ON THE NUMBER OF SIMILAR INSTANCES OF A
PATTERN IN A FINITE SET

BERNARDO ÁBREGO, SILVIA FERNÁNDEZ-MERCHANT, DANIEL J. KATZ,
AND LEVON KOLESNIKOV

Abstract. New bounds on the number of similar or directly simi-
lar copies of a pattern within a finite subset of the line or the plane
are proved. The number of equilateral triangles whose vertices all lie
within an \( n \)-point subset of the plane is shown to be no more than
\( \left\lfloor \frac{(4n-1)(n-1)}{18} \right\rfloor \). The number of \( k \)-term arithmetic progressions
that lie within an \( n \)-point subset of the line is shown to be at most
\( (n-r)(n+r-k+1)/(2k-2) \), where \( r \) is the remainder when \( n \) is
divided by \( k-1 \). This upper bound is achieved when the \( n \) points
themselves form an arithmetic progression, but for some values of \( k \)
and \( n \), it can also be achieved for other configurations of the \( n \) points, and a
full classification of such optimal configurations is given. These results
are achieved using a new general method based on ordering relations.

1. Introduction

Erdős and Purdy [18, 19, 20] raised the question of finding the maxi-
mum number of equilateral triangles that can be determined by \( n \) points
in the plane, where we say that a triangle is determined by a set \( V \) when
the three vertices of the triangle lie in \( V \). This problem is also mentioned
in the compendia of unsolved problems in geometry by Croft, Falconer,
and Guy [13] and by Brass, Moser, and Pach [9], and is discussed by Pach
and Sharir in [27]. There are many other variations of this problem in-
volving other patterns, constraints on the \( n \)-set in which instances of the
pattern are sought, higher-dimensional ambient spaces, different equiva-
lence relations for the patterns, and different optimization objectives, e.g.,
[1, 3, 4, 5, 6, 7, 11, 12, 25, 26, 28, 29]. These problems trace their inspira-
tion to Erdős’ question [17] about the maximum number of pairs of points
at unit distance that can be determined by an \( n \)-subset of the plane, and
other related questions, which have led to a rich literature (see [9, 24] for
an overview). Apart from being a well-known and important question in
discrete geometry, the problem of determining \( S_\mathcal{P}(n) \) is relevant to the
problem of pattern recognition in data from scanners, cameras, and telescopes
[8, 9, 10].

In all these problems, we have a universe, which is a set \( U \), an equivalence
relation \( \sim \) on the power set of \( U \), and a pattern \( \mathcal{P} \), which is an equivalence
class of $\sim$. If $P \in \mathcal{P}$, we say that $P$ is an instance of pattern $\mathcal{P}$. We call $\mathcal{P}$ a $k$-pattern if its instances are $k$-sets. For the rest of this paper, the pattern $\mathcal{P}$ is always assumed to be a finite set. For $V \subseteq U$, we let
\[ S_\mathcal{P}(V) = |\{P \subseteq V : P \in \mathcal{P}\}|. \]

We are concerned with finite subsets of our universe $U$. For an integer $n$, we define
\[ S_\mathcal{P}(n) = \max_{V \subseteq U} S_\mathcal{P}(V), \]
that is, the largest number of instances of $\mathcal{P}$ that can be found in an $n$-subset of $U$.

The results of this paper concern the case when $U$ is the line $\mathbb{R}$ or the plane $\mathbb{R}^2$ (identified with $\mathbb{C}$), where $\sim$ is the geometric relation of similarity or direct similarity, and where the patterns $\mathcal{P}$ are classes of finite subsets, such as arithmetic progressions in the line, or equilateral triangles in the plane.

If our relation $\sim$ is similarity or direct similarity, and our pattern $\mathcal{P}$ has more than one point, and we are counting the number of instances of $\mathcal{P}$ in subsets of the plane, then it is not hard to show that $S_\mathcal{P}(n)$ is $O(n^2)$. Elekes and Erdős [16] proved a slightly subquadratic lower bound for general patterns $\mathcal{P}$ and a quadratic lower bound if there is an instance of $\mathcal{P}$ whose points all have algebraic numbers for their coordinates. Laczkovich and Ruzsa [23] later showed that $S_\mathcal{P}(n) = \Theta(n^2)$ if and only if the cross ratio of every quadruplet of distinct points in an instance of $\mathcal{P}$ is algebraic. However, before this work there were no patterns $\mathcal{P}$ (with at least three points) for which the quadratic coefficient was known.

Let $\mathcal{E}$ be the pattern of the vertices of an equilateral triangle, that is, $\mathcal{E}$ contains all 3-point subsets of the plane such that all three pairs of points within the set have the same distance. Then the results of Laczkovich and Ruzsa show that $S_\mathcal{E}(n) = \Theta(n^2)$. Prior to this paper, the best known bounds [2] for $S_\mathcal{E}(n)$ were
\[ \left( \frac{1}{3} - \frac{\sqrt{3}}{4\pi} \right) n^2 + O(n^{3/2}) \leq S_\mathcal{E}(n) \leq \left\lfloor \frac{(n-1)^2}{4} \right\rfloor, \]
so that
\[ 0.1955 < \frac{1}{3} - \frac{\sqrt{3}}{4\pi} \leq \liminf_{n \to \infty} \frac{S_\mathcal{E}(n)}{n^2} \leq \limsup_{n \to \infty} \frac{S_\mathcal{E}(n)}{n^2} \leq \frac{1}{4}. \]
The lower bound is obtained from the points in the equilateral triangle lattice contained in a disk of a suitable radius. It is conjectured [2, Conjecture 1] that this construction is asymptotically optimal and that $\lim_{n \to \infty} S_\mathcal{E}(n)/n^2$ exists and equals $\frac{1}{3} - \frac{\sqrt{3}}{4\pi}$, the lower endpoint of the interval in which the limits inferior and superior for $S_\mathcal{E}(n)/n^2$ are known to lie.

The main result of this paper is an improved upper bound on $S_\mathcal{E}(n)$ that cuts this interval to less than half the previous length.
Theorem 1.1. If $E$ is the pattern of the vertices of equilateral triangles in $\mathbb{R}^2$, then $S_E(n) \leq \left\lfloor (4n-1)(n-1)/18 \right\rfloor$.

Therefore $\limsup_{n \to \infty} S_E(n)/n^2 \leq 2/9$. The proof is given in Section 8 and is based on a new and highly general order-theoretic methodology, developed in Sections 2–4 for attacking pattern-counting problems.

Our new methodology was first developed while exploring the simpler problem of counting instances of finite patterns (such as arithmetic progressions) in finite subset of the line $\mathbb{R}$, and was used to discover interesting results that are presented in Sections 5–7. After settling preliminaries in Section 5 we calculate the precise maximum number of $k$-term arithmetic progressions that can occur in an $n$-point subset of the line in Section 6.

Theorem 1.2. Let $k > 1$, and let $A_k$ be the pattern of $k$-term arithmetic progressions on $\mathbb{R}$. If $n \in \mathbb{N}$ and $r$ is the remainder when $n$ is divided by $k-1$, then $S_{A_k}(n) = (n-r)(n+r-k+1)/(2k-2)$.

For each $k > 2$, we completely classify all $n$-point subsets of the line that contain as many $k$-term arithmetic progressions as possible. These optimal sets include $n$-term arithmetic progressions, but there are some other optimal sets: when $k > 3$ and $k-1$ divides $n$, sets obtained by deleting the second or penultimate point from an arithmetic progression are also optimal (see Proposition 6.5). When $k = 3$, the variety of the optimal sets is considerably richer (see Proposition 6.3). This work complements the asymptotic approach of Elekes [15], who studied the structure of $n$-subsets of the line that contain an asymptotically large number of $k$-term arithmetic progressions, whereas we determine the precise maximum and those sets that achieve it.

In Section 7 we show that our method can be used to obtain both upper and lower bounds on $S_P(n)$ when $P$ is a commensurable pattern in the line, that is, where the ratios of the distances between points in $P$ are all rational. We count the number of directly similar copies of $P = \{0,1,3\}$ in finite subsets of the line, so if $P$ is the pattern for $P$, we show that

\begin{equation}
\frac{1}{6} \leq \liminf_{n \to \infty} \frac{S_P(n)}{n^2} \leq \limsup_{n \to \infty} \frac{S_P(n)}{n^2} \leq \frac{1}{4},
\end{equation}

and we construct an interesting family of $n$-point subsets $V_n$ of the line with $\lim_{n \to \infty} S_P(V_n)/n^2 = 3/16$, which is the highest known value and, interestingly, lies strictly between the bounds of (2).

