

Everywhere unbalanced configurations

David Conlon* Jeck Lim†

Abstract

An old problem in discrete geometry, originating with Kupitz, asks whether there is a fixed natural number k such that every finite set of points in the plane has a line through at least two of its points where the number of points on either side of this line differ by at most k . We give a negative answer to a natural variant of this problem, showing that for every natural number k there exists a finite set of points in the plane together with a pseudoline arrangement such that each pseudoline contains at least two points and there is a pseudoline through any pair of points where the number of points on either side of each pseudoline differ by at least k . Moreover, we may find such a configuration with at most $2^{2^{ck}}$ points, which, by a result of Pinchasi, is best possible up to the value of the constant c .

1 Introduction

Does there exist a positive integer k such that for every finite set of points in the plane there is a line containing at least two of these points where the number of points on either side of this line differ by at most k ? This basic problem was first raised by Kupitz [9] in the late 1970s and has since been reiterated by many authors, including Alon [1], Erdős [3], Green [7], Kalai [8], Pach [10] and Pinchasi [11]. It was also singled out for inclusion in the book of Brass, Moser and Pach [2] that surveys many of the most interesting open problems in discrete geometry.

If a point set has an odd number of points with no three of them on a line, then the number of points on either side of each line through two of the points must differ by at least one. Kupitz's original conjecture [9] was that this should be the extremal case, that is, that there should always be a line containing at least two points of any finite point set where the number of points on either side differ by at most one. This was disproved by Alon [1], who showed that there are finite point sets where the number of points on either side of each line determined by the set differ by at least two. No better example is known, though in Figure 1 we give a different example to Alon's with the same property, but, unlike his examples, containing an odd number of points. This example can easily be extended to any larger odd number of points by adding an equal number of points on both ends of some line.

In the opposite direction, improving earlier unpublished results of Alon and Perles, it was shown by Pinchasi [11] that there exists an absolute constant C such that every n -point set determines a line where the number of points on either side of the line differ by at most $C \log \log n$. In fact, his

*Department of Mathematics, Caltech, Pasadena, CA 91125, USA. Email: dconlon@caltech.edu. Research supported by NSF Awards DMS-2054452 and DMS-2348859.

†Department of Mathematics, Caltech, Pasadena, CA 91125, USA. Email: jlim@caltech.edu. Research partially supported by an NUS Overseas Graduate Scholarship.

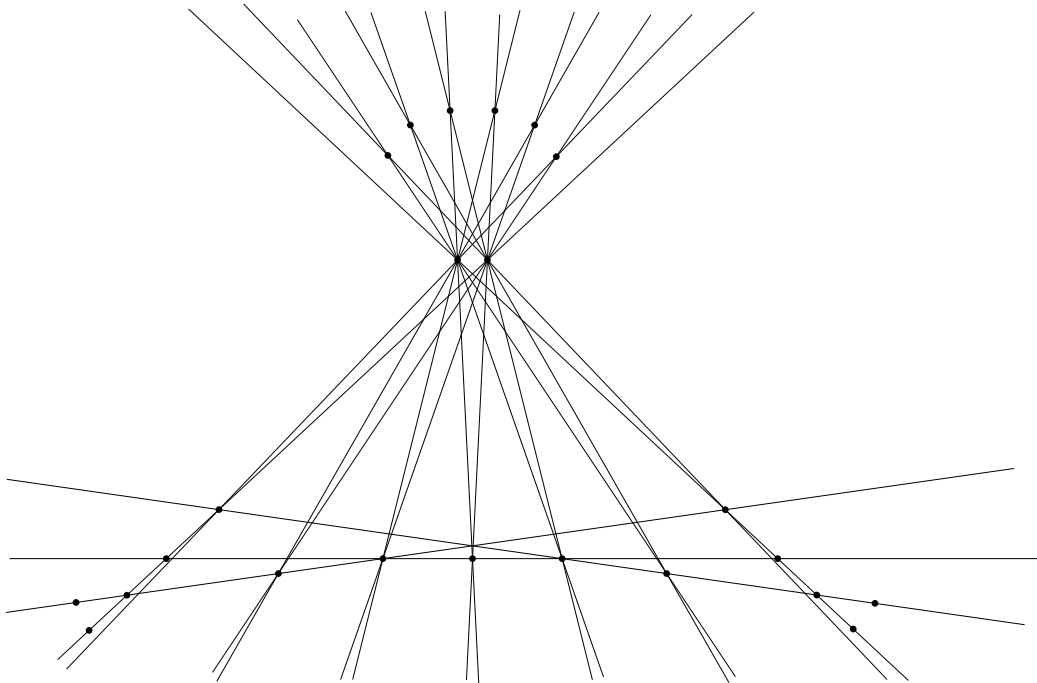


Figure 1: A set of 23 points where the number of points on either side of each line determined by the set differ by at least two.

result applies in the much broader context of generalised configurations. Recall that a *pseudoline arrangement* is a collection of two-way unbounded simple curves any two of which meet in at most one point. A *generalised configuration* is then a finite set of points in the plane together with a pseudoline arrangement such that each pseudoline contains at least two points and there is a pseudoline through any pair of points. With this terminology, we may now state Pinchasi's result, of which the result stated above is clearly a special case.

Theorem 1.1 (Pinchasi). *There exists an absolute constant C such that every generalised configuration with n points contains a pseudoline where the number of points on either side of the pseudoline differ by at most $C \log \log n$.*

Our main result says that Pinchasi's result is tight up to the constant.

Theorem 1.2. *There exists a positive constant c such that for every sufficiently large natural number n there is a generalised configuration with n points where the number of points on either side of each pseudoline differ by at least $c \log n$.*

In particular, for any positive integer k , there is a generalised configuration where the number of points on either side of each pseudoline differ by at least k .

The reason why Pinchasi's result applies to pseudolines as well as lines is that he works throughout with *allowable sequences of permutations*. This is a sequence of permutations of $[n] := \{1, 2, \dots, n\}$ starting with the identity permutation $1, 2, \dots, n$ and ending with its reverse $n, n-1, \dots, 1$ where every permutation in the sequence arises from its predecessor by flipping the elements in one or more non-overlapping increasing subsequences of consecutive elements. For example, the sequence

$$12345 \rightarrow 21354 \rightarrow 25314 \rightarrow 52341 \rightarrow 54321$$

is an allowable sequence of permutations of $[5]$.

Introduced by Goodman and Pollack [5], the importance of allowable sequences is that every generalised configuration of points gives rise to such a sequence and, conversely [6], that every allowable sequence can be realised as a generalised configuration. Some intuition can be gained by thinking about the generalised configuration consisting of a point set and the line arrangement determined by that point set. If we fix another line ℓ such that the orthogonal projection of the points of our set onto ℓ are all distinct, we can label the points as $1, 2, \dots, n$ in that order. If we now rotate ℓ counterclockwise about a fixed point, then the order of the orthogonal projections of our points onto ℓ shifts as we rotate, jumping to a different permutation every time ℓ moves through the perpendicular to any line in our collection, ultimately arriving at $n, n-1, \dots, 1$ after rotating through 180° . The construction of an allowable sequence when we instead have a pseudoline arrangement is similar, though we refer the reader to [4] for more details.

In this language, Pinchasi's result may be stated as saying that there is an absolute constant C such that every allowable sequence of permutations of $[n]$ uses a flip $[a, b]$ where $|\frac{a+b}{2} - \frac{n+1}{2}| \leq C \log \log n$, where here the *flip* $[a, b]$ is understood as taking a permutation $\pi_1, \pi_2, \dots, \pi_n$ of $[n]$ and reversing the elements in the block $\pi_a, \pi_{a+1}, \dots, \pi_b$. Similarly, the result that we will actually prove can be stated as follows.

Theorem 1.3. *There exists a positive constant c such that for every sufficiently large natural number n there is an allowable sequence of permutations of $[n]$ where each flip $[a, b]$ has $|\frac{a+b}{2} - \frac{n+1}{2}| \geq c \log n$.*

As already mentioned, a result of Goodman and Pollack [6] says that any allowable sequence of permutations of $[n]$ can be realised as a generalised configuration with n points. Moreover, when we flip $[a, b]$, the number of elements on either side of the corresponding pseudoline must differ by at least $|n - b - a + 1|$, so Theorem 1.2 follows as an immediate corollary of Theorem 1.3. As such, the remainder of the paper will be concerned with proving this latter theorem. We will prove the theorem for sufficiently large odd n , but it trivially extends to all large n by removing a single point when necessary.

2 Centred sequences, flips and blocks

We begin our proof of Theorem 1.3 by introducing some notation and justifying some standard procedures that will be used repeatedly in the construction. We first fix a positive integer t that will appear throughout the argument.

Definition 2.1. A *centred sequence* \mathbf{A} is an injective map $\mathbf{A} : [a, b] \rightarrow \mathbb{Z}$ for some integers $a \leq b$. For $a \leq n \leq b$, we denote $\mathbf{A}(n)$ by \mathbf{A}_n . The *underlying sequence* of \mathbf{A} is then $\mathbf{A}_a, \mathbf{A}_{a+1}, \dots, \mathbf{A}_b$.

Given such a centred sequence \mathbf{A} , a *flip* F is an interval $[c, d]$, where $a \leq c \leq d \leq b$. The *size* of F is $d - c + 1$. Performing the flip F on the sequence \mathbf{A} gives a new centred sequence $\mathbf{A}' : [a, b] \rightarrow \mathbb{Z}$ where the subsequence $[c, d]$ of \mathbf{A} is reversed.

The flip F is *valid* if the centred subsequence $\mathbf{A}|_{[c, d]}$ is increasing and $\frac{c+d}{2}$ does not lie in the real interval $[-t, t]$.

Our aim will be to find some $n > t$ and, starting from the identity centred sequence $\mathbf{A} : [-n, n] \rightarrow \mathbb{Z}$, to perform a sequence of valid flips, ending at the reverse of \mathbf{A} .

Definition 2.2. A *block* B is a sequence B_1, \dots, B_n of distinct integers. The *size* of B , which we denote by $|B|$, is n .

We can again define a *flip* $[c, d]$ of B , where performing this flip will reverse the subsequence B_c, B_{c+1}, \dots, B_d . We say that a flip is *valid* if the subsequence B_c, \dots, B_d is initially increasing.

