamma_q(l_1, l_2)$ has the shape $\Delta = \begin{pmatrix} 0 & 4 & 4 & 2 \\ 4 & 0 & 2 & 4 \\ 4 & 2 & 0 & \star \\ 2 & 4 & \star & 0 \end{pmatrix}$ where the $\star$ stands for an arbitrary value. In general the distance matrix Δ is obtained from l_1 and l_2 by appending the last row and the last column of l_2 to l_1. The proper values of $\star$ can be calculated by a little computation in the Euclidean metric. It is easy to see that $\star$ can take at most two different values. Additionally we require that the values of $\star$ are integers and that $\Gamma_q(l_1, l_2)$ is ordered by $\prec$. In the above example we have

$$\Gamma_q(l_1, l_2) = \left(\begin{pmatrix} 0 & 4 & 4 & 2 \\ 4 & 0 & 2 & 4 \\ 4 & 2 & 0 & 3 \\ 2 & 4 & 3 & 0 \end{pmatrix} \right).$$

The last ingredient for an orderly generation algorithm is a definition of canonicity. In each equivalence class we have to mark exactly one element. This element is called canonical. There are several ways to define canonical elements. We say that an element $l \in L_q$ is canonical if $l \succeq \sigma(l)\ \forall\, \sigma \in S_q$, i.e. it is the largest object in its equivalence class. For our variant of orderly generation we need also another definition. We call an element $l \in L_q$ semi-canonical if $\downarrow l \succeq \downarrow \sigma(l)\ \forall \sigma \in S_q$. So each canonical element is also semi-canonical. For algorithmic purposes we define a function χ which maps an element $l \in L$ into the set $\{canonical, semi\text{-}canonical, none\}$ where $\chi(l) = semi\text{-}canonical$ if l is semi-canonical and not canonical. Our aim is to construct complete lists $\mathcal{L}_q$ of the semi-canonical plane integral point sets in semi-general position consisting of q points. We suppose that we have already a list $\mathcal{L}_3$ of all integral semi-canonical triangles with given diameter which can be obtained by a simple double loop. To recursively construct the lists $\mathcal{L}_{q+1}$ we apply the following orderly algorithm:

Algorithm 1.
Generation of semi-canonical integral point sets in semi-general position
input: $\mathcal{L}_q$, Γ_q, $\downarrow$, $\prec$
output: $\mathcal{L}_{q+1}$
begin
    $\mathcal{L}_{q+1} = \emptyset$
    loop over $l_1 \in \mathcal{L}_q$, $\chi(l_1) = canonical$ **do**
        loop over $l_2 \preceq l_1$, $l_2 \in \mathcal{L}_q$, $\downarrow l_2 = \downarrow l_1$ **do**
            loop over $y \in \Gamma_q(l_1, l_2)$ **do**
                if $\chi(y) \neq none$ **then** *append* y *to* $\mathcal{L}_{q+1}$ **end**
            end
        end
    end
end

The plane integral point set in semi-general position given by the distance matrix $\Delta = \begin{pmatrix} 0 & 100 & 89 & 21 \\ 100 & 0 & 21 & 89 \\ 89 & 21 & 0 & 82 \\ 21 & 89 & 82 & 0 \end{pmatrix}$ shows why semi-canonical elements are needed for an exhaustive generation. We notice that Δ is canonical and can be only combined with a canonical and a semi-canonical triangle. We leave the proof of the correctness of Algorithm 1 to the reader and also refer to [12], as it is a bit technical but not difficult.

For the canonicity check χ we use backtracking with isomorphism pruning in the general case. Because we have to check 4×4-matrices very often we have developed a fast algorithm for this special case, see [14]. It needs at most 6 integer comparisons to decide whether a given 4×4-matrix is canonical, semi-canonical, or none of both. In our special case of integral point sets we can improve Algorithm 1 by using the characteristic of an integral point set, see Section 3.