As noted, our order-theoretic method is a new and highly general tool for counting patterns in finite sets. It combines nicely with other techniques of discrete geometry, for example, the proof of Theorem 1.1 combines our tools with the topological technique of finding three coincident halving lines of a set (cf. [22] Lemma 2, [21] Lemma 3, and [14] Lemma 2). Further applications will be explored in future works.
2. Order and Partitions

A key theme in this work is that one can place an order relation on the universe $U$, and use it to create bounds on $S_p(n)$. So henceforth, we assume that our universe $U$ has a total ordering relation $\preceq$. If $V \subseteq U$, then we say that a point $v \in V$ is the $i$th point of $V$ with respect to $\preceq$ to mean that there are precisely $i-1$ points $w \in V$ with $w \prec v$. If $V, W \subseteq U$, then we write $V \prec W$ to mean that $v \prec w$ for every $v \in V$ and $w \in W$.

If $V$ is a set and $\ell$ a positive integer, then an $\ell$-partition of $V$ is a partition of $V$ into $\ell$ subsets. An $\ell$-partition of a finite set $V$ with $n$ points is said to be balanced if each set in the partition has either $\lceil n/\ell \rceil$ or $\lfloor n/\ell \rfloor$ points in it. An $\ell$-partition $\mathcal{V}$ of $V$ is called $\preceq$-orderly if, whenever $V, V'$ are distinct elements of $\mathcal{V}$, then either $V \prec V'$ or $V' \prec V$. If $\mathcal{V}$ is an orderly $\ell$-partition of $V$, then we say that $V \in \mathcal{V}$ is the $i$th set in $\mathcal{V}$ to mean that there are precisely $i-1$ sets $W \in \mathcal{V}$ with $W \prec V$.

If $k \geq 2$ and there is a $\preceq$-orderly $(k-1)$-partition $\mathcal{V}$ of $V$, and a $k$-subset $P$ of $V$, then we say that $P$ is of echelon $j$ in $\mathcal{V}$ if the $j$th and $(j+1)$th points of $P$ lie in the $j$th set of $\mathcal{V}$. That is, if $V = \{V_1 < \cdots < V_{k-1}\}$ and $P = \{p_1, \cdots, p_k\}$, then $P$ is of echelon $j$ in $\mathcal{V}$ if $p_j, p_{j+1} \in V_j$. It is a key fact that $P$ must be of echelon $j$ in $\mathcal{V}$ for at least one $j$.

**Lemma 2.1.** Let $k \geq 2$ and let $V$ be a set with $\preceq$-orderly $(k-1)$-partition $\mathcal{V}$. If $P$ is a $k$-subset of $V$, then $P$ is of echelon $j$ in $\mathcal{V}$ for some $j \in \{1, \ldots, k-1\}$.

**Proof.** By induction on $k$, with the $k = 2$ case trivial. For $k > 2$, suppose that $P = \{p_1, \cdots, p_k\}$ is not of echelon $(k-1)$ in $\mathcal{V} = \{V_1 < \cdots < V_{k-1}\}$. Then $p_{k-1} \not\in V_{k-1}$, and thus $P \setminus \{p_k\}$ is a $(k-1)$-subset of $V \setminus V_{k-1}$ with $\preceq$-orderly $(k-2)$-partition $\{V_1, \ldots, V_{k-2}\}$, and hence there is some $j \in \{1, \ldots, k-2\}$ such that $p_j, p_{j+1} \in V_j$ by induction. □

Of course, if $k \geq 4$, and $V$ is a set with $\preceq$-orderly $(k-1)$-partition $\mathcal{V}$, then a $k$-subset $P$ of $V$ may be of echelon $j$ in $\mathcal{V}$ for more than one $j$.

3. Reconstructibility

Let $\mathcal{P}$ be a $k$-pattern in our universe $U$ with order relation $\preceq$. Suppose we are given some $j \in \{1, \ldots, k-1\}$ and two points $u, v \in U$ with $u \prec v$. Any $P \in \mathcal{P}$ that has $u$ and $v$ respectively as its $j$th and $(j+1)$th points is called a reconstruction of $\mathcal{P}$ with the prescribed points as its $j$th and $(j+1)$th points.

If, for every $j \in \{1, \ldots, k-1\}$ and $u \prec v \in U$, there is at least one reconstruction of $\mathcal{P}$ with $u$ and $v$ as $j$th and $(j+1)$th points, then we say that $\mathcal{P}$ admits at least one reconstruction from $\preceq$-consecutive points. If there is at most one reconstruction for every $j \in \{1, \ldots, k-1\}$ and $u \prec v \in U$, we say that $\mathcal{P}$ admits at most one reconstruction from $\preceq$-consecutive points. If both of these hold, then we say that $\mathcal{P}$ admits a unique reconstruction from $\preceq$-consecutive points, or is uniquely reconstructible from $\preceq$-consecutive points.
4. A General Upper Bound on $S_P(n)$

**Theorem 4.1.** Let $k > 1$ and let $P$ be a $k$-pattern in $U$ that admits at most one reconstruction from $\leq$-consecutive points. If $n \in \mathbb{N}$ and $r$ is the remainder when $n$ is divided by $k - 1$, then $S_P(n) \leq (n - r)(n + r - k + 1)/(2k - 2)$.

**Proof.** Let $V$ be an $n$-subset of $U$, and let $V = \{V_1 < \cdots < V_{k-1}\}$ be a $\leq$-orderly, balanced $(k-1)$-partition of $V$. By Lemma 2.1, each $P \subseteq V$ with $P \in \mathcal{P}$ is of echelon $j$ in $V$ for some $j \in \{1, \ldots, k-1\}$. Each subset of echelon $j$ has its $j$th and $(j+1)$st elements in $V_j$, and since no two instances of $P$ may have the same $j$th and $(j + 1)$th points, this means that there are at most $\binom{|V_j|}{2}$ instances of $P$ that are of echelon $j$. Thus there are at most $\sum_{j=0}^{k-1} \binom{|V_j|}{2}$ instances of $P$ in $V$. If we write $n = q(k-1) + r$, then our partition, being balanced, has $|V_j| = q+1$ for $r$ values of $j$ and $|V_j| = q$ for $k-1-r$ values of $j$, so that

$$S_P(V) \leq r\binom{q+1}{2} + (k-1-r)\binom{q}{2} = (n-r)(n+r-k+1)/(2k-2).$$

The method of counting used in this proof provides a criterion for when this upper bound is achieved.

**Lemma 4.2.** Let $k > 1$ and let $P$ be a $k$-pattern in $U$ that admits at most one reconstruction from $\leq$-consecutive points. Let $n \in \mathbb{N}$ and let $r$ be the remainder when $n$ is divided by $k - 1$. Let $V$ be a $n$-subset of $U$ with $\leq$-orderly, balanced $(k-1)$-partition $V = \{V_1, \ldots, V_{k-1}\}$. Then $S_P(V) = (n - r)(n + r - k + 1)/(2k - 2)$ if and only if for each $j \in \{1, \ldots, k-1\}$ and every pair of distinct points $v, w \in V_j$, there exists a reconstruction $P$ of $P$ with $v$ and $w$ as $j$th and $(j + 1)$th points, and $P$ is a subset of $V$ that is not of echelon $i$ in $V$ for any $i \neq j$.

**Proof.** Examining the proof of the preceding theorem, we see that if any of the reconstructions mentioned in the statement of this proposition did not exist, or did not lie in $V$, then $\binom{|V_j|}{2}$ would be an overestimate of the number of instances of $P$ that are of echelon $j$. And if any instance were of two distinct echelons in $V$, then we would be counting it twice in $\sum_{j=0}^{k-1} \binom{|V_j|}{2}$, thus making our bound an overestimate. \hfill $\Box$

5. General Patterns in the Line

In this section, our universe $U$ is the line $\mathbb{R}$, our equivalence relation $\sim$ is direct similarity, and our pattern $P$ can be the equivalence class for any finite subset of $\mathbb{R}$. Our order relation is the usual order relation $\leq$ for $\mathbb{R}$. We note that any pattern $P$ is uniquely reconstructible from $\leq$-consecutive points, so Theorem 4.1 applies, amounting to the following.

**Theorem 5.1.** Let $k > 1$ and let $P$ be any $k$-pattern in $\mathbb{R}$. If $n \in \mathbb{N}$ and $r$ is the remainder when $n$ is divided by $k - 1$, then $S_P(n) \leq (n - r)(n + r - k + 1)/(2k - 2)$. 

This upper bound is tight in the sense that for each $k > 1$, when $\mathcal{P}$ is the
pattern of $k$-term arithmetic progressions, then $S_{\mathcal{P}}(n)$ equals the bound, as
we shall now show.

6. Arithmetic Progressions in the Line

Here the universe $U$ is the line $\mathbb{R}$ and our pattern is $\mathcal{A}_k$, the class of $k$-
term arithmetic progressions. Because $\mathcal{A}_k$ is invariant under reflection, the
results are the same whether we make $\sim$ direct similarity or similarity. This
enables an exact computation of $S_{\mathcal{A}_k}(n)$: we recall and prove Theorem 1.2.