Denote by \bar{B} the block which is the reverse of B . If $|B| = m$, write \mathbf{B} for the centred sequence $\mathbf{B} : [t - m + 1, t] \rightarrow \mathbb{Z}$ with the same underlying sequence as B . Conversely, if \mathbf{A} is a centred sequence, we write A for the block corresponding to the underlying sequence of \mathbf{A} .

Definition 2.3. We say that a centred sequence \mathbf{A}' can be *obtained* from a centred sequence \mathbf{A} if there is a sequence of valid flips sending \mathbf{A} to \mathbf{A}' . We say that a block B' can be *obtained* from a block B if there is a sequence of valid flips sending B to B' . We also say that we can *go from* \mathbf{A} to \mathbf{A}' or that we can *go from* B to B' .

Note that a valid flip of a block is equivalent to making several valid flips of size 2 each. Thus, to see if B' can be obtained from B , it suffices to consider only valid flips of size 2.

Definition 2.4. Given two blocks B, C of sizes m, n , their *concatenation* BC is the block of size $m + n$ with sequence $B_1, \dots, B_m, C_1, \dots, C_n$.

Given a centred sequence $\mathbf{A} : [a, b] \rightarrow \mathbb{Z}$, the *concatenation* $\mathbf{A}B$ is the centred sequence $\mathbf{A}' : [a, b + m] \rightarrow \mathbb{Z}$ with underlying sequence $\mathbf{A}_a, \dots, \mathbf{A}_b, B_1, \dots, B_m$. We can similarly concatenate a block to the left, writing it as BA .

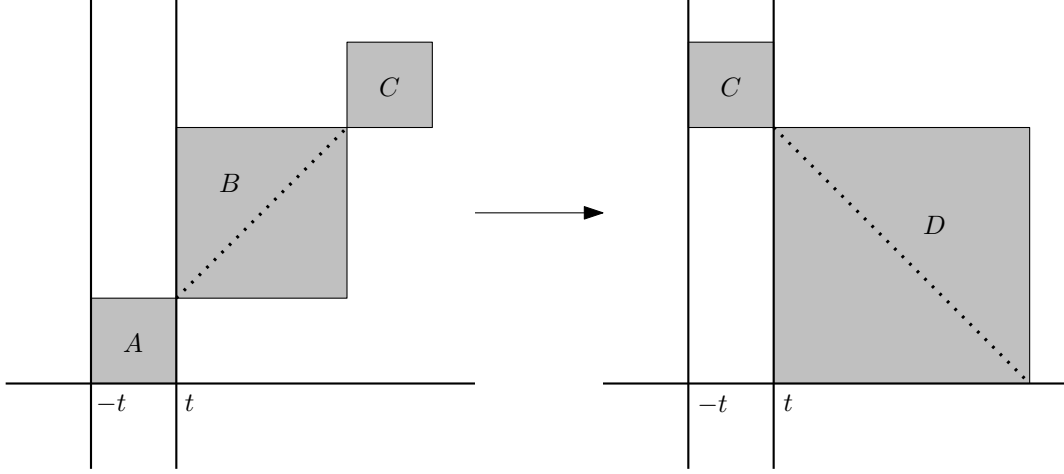


Figure 2: A graphical depiction of Lemma 2.6, Shifting.

Definition 2.5. Given blocks or centred sequences B, C , we write $B \prec C$ if $\max_i B_i < \min_j C_j$. We also write $B \succ 0$ if $B_i > 0$ for all i and similarly for $B \prec 0$.

We say that a block or centred sequence B is *increasing* (resp., *decreasing*) if its underlying sequence is increasing (resp., decreasing).

Observe that if B, C are blocks such that $B \prec C$, then we can easily go from BC to CB . We will constantly use this simple operation in what follows. We now describe another basic operation that we use repeatedly.

Lemma 2.6 (Shifting). *For all $n \geq 3^{2t}$, the following holds. Let $\mathbf{A} : [-t, t] \rightarrow \mathbb{Z}$ be a centred sequence and B, C be blocks such that*

- $|B| = n$, $|C| = 2t + 1$,
- B is increasing,
- $\mathbf{A} \prec B \prec C$.

Then one can go from $\mathbf{A}BC$ to $\mathbf{C}D$, where D is decreasing and, in keeping with Definition 2.2, $\mathbf{C} : [-t, t] \rightarrow \mathbb{Z}$ is the centred sequence with the same underlying sequence as the block C .

The idea here is that shifting a block C into the forbidden region is difficult and requires “ammunition” in the form of an increasing block B . Moreover, after we “use up” the ammunition B , the resulting decreasing block D is less useful to us.

A graphical depiction of this lemma can be seen in Figure 2. Throughout the paper, we will make extensive use of such figures to illustrate the steps in our construction. These figures will also accurately depict which blocks lie above which other blocks.

Proof. We will show by induction on $k = t, t - 1, \dots, -t$ that there is some N_k such that, for all $n \geq N_k$, the following holds. Let $\mathbf{A} : [k, t] \rightarrow \mathbb{Z}$ be a centred sequence and B, C be blocks such that

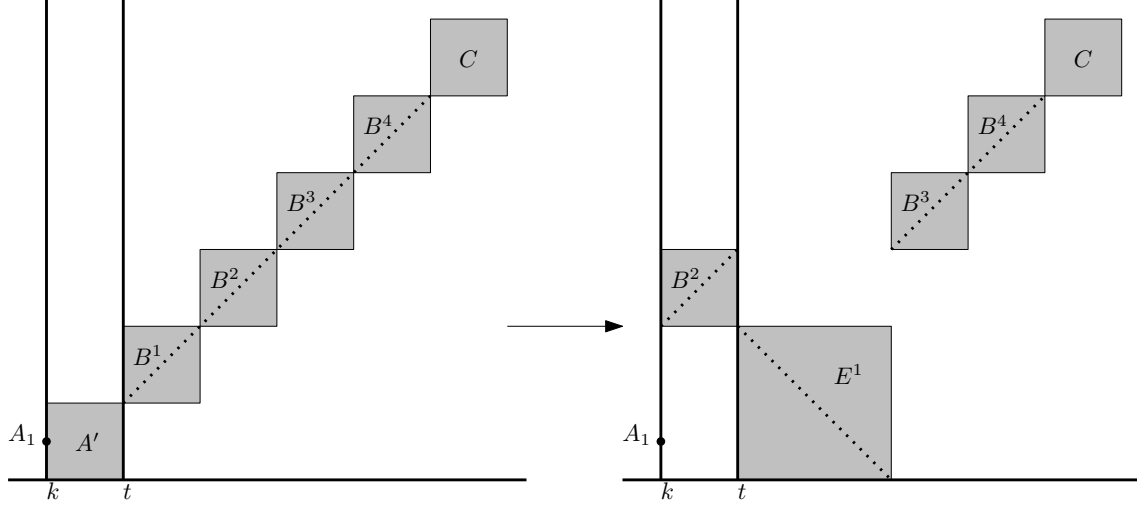


Figure 3: Proof of Lemma 2.6, part 1.

- $|B| = n$, $|C| = t - k + 1$,
- B is increasing,
- $\mathbf{A} \prec B \prec C$.

Then one can go from $\mathbf{A}BC$ to $\mathbf{C}D$, where D is decreasing and $\mathbf{C} : [k, t] \rightarrow \mathbb{Z}$ is the centred sequence with the same underlying sequence as the block C .

The base case $k = t$ is simple: we may take $N_t = 0$ and perform a single flip on the whole centred sequence $\mathbf{A}BC$.

For $-t \leq k < t$, set $N_k = 2(N_{k+1} + t - k)$. Write $\mathbf{A} = A_1 \mathbf{A}'$, where A_1 is viewed as a block of size 1 and \mathbf{A}' is the remaining centred sequence. Write $C = C_1 C'$, where C_1 is a block of size 1 and C' is the remaining block, and $B = B^1 B^2 B^3 B^4$, where $|B^1| = N_{k+1}$ and $|B^2| = |B^3| = t - k$, so that $|B^4| \geq N_{k+1}$. Apply the induction hypothesis for $k + 1$ on $\mathbf{A}' B^1 B^2$ to obtain $\mathbf{B}^2 E^1$, where E^1 is decreasing (see Figure 3).

Since $E^1 \prec B^3$, we may swap them and go from $E^1 B^3$ to $B^3 E^1$. By continuing to freely swap blocks this way, we can go from $A_1 \mathbf{B}^2 E^1 B^3 B^4 C$ to $A_1 \mathbf{B}^2 B^3 C_1 E^1 B^4 C'$. We then perform the flip $[k, 2t - k + 1]$ to get

$$C_1 \overline{\mathbf{B}^3} \overline{B^2} A_1 E^1 B^4 C'$$

(see Figure 4).

We can go from $\overline{B^2} A_1 E^1$ to E^2 , where E^2 is decreasing. Then we go from $E^2 B^4 C'$ to $B^4 C' E^2$ to obtain

$$C_1 \overline{\mathbf{B}^3} B^4 C' E^2.$$

Finally, we again apply the induction hypothesis for $k + 1$ to $\overline{\mathbf{B}^3} B^4 C'$ to obtain

$$\mathbf{C} E^3 E^2,$$

so we can set $D = E^3 E^2$, which is decreasing (see Figure 5).

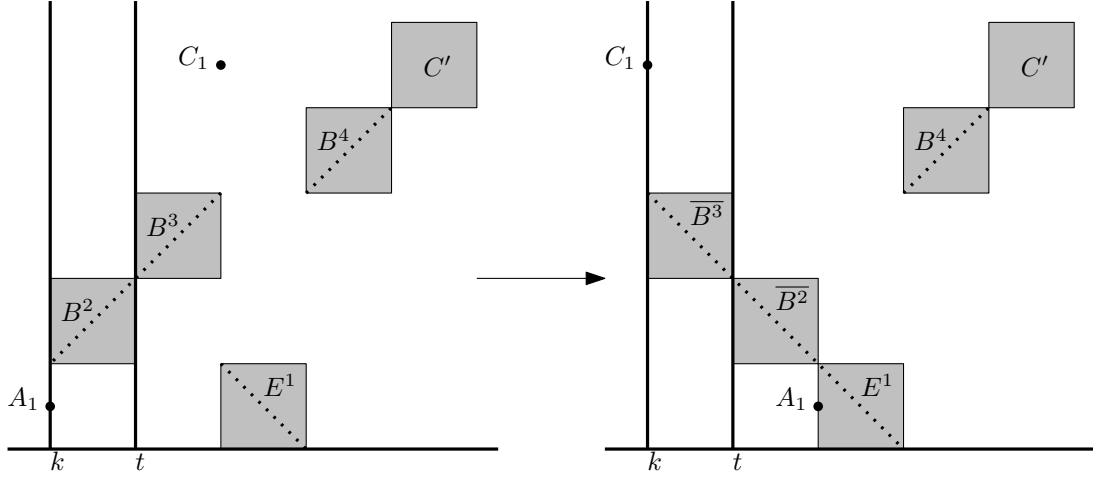


Figure 4: Proof of Lemma 2.6, part 2.