2.2. **Clique search.** In this subsection we will present a hybrid construction algorithm that combines the orderly algorithm of the previous subsection and clique search to search for integral point sets with large diameter. We remark that clique search is a common technique in extremal combinatorics and was also utilized before in the construction of 3-dimensional integral point sets, see [20]. Suppose we are given an integral triangle Δ. Via orderly generation we can construct all integral point sets $\mathcal{P}_i$ in semi-general position consisting of 4 points with $\downarrow\mathcal{P}_i = \Delta$. This means that each point set $\mathcal{P}_i$ consists of Δ and a further point v_i. The next step is to build up a graph $\mathcal{G}_\Delta$ with v_i as its vertices. We define $\{v_i, v_j\}$ to be an edge in $\mathcal{G}_\Delta$ iff the distance between v_i and v_j is integral and v_i, v_j are not collinear with a point of Δ.

Obviously for every plane integral point set $\mathcal{P}$ in semi-general position there exists an integral triangle Δ such that $\mathcal{P}$ corresponds to a clique of $\mathcal{G}_\Delta$. In the other direction every clique of $\mathcal{G}_\Delta$ corresponds to a plane integral point set but it could happen that three points v_h, v_i, and v_j are collinear. So we produce only candidates of plane integral point sets which have to be checked whether they are in semi-general position.

Our hybrid algorithm works as follows. For a given diameter d and a lower bound b for the number of points we loop over all integral triangles with diameter d. Then we determine the vertices of the graph $\mathcal{G}_\Delta$ by orderly generation and insert the edges. Here, we can use the lower bound b to shrink the graph by deleting edges with at most $b-5$ neighbors. On the resulting graph we perform a clique search using CLIQUER [18, 19] or an implementation of the Bron-Kerbosch algorithm [1] to generate the cliques of size at least $b-3$. As a last step we reconstruct from the vertices of each clique and the triangle Δ a plane integral point set $\mathcal{P}$ and check if it is in semi-general position.

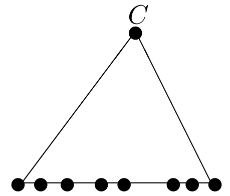

FIGURE 2. Integral points on the side of a triangle.

If we want to generate plane integral point sets in arbitrary position we have to modify our approach slightly since it does not produce all possibilities of plane integral point sets consisting of 4 points which can be seen as follows. Suppose we are given a triangle Δ as in Figure 2 where $\overline{AB}$ is the largest side. Algorithm 1 joins Δ with all other possible triangles along the side $\overline{AB}$ so it cannot generate point sets with further points on the line l through A and B.

The situation can be cleared easily. We may simply test all points on l with integral distances to the endpoints of l whether their distance to the third point C of Δ is also integral. In Section 5 we will describe a more sophisticated algorithm for this purpose. So we simply add those points on l and the corresponding edges to $\mathcal{G}_\Delta$.

We remark that in our computer calculations the major part of the running time of our hybrid algorithm is used for the orderly generation. This part could be replaced by a more direct algorithm which construct all points v_i which have integral distances to the three corners of a Δ without the restriction of the diameter, see [9] for details. We do not use this algorithm for our purpose because it runs more slowly.

3. CHARACTERISTIC OF A PLANE INTEGRAL POINT SETS

Due to the Heron formula

$$A_\Delta(a,b,c) = \frac{\sqrt{(a+b+c)(a+b-c)(a-b+c)(-a+b+c)}}{4}$$

for the area A_Δ of a triangle with side lengths a, b, c we can write the area of a non-degenerate triangle with integral side lengths uniquely as $A = q\sqrt{k}$ with a rational number q and a squarefree integer k. The number k is called the characteristic char(Δ) of the triangle. The next theorem allows us to talk about the characteristic char$(\mathcal{P})$ of a point set $\mathcal{P}$.

Theorem 5. *Each non-degenerate triangle Δ in a plane integral point $\mathcal{P}$ set has the same characteristic char$(\mathcal{P})$ = char(Δ).*

Proof. See [7] or [12, 13] □

Clearly, we modify Algorithm 1 and combine only integral point sets with equal characteristic. To be able to measure the complexity of the modified algorithm and for some number theoretical insight we introduce the function $\psi(d,k) =$

$$\left|\left\{(a,b)\;\middle|\; \begin{array}{l} a,b\in\{1,2,\dots,d\},\ a+b>d,\ \exists w\in\mathbb{N}: kw^2 = \\ (a+b+d)(a+b-d)(a-b+d)(-a+b+d) \end{array}\right\}\right|$$

which counts the number of integral semi-canonical triangles with diameter d and characteristic k if we set $\psi(d,k)=0$ for non-squarefree numbers k. We remark that the the running time of the modified orderly algorithm for the generation of integral point sets consisting of 4 points is given by $\sum_k \binom{\psi(d,k)+1}{2}$.