Theorem 6.1. Let $k > 1$, and let $\mathcal{A}_k$ be the pattern of $k$-term arithmetic
progressions on $\mathbb{R}$. If $n \in \mathbb{N}$ and $r$ is the remainder when $n$ is divided by
$k - 1$, then $S_{\mathcal{A}_k}(n) = (n - r)(n + r - k + 1)/(2k - 2)$.

Proof. This is an immediate consequence of Theorem 5.1 above and Lemma
6.2 below. □

Lemma 6.2. If $k > 1$ and $n \in \mathbb{N}$ with $r$ the remainder when $n$ is divided by
$k - 1$, then the number of $k$-term arithmetic progressions in $\{0, 1, \ldots, n-1\}$
is $(n - r)(n + r - k + 1)/(2k - 2)$.

Proof. For each positive integer $s \leq n/(k - 1)$, our set $\{0, 1, \ldots, n - 1\}$
contains precisely $n - s(k - 1)$ arithmetic progressions having $k$ points with
distance $s$ between consecutive points. The result follows by adding these
quantities for $1 \leq s \leq n/(k - 1)$. □

Now that we know the precise value of $S_{\mathcal{A}_k}(n)$, we would like to completely
classify the $n$-subsets $V$ of $\mathbb{R}$ achieving $S_{\mathcal{A}_k}(V) = S_{\mathcal{A}_k}(n)$. We call such $n$-sets
optimal for $k$-term arithmetic progressions. Lemma 6.2 above shows that
$n$-term arithmetic progressions are always optimal for $k$-term arithmetic
progressions, but in many cases there are other optimal sets, and we now
classify them (up to similarity).

All sets are trivially optimal for 1- or 2-term arithmetic progressions, and
any $n$-set with $n < k$ is trivially optimal for $k$-term arithmetic progressions.
The barycenter of an arithmetic progression is the arithmetic mean of its
points. Two progressions are said to be concentric if they have the same
barycenter, or nearly concentric if their barycenters differ by 1.

Proposition 6.3. For $n \geq 3$, an $n$-subset of $\mathbb{R}$ is optimal for 3-term arith-
metic progressions if and only if, up to similarity, it is the union $E \cup O$
of two nonempty arithmetic progressions, where $E$ consists of consecutive
even integers and $O$ consists of consecutive odd integers and $E$ and $O$ are
concentric if $n$ is odd, or nearly concentric if $n$ is even.

Remark 6.4. Before we embark upon the proof, we pause to note that the
set $E \cup O$ described here is an arithmetic progression when $|E|$ and $|O|$ are
either equal, or one differs from the other by 1.
Proof of Proposition 6.3. In view of the value of \( S_{A_3}(n) \) from Theorem 6.1, we may use Lemma 4.2 as the criterion for optimality. Suppose that \( V \) is an \( n \)-subset of \( \mathbb{R} \) and \( \mathcal{V} = \{ V_1 < V_2 \} \) is a \( \leq \)-orderly, balanced 2-partition of \( V \). Because arithmetic progressions are uniquely reconstructible from consecutive points, and since no 3-term progression may be both of echelon 1 and of echelon 2 (since that would require it to have more than three points), Lemma 4.2 shows that \( V \) is optimal for 3-term arithmetic progressions if and only if for every \( j \in \{1, 2\} \) and every pair of distinct \( v, w \in V_j \), the reconstruction of \( A_3 \) with \( v \) and \( w \) as \( j \)th and \( (j+1) \)th points is contained in \( V \).

First, suppose \( V \) has the form \( E \cup O \) as described in the statement of this proposition. If \( |E| \) and \( |O| \) are equal or differ by 1, then \( V \) is itself an arithmetic progression, and hence optimal by Lemma 4.2, so we may assume that \( |E| \neq |O| \). We may then view \( V \) as \( X \cup Y \), where \( X \) is the larger of \( E \) and \( O \), while \( Y \) contains the smaller of \( E \) and \( O \), along with any points from the larger of \( E \) and \( O \) that lie between the points of the smaller of \( E \) and \( O \). Thus \( X \) is an arithmetic progression with distance 2 between its consecutive points, while \( Y \) is an arithmetic progression with distance 1 between its consecutive points. We let \( \mathcal{V} = \{ V_1 < V_2 \} \) be an \( \leq \)-orderly, balanced 2-partition of \( V \). Then \( \mathcal{V} \) induces \( \leq \)-orderly, balanced 2-partitions of \( X \) and \( Y \), which are \( X = \{ X \cap V_1, X \cap V_2 \} \) and \( Y = \{ Y \cap V_1, Y \cap V_2 \} \), respectively. Now suppose that \( v_1 < v_2 \) are points in \( V_1 \), and let \( v_3 \) be the point that makes \( v_1 < v_2 < v_3 \) a 3-term arithmetic progression. If \( v_1 \) and \( v_2 \) are both in \( X_1 \) (resp., \( Y_1 \)), then since \( X \) (resp., \( Y \)) is an arithmetic progression, and hence optimal for 3-term arithmetic progressions, we may apply Lemma 4.2 to conclude that \( v_3 \in X \) (resp., \( Y \)), and hence \( v_3 \in V \).

In the remaining case, where \( v_1 \in X_1 \) and \( v_2 \in Y_1 \), we note that \( v_3 \) cannot be larger than the largest point in \( V \) and has the same parity as \( v_1 \), hence \( v_3 \in X \subseteq V \). By the same argument, if \( v_2 < v_3 \) are two points in \( V_2 \), then the point \( v_1 \) that makes \( v_1 < v_2 < v_3 \) a 3-point arithmetic progression is always in \( V \). Thus our set \( V \) satisfies the sufficient criteria of Lemma 4.2 for optimality.

Now suppose that \( V \) is an \( n \)-subset of \( \mathbb{R} \) that is optimal for 3-term arithmetic progressions. We want to show that \( V \) has the form described in the statement of this proposition. First we deal with the case where \( n \) is odd, and without loss of generality, we may apply a similarity transformation so that the middle point of \( V \) is 0, and the next point in \( V \) is 1. We let \( \mathcal{V} = \{ V_1 < V_2 \} \) be an \( \leq \)-orderly, balanced 2-partition of \( V \) with the negative points in \( V_1 \), and the nonnegative points in \( V_2 \). By Lemma 4.2, for every positive \( v_2 \in V_2 \), the 3-term arithmetic progression \(-v_2 < v < v_2 \) must lie in \( V \), so the point \(-v_2 \) must lie in \( V_1 \). Thus \( V \) has reflection symmetry about 0.

If \( u < v < w \) are consecutive points in \( V_2 \), then by Lemma 4.2, the 3-term progression \( 2v - w < v < w \) must lie in \( V \), and so \( 2v - w \leq u \), that
is, \( w - v \geq v - u \). So the spacings between consecutive points of \( V_2 \) are nondecreasing as we proceed to the right.

If \( 1 \leq u < v \) are consecutive points in \( V_2 \), then by Lemma 4.2, the 3-term progression \( 2 - v < 1 < v \) lies in \( V \), so that \( v - 2 \) lies in \( V \) by reflection symmetry, and so \( u \geq v - 2 \). So the spacing between any two consecutive points of \( V_2 \) is no greater than 2.

We also claim that all the points of \( V_2 \) are integers: we proceed by induction. The first two elements of \( V_2 \) are 0 and 1, so suppose that \( v \in V_2 \) with \( v > 1 \), and we know that all elements of \( V_2 \) less than \( v \) are integers. By Lemma 4.2, the arithmetic progression \( 2 - v < 1 < v \) lies in \( V \), so that \( |v - 2| \) lies in \( V \) by reflection symmetry, so \( |v - 2| \) must be an element of \( V_2 \) less than \( v \), hence an integer. So \( v \) is an integer.

Thus spacings between any consecutive points of \( V_2 \) is always 1 or 2, and these spacings are nondecreasing as we proceed to the right. Since \( V \) has reflection symmetry about 0, we know that \( V \) has the form described in this proposition. This completes the proof when \( n \) is odd.

Now suppose that \( n \) is even and \( V \) is an \( n \)-subset of \( \mathbb{R} \) optimal for 3-term arithmetic progressions. Then the technical Lemma 6.5 whose proof is delayed to the end of this section, shows that an \((n - 1)\)-set \( V' \) obtained from \( V \) by removing either the leftmost or rightmost point is optimal for 3-term arithmetic progression. By applying similarity transformations, we may assume that the rightmost point \( r \) is removed and that \( V' \) has the form described in the statement of this proposition, with 0 being the middle point of \( V' \). Then \( V' \) contains the point 1 since the set of odd points in \( V' \) is nonempty and centered at 0. Let \( r' \) be the rightmost point of \( V' \). A \( \leq \)-orderly, balanced 2-partition \( V = \{ V_1 < V_2 \} \) of \( V \) places all the nonpositive points into \( V_1 \) and all the positive points into \( V_2 \).