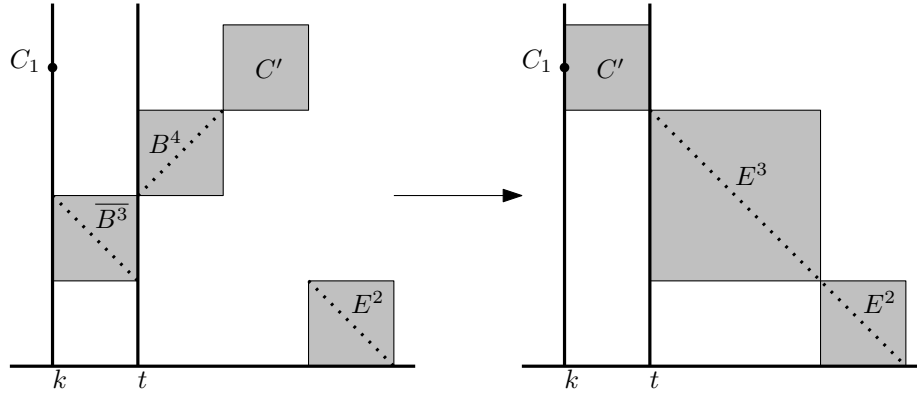


Figure 5: Proof of Lemma 2.6, part 3.

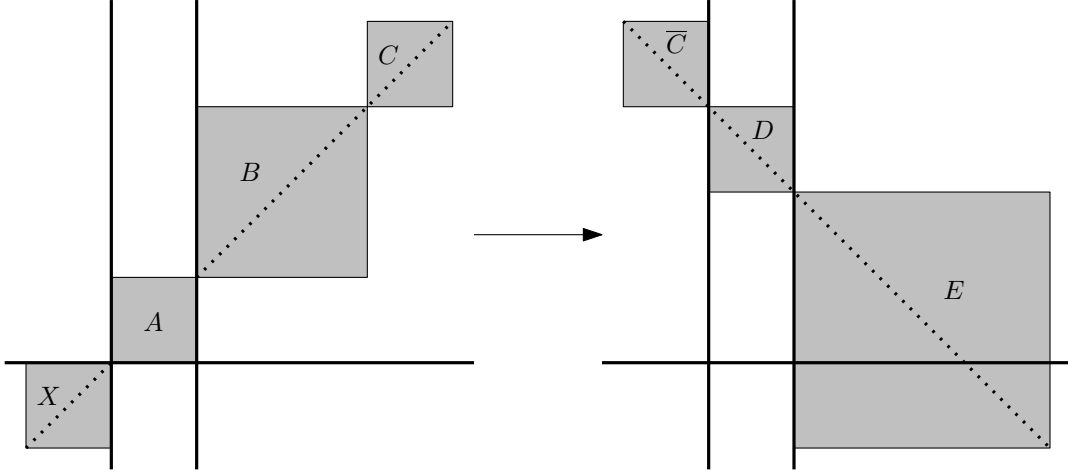


Figure 6: Reflection.

To finish the proof, note that a simple induction shows that $N_k \leq 3^{t-k}$. In particular, $N_{-t} \leq 3^{2t}$, as required. \square

When we apply Lemma 2.6, we will simply say that we *shift* C . With one additional step, we also obtain a procedure for bringing a block across 0.

Lemma 2.7 (Reflection). *For all $n \geq 3^{2t} + 4t + 2$, the following holds. Let $\mathbf{A} : [-t, t] \rightarrow \mathbb{Z}$ be a centred sequence and B, C, X be blocks such that*

- $|B| = n$,
- $|C| = |X|$,
- B, C, X are increasing,
- $X \prec \mathbf{A} \prec B \prec C$.

Then one can go from $X\mathbf{A}BC$ to $\overline{C}\mathbf{D}E$, where $\mathbf{D} : [-t, t] \rightarrow \mathbb{Z}$ is a decreasing centred sequence (with the same underlying sequence as the last $2t + 1$ elements of B in reverse), E is decreasing and $\mathbf{D} \succ E$ (see Figure 6).

Proof. Decompose $B = B^1 B^2 B^3$, where $|B^2| = |B^3| = 2t + 1$. Then, $|B^1| \geq 3^{2t}$, so, by shifting B^2 using Lemma 2.6, we can go from $\mathbf{A}BC$ to

$$\mathbf{B}^2 D' B^3 C,$$

where $\mathbf{B}^2 : [-t, t] \rightarrow \mathbb{Z}$ and D' is decreasing (see Figure 7).

Thus, after moving D' to the right, we can go from $X\mathbf{A}BC$ to $X\mathbf{B}^2 B^3 C D'$. Now perform the flip $[-t - |C|, 3t + 1 + |C|]$ to get

$$\overline{C} \overline{\mathbf{B}^3} \overline{B^2} \overline{X} D',$$

from which we can go to $\overline{C}\mathbf{D}E$, where $\mathbf{D} = \overline{\mathbf{B}^3}$ and $E = \overline{B^2 D' X}$ (see Figure 8). \square

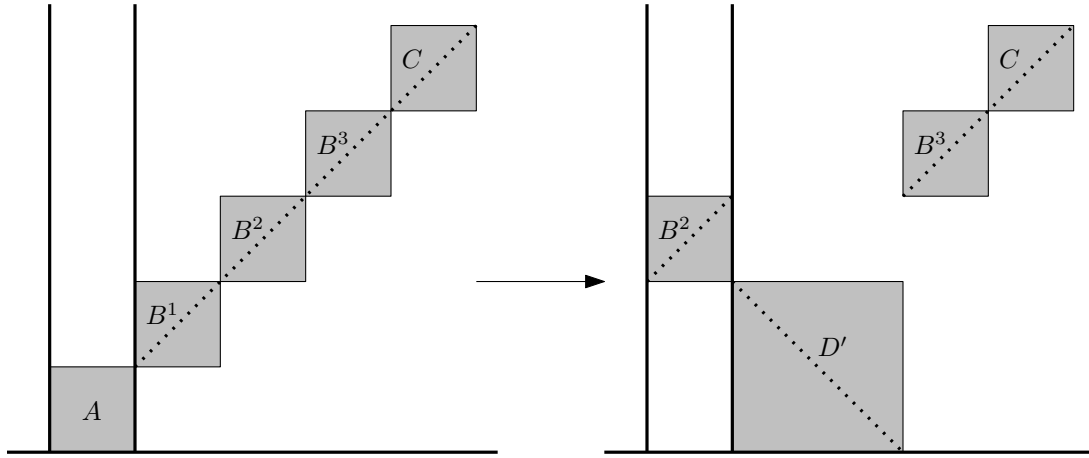


Figure 7: Proof of Lemma 2.7, part 1.

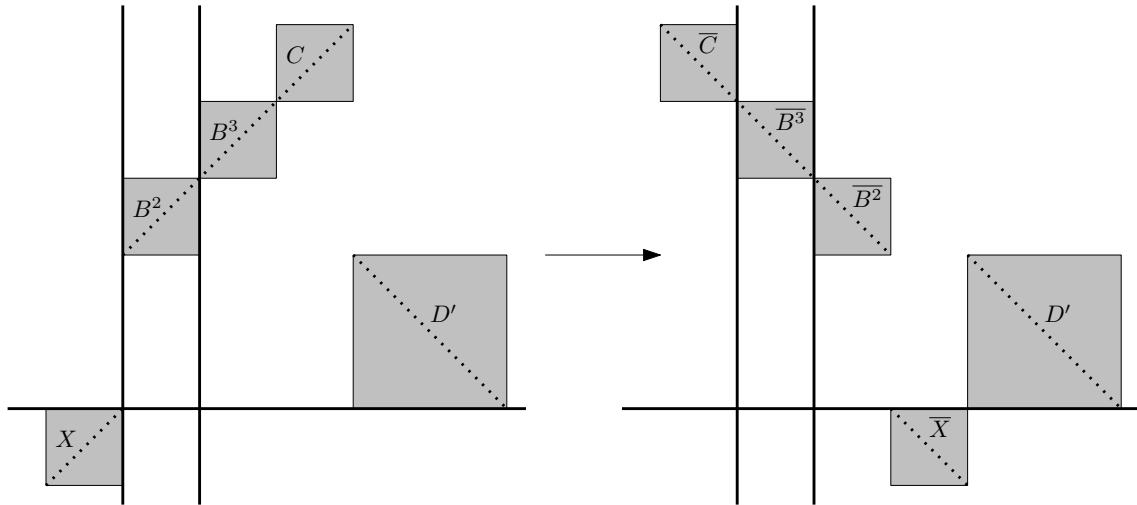


Figure 8: Proof of Lemma 2.7, part 2.

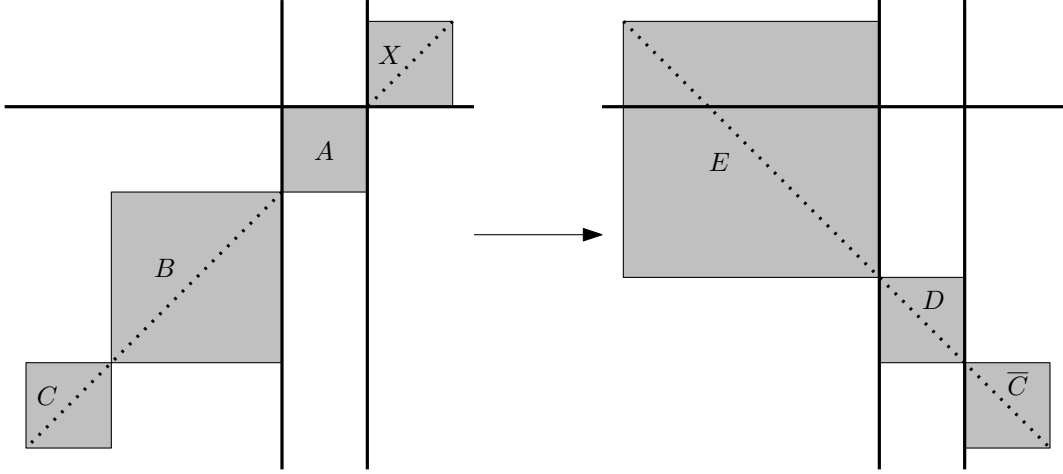


Figure 9: The mirrored version of Reflection.