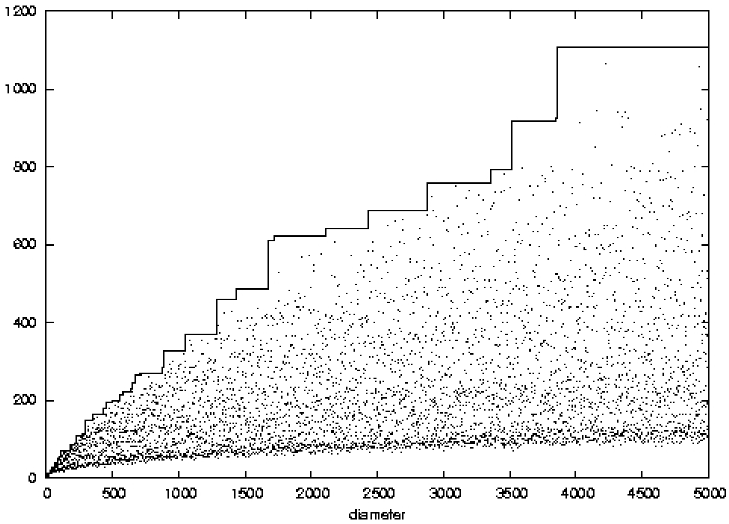

FIGURE 3. Occurring values $\psi(d,k)$ for $d\le 5000$.

In Figure 3 we have depicted the occurring values of $\psi(d,k)$ for $d\le 5000$. We introduce the function $\tilde{\psi}(d)=\max_k\{\psi(d,k)\mid k\in\mathbb{N}\}$. A parameter solution for the set of all integral triangles with characteristic k given in [7] can be rewritten to:

Theorem 6. *For all integral triangles with side lengths a, b, c and characteristic k there exists at least one integer tuple (p,q,h,i,j) fulfilling*

$$\begin{aligned} a &= \frac{ph(i^2+kj^2)}{q}, \\ b &= \frac{pi(h^2+kj^2)}{q}, \\ c &= \frac{(i+h)(ih-kj^2)}{q}, \end{aligned}$$

$\gcd(p,q)=\gcd(h,i,j)=1$, $i\ge h$, *and* $ih>kj^2$.

Using Theorem 6 and some cumbersome technical computation one can deduce $\tilde{\psi}(d)\in O\!\left(d^{1+\frac{c}{\log\log d}}\right)$ for a suitable constant c. For the details we refer the interested reader to [12], where also more detailed considerations on the running time of Algorithm 1 and some more numerical data can be found.

In the other direction we have:

Lemma 1. *The number of different characteristics of integral triangles with diameter d is in $O(d^2)$ and in $\Omega\!\left(\frac{d^2}{(\log d)^2}\right)$.*

Proof. Since there are $O(d^2)$ integral triangles with diameter d we have the stated trivial upper bound. For the other direction we choose for suitable large d two primes p_1,p_2

fullfilling $\frac{9}{4}d < p_1 < \frac{10}{4}d$ and $\frac{5}{4}d < p_2 < \frac{6}{4}d$. With this we set $a = d$, $b = \frac{p_1+p_2}{2} - d$, and $c = \frac{p_1-p_2}{2}$. Thus

$$a+b+c = p_1,\ a+b-c = p_2,\ \frac{3}{4}d < b < d,\ \text{ and }\ \frac{3}{8}d < c < \frac{5}{8}d.$$

Because p_1 and p_2 have to be odd for big enough d and because $b+c = p_1 - d > d = a$ the values $a > b > c$ are integers and fulfill the triangle conditions. Since $\frac{1}{2}d < a-b+c < \frac{3}{4}d$ and $\frac{1}{4}d < -a+b+c < \frac{1}{2}d$ the characteristic of the triangle with side lengths a, b, and c is divisible by p_1p_2. Due to the prime number theorem we have $\Omega\left(\frac{d}{\log d}\right)$ choices for p_1 and p_2 each. Thus there are at least $\Omega\left(\frac{d^2}{(\log d)^2}\right)$ different characteristics. □

We would like to mention that the first author has recently generalized the definition of the characteristic of a plane integral point set to m-dimensional integral point sets and has proven an analog theorem to Theorem 5, see [12, 13].