In view of Lemma 4.2 the 3-term progression \( 2 - r < 1 < r \) lies in \( V \), so \( 2 - r \in V' \). Thus \( r \) must be an integer, and by reflection symmetry in \( V' \), we must have \( r - 2 \leq r' \), that is \( r \in \{ r' + 1, r' + 2 \} \). It is immediate that \( V' \cup \{ r' + 2 \} \) is of the form described in the proposition, so we are done if \( r = r' + 2 \). If \( r = r' + 1 \), then by Lemma 4.2, the 3-term progression \( r' - 1 < r' < r \) is contained in \( V \), so then \( V' \) must consist of all consecutive integers from \(-r'\) to \( r' \), and then it is clear that \( V = V' \cup \{ r \} \) is of the form described in the proposition.

**Proposition 6.5.** For \( n \geq k \geq 4 \), an \( n \)-subset \( V \) of \( \mathbb{R} \) is optimal for \( k \)-term arithmetic progressions if and only if, up to similarity, \( V \) is an arithmetic progression or \( k - 1 \) divides \( n \) and \( V \) is an arithmetic progression with the second point deleted.

**Proof.** We first analyze all optimal sets \( V \) for four-term arithmetic progressions. Suppose that \( V \) is such an \( n \)-subset of \( \mathbb{R} \) and \( n \geq 4 \). Let \( V = \{ V_1 < V_2 < V_3 \} \) with \( |V_2| \geq 2 \) be an \( \leq \)-orderly, balanced 3-partition of \( V \). Lemma 6.7 below shows that if we remove \( V_3 \) from \( V \), then we are left with an optimal set for three-term arithmetic progressions. Since \( V_1 \cup V_2 \) is
ON THE NUMBER OF SIMILAR INSTANCES OF A PATTERN IN A FINITE SET

optimal for three-term arithmetic progressions, Proposition 6.3 tells us that without loss of generality, we may take $V_1 \cup V_2$ to be the union of a set $E$ of consecutive even integers and a set $O$ of consecutive odd integers, which either have a common center $c$ (if $|E \cup O|$ is odd) or $E$ and $O$ have distinct centers $d$ and $d + 1$ in some order.

We now prove that the distance between the leftmost two points of $V_2$ is 1. Since $V_1 \cup V_2 = E \cup O$ and $E$ and $O$ are nonempty, $V_1 \cup V_2$ contains a set of three consecutive integers. The integers are $c - 1, c, c + 1$ if $|E \cup O|$ is odd or one of $d - 1, d, d + 1$ or $d, d + 1, d + 2$ if $|E \cup O|$ is even. Then the three consecutive points are either the two rightmost of $V_1$ and the one leftmost of $V_2$ or the one rightmost of $V_1$ and the two leftmost of $V_2$. The first case implies the second one, since if the rightmost two points of $V_1$ have a spacing 1, then the leftmost two points of $V_2$ will complete a four-term arithmetic progression with the former two points. So we know that distance between the leftmost two points of $V_2$ is 1.

By Lemma 6.7, the union of $V_2$ and $V_3$ is optimal for three-term arithmetic progressions. Now since the leftmost two points of $V_2 \cup V_3$ are at distance 1, it follows from Proposition 6.3 that $V_2 \cup V_3$ is either an arithmetic progression, or an arithmetic progression with the second-to-last point removed. Similarly, $V_1 \cup V_2$ is either an arithmetic progression or an arithmetic progression with the second point removed. By Lemma 6.8, $V$ is either an arithmetic progression or else an arithmetic progression with either the second point or the second-to-last point (but not both) removed, and these latter two cases can only occur if $3 \mid n$. Conversely, Lemma 6.8 shows that sets $V$ with these forms really are optimal for four-term arithmetic progressions.

Now we analyze all optimal sets for $k$-term arithmetic progressions where $k \geq 5$. For $n \geq k$, let $\mathcal{V} = \{V_1 < V_2 < \ldots < V_{k-1}\}$ be an $\leq$-orderly, balanced $(k - 1)$-partition of an $n$-subset $V$ of $\mathbb{R}$, where we insist $|V_2| \geq 2$. Then $V' = V_1 \cup \ldots \cup V_{k-2}$ and $V'' = V_2 \cup \ldots \cup V_{k-1}$ are optimal for $(k - 1)$-term arithmetic progressions by Lemma 6.7. Hence, without loss of generality $V'$ is either an arithmetic progression of consecutive integers, or what one obtains by removing either the second or second-to-last point (but not both) from such an arithmetic progression. If the rightmost two points of $V_{k-2}$ had spacing 2, then $V''$ would begin with two points at spacing 1, then later would have a spacing of 2 between consecutive points, neither of which would be endpoints, and this would contradict the optimality of $V''$ for $(k - 1)$-term arithmetic progressions. So any spacing of 2 between consecutive points in $V'$ must be between the leftmost two points. And by a similar argument with $V''$, we see that all consecutive points of $V''$ have spacing 1, except possibly the two rightmost, which can have spacing 2. Thus all consecutive points of $V$ must have spacing 1, except for possible spacings of 2 between the leftmost pair and the rightmost pair. By Lemma 6.8, $V$ is either an arithmetic progression or else an arithmetic progression with either the second point or the second-to-last point (but not both) removed, and these latter two cases
can only occur if \( k - 1 \mid n \). Conversely, Lemma 6.8 shows that sets \( V \) with these forms really are optimal for \( k \)-term arithmetic progressions.

We close with the technical lemmata use to prove the two propositions above.

**Lemma 6.6.** Let \( V \) be a \( n \)-subset of \( \mathbb{R} \) that is optimal for 3-term arithmetic progressions. Then either the \((n - 1)\)-set obtained by removing the leftmost point of \( V \) or the \((n - 1)\)-set obtained by removing the rightmost point of \( V \) is optimal for 3-term arithmetic progressions.

**Proof.** Let \( V = \{V_1 < V_2\} \) be a \( \leq \)-orderly, balanced 2-partition of \( V \), with \(|V_1| \geq |V_2|\). By Lemma 4.2, we see that any pair of points in \( V_1 \) (resp., \( V_2 \)) are the leftmost (resp., rightmost) two points of a 3-term progression in \( V \).

Suppose that the leftmost point \( \ell \) of \( V \) is such that there is no 3-term progression \( \ell < u < v \) with \( u, v \in V_2 \). Then let \( V' = V \setminus \{\ell\} \) with \( \leq \)-orderly, balanced 2-partition \( V' = \{V_1 \setminus \{\ell\}, V_2\} \), and we see that any pair of points in \( V_1 \setminus \{\ell\} \) (resp., \( V_2 \)) are the leftmost (resp., rightmost) two points of a 3-term progression in \( V' \), so \( V' \) is optimal by Lemma 4.2.

Now suppose that the leftmost point \( \ell \) of \( V \) is such that \( \ell < m < r \) is a 3-term progression with \( m, r \in V_2 \). Then \( m \) and \( r \) must respectively be the leftmost and rightmost points of \( V_2 \), for otherwise Lemma 4.2 would dictate that the leftmost and rightmost points of \( V_2 \) would be a part of a 3-term progression in \( V \), and that progression would need to involve a point to the left of \( \ell \), which is absurd.

We set \( V' = V \setminus \{r\} \), and claim that it is optimal for 3-term arithmetic progressions. For any \( u < v \in V_1 \), we claim that the point \( w \) that makes \( u < v < w \) a 3-term arithmetic progression is in \( V' \). By Lemma 4.2, \( w \in V \), and \( w = 2v - u \), with \( v < m \) and \( u \geq \ell \), so that \( w < 2m - \ell = r \) (since \( \ell < m < r \) is an arithmetic progression).

If \( n \) is even, then set \( V' = \{V_1, V_2 \setminus \{r\}\} \), which is a \( \leq \)-orderly, balanced 2-partition of \( V' \), and we see that any pair of points in \( V_1 \) (resp., \( V_2 \setminus \{r\}\)) are the leftmost (resp., rightmost) two points of a 3-term progression in \( V' \), so \( V' \) is optimal by Lemma 4.2.

If \( n \) is odd, let \( y \) be the rightmost point of \( V_1 \), and set \( W = \{V_1 \setminus \{y\}\} \) with \( W_1 = V_1 \setminus \{y\} \) and \( W_2 = V_2 \cup \{y\} \), which is a \( \leq \)-orderly, balanced 2-partition of \( V \) with \( |W_2| > |W_1| \). Then set \( W' = \{W_1, W_2 \setminus \{r\}\} \), which is a \( \leq \)-orderly, balanced 2-partition of \( V' = V \setminus \{r\} \). By Lemma 4.2, any pair \( v < w \) of points in \( W_2 \) form the rightmost two points of a 3-term arithmetic progression in \( V \), so any pair \( v < w \) of points in \( W_2 \setminus \{r\} \) for the rightmost two points of a 3-term arithmetic progression in \( V' \). By the paragraph before the previous one, we know that every pair of points in \( W_1 = V_1 \setminus \{y\} \) form the leftmost two points of a 3-term arithmetic progression in \( V' \). Thus Lemma 4.2 shows that \( V' \) is optimal for 3-term arithmetic progressions. \( \□ \)
Lemma 6.7. If \( V \) is a finite subset of \( \mathbb{R} \) that is optimal for \( k \)-term arithmetic progressions, and \( \mathcal{V} = \{V_1 < \cdots < V_{k-1}\} \) is a \( \leq \)-orderly, balanced \((k-1)\)-partition of \( V \), then both \( \bigcup_{j=1}^{k-2}V_j \) and \( \bigcup_{j=2}^{k-1}V_j \) are optimal for \((k-1)\)-term arithmetic progressions.