When we apply Lemma 2.7, we will simply say that we *reflect* C .

To close this section, we note that the operations above also hold if we mirror the blocks and centred sequences horizontally and vertically. For example, by mirroring Reflection, we obtain the following result: for $n \geq 3^{2t} + 4t + 2$, let $\mathbf{A} : [-t, t] \rightarrow \mathbb{Z}$ be a centred sequence and B, C, X be blocks such that

- $|B| = n$,
- $|C| = |X|$,
- B, C, X are increasing,
- $C \prec B \prec \mathbf{A} \prec X$.

Then one can go from $CBA X$ to $ED\overline{C}$, where $\mathbf{D} : [-t, t] \rightarrow \mathbb{Z}$ is a decreasing centred sequence, E is decreasing and $\mathbf{D} \prec E$ (see Figure 9).

3 The recursive step

In this section, we show how to perform a certain recursive step which lies at the heart of our construction and is arguably its most difficult part. To state the result, we need some further definitions.

Definition 3.1. Given a block B of size n with sequence B_1, \dots, B_n , the *width* of B , denoted by $\text{width}(B)$, is the largest integer k such that there exist $i_1 < i_2 < \dots < i_k$ with $B_{i_1} > B_{i_2} > \dots > B_{i_k}$.

Note that, by Dilworth's theorem, this is the same as the smallest integer k such that B_1, \dots, B_n can be partitioned into k disjoint increasing subsequences.

Definition 3.2. Given a block B , denote by B^+ the block corresponding to the subsequence of B consisting of all the positive values. Denote by B^- the block corresponding to the subsequence of B consisting of all the negative values.

Definition 3.3. Let B be a block of size n and $r \geq 0$ a real number. We say that B is r -balanced if

1. the block B^- is increasing,
2. for each $1 \leq k \leq n$, the initial segment $B' = B|_{[k]}$ has $|B'^-| \geq r \cdot \text{width}(B'^+)$.

The following key property of r -balanced blocks will be crucial to our main construction.

Lemma 3.4. Suppose B is a block that is r -balanced for some real $r \geq 0$. Then there is some k such that a block of the form

$$C_1 \cdots C_k$$

can be obtained from B , where each block C_i is increasing, $|C_i^-| \geq \lfloor r \rfloor$ and $C_1^- \prec C_2^- \prec \cdots \prec C_k^-$.

Proof. Note that if B is r -balanced, then it is r' -balanced for any $r' \leq r$. Hence, without loss of generality, we can assume that r is an integer.

Let A be a block of size n . We shall decompose A into a collection of increasing subsequences. First set $I_0 = [n]$. We define an increasing subsequence $i_{1,1}, \dots, i_{1,n_1}$ from I_0 as follows. Let $i_{1,1}$ be the smallest element of I_0 . Given $i_{1,k}$, define $i_{1,k+1}$ to be the smallest element in I_0 larger than $i_{1,k}$ such that $A_{i_{1,k+1}} > A_{i_{1,k}}$. If no such $i_{1,k+1}$ exists, then we terminate the sequence and set $I_1 = I_0 \setminus \{i_{1,1}, \dots, i_{1,n_1}\}$. We then define $i_{2,1}, \dots, i_{2,n_2}$ similarly in terms of I_1 and set $I_2 = I_1 \setminus \{i_{2,1}, \dots, i_{2,n_2}\}$. Repeating this procedure until we reach d with $I_d = \emptyset$, we see that we have partitioned A into d increasing subsequences $A_{i_{k,1}}, \dots, A_{i_{k,n_k}}$ for $k = 1, \dots, d$.

We claim that $d = \text{width}(A)$. In fact, we shall show that for any $1 \leq m \leq n$, if k is the largest integer for which $i_{k,1} \leq m$, then $\text{width}(A|_{[m]}) = k$. Indeed, $\text{width}(A|_{[m]}) \leq k$, since we can cover $A|_{[m]}$ with k increasing subsequences $A_{i_{j,1}}, \dots, A_{i_{j,n_j}}$ for $j = 1, \dots, k$. To show that $\text{width}(A|_{[m]}) \geq k$, we find a decreasing subsequence of length k . Let $j_k = i_{k,1}$. Since $i_{k,1} \notin \{i_{k-1,1}, \dots, i_{k-1,n_{k-1}}\}$, there is some l such that $i_{k-1,l} < j_k$ and $A_{i_{k-1,l}} > A_{j_k}$. Set $j_{k-1} = i_{k-1,l}$. By repeating this step, we eventually obtain a sequence $j_k > j_{k-1} > \cdots > j_1$ with $A_{j_k} < A_{j_{k-1}} < \cdots < A_{j_1}$, so we have the required decreasing subsequence of length k .

Now consider the above decomposition into increasing subsequences applied to B^+ . If $k = \text{width}(B^+)$, then we can partition B^+ into k increasing subsequences S_1, \dots, S_k , the j th of which starts from B_{i_j} for some i_j . Set C_1 to be the concatenation of the first r elements of B^- and S_1 . Set C_2 to be the concatenation of the next r elements of B^- and S_2 and so on, until C_k , which we take to be the concatenation of the remaining elements of B^- and S_k . Note that, for each $j \leq k$, $\text{width}((B|_{[i_j]})^+) = j$, so, since B is r -balanced, there are at least jr elements of B^- that come before S_j . Since B^- is also increasing by the definition of r -balancedness, we can go from B to $C_1 \cdots C_k$, where each C_i is increasing and $|C_i^-| \geq r$. Furthermore, $C_1^- \prec C_2^- \prec \cdots \prec C_k^-$, since they are the elements of B^- in the right order. \square

The following lemma is the main recursive step in our construction (see Figure 10).

Lemma 3.5. Let $d, n \geq 1$, set $T = 3^{2^t}$ and define sequences $\alpha_0, \alpha_1, \dots$ and β_0, β_1, \dots inductively by

- $\alpha_0 = T + 4t + 2$,
- $\beta_0 = 0$,
- $\alpha_{i+1} = d\alpha_i + 2Td^i + d$,
- $\beta_{i+1} = d\beta_i + \frac{d^{i+1}}{3T}$.

Suppose that $d \geq 9T$. Then, for each $k \geq 0$ and any centred sequence $\mathbf{I} : [-t, t] \rightarrow \mathbb{Z}$, there are increasing blocks X, Y of some sizes with $X \prec \mathbf{I} \prec Y$ and $X \prec 0 \prec Y$ such that one can go from $X\mathbf{I}Y$ to

$$L_k W_k \mathbf{A}_k B_k R_k,$$

where:

1. $\mathbf{A}_k : [-t, t] \rightarrow \mathbb{Z}$,
2. $W_k, L_k \succ 0 \succ R_k$,
3. if $k > 0$, then $\mathbf{A}_k \prec B_k$, while if $k = 0$, then $\mathbf{A}_k \succ B_k$,
4. W_k only contains elements from Y ,
5. W_k is of the form

$$K_1 \cdots K_m M_m \cdots M_1,$$

where $m = d^k$, each K_i is decreasing, $|K_i| = n$, $|M_i| = 1$ and

$$K_m \succ M_m \succ K_{m-1} \succ M_{m-1} \succ \cdots \succ K_1 \succ M_1,$$

6. $\text{width}(B_k^+) \leq \alpha_k$,
7. $|B_k^-| \geq \beta_k$,
8. B_k is (β_k/α_k) -balanced.

Furthermore, one can take $|X|, |Y| \leq 10d^{2k+1}n$.

Though somewhat hidden in the detail of the induction hypothesis, the key point here is that one can make B_k r -balanced for any given r by choosing k and d appropriately. It is also worth noting that we will only make use of the lemma with $n = 1$, but our induction argument requires that we consider larger values of n .

Proof. We proceed by induction on k . We first prove the base case $k = 0$. Set $|X| = n + 1$ and $Y = C_1 C_2 C_3$, where $|C_1| = T + 2t + 1$, $|C_2| = 2t + 1$ and $|C_3| = n + 1$. By reflecting C_3 , we can go from $X\mathbf{I}Y$ to

$$\overline{C_3} \overline{\mathbf{C}_2} \overline{C_1} D \overline{X},$$

where D is the block we get by sorting \mathbf{I} in decreasing order (see Figure 11). But this completes the base case by setting $L_0 = \emptyset$, $W_0 = \overline{C_3}$, $\mathbf{A}_0 = \mathbf{C}_2$, $B_0 = \overline{C_1} D^+$ and $R_0 = D^- \overline{X}$ and noting that $\text{width}(B_0) = |B_0| \leq T + 4t + 2 = \alpha_0$.

We now move on to the induction proper, assuming that the result holds for k and then showing that it also holds for $k + 1$. We first describe the overall strategy, which is driven by the need to

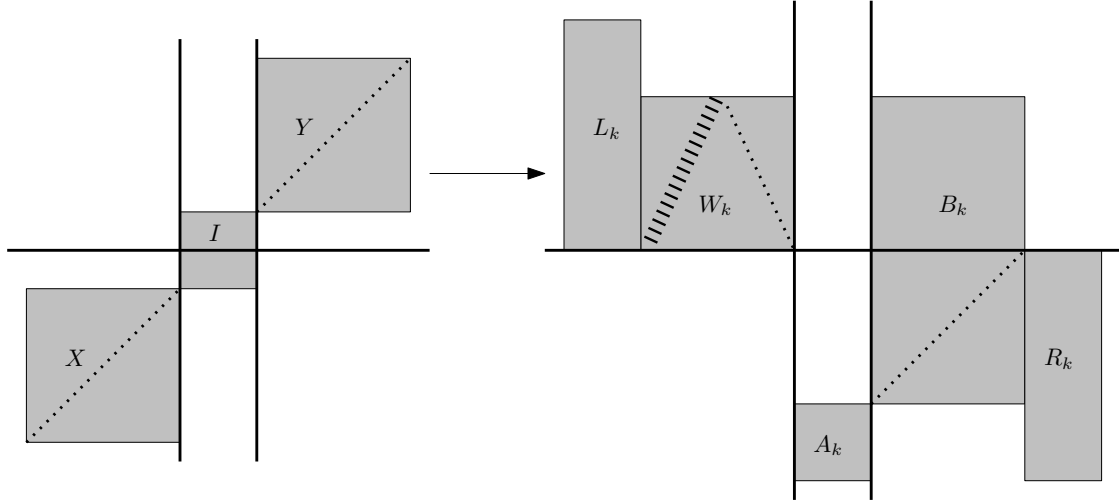


Figure 10: The main recursive step.