4. Minimum diameter of plane integral point sets

With the algorithms of Section 2 at hand we are now state the results of our exhaustive search for plane integral point sets with minimum diameter.

4.1. Plane integral point sets in general position. To construct plane integral point sets in general position we need a check to decide whether three points lie on a line or four points are located on a circle. For the first check we can use the triangle inequalities in the degenerated case and for the second check we can use Ptolemy's theorem.

We have implemented the algorithm described in Subsection 2.1 and our computers have constructed all plane integral point sets in general position with diameter at most 30 000. We have only found one such point set consisting of seven points, which proves Theorem 2. For details, pictures, and a second example of diameter 66 810 we refer to [10].

4.2. Plane integral point sets in semi-general position. We have also applied our orderly algorithm on the construction of plane integral point sets in semi-general position. Our available computer power has allowed us to generate all such point sets with diameter at most 5 000 and thus to determine $\overline{d}(2, n)$ up to $n = 24$. With the hybrid construction algorithm of Subsection 2.2 we were able to enumerate all plane integral point sets in semi-general position with minimum diameter $\overline{d}(2, n)$ up to $\text{diam}(\mathcal{P}) = 20\,000$ leading to

$$\begin{aligned}\left(\overline{d}(2,n)\right)_{n=10,\dots,36} = {} & 105, 105, 105, 532, 532, 735, 735, 735, 735,\\ & 1995, 1995, 1995, 1995, 1995, 1995, 9555, 9555, 9555, 10672,\\ & 13975, 13975, 13975, 13975, 13975, 13975, 13975, 13975.\end{aligned}$$

We remark that our hybrid algorithm of Subsection 2.2 did not produce any candidates with collinear triples.

With the minimal examples up to to $n = 36$ at hand we could check Theorem 1 which states that the points of integral point sets in semi-general position with minimum diameter are located on circles with radius $r = \frac{z}{k}\sqrt{k}$ where $k \in \{3, 7, 15\}$ and z is an integer with many prime factors. With this we are motivated to conjecture this pattern in general, see Conjecture 2. Searching for these special types of integral point sets yields the upper bounds given in Table 1.

Conjecture 3. *The upper bounds from Table 1 are sharp.*

We remark that the authors of [4] describe a construction, for $k = 3$ and $z = \prod_{i=1}^{r} p_i^{v_i}$ with $p_i \equiv 1 \mod 6$ being primes, directly yielding the integral point sets of Table 1 in these cases. Their construction relies on calculations over the ring $\mathbb{Z}\left[\frac{-1+\sqrt{-3}}{2}\right]$. It is possible that there are similar constructions for squarefree k or at least $k = 7, 15$.

z	k	$\|\mathcal{P}\|$	diam($\mathcal{P}$)	z	k	$\|\mathcal{P}\|$	diam($\mathcal{P}$)
1	3	3	1	118864	15	40	61 375
8	15	4	4	53599	3	48	61 880
7	3	6	8	157339	3	54	181 675
64	15	7	33	475456	15	56	245 518
49	3	9	56	375193	3	72	433 225
91	3	12	105	3684784	15	80	1 902 813
704	7	14	532	1983163	3	96	2 289 957
637	3	18	735	4877509	3	108	5 632 056
1729	3	24	1 995	14739136	15	112	7 611 252
8281	3	27	9 555	13882141	3	144	16 029 704
20672	15	28	10 672	85276009	3	192	98 468 151
12103	3	36	13 975				

TABLE 1. Upper bounds for $\overline{d}(2,n)$ from point sets on circles with radius $r = \frac{z}{k}\sqrt{k}$.