Proof. Let \( V' = \bigcup_{j=1}^{k-2} \) with \( \leq \)-orderly, balanced \((k-2)\)-partition \( \mathcal{V}' = \{V_1 < \cdots < V_{k-2}\} \). By Lemma 4.2 for a pair of points \( p \) and \( q \) in the subset \( V_j \) of \( \mathcal{V}' \), there is a \( k \)-term arithmetic progression \( P \) in \( \mathcal{V} \) that contains these points and is only of echelon \( i \) in \( \mathcal{V} \) for \( i = j \). Note that \( P' = P \setminus (P \cap V_{k-1}) \) is an arithmetic progression of some length that of echelon \( j \) in \( \mathcal{V}' \). If \( P' \) has fewer than \( k-1 \) points, then it must be true that more than one point of \( P \) is in \( V_{k-1} \), which implies that \( P \) is both of echelon \( k-1 \) and echelon \( j \) in \( \mathcal{V} \), contradicting Lemma 4.2. Hence \( P' \) is an arithmetic progression of at least \( k-1 \) terms contained in \( \mathcal{V}' \). By Lemma 4.2, \( V' \) is optimal for \( k-1 \)-term arithmetic progressions. By the same reasoning, \( \bigcup_{j=2}^{k-1}V_j \) is also optimal for \((k-1)\)-term arithmetic progressions. \( \square \)

Lemma 6.8. Let \( n \geq k \geq 4 \). Let \( V \) be an \( n \)-subset of \( \mathbb{R} \) that is either an arithmetic progression, or else an arithmetic progression with the second point \( p \) or the second-to-last point \( q \) (or both) removed. Then \( V \) is optimal for \( k \)-term arithmetic progressions if and only if (i) \( V \) is an arithmetic progression or (ii) \( k-1 \) divides \( n \) and \( V \) is an arithmetic progression with only one of \( p \) or \( q \) removed.

Proof. We can assume, without loss of generality, that \( V \subseteq \mathbb{Z} \), that the first element of \( V \) is 0, and the smallest spacing between any two consecutive integers is 1. Write \( n = (k-1)q + r \) with \( 0 \leq r < k-1 \).

We already know from Theorem 6.1 and Lemma 6.2 that if \( V \) is an arithmetic progression, then it is optimal for \( k \)-term arithmetic progressions.

Now we assume that \( V \) is obtained from an arithmetic progression by removing only the second point, so that \( V = \{0, 2, 3, \ldots, n\} \). Assume \( V \) is optimal for \( k \)-term arithmetic progressions.

If \( 0 < r < k-1 \), let \( \mathcal{V} = \{V_1 < \cdots < V_{k-1}\} \) be a \( \leq \)-orderly, balanced \((k-1)\)-partition of \( V \), such that \( V_1 = \{0, 2, 3, \ldots, q+1\} \). If we consider the \( k \)-term arithmetic progression beginning with 0 and \( q+1 \), its rightmost point is \( (k-1)(q+1) = (k-1)q + (k-1) > (k-1)q + r = n \), which must be contained in \( V \) by Lemma 4.2 which is absurd.

On the other hand, if \( r = 0 \), then it is not difficult to use Lemma 4.2 to show that \( V = \{0, 2, 3, \ldots, (k-1)q\} \) is in fact optimal for \( k \)-term arithmetic progressions. By symmetry, we have also covered the cases where \( V \) is obtained from an arithmetic progression by only removing the second-to-last point.

Now we assume that \( v \) is obtained from an arithmetic progression by removing both the second point and the second-to-last point, so that \( V = \{0, 2, 3, \ldots, n-2, n-1, n+1\} \). Assume \( V \) is optimal for \( k \)-term arithmetic progressions. Recall that we are writing \( n = (k-1)q + r \) with \( 0 \leq r < k-1 \).
If \( r = 0 \), let \( V = \{ V_1 < \cdots < V_{k-1} \} \) be a \( \leq \)-orderly, balanced \((k-1)\)-partition of \( V \), in which \( V_1 = \{0, 2, 3, \ldots, q\} \). If we consider the \( k \)-term arithmetic progression beginning with 0 and \( q \), then its rightmost point is \((k - 1)q = n\), so Lemma 4.2 says that \( n \) must lie in \( V \), which is absurd.

If \( 0 < r < k - 1 \), let \( V = \{ V_1 < \cdots < V_{k-1} \} \) be a \( \leq \)-orderly, balanced \((k - 1)\)-partition of \( V \), such that \( V_1 = \{0, 2, 3, \ldots, q\} \) and \( V_2 = \{q + 1, q + 2, \ldots, 2q + 1\} \). If we consider the \( k \)-term arithmetic progression whose second and third points are \( q + 1 \) and \( 2q + 1 \), then the first point must be 1, so Lemma 4.2 says that 1 must lie in \( V \), which is absurd. \( \square \)

7. Commensurable Patterns in the Line

Here the universe \( U \) is the line \( \mathbb{R} \), and our equivalence relation \( \sim \) is direct similarity. A commensurable pattern \( \mathcal{P} \) in \( \mathbb{R} \) is one such that for \( P \in \mathcal{P} \), all the distances between pairs in \( P \) are commensurable, that is, are related by rational ratios. Equivalently, \( \mathcal{P} \) is commensurable if it contains some \( P \subseteq \mathbb{Z} \). Or yet again, \( \mathcal{P} \) is commensurable if its instances are subsets of arithmetic progressions. Indeed if \( P \in \mathcal{P} \), there is a unique arithmetic progression \( A \) of minimum cardinality such that \( P \subseteq A \). We call this \( A \) the enveloping arithmetic progression for \( P \), and of course the set of all enveloping arithmetic progressions of elements of \( \mathcal{P} \) is itself a pattern, called the enveloping pattern for \( \mathcal{P} \). For a positive \( \ell \), we let \( A_\ell \) be the pattern consisting of \( \ell \)-term arithmetic progressions.

**Theorem 7.1.** Let \( \mathcal{P} \) be a commensurable \( k \)-pattern on \( \mathbb{R} \), and suppose that \( A_\ell \) is the enveloping pattern for \( \mathcal{P} \). If \( n \in \mathbb{N} \) and \( r \) and \( s \) are respectively the remainders when \( n \) is divided by \( k - 1 \) and \( \ell - 1 \), then

\[
\frac{(n - s)(n + s - \ell + 1)}{2(\ell - 2)} \leq S_\mathcal{P}(n) \leq \frac{(n - r)(n + r - k + 1)}{2(k - 2)}.
\]

*Proof.* The right side of the inequality follows directly from Lemma 4.2, which gives an upper bound on the maximum number of instances of a \( k \)-pattern \( \mathcal{P} \) in an \( n \)-subset of \( \mathbb{R} \).

Now let \( A = \{p, \ldots, q\} \) and \( A' = \{p', \ldots, q'\} \) be distinct \( l \)-term arithmetic progressions in \( R \). Then they are enveloping arithmetic progressions for distinct instances of the commensurable pattern \( \mathcal{P} \): \( P = \{p, \ldots, q\} \) and \( P' = \{p', \ldots, q'\} \). We see that in a given set, there are at least as many instances of \( \mathcal{P} \) as there are \( l \)-term arithmetic progressions, so that \( S_{A_\ell}(n) = (n - s)(n + s - \ell + 1)/(2\ell - 2) \leq S_\mathcal{P}(n) \). \( \square \)

We now explore the commensurable pattern \( \mathcal{P} \) containing \( \{0, 1, 3\} \). The enveloping pattern is \( A_4 \). We know from Theorem 7.1 that \( S_{A_4}(n) \leq S_\mathcal{P}(n) \leq S_{A_3}(n) \). Note from Theorem 6.4 that \( \lim_{n \to \infty} S_{A_4}(n)/n^2 = 1/6 \) and \( \lim_{n \to \infty} S_{A_3}(n)/n^2 = 1/4 \).