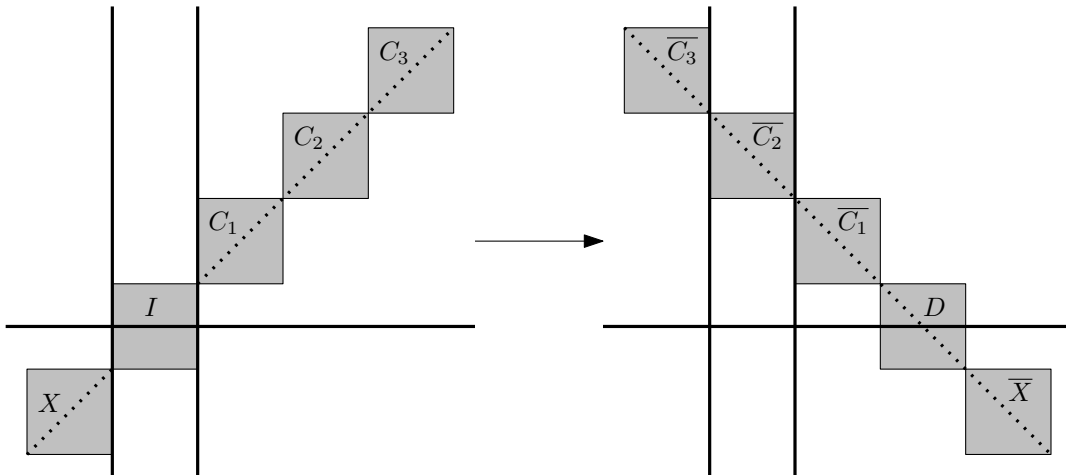


Figure 11: The case $k = 0$.

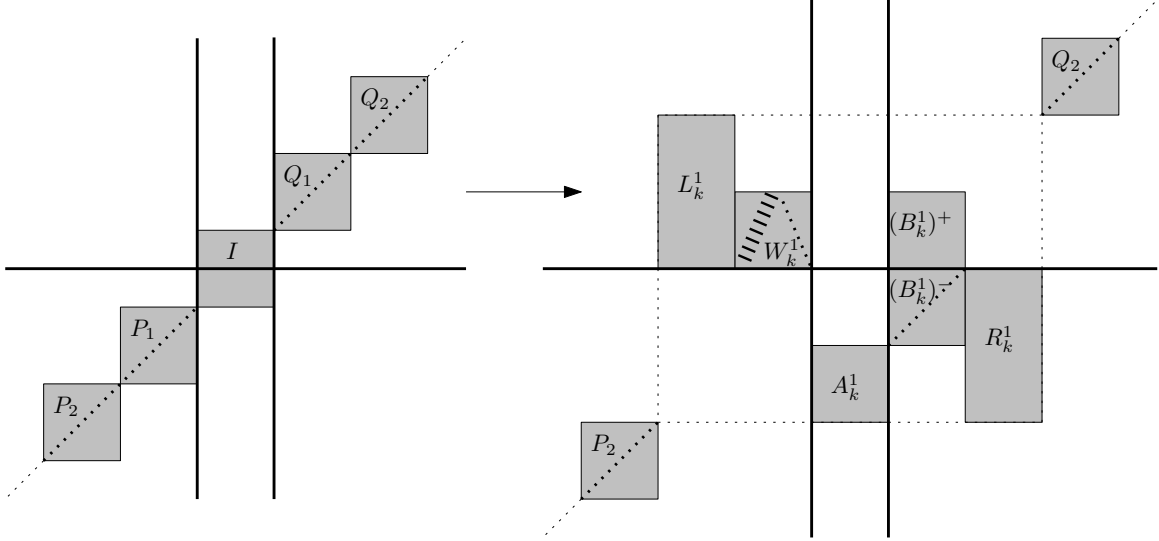


Figure 12: Step 1, part 1.

maintain the structure of the block W_{k+1} , which has a head (the part containing the K_i) and a tail (the part containing the M_i). The first step is to perform the construction for $(d, n+1, k)$ d times, thereby obtaining blocks W_k^1, \dots, W_k^d . Next, we move the tails of the W_k^i to the right. These will eventually be part of B_{k+1} and their structure will allow us to boost the balancedness of that block. Finally, we peel off one layer from the heads of the W_k^i and use it as ammunition to move elements into B_{k+1}^- . Once this is used up, it then forms the tail of W_{k+1} .

Step 1: Repeating the induction hypothesis d times.

Write

$$XIY = X'P_dP_{d-1} \cdots P_1 \mathbf{I} Q_1 Q_2 \cdots Q_d Y',$$

where P_i, Q_i are blocks of suitable lengths for what follows. By our induction hypothesis (using k and $n+1$ instead of n), we can go from $P_1 \mathbf{I} Q_1$ to $L_k^1 W_k^1 \mathbf{A}_k^1 B_k^1 R_k^1$. Since $P_i \prec P_1 \mathbf{I} Q_1 \prec Q_i$ for $i > 1$, we can therefore go from XIY to

$$L_k^1 X' W_k^1 P_d \cdots P_2 \mathbf{A}_k^1 Q_2 \cdots Q_d B_k^1 Y' R_k^1$$

(see Figure 12).

Repeating our induction hypothesis on $P_2 \mathbf{A}_k^1 Q_2$ to get $L_k^2 W_k^2 \mathbf{A}_k^2 B_k^2 R_k^2$ and so on, we can eventually go to

$$L' X' W_k^1 \cdots W_k^d \mathbf{A}_k^d B' Y' R',$$

where $L' = L_k^1 \cdots L_k^d$, $R' = R_k^d \cdots R_k^1$ and $B' = B_k^d \cdots B_k^1$ (see Figure 13). Since W_k^i only contains elements from Q_i , we have $W_k^1 \prec W_k^2 \prec \cdots \prec W_k^d$, as depicted in Figure 13.

If $k = 0$, then $\mathbf{A}_k^{i-1} \succ B_k^{i-1}$, $(B_k^i)^+$ only contains elements from $\mathbf{A}_k^{i-1} Q_i$ and $(B_k^i)^-$ only contains elements from P_i . Otherwise, if $k > 0$, then $\mathbf{A}_k^{i-1} \prec B_k^{i-1}$, $(B_k^i)^+$ only contains elements from Q_i and

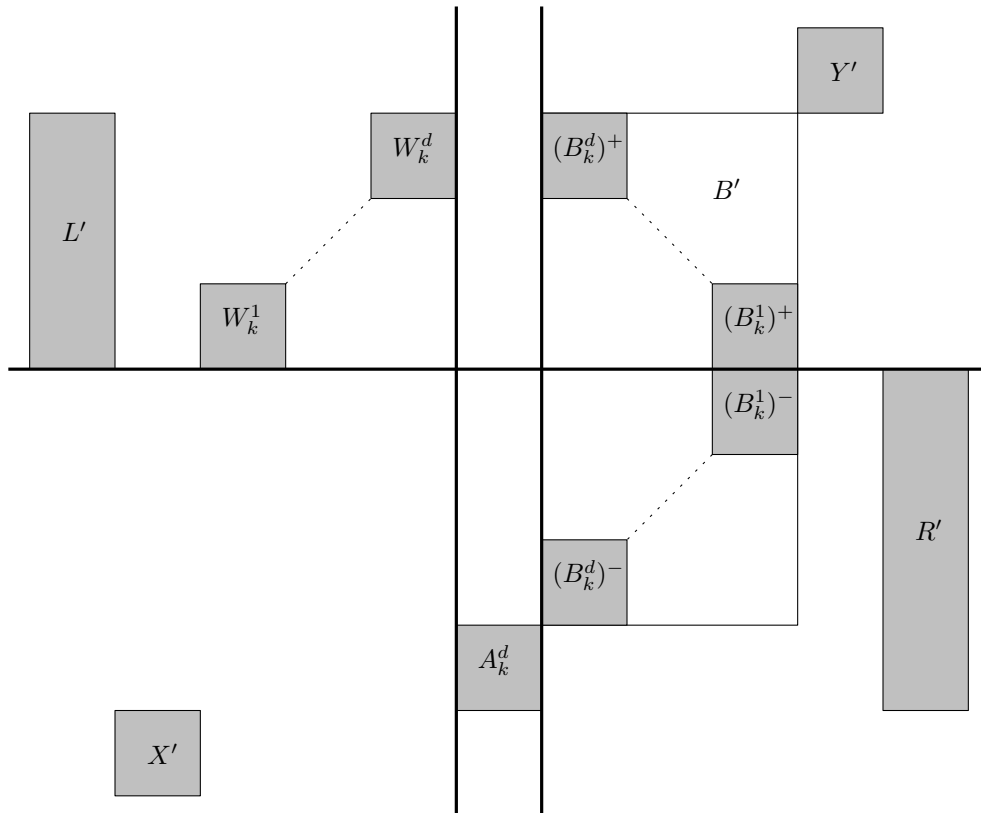


Figure 13: Step 1, part 2.

$(B_k^i)^-$ only contains elements from $P_i \mathbf{A}_k^{i-1}$. In either case, we have $(B_k^d)^+ \succ (B_k^{d-1})^+ \succ \dots \succ (B_k^1)^+$ and $(B_k^d)^- \prec \dots \prec (B_k^1)^-$, so $(B')^-$ is increasing. Furthermore, $|B'^-| = |(B_k^1)^-| + \dots + |(B_k^d)^-| \geq d\beta_k$, $\text{width}(B'^+) \leq \text{width}((B_k^1)^+) + \dots + \text{width}((B_k^d)^+) \leq d\alpha_k$ and, since each B_k^i is (β_k/α_k) -balanced, B' is (β_k/α_k) -balanced.

Step 2: Moving the tails of each W_k to the right.