4.3. **Plane integral point sets in arbitrary position.** The orderly generation by combination approach is limited by the number of substructures as n grows. For example we consider a plane integral point set $\mathcal{P}$ consisting of 89 points where 88 points are collinear. We will determine $d(2,89)$ and it will turn out that the corresponding point set has the shape of Figure 2. The convex hull of $\mathcal{P}$ is formed by three points A, B, and C. Let us assume that the 86 other points of $\mathcal{P}$ are located on the line through A and B. Now we consider all point sets consisting of A, B, C, and 43 of the other points. Then we have the maximum number of $\binom{86}{43} \approx 6.6 \cdot 10^{24}$ possibilities. Because orderly generation has to generate all these point sets to finish in $\mathcal{P}$ it is beyond our means to determine $d(2,89)$ with this method.

So again we had to use the hybrid algorithm of Subsection 2.2. By doing an exhaustive search up to diameter 10 000 we were able to determine the minimum diameter $d(2,n)$ up to $n = 122$ points. Checking the minimal examples uncovers that for $9 \le n \le 122$ points in each case $n-1$ points are collinear, which proves Theorem 3 and motivates Conjecture 2.

5. PLANE INTEGRAL POINT SETS WITH MANY POINTS ON A LINE

We have seen in the last subsection that plane integral point sets with many collinear points are interesting objects to study because plane integral point sets with minimum diameter seem to belong to this class. At first we present an important link between plane integral point sets with $n-1$ collinear points and the decompositions of a certain integer, the decomposition number, into two factors.

Definition 1. *The decomposition number D of an integral triangle with side lengths a, b, and c is given by*

$$D = \frac{(a+b+c)(a+b-c)(a-b+c)(-a+b+c)}{\gcd(b^2-c^2+a^2,2a)^2}.$$

Lemma 2. (Decomposition Lemma)
The distances of a plane integral point set $\mathcal{P}$ consisting of n points where a subset of $n-1$ points is collinear correspond to decompositions of the decomposition number D of the largest triangle of $\mathcal{P}$ into two factors.

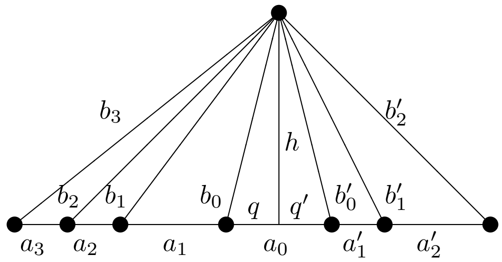

FIGURE 4. Plane integral point set $\mathcal{P}$ with $n-1$ points on a line.

Proof. We use the notation of Figure 4, let s be the largest index of a_i, let t be the largest index of a_i', and set

$$c_i = q + \sum_{j=1}^{i} a_j \quad \text{for} \quad 0 \le i \le s,\ c_i' = q' + \sum_{j=1}^{i} a_j' \quad \text{for} \quad 0 \le i \le t.$$

Let us associate b with b_s, c with b_t', and a with $\sum_{i=1}^{s} a_i + a_0 + \sum_{i=1}^{t} a_i'$. With this we have

$$c_s = q + \sum_{j=1}^{s} a_j = \frac{b^2 - c^2 + a^2}{2a}.$$

By defining $g := \frac{2a}{\gcd(b^2-c^2+a^2,2a)}$ we obtain $gc_s \in \mathbb{N}$ and $gq \in \mathbb{N}$. From $q+q' = a_0 \in \mathbb{N}$ we conclude $gq' \in \mathbb{N}$. Thus $gc_i, gc_j' \in \mathbb{N}$ for all possible indices. An application of Pythagoras' Theorem yields $c_{i+1}^2 + h^2 = b_{i+1}^2$ and $c_i^2 + h^2 = b_i^2$ for $0 \le i < s$. We conclude

$$g^2h^2 = (gb_i + gc_i)(gb_i - gc_i) = (gb_j' + gc_j')(gb_j' - gc_j').$$

Since the b_i, b_j' are integers we can obtain the possible values for c_i and c_i' by decomposing g^2h^2 into two factors.