We now construct a family of sets \( V_n \) where each \( V_n \) is a set with \( n \) points containing \( S_\mathcal{P}(V_n) = (3n^2 - 8n)/16 \) directly similar copies of \( \{0, 1, 3\} \). This makes \( \lim_{n \to \infty} S_\mathcal{P}(V_n)/n^2 = 3/16 \), which is strictly between the two limits.
ON THE NUMBER OF SIMILAR INSTANCES OF A PATTERN IN A FINITE SET

we computed in the previous paragraph using the lower and upper bounds on \(S_P(V_n)\) furnished by Theorem 7.1. We now describe our construction \(V_n\), and then prove our claim that \(S_P(V_n) = (3n^2 - 8n)/16\).

**Construction 7.2.** If \(n = 96k\) for any positive integer \(k\), we let \(V_n\) be the union of the following four sets:

\[
M_0 = 6\mathbb{Z} \cap [0, 108k] \\
M_1 = (1 + 6\mathbb{Z}) \cap [72k + 1, 144k - 5] \\
M_3 = (3 + 6\mathbb{Z}) \cap [3, 324k - 9] \\
M_5 = (5 + 6\mathbb{Z}) \cap [72k - 1, 144k - 7],
\]

or equivalently

\[
M_0 = \{6a : 0 \leq a \leq 18k\}, \\
M_1 = \{6a + 1 : 12k \leq a \leq 24k - 1\}, \\
M_3 = \{6a + 3 : 0 \leq a \leq 54k - 2\}, \\
M_5 = \{6a - 1 : 12k \leq a \leq 24k - 1\}.
\]

**Proposition 7.3.** If \(P\) is the pattern containing \(\{0, 1, 3\}\), and \(V_n\) is the \(n\)-point set \(V_n\) described in Construction 7.2, then \(S_P(V_n) = (3n^2 - 8n)/16\).

**Proof.** To compute \(S_P(V_n)\) we first note that by choosing the congruence classes modulo 6 of the first two points of any \(P \in \mathcal{P}\), the class of the third point of \(P\) is fixed. For example, if we choose the first two points congruent to 5 (mod 6), then the third point must also be of the same class. Hence the \(4 \times 4\) cases for the respective congruence classes of the three points are \((0, 0, 0), (0, 1, 3), (0, 3, 3), (0, 5, 3), (1, 0, 4), (1, 1, 1), (1, 3, 1), (1, 5, 1), (3, 0, 0), (3, 1, 3), (3, 3, 3), (3, 5, 3), (5, 0, 2), (5, 1, 5), (5, 3, 5), \) and \((5, 5, 5)\). Note that we omit \((1, 0, 4)\) and \((5, 0, 2)\) as none of the points in \(V_n\) are congruent to 4 or to 2 (mod 6).

To compute the number of instances \(\{p_1 < p_2 < p_3\}\) of \(P\) in \(V_n\) with \((p_1, p_2, p_3) \equiv (a, b, c) \pmod{6}\) for a given triple \((a, b, c)\), one counts lattice points in regions of \(\mathbb{Z}^2\) defined by a system of inequalities. For example, if \(n = 96k\) and \((a, b, c) = (0, 5, 3)\), we could represent each instance \(\{p_1 < p_2 < p_3\}\) of \(P\) with

\[
p_1 = 6x \\
p_2 = 6y - 1 \\
p_3 = 3p_2 - 2p_1 = 6(3x - 2y - 1) + 3,
\]

subject to the inequalities

\[
x < y \\
0 \leq x \leq 18k \\
12k \leq y \leq 24k - 1 \\
0 \leq 3y - 2x - 1 \leq 54k - 2.
\]
The first inequality makes sure that $p_1 < p_2$, and the remaining three inequalities make sure that $p_1$, $p_2$, and $p_3$ are in the ranges prescribed for their respective congruence classes, as described in Construction \[122\].

In this way we calculate the following number of instances \(\{p_1 < p_2 < p_3\}\) of \(\mathcal{P}\) in \(V_n\):

\[
\begin{align*}
54k^2 - 3k & \text{ instances with } (p_1, p_2, p_3) \equiv (0, 0, 0) \pmod{6}, \\
171k^2 + 6k & \text{ instances with } (p_1, p_2, p_3) \equiv (0, 1, 3) \pmod{6}, \\
270k^2 + 9k & \text{ instances with } (p_1, p_2, p_3) \equiv (0, 3, 3) \pmod{6}, \\
171k^2 + 3k - 1 & \text{ instances with } (p_1, p_2, p_3) \equiv (0, 5, 3) \pmod{6}, \\
24k^2 - 6k & \text{ instances with } (p_1, p_2, p_3) \equiv (1, 1, 1) \pmod{6}, \\
24k^2 + 2k & \text{ instances with } (p_1, p_2, p_3) \equiv (1, 3, 1) \pmod{6}, \\
24k^2 - 2k & \text{ instances with } (p_1, p_2, p_3) \equiv (1, 5, 1) \pmod{6}, \\
54k^2 + 3k & \text{ instances with } (p_1, p_2, p_3) \equiv (3, 0, 0) \pmod{6}, \\
189k^2 - 6k & \text{ instances with } (p_1, p_2, p_3) \equiv (3, 1, 3) \pmod{6}, \\
486k^2 - 45k + 1 & \text{ instances with } (p_1, p_2, p_3) \equiv (3, 3, 3) \pmod{6}, \\
189k^2 - 3k & \text{ instances with } (p_1, p_2, p_3) \equiv (3, 5, 3) \pmod{6}, \\
24k^2 + 2k & \text{ instances with } (p_1, p_2, p_3) \equiv (5, 1, 5) \pmod{6}, \\
24k^2 - 2k & \text{ instances with } (p_1, p_2, p_3) \equiv (5, 3, 5) \pmod{6}, \ \text{ and} \\
24k^2 - 6k & \text{ instances with } (p_1, p_2, p_3) \equiv (5, 5, 5) \pmod{6}.
\end{align*}
\]

If we add up all the counts, we get $1728k^2 - 48k$, or equivalently $(3n^2 - 8n)/16$ for the the total number of instances of $\mathcal{P}$ in $V_n$. \qed

8. Equilateral Triangles in the Plane

In this section we explore $S_\mathcal{E}(n)$, where $U$ is the Euclidean plane (identified with $\mathbb{C}$), $V$ is a finite subset, and $\mathcal{E}$ is the pattern of vertices of an equilateral triangle. Our order relation on $\mathbb{C}$ is lexicographic ordering: $y < z$ means that either (i) $\Re(y) < \Re(z)$ or (ii) $\Re(y) = \Re(z)$ and $\Im(y) < \Im(z)$. This order relation respects addition, so if $y < y'$ and $z \leq z'$, then $y + z < y' + z'$.

If $u$ and $v$ are distinct points in $\mathbb{C}$, they are vertices of only two equilateral triangles, and if we let $w$ and $w'$ be the respective third vertices of these two triangles, then $u$, $w$, $v$, and $w'$ are vertices of a parallelogram with $u + v = w + w'$. Then the compatibility of our order relation with addition shows that $\mathcal{E}$ admits at most one reconstruction from any pair of consecutive points, for neither $u, v < w, w'$ nor $u, v > w, w'$ is consistent with $u + v = w + w'$.

Since $\mathcal{E}$ admits at most one reconstruction from any pair of consecutive points, Theorem \[14\] immediately tells us that $S_\mathcal{E}(n) \leq n(n-2)/4$ if $n$ is even and $S_\mathcal{E}(n) \leq (n-1)^2/4$ if $n$ is odd, that is, $S_\mathcal{E}(n) \leq [(n-1)^2/4]$ in
any case. This observation recovers upper bound on $S_{\xi}(n)$ of [2] Theorem 1] (given here in [1]), which was previously obtained using a much more sophisticated geometric proof.

We shall get an improved upper bound on $S_{\xi}(n)$ by noting that sometimes $\xi$ admits no reconstruction from a pair of consecutive points. To see this, we define the usual principal value of the argument function $\text{Arg}: \mathbb{C}^* \rightarrow (-\pi, \pi]$.

**Lemma 8.1.** Let $u, v$ be distinct points in $\mathbb{C}$. Then these are the first two vertices (under $\preceq$) of precisely one equilateral triangle and last two vertices of precisely one equilateral triangle if $\text{Arg}(v-u) \in (-5\pi/6, -\pi/6] \cup [\pi/6, 5\pi/6]$. Otherwise, they are not the first two nor the last two vertices of any equilateral triangle.

**Proof.** Let $\zeta_6 = e^{\pi i/3}$. Then $u$ and $v$ are vertices of only two equilateral triangles, with the third point being either (i) $w = \zeta_6 u + \zeta_6 v$ or (ii) $w' = \zeta_6 u + \zeta_6 v$.