Write $W_k^i = K_1^i K_2^i \dots K_m^i M_m^i \dots M_1^i$, where $m = d^k$ and $|K_j^i| = n + 1$. Decompose Y' into $S_1 \dots S_m Y''$, where $|S_i| = T + d + 4t + 2$. Further decompose each S_i into $T_i U_i$ where $|T_i| = T + 2t + 1$ and $|U_i| = d + 2t + 1$. Go to

$$L' X' W_k^1 \dots W_k^d \mathbf{A}_k^d T_1 U_1 B' S_2 \dots S_m Y'' R'.$$

By performing suitable transpositions to move the points M_j^i , we can go to

$$L' X' W' M_m^1 M_m^2 \dots M_m^d M_{m-1}^1 \dots M_1^d \mathbf{A}_k^d T_1 U_1 B' S_2 \dots S_m Y'' R',$$

where $W' = K_1^1 K_2^1 \dots K_m^1 K_1^2 \dots K_m^d$ (see Figure 14).

We would now like to move all the M_j^i to the right. Write $T_1 = T'_1 J_1$, where $|J_1| = 2t + 1$ and $|T'_1| = T$. By shifting J_1 , we can go to

$$L' X' W' M_m^1 \dots M_1^d \mathbf{J}_1 C_1 U_1 B' S_2 \dots S_m Y'' R',$$

where $\mathbf{J}_1 : [-t, t] \rightarrow \mathbb{Z}$ is increasing and C_1 is decreasing, noting that $|C_1| = |T'_1| + |A_k^d| = T + 2t + 1$. If $k = 0$, then $\mathbf{A}_0^d \succ 0$ implies that $C_1^- = \emptyset$. Otherwise, $\mathbf{A}_k^d \prec B'$, so $C_1^- \prec B'$ and we can move C_1^- all the way to the right to be part of R' . Thus, we can assume that $C_1 \succ 0$ and $|C_1| \leq T + 2t + 1$ (see Figure 15).

We can go from $C_1 U_1$ to $U_1 C_1$. Decompose $U_1 = U'_1 U''_1$, where $|U'_1| = 2t + 1$ and $|U''_1| = d$. Then, perform the flip $[-d - t, d + 3t + 1]$ to go to

$$L' X' W' M_m^1 \dots M_2^d \overline{U'_1} \overline{U''_1} \overline{J_1} N_1 C_1 B' S_2 \dots S_m Y'' R',$$

where $N_1 = M_1^d M_1^{d-1} \dots M_1^1$. We can then move $\overline{U''_1}$ to the left to be absorbed by L' (see Figure 16).

Repeating this process, we can go to

$$L' X' W' \overline{\mathbf{U}_m} C B' Y'' R',$$

where $C = \overline{J_m} N_m C_m \dots \overline{J_1} N_1 C_1 \succ 0$ and $N_i = M_i^d M_i^{d-1} \dots M_i^1$ (see Figure 17).

Before continuing with the construction, let us analyse the width of the block C . Observe that the width of interlaced blocks is at most the sum of the widths of the individual blocks, so

$$\text{width}(C) \leq \sum_{i=1}^m \text{width}(\overline{J_i} C_i) + \text{width}(N_m \dots N_1).$$

We have $\text{width}(\overline{J_i} C_i) \leq |J_i| + |C_i| \leq T + 4t + 2$. Moreover, $N_m \dots N_1$ can be covered by $m + d - 1$ increasing subsequences, so that $\text{width}(N_m \dots N_1) \leq m + d - 1$. Overall, we have

$$\text{width}(C) \leq m(T + 4t + 2) + m + d - 1 \leq 2mT + d.$$

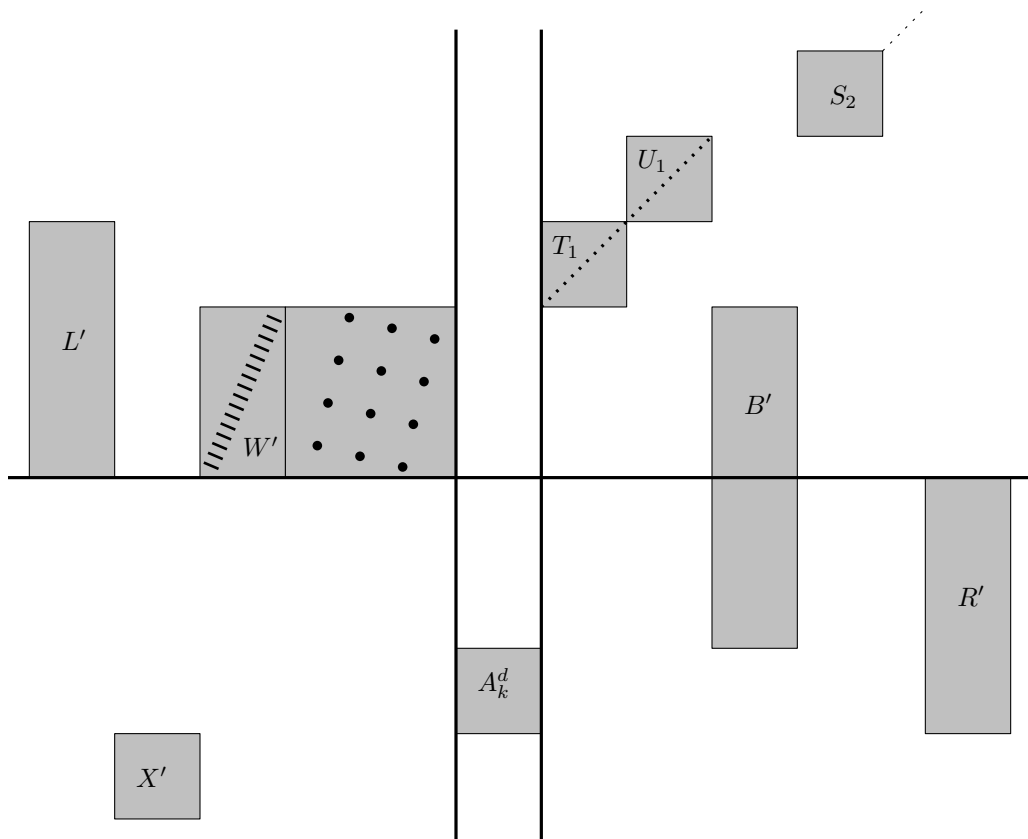


Figure 14: Step 2, part 1.

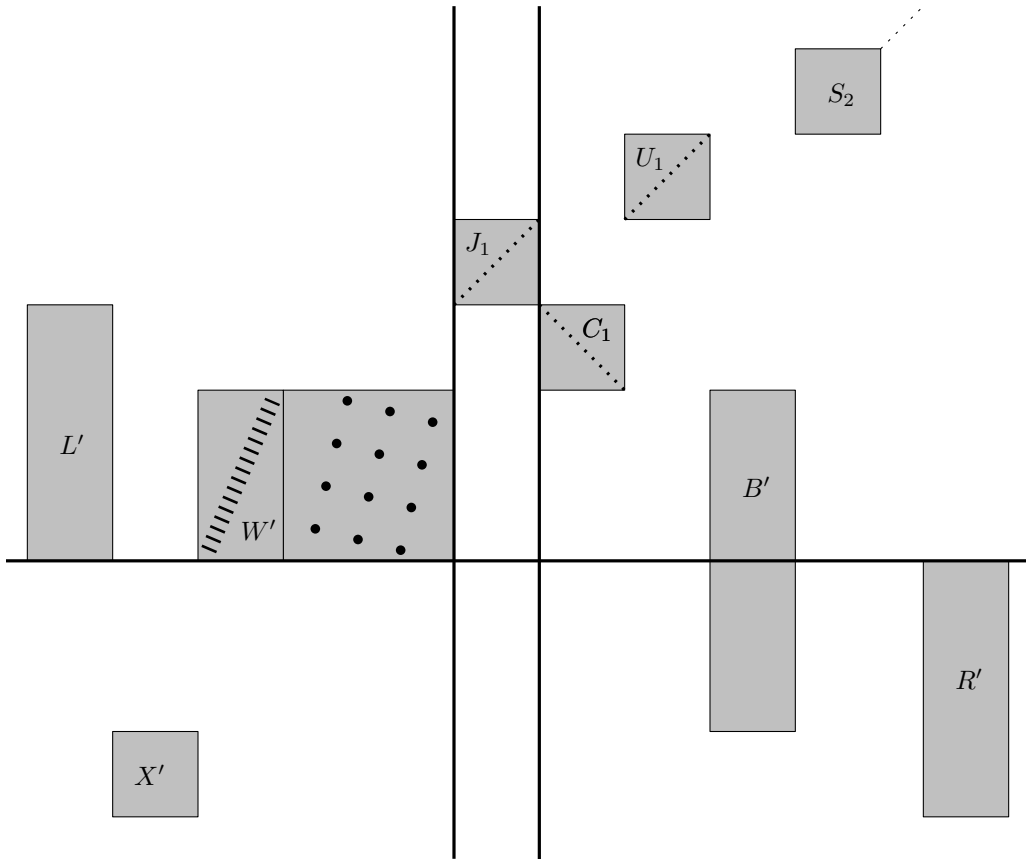


Figure 15: Step 2, part 2.

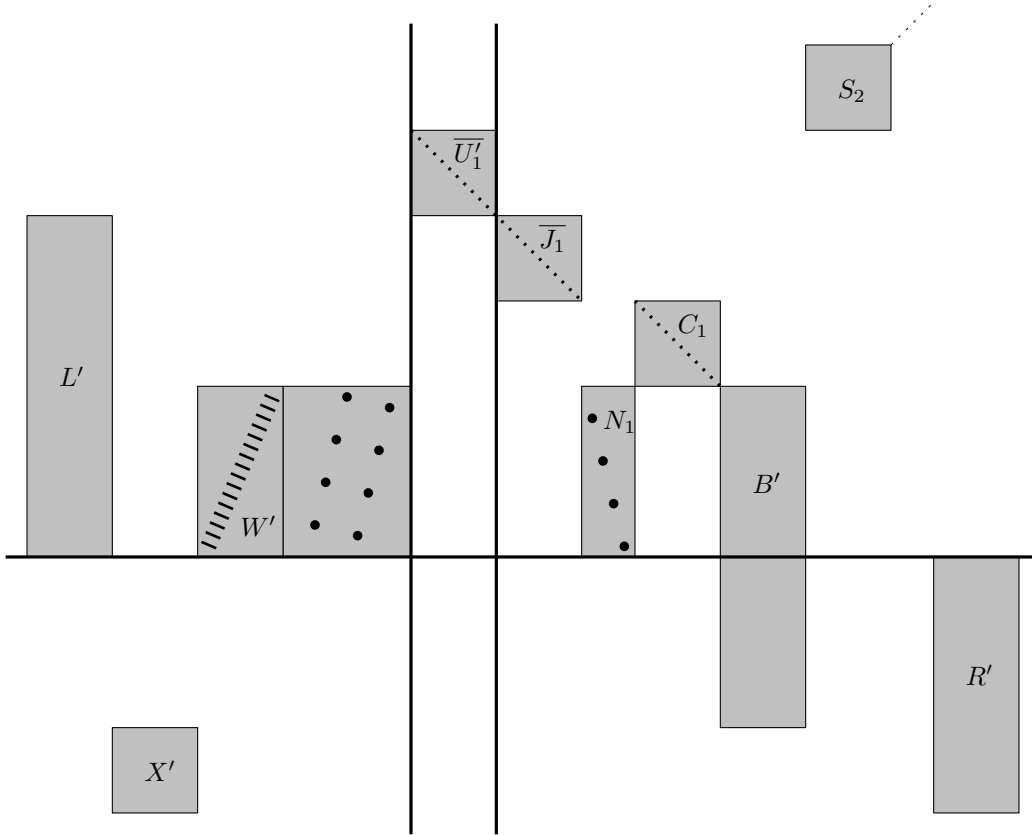


Figure 16: Step 2, part 3.