Due to the Heron formula $16A_\Delta^2 = (a+b+c)(a+b-c)(a-b+c)(-a+b+c)$ and the formula for the area of a triangle $2A_\Delta = ah$ we finally get

$$\begin{aligned} g^2h^2 &= \frac{g^2(a+b+c)(a+b-c)(a-b+c)(-a+b+c)}{4a^2} = \\ &= \frac{(a+b+c)(a+b-c)(a-b+c)(-a+b+c)}{\gcd(b^2-c^2+a^2,2a)^2} = D\,. \end{aligned}$$

□

With the aid of the Decomposition Lemma and Theorem 3 we are able to describe the plane integral points with minimum diameter for $9 \le n \le 122$ points in a very compact manner by giving n, $D = g^2h^2$, and g. For $9 \le n \le 20$ see Table 2 and for the complete listing see [11]. We remark that for $n \le 122$ we only have $g = 2$ for $n = 9, 10, 11$ otherwise we have $g = 1$. The Decomposition Lemma can also be used as a heuristic to determine good upper bounds for $d(2,n)$. One only has to loop over Decompositions numbers D with many divisors and apply Lemma 2 for the construction of an integral point set. For $n \le 148$ the results of this heuristic are given in [11].

We can also utilize the Decomposition Lemma for a fast algorithm to determine the integral points on a side of an integral triangle Δ. Suppose we are given a triangle with side lengths a, b, c and diameter d. We also assume that we have tabulated the prime factorizations of all integers smaller equal $3d$ in a pre-calculation. So we can use the formula in the proof of Lemma 2 to obtain the prime factorization of g^2h^2 and loop through

n	$d(2,n)$	$D=g^2h^2$	g	n	$d(2,n)$	$D=g^2h^2$	g
9	29	$3^2\cdot 5\cdot 7$	2	15	104	$2^5\cdot 3^2\cdot 5$	1
10	40	$3^2\cdot 5\cdot 7$	2	16	121	$2^5\cdot 3^2\cdot 5$	1
11	51	$3^2\cdot 5\cdot 7$	2	17	134	$2^5\cdot 3^2\cdot 5$	1
12	63	$2^5\cdot 3\cdot 5$	1	18	153	$2^5\cdot 3^2\cdot 5$	1
13	74	$2^5\cdot 3\cdot 5$	1	18	153	$2^6\cdot 3^2\cdot 5$	1
13	74	$2^5\cdot 3^2\cdot 5$	1	19	164	$2^6\cdot 3^2\cdot 5$	1
14	91	$2^5\cdot 3^2\cdot 5$	1	20	196	$2^6\cdot 3^2\cdot 5$	1

TABLE 2. Parameters for plane integral point sets with minimum diameter and $9 \le n \le 20$.

the divisors and determine the suitable points on a side of Δ in time $O\left(d^{\frac{c'}{\log\log d}}\right)$ where c' is a suitable constant.

The last application of the Decomposition Lemma is the proof of Theorem 4. Therefore we need the following theorem from number theory:

Theorem 7. *(**Theorem 317** [5]) There exists a function $m_0(\varepsilon)$ such that for each $\varepsilon > 0$ and all $m > m_0(\varepsilon)$ we have*

$$\tau(m) < 2^{(1+\varepsilon)\frac{\log m}{\log\log m}}$$

where $\tau(m)$ denotes the number of divisors of m.

Proof of Theorem 4.
Due to the Decomposition lemma and $\max(s+1,t+1) \ge \frac{n^\delta}{2}$ we need at least $\frac{n^\delta}{2}$ different factorizations of g^2h^2 into two factors, and so we have the condition $\tau(g^2h^2) \ge \frac{n^\delta}{2}$. With $g^2h^2 \le 4\,\mathrm{diam}(\mathcal{P})^4$ and Theorem 7 we conclude

$$2^{\frac{(1+\varepsilon')\log(4\cdot\mathrm{diam}(\mathcal{P})^4)}{\log\log(4\cdot\mathrm{diam}(\mathcal{P})^4)}} \ge \frac{n^\delta}{2}$$

for $n \ge n_0'(\varepsilon')$ and $\varepsilon' > 0$. Because $\mathrm{diam}(\mathcal{P}) \ge \mathrm{d}(2,n) \ge c_1 n$ and with a change from ε' to $\epsilon > 0$ we have

$$2^{\frac{(1+\varepsilon)4\log\mathrm{diam}(\mathcal{P})}{\log\log n}} \ge n^\delta$$

for $n \ge n_0(\varepsilon)$, and the theorem follows by an elementary transformation. $\square$