Note that $u, v \prec w$ requires $\text{Re}(w) \geq \text{Re}(u)$ (with $\text{Im}(w) > \text{Re}(u)$ in case of equality) and $\text{Re}(w) \geq \text{Re}(v)$ (with $\text{Im}(w) > \text{Im}(v)$ in case of equality). Since $\zeta_6 + \zeta_6 = 1$, this is equivalent to $\text{Re}(\zeta_6 (v-u)) \geq 0$ and $\text{Re}(-\zeta_6 (v-u)) \geq 0$, with the corresponding imaginary parts strictly positive in cases of equality. These conditions are fulfilled if and only if $\text{Arg}(v-u) \in (\pi/6, 5\pi/6]$. One can similarly show that $u, v \succ w'$ under the same conditions, and that $u, v \prec w'$ if and only if $\text{Arg}(v-u) \in (-5\pi/6, -\pi/6]$, with $u, v \succ w$ under the same conditions. \(\Box\)

Let $A = (-5\pi/6, -\pi/6] \cup (\pi/6, 5\pi/6]$, the set of values of $\text{Arg}(v-u)$ that admit a reconstruction of $\xi$ with $u$ and $v$ as consecutive points. To get a good upper bound for $S_{\xi}(V)$ in an $n$-set $V$, we rotate $V$ in such a way that the first $\lfloor n/2 \rfloor$ points of $V$ (with respect to ordering $\preceq$) have many pairs $(u, v)$ of points with $\text{Arg}(v-u) \notin A$, and the last $\lceil n/2 \rceil$ points of $V$ have many pairs $(u, v)$ of points with $\text{Arg}(u-v) \notin A$. This requires some care, for as we rotate, points change from being among the first $\lfloor n/2 \rfloor$ to the last $\lceil n/2 \rceil$.

We define a direction $\zeta$ to be complex number of unit modulus, so the set of directions is the complex unit circle. A directed line with direction $\zeta$ is line with parameterization $t \mapsto \zeta t + \eta$ for some $\eta \in \mathbb{C}$, where the parameter $t$ ranges over $\mathbb{R}$. Every directed line with direction $\zeta$ may be uniquely written as $L(t) = \zeta t + s(-i\zeta)$, where $s$ is a scalar uniquely determined by the line $L$; we call this the canonical form of the directed line $L$. Note that $|s|$ is the distance from $L$ to the origin, and the sign of $s$ tells us which side of $L$ the origin lies on. We may thus identify the set of directed lines with the Cartesian product of the unit circle (for $\zeta$) and the real line (for $s$), and give the directed lines the topology of this product space.

For each direction $\zeta$, we define a total ordering relation $\preceq_\zeta$ on $\mathbb{C}$ with $y \preceq_\zeta z$ if and only if $i\zeta y < i\zeta z$. Note that $\preceq_\zeta$ is the same as $\preceq$, and for any direction $\zeta$, we have $y \succeq_{-\zeta} z$ if and only if $y \preceq z$. If $L(t)$ is a directed line
with direction $\zeta$, then we say $z \in \mathbb{C}$ lies to the left (resp., to the right) of $L(t)$ if $z \prec_{\zeta} L(t)$ for all $t \in \mathbb{R}$ (resp., $z \succ_{\zeta} L(t)$ for all $t \in \mathbb{R}$). If $y, z \in \mathbb{C}$ are on the line $L(t)$, say $y = L(t_y)$ and $z = L(t_z)$, then we say that $y$ is below (resp., above) $z$ on $L$ if $t_y < t_z$ (resp., $t_y > t_z$). Note that these concepts of left and right, below and above, all correspond to the usual notions when $L(t) = \text{it}$ is the positively directed $y$-axis, or indeed if $L$ is any translate thereof.

If $\zeta$ is a direction and $V = \{v_1 \prec_{\zeta} \cdots \prec_{\zeta} v_n\}$ is a finite set, then the $\zeta$-median of $V$ is $v_{(n+1)/2}$ if $n$ is odd, and is $\frac{\sum_{i=1}^{n/2} v_{2i-1} + \sum_{i=1}^{n/2} v_{2i}}{n}$ if $n$ is even. If $\zeta$ is a direction and $V$ is a finite set of points, then the $\zeta$-halving line of $V$ is the directed line $L_{V,\zeta}$ with direction $\zeta$ that passes through the $\zeta$-median of $V$. The line is called a halving line because $\lfloor n/2 \rfloor$ of the points of $V$ are (with respect to $\preceq_{\zeta}$) less than the median, and so lie to the left of the line or below the $\zeta$-median of $V$ on the directed line, while $\lceil n/2 \rceil$ of the points of $V$ are greater than the median, and so lie to the right of the line or above the median on the directed line. Our definition of halving line makes a unique halving line for $V$ in each direction, and is identical to the definition used by Erickson, Hurtado, and Morin [21]. If $c$ is the $\zeta$-median of $V$, we call any point $z$ with $z \preceq_{\zeta} c$ formally to the left of $L_{V,\zeta}$ and any point $z$ with $z \succ_{\zeta} c$ formally to the right of $L_{V,\zeta}$. Thus $\lfloor n/2 \rfloor$ of the points of $V$ lie formally to the left of $L_{V,\zeta}$ and $\lceil n/2 \rceil$ of the points of $V$ lie formally to the right of $L_{V,\zeta}$.

With the topology for the space of directed lines introduced above, we claim that $L_{V,\zeta}$ is continuous in $\zeta$. Indeed, $L_{V,\zeta}$ evolves by rotating about a particular median for a segment of values of $\zeta$, changing from median $v$ to $w$ precisely for the value of $\zeta$ that makes $v$ and $w$ both lie on the line $L_{V,\zeta}$. Note that the $\zeta$-median and the $(-\zeta)$-median of $V$ are the same, so that $L_{V,\zeta}$ and $L_{V,-\zeta}$ are the same line, but with opposite directions. These considerations lead to the following useful observation, first proved in this precise form in [21] Lemma 3 (see also [22] Lemma 2 and [14] Lemma 2).

**Lemma 8.2.** Let $V$ be a finite subset of $\mathbb{C}$. Then there is some direction $\zeta$ such that the halving lines of $V$ in directions $\zeta$, $e^{2\pi i/3} \zeta$, and $e^{4\pi i/3} \zeta$ are concurrent.

**Proof.** Let $\omega = e^{2\pi i/3}$. If the point $L_{V,\omega} \cap L_{V,\omega^2}$ is on or to the left (resp., right) of the directed line $L_{V,1}$, then the same point $L_{V,-\omega} \cap L_{V,-\omega^2}$ must be on or to the right (resp., left) of the same but oppositely directed line $L_{V,-1}$. As $\zeta$ traverses the unit circle, the lines $L_{V,\zeta}$, $L_{V,\omega \zeta}$, and $L_{V,\omega^2 \zeta}$ evolve continuously, and so the intersection point $L_{V,\omega \zeta} \cap L_{V,\omega^2 \zeta}$ also evolves continuously. Thus there must be some $\zeta$ such that $L_{V,\omega \zeta} \cap L_{V,\omega^2 \zeta}$ crosses $L_{V,\zeta}$. $\square$

Now we are ready to recall and prove our improved bound (Theorem 1.1) on $S_{\zeta}(n)$.

**Theorem 8.3.** If $\mathcal{E}$ is the pattern of the vertices of equilateral triangles in $\mathbb{R}^2$, then $S_{\zeta}(n) \leq \lfloor (4n - 1)(n - 1)/18 \rfloor$. 
Proof. Let \( V \) be an \( n \)-subset of \( \mathbb{R}^2 \) (identified with \( \mathbb{C} \)). Obtain three concurrent halving lines of \( V \), say \( L, M, \) and \( N \), as described in Lemma 8.2, where the directions of \( M \) and \( N \) are obtained from that of \( L \) by rotating by \( 2\pi/3 \) and \( 4\pi/3 \), respectively. This means that a point can only be formally left (or formally right) of all three lines if it lies on all three. So we may partition \( V \) into six or seven classes of points, where a point is classified according to its position (formally left or right) relative to the three lines.

| Class \( V_j \) | Position Relative to Line |
|-----------------|---------------------------|
| \( V_1 \)       | formally left | formally right | formally left |
| \( V_2 \)       | formally left | formally right | formally right |
| \( V_3 \)       | formally left | formally left  | formally right |
| \( V_4 \)       | formally right| formally left  | formally right |
| \( V_5 \)       | formally right| formally left  | formally left  |
| \( V_6 \)       | formally right| formally right| formally left  |
| \( V_7 \)       | on            | on            | on            |

Note that \( V_7 \) is nonempty if and only if a point of \( V \) happens to lie on the intersection of the three lines and if said point happens to be formally left of all three lines or formally right of all three lines.