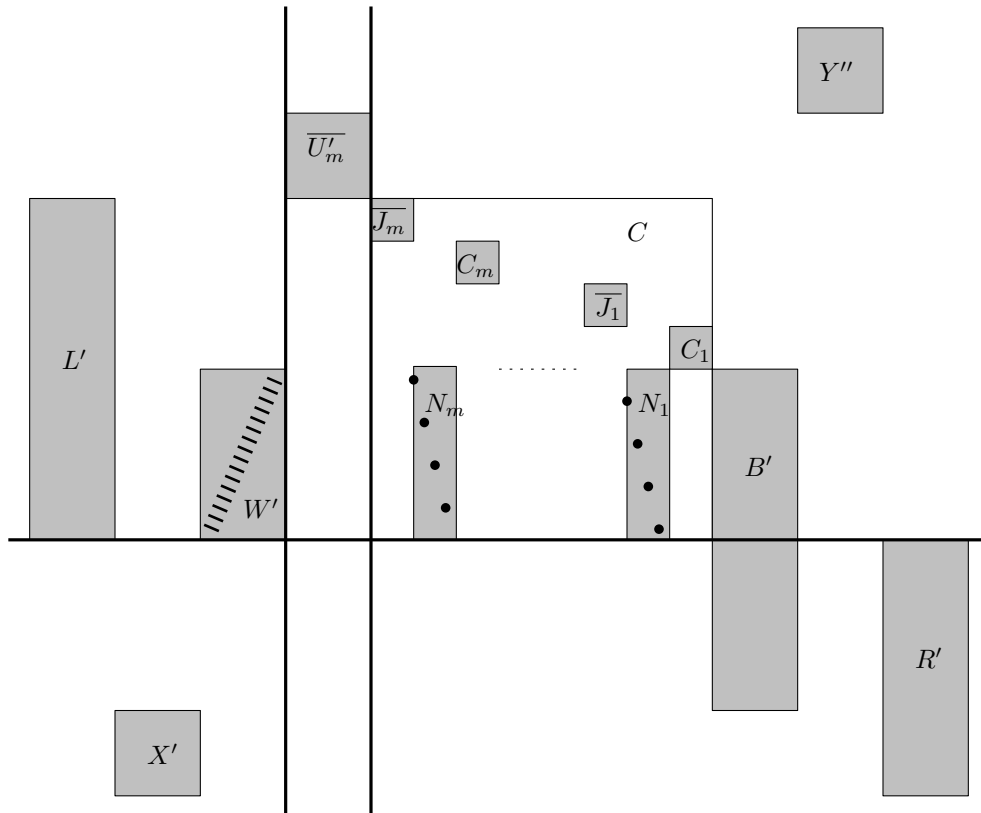


Figure 17: Step 2, part 4.

Step 3: Forming the tail of W_{k+1} .

Move X' and Y'' towards the centre. In this step, we will largely focus on the centred subsequence $W'X'\overline{U'_m}Y''$, only coming back to the full sequence towards the end.

Expand $W' = K_1 \cdots K_{md}$, where each K_i is decreasing and $|K_i| = n + 1$. Write $K_i = K'_i O_i$, where $|K'_i| = n$ and $|O_i| = 1$. Then we can go from W' to

$$K'_1 \cdots K'_{md} O_1 \cdots O_{md}.$$

Write $X' = X'' P_p P_{p-1} \cdots P_1$ and $Y'' = Q_1 \cdots Q_p$, where $|P_i| = |Q_i| = 1$, $|X''| = 2t + 1$ and $p = \left\lfloor \frac{md-T}{T+4t+2} \right\rfloor \geq 1$, by using $d \geq 9T$ (note that we have now fixed the original sizes of X and Y). Set $W'' = K'_1 \cdots K'_{md}$. We shall group the O_i as

$$O_1 \cdots O_{md} = F G_p H_p G_{p-1} H_{p-1} \cdots G_1 H_1,$$

where $|G_i| = 2t + 1$, $|H_i| = T + 2t + 1$ and F is what remains, so that $|F| \geq T$ from the definition of p .

By rearranging the blocks, we can go to

$$W'' X'' F P_p G_p H_p \cdots P_1 G_1 H_1 \overline{U'_m} Q_1 \cdots Q_p.$$

By reflecting P_1 in the mirrored sense, we can go from $P_1 G_1 H_1 \overline{U'_m} Q_1$ to $Q_1 \overline{U'_m} \overline{H_1} \overline{G_1} P_1$. Then, moving $Q_1 \overline{U'_m}$ to the left, we can go to

$$Q_1 \overline{U'_m} W'' X'' F P_p G_p H_p \cdots P_2 G_2 H_2 \overline{H_1} \overline{G_1} P_1 Q_2 \cdots Q_p$$

(see Figure 18).

Move $\overline{H_1}$ to be between W'' and X'' and move P_1 to the right. Reflecting P_2 similarly, we can go from $P_2 G_2 H_2 \overline{G_1} Q_2$ to $Q_2 \overline{G_1} \overline{H_2} \overline{G_2} P_2$ (see Figure 19). Rearranging the blocks, we obtain

$$Q_1 Q_2 \overline{U'_m} W'' \overline{H_1} \overline{G_1} \overline{H_2} X'' F P_p G_p H_p \cdots P_3 G_3 H_3 \overline{G_2} Q_3 \cdots Q_p P_2 P_1.$$

Repeating all the way to $P_p G_p H_p$, we obtain

$$Q \overline{U'_m} W'' \overline{H_1} \overline{G_1} \cdots \overline{H_p} X'' F \overline{G_p} P,$$

where $P = P_p \cdots P_1$ and $Q = Q_1 \cdots Q_p$ are increasing. By shifting X'' in the mirrored sense, we obtain

$$Q \overline{U'_m} W'' \overline{H_1} \overline{G_1} \cdots \overline{H_p} \overline{G_p} \overline{F} X'' P$$

(see Figure 20).

Overall, we managed to go from the original sequence XIY to

$$L' Q \overline{U'_m} W'' O_{md} \cdots O_1 X'' P C B' R'.$$

Now set $L_{k+1} = L' Q \overline{U'_m}$, $W_{k+1} = W'' O_{md} \cdots O_1$, $\mathbf{A}_{k+1} = X''$, $B_{k+1} = P C B'$ and $R_{k+1} = R'$. This concludes the construction, but we still need to check that the required conditions hold. Conditions (1), (2), (3), (4) and (5) are easy to check, so it only remains to show that B_{k+1} satisfies conditions (6), (7) and (8).

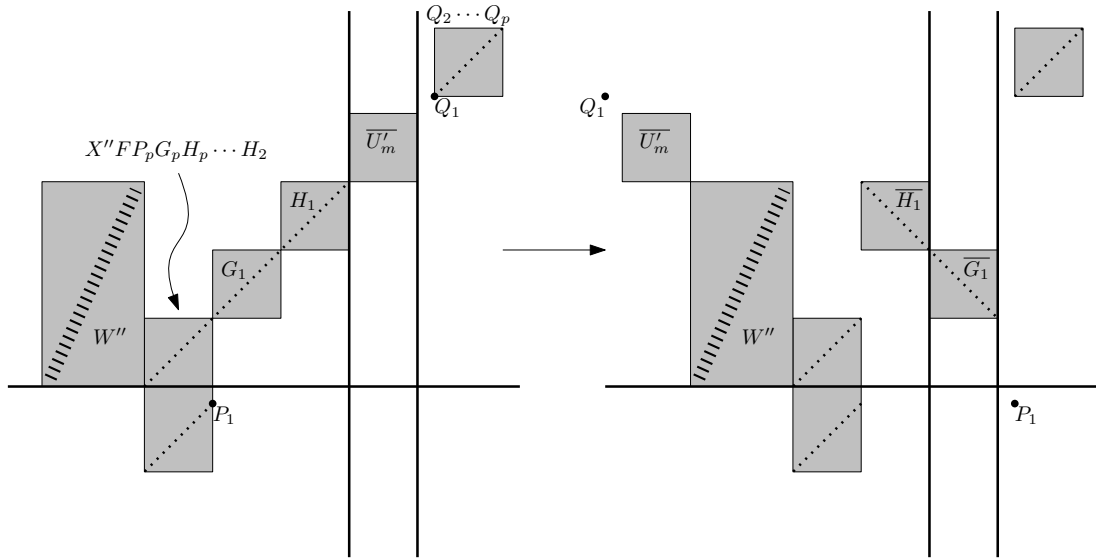


Figure 18: Step 3, part 1.