6. SOME REMARKS ON LOWER BOUNDS FOR $\mathbf{d(2,n)}$

So far we have derived a good lower bound (compared to the upper bound) only in the case when the point set contains *many* collinear points. It would be nice to get rid of this last condition. We would like to mention a possible reason why plane integral point sets with minimum diameter seem to have subsets of many collinear points. Theorem 5 states that each non-degenerated triangle of a plane integral point set $\mathcal{P}$ has the same characteristic. In Section 3 we have observed that integral triangles with equal characteristic and equal diameter are somewhat rare. So a plane integral point set with small diameter is forced to either use isomorphic triangles several times or to have many degenerated triangles. For the second possibility we have the following analysis. The maximum number of non-degenerate triangles is achieved by a point set in semi-general position and consists of $\binom{n}{3} = \frac{n(n-1)(n-2)}{6}$ non-degenerate triangles where n is the number of points. The minimum number $\binom{n-1}{2} = \frac{(n-1)(n-2)}{2}$ is attained by a point set where $n-1$ points are collinear. The first possibility is ruled out by the following lemma.

Lemma 3. *Each plane integral point set consisting of n points with at most $\frac{n}{2}$ points on a line contains a set of at least $\frac{n}{8}$ different integral triangles with equal diameter and equal characteristic.*

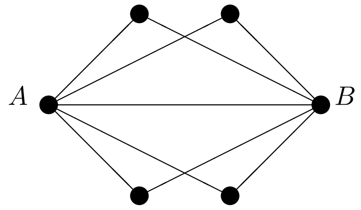

FIGURE 5. Equivalent triangles sharing a common side.

Proof. We can choose an arbitrary pair (A, B) of points out of the n points. Because at most $\frac{n}{2}$ points lie on a line there are at least $\frac{n}{2}$ points C not on the line through A and B. The proof is finished by the fact that at most 4 equivalent triangles can share the side between A and B, see Figure 5. $\square$

With $\tilde{\psi}(d) \in O\!\left(d^{1+\frac{c}{\log\log d}}\right)$, Lemma 3, and Theorem 4 one could conclude

$$d(2, n) \geq cd^{1-\varepsilon}$$

for arbitrary $\varepsilon > 0$ and a suitable constant c, see [12] for the details. Unfortunately the bound is slightly less than the bound of Solymosi [22]. So we propose a similar strategy along the same lines. Consider a fix integral triangle Δ with diameter d and define $\Upsilon(\Delta, d)$ as the number of canonical integral point sets $\mathcal{P}$ in semi-general position with $\downarrow\mathcal{P} = \Delta$. By $\overline{\Upsilon}(d)$ we denote the maximum of $\Upsilon(\Delta, d)$ over all integral triangles Δ with diameter d. Similar to Lemma 3 we can prove that a plane integral point set consisting of n points with *few* collinear points and diameter d contains $\Omega(n)$ point sets of the type counted by $\Upsilon(\Delta, d)$. So we have $\frac{n}{c'} \geq \overline{\Upsilon}^{-1}(d)$ for a suitable constant c'. There is some numerical evidence that $\overline{\Upsilon}(d) \in O\!\left(d^{\frac{c}{\log\log d}}\right)$ for a suitable constant c. If that could be proven we would have a lower bound of $d(2, n) \geq \tilde{c} \cdot n^{c' \log\log n}$ for suitable constants $\tilde{c}, c'$, see [12] for the details.

7. CONCLUSION

We have presented some new exact values for the minimum diameters $d(2, n)$, $\overline{d}(2, n)$ and for $\dot{d}(2, 7)$ obtained by exhaustive search with custom-made algorithms. Having these new values and the corresponding integral point sets at hand we may speculate about a structure theorem for integral point sets with minimum diameter in arbitrary or semi-general position. We have formulated this as Conjecture 2 and Conjecture 1, respectively. In Lemma 2 we have presented an important link between decompositions of a certain number into two factors and the distances of a plane integral point set consisting of n points with a subset of $n - 1$ collinear points. It also seems that the minimum diameter $\overline{d}(2, n)$ is dominated by decompositions of certain numbers. Here, [4] gives a first glance at the possibly underlying rich number theoretic structure, but more research has to be done. The derivation of tight bounds for the minimum diameter $d(2, n)$ is a challenging task for the future. Our contribution was the resolution of the special case of point sets with many collinear points.