Let \( A = (-5\pi/6, -\pi/6] \cup (\pi/6, 5\pi/6] \), the set of values of \( \text{Arg}(v - u) \) that (by Lemma 8.1) admit a reconstruction of \( E \) with \( u \) and \( v \) as consecutive points. We can rotate \( V \) and the directed lines \( L, M, \) and \( N \) so that one of them (without loss of generality, \( L \)), is the positively directed \( y \)-axis, and so that at least one-third of the pairs \( (u, v) \) of points in \( \bigcup_{j=1}^{6} V_j \times V_j \) have \( \text{Arg}(v - u) \notin A \). Every instance \( \{v_1 \prec v_2 \prec v_3\} \) of \( E \) in \( V \) either has \( v_1 \) and \( v_2 \) formally to the left of \( L \) or has \( v_2 \) and \( v_3 \) formally to the right of \( L \). Furthermore, Lemma 8.1 shows that in the former case \( \text{Arg}(v_2 - v_1) \in A \) and \( v_3 \) is uniquely determined by \( v_1 \) and \( v_2 \), while in the latter case \( \text{Arg}(v_3 - v_2) \in A \) and \( v_1 \) is uniquely determined by \( v_2 \) and \( v_3 \). Since \( L \) is a halving line, there are \( \lfloor n/2 \rfloor \) points formally to the left of \( L \) and \( \lfloor n/2 \rfloor \) points formally to the right of \( L \). So we count the number of pairs \( (u, v) \) in \( V \) with \( u \) and \( v \) formally to the same side of \( L \), and deduct the number of pairs that we deliberately rotated to prevent them from admitting a reconstruction of \( E \) to obtain

\[
S_{E}(V) \leq \binom{n/2}{2} + \binom{n/2}{2} - \frac{1}{3} \sum_{j=1}^{6} \binom{|V_j|}{2}.
\]

Now \( \sum_{j=1}^{6} |V_j| \) is either \( n \) or \( n - 1 \) (depending on whether \( V_7 \) is empty or not), and by convexity, the sum of binomial coefficients in (3) is minimized when the various values \( |V_j| \) are as close to equal as possible. So if \( n - 1 = 6q + 2r + s \)
with \( q \in \mathbb{Z}, r \in \{0, 1, 2\} \) and \( s \in \{0, 1\} \), then

\[
\sum_{j=1}^{6} \binom{|V_j|}{2} \geq (2r+s)\binom{q+1}{2} + (6-2r-s)\binom{q}{2} \]

\[= q(3q + 2r + s - 3). \]

In this case, we see that the first two terms of (3) become

\[
\left(\left\lfloor \frac{n}{2} \right\rfloor \right)^2 + \left(\left\lceil \frac{n}{2} \right\rceil \right)^2 = (s+1)\left(3q+r+1\right) + (1-s)\left(3q+r\right) \]

and substituting (4) and (5) into (3), we obtain

\[
S_{\mathcal{E}}(V) \leq \frac{8q(q + 2r + s) + 3q + 3r(r + s)}{3},
\]

which works out to

\[
S_{\mathcal{E}}(V) \leq \begin{cases} 
\frac{2n^2}{3} - \frac{5}{18}n - \frac{1}{3} & \text{if } n \equiv 0 \text{ or } 2 \pmod{6}, \\
\frac{2n^2}{3} - \frac{5}{18}n + \frac{1}{15} & \text{if } n \equiv 1 \pmod{6}, \\
\frac{2n^2}{3} - \frac{5}{18}n - \frac{1}{6} & \text{if } n \equiv 3 \text{ or } 5 \pmod{6}, \\
\frac{2n^2}{3} - \frac{5}{18}n - \frac{4}{9} & \text{if } n \equiv 4 \pmod{6},
\end{cases}
\]

so that \( S_{\mathcal{E}}(V) \leq \lfloor (4n - 1)(n - 1)/18 \rfloor \) for arbitrary \( n \). Since \( V \) was (up to our rotation) an arbitrary \( n \)-subset of \( \mathbb{C} \), this upper bound holds for \( S_{\mathcal{E}}(n) \).

\[\square\]

References

[1] B. M. Ábrego, G. Elekes, and S. Fernández-Merchant. Structural results for planar sets with many similar subsets. *Combinatorica*, 24(4):541–554, 2004.

[2] B. M. Ábrego and S. Fernández-Merchant. On the maximum number of equilateral triangles. *Discrete Comput. Geom.*, 23(1):129–135, 2000.

[3] B. M. Ábrego and S. Fernández-Merchant. Convex polyhedra in \( \mathbb{R}^3 \) spanning \( \Omega(n^{4/3}) \) congruent triangles. *J. Combin. Theory Ser. A*, 98(2):406–409, 2002.

[4] B. M. Ábrego, S. Fernández-Merchant, and B. Llano. On the maximum number of translates in a point set. *Discrete Comput. Geom.*, 43(1):1–20, 2010.

[5] P. K. Agarwal, R. Apfelbaum, G. Purdy, and M. Sharir. Similar simplices in a \( d \)-dimensional point set. In *Computational Geometry (SCG’07)*, pages 232–238. ACM, New York, 2007.

[6] R. Apfelbaum and M. Sharir. Repeated angles in three and four dimensions. *SIAM J. Discrete Math.*, 19(2):294–300, 2005.

[7] P. Brass. Exact point pattern matching and the number of congruent triangles in a three-dimensional pointset. In *Algorithms—ESA 2000 (Saarbrücken)*, volume 1879 of *Lecture Notes in Comput. Sci.*, pages 112–119. Springer, Berlin, 2000.

[8] P. Brass. Combinatorial geometry problems in pattern recognition. *Discrete Comput. Geom.*, 28(4):495–510, 2002.

[9] P. Brass, W. Moser, and J. Pach. *Research problems in discrete geometry*. Springer, New York, 2005.
ON THE NUMBER OF SIMILAR INSTANCES OF A PATTERN IN A FINITE SET

[10] P. Brass and J. Pach. Problems and results on geometric patterns. In Graph Theory and Combinatorial Optimization, volume 8 of GERAD 25th Anniv. Ser., pages 17–36. Springer, New York, 2005.

[11] G. R. Burton and G. B. Purdy. The directions determined by n points in the plane. J. London Math. Soc. (2), 20(1):109–114, 1979.

[12] V. Cortier, X. Goaoc, M. Lee, and H.-S. Na. A note on maximally repeated sub-patterns of a point set. Discrete Math., 306(16):1965–1968, 2006.

[13] H. T. Croft, K. J. Falconer, and R. K. Guy. Unsolved problems in geometry. Problem Books in Mathematics. Springer-Verlag, New York, 1991.

[14] A. Dumitrescu, J. Pach, and G. Tóth. Drawing Hamiltonian cycles with no large angles. Electron. J. Combin., 19(2):Paper 31, 13, 2012.

[15] G. Elekes and P. Erdős. Similar configurations and pseudo grids. In Intuitive Geometry (Szeged, 1991), volume 63 of Colloq. Math. Soc. János Bolyai, pages 85–104. North-Holland, Amsterdam, 1994.

[16] P. Erdős. On sets of distances of n points. Amer. Math. Monthly, 53(5):248–250, 1946.

[17] P. Erdős and G. Purdy. Some extremal problems in geometry. I. Combinatorial Theory Ser. A, 10:246–252, 1971.

[18] P. Erdős and G. Purdy. Some extremal problems in geometry. III. In Proceedings of the Sixth Southeastern Conference on Combinatorics, Graph Theory and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1975), Congressus Numerantium, No. XIV, pages 291–308. Utilitas Math., Winnipeg, Man., 1975.

[19] P. Erdős and G. Purdy. Some extremal problems in geometry. IV. In Proceedings of the Seventh Southeastern Conference on Combinatorics, Graph Theory, and Computing (Louisiana State Univ., Baton Rouge, La., 1976), Congressus Numerantium, No. XVII, pages 307–322. Utilitas Math., Winnipeg, Man., 1976.

[20] J. Erickson, F. Hurtado, and P. Morin. Centerpoint theorems for wedges. Discrete Math. Theor. Comput. Sci., 11(1):45–54, 2009.

[21] S. P. Fekete and H. Meijer. On minimum stars and maximum matchings. Discrete Comput. Geom., 23(3):389–407, 2000.

[22] M. Laczkovich and I. Z. Ruzsa. The number of homothetic subsets. In The Mathematics of Paul Erdős, II, volume 14 of Algorithms Combin., pages 294–302. Springer, Berlin, 1997.

[23] J. Pach and P. K. Agarwal. Combinatorial geometry. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, 1995.

[24] J. Pach and R. Pinchasi. How many unit equilateral triangles can be generated by N points in convex position? Amer. Math. Monthly, 110(5):400–406, 2003.

[25] J. Pach and M. Sharir. Repeated angles in the plane and related problems. J. Combin. Theory Ser. A, 59(1):12–22, 1992.

[26] J. Pach and M. Sharir. Combinatorial geometry and its algorithmic applications, volume 152 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2009.

[27] M. J. van Kreveld and M. T. de Berg. Finding squares and rectangles in sets of points. BIT, 31(2):202–219, 1991.

[28] C. Xu and R. Ding. The number of isosceles right triangles determined by n points in convex position in the plane. Discrete Comput. Geom., 31(3):491–499, 2004.

DEPARTMENT OF MATHEMATICS, CALIFORNIA STATE UNIVERSITY, NORTH RIDGE, UNITED STATES