REFERENCES

[1] C. Bron and J. Kerbosch. Finding all cliques of an undirected graph. *Commun. ACM*, 16:575–577, 1973.
[2] I.A. Faradzev. Constructive enumeration of combinatorial objects. In *Problémes combinatoires et théorie des graphes, Orsay 1976, Colloq. int. CNRS No.260*, pages 131–135. 1978.
[3] H. Harborth. Integral distances in point sets. In P. L. Butzer et al., editors, *Karl der Grosse und sein Nachwirken. 1200 Jahre Kultur und Wissenschaft in Europa. Band 2: Mathematisches Wissen. Turnhout: Brepols*, pages 213–224. 1998.

[4] H. Harborth, A. Kemnitz, and M. Möller. An upper bound for the minimum diameter of integral point sets. *Discrete Comput. Geom.*, 9(4):427–432, 1993.
[5] G.H. Hardy and E.M. Wright. *An introduction to the theory of numbers. 5th ed.* Oxford etc.: Oxford at the Clarendon Press. XVI, 426 p. , 1979.
[6] P. Kaski and P.R.J. Östergård. *Classification algorithms for codes and designs.* Algorithms and Computation in Mathematics 15. Berlin: Springer. xii, 412 p. with CD-ROM., 2006.
[7] A. Kemnitz. Punktmengen mit ganzzahligen Abständen. Habilitationsschrift, TU Braunschweig, 1988.
[8] A. Kerber. *Applied finite group actions. 2nd, rev. and exp. ed.* Algorithms and Combinatorics. 19. Berlin: Springer. xxvi, 454 p., 1999.
[9] A. Kohnert and S. Kurz. A note on Erdös-Diophantine graphs and Diophantine carpets. *Mathematica Balkanica*, 21(1-2), 2007.
[10] T. Kreisel and S. Kurz. There are integral heptagons, no three points on a line, no four on a circle. *Discrete Comput. Geom.*, (to appear).
[11] S. Kurz. Minimum diameters of plane integral point sets. http://www.wm.uni-bayreuth.de/index.php?id=243.
[12] S. Kurz. *Konstruktion und Eigenschaften ganzzahliger Punktmengen*. PhD thesis, Bayreuth. Math. Schr. 76. Universität Bayreuth, 2006.
[13] S. Kurz. On the characteristic of integral point sets in $\mathbb{E}^m$. *Australas. J. Comb.*, 36:241–248, 2006.
[14] S. Kurz. Enumeration of integral tetrahedrons. *J. Integer Seq.*, 2007. Artikel 07.9.3.
[15] S. Kurz and R. Laue. Bounds for integral point sets. *Australas. J. Comb.*, 39:233–240, 2007.
[16] R. Laue. Construction of combinatorial objects: A tutorial. *Bayreuther Math. Schr.*, 43:53–96, 1993.
[17] B.D. McKay. Isomorph-free exhaustive generation. *J. Algorithms*, 26(2):306–324, 1998.
[18] S. Niskanen and P.R.J. Östergård. Cliquer user's guide, version 1.0. Technical Report T48, Communications Laboratory, Helsinki University of Technology, Espoo, Finland, 2003.
[19] P.R.J. Östergård. A fast algorithm for the maximum clique problem. *Discrete Appl. Math.*, 120(1-3):197–207, 2002.
[20] L. Piepmeyer. Räumliche ganzzahlige Punktmengen. Master's thesis, TU Braunschweig, 1988.
[21] R.C. Read. Every one a winner or how to avoid isomorphism search when cataloguing combinatorial configurations. *Ann. Discrete Math.*, 2:107–120, 1978.
[22] J. Solymosi. Note on integral distances. *Discrete Comput. Geom.*, 30(2):337–342, 2003.

SASCHA KURZ, DEPARTMENT OF MATHEMATICS, PHYSIC AND INFORMATICS, UNIVERSITY OF BAYREUTH, GERMANY
E-mail address: sascha.kurz@uni-bayreuth.de

ALFRED WASSERMANN, DEPARTMENT OF MATHEMATICS, PHYSIC AND INFORMATICS, UNIVERSITY OF BAYREUTH, GERMANY
E-mail address: alfred.wassermann@uni-bayreuth.de