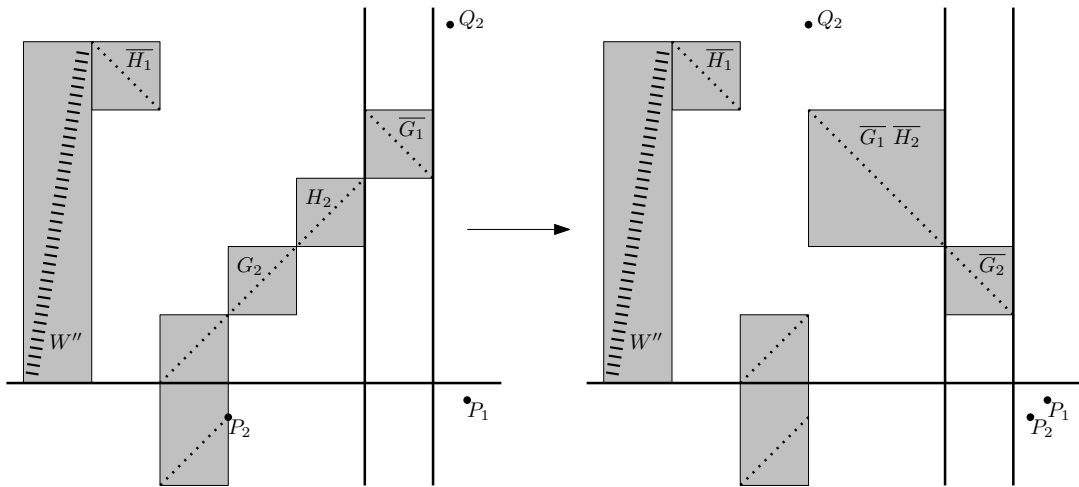


Figure 19: Step 3, part 2.

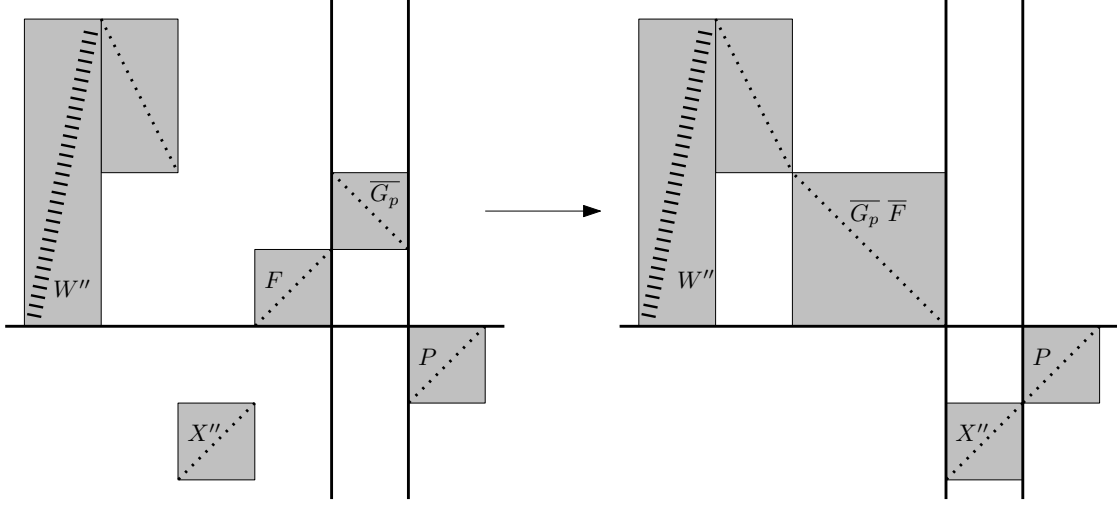


Figure 20: Step 3, part 3.

For $l \geq 0$, let $\beta'_l = \frac{d^{l+1}}{3T}$ and $\alpha'_l = 2d^lT + d$. In particular, $\beta'_k = \frac{md}{3T} \leq \left\lfloor \frac{md-T}{T+4t+2} \right\rfloor = p$ and $\alpha'_k = 2mT + d$, where the inequality holds since $d \geq 9T$. We have $\text{width}(B_{k+1}^+) \leq \text{width}(C) + \text{width}(B'^+) \leq \alpha'_k + d\alpha_k = \alpha_{k+1}$, proving (6). We also have that $|B_{k+1}^-| = |P| + |B'^-| \geq \beta'_k + d\beta_k = \beta_{k+1}$, proving (7). To prove (8), we need the following claim.

Claim. For all $l \geq 0$, $\beta'_l/\alpha'_l > \beta_{l+1}/\alpha_{l+1} > \beta_l/\alpha_l$.

Proof. We will show by induction on l that $\beta'_l/\alpha'_l > \beta_l/\alpha_l$ and the full result will fall out as a consequence. The base case $l = 0$ is trivial, since $\beta_0 = 0$. To show the result for $l + 1$, note that, since $\frac{\beta_{l+1}}{\alpha_{l+1}} = \frac{\beta'_l + d\beta_l}{\alpha'_l + d\alpha_l}$ and $\beta'_l/\alpha'_l > \beta_l/\alpha_l$, we have $\beta'_l/\alpha'_l > \beta_{l+1}/\alpha_{l+1} > \beta_l/\alpha_l$. But $\frac{\beta'_{l+1}}{\alpha'_{l+1}} = \frac{d^{l+2}}{3T(2d^{l+1}T+d)} \geq \frac{d^{l+1}}{3T(2d^lT+d)} = \frac{\beta'_l}{\alpha'_l}$, so that $\beta'_{l+1}/\alpha'_{l+1} > \beta_{l+1}/\alpha_{l+1}$, as required. \square

Suppose now that J is an initial segment of B_{k+1} . If J is a subblock of P , then $J^+ = \emptyset$. If J contains P and is a subblock of PC , then $|J^-| = p \geq \beta'_k$ and $\text{width}(J^+) \leq \text{width}(C) \leq \alpha'_k$, so, by the claim, $|J^-| \geq (\beta_{k+1}/\alpha_{k+1}) \text{width}(J^+)$. Suppose then that $J = PCJ'$, where J' is a non-empty initial segment of B' . Since B' is (β_k/α_k) -balanced, we have $|J'^-| \geq (\beta_k/\alpha_k) \text{width}(J'^+)$. But $|J^-| \geq \beta'_k + |J'^-|$ and $\text{width}(J^+) \leq \alpha'_k + \text{width}(J'^+)$, so, since $\text{width}(J'^+) \leq \text{width}(B'^+) \leq d\alpha_k$, we have

$$\frac{|J^-|}{\text{width}(J^+)} \geq \frac{\beta'_k + (\beta_k/\alpha_k) \text{width}(J'^+)}{\alpha'_k + \text{width}(J'^+)} \geq \frac{\beta'_k + d\beta_k}{\alpha'_k + d\alpha_k} = \frac{\beta_{k+1}}{\alpha_{k+1}}.$$

Note that here, in the second inequality, we used that, since $\beta'_k/\alpha'_k > \beta_k/\alpha_k$, the function $(\beta'_k + (\beta_k/\alpha_k)x)/(\alpha'_k + x)$ is decreasing in x . Since B_{k+1}^- is increasing, this implies that B_{k+1} is $(\beta_{k+1}/\alpha_{k+1})$ -balanced, proving (8).

Finally, we give bounds on $|X|$ and $|Y|$. Suppose, for given d, n, k , that the construction gives $|X| = x_{d,n,k}$ and $|Y| = y_{d,n,k}$. Then $x_{d,n,0} = n + 1$ and $y_{d,n,0} = T + 4t + 3 + n$. Moreover, following

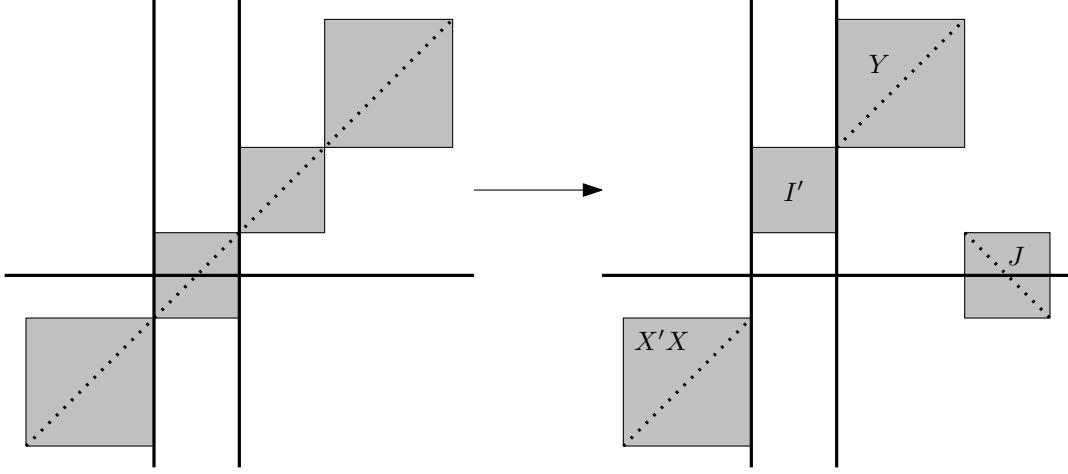


Figure 21: Starting the construction.

the construction, we have that, for $k \geq 0$,

$$\begin{aligned} x_{d,n,k+1} &= dx_{d,n+1,k} + p + 2t + 1 \leq dx_{d,n+1,k} + \frac{d^{k+1}}{2T} + 2t + 1 \\ y_{d,n,k+1} &= dy_{d,n+1,k} + m(T + d + 4t + 2) + p \leq dy_{d,n+1,k} + 2d^{k+1}, \end{aligned}$$

so we get the (loose) bounds $x_{d,n,k}, y_{d,n,k} \leq 10d^{2k+1}n$. \square

4 The full construction

We now proceed to the full construction. Set $T = 3^{2t}$. Start with the identity centred sequence $\mathbf{I} : [a, b] \rightarrow \mathbb{Z}$ for suitable $a < 0 < b$ to be decided later. Perform the flip $[-t, 3t + 1]$ to obtain

$$X'X\mathbf{I}'JY,$$

where $\mathbf{I}' : [-t, t] \rightarrow \mathbb{Z}$, $|J| = 2t + 1$ and X', X are of suitable lengths to be decided later. J is an important piece that we will have to bring back to the centre at the end. For now, we set it aside by moving it to the right to get

$$X'X\mathbf{I}'YJ$$

(see Figure 21).

Applying Lemma 3.5 to $X\mathbf{I}'Y$ with $n = 1$ and some d (with suitably chosen a, b), we can go to

$$X'L_kW_k\mathbf{A}_kB_kR_kJ,$$

where B_k is $r = (\beta_k/\alpha_k)$ -balanced.

We now show that for a suitable choice of d and k , we can make r arbitrarily large. Indeed, we have $\alpha_1 \leq 3dT$,

$$\alpha_{i+1} \leq d\alpha_i + 3Td^i$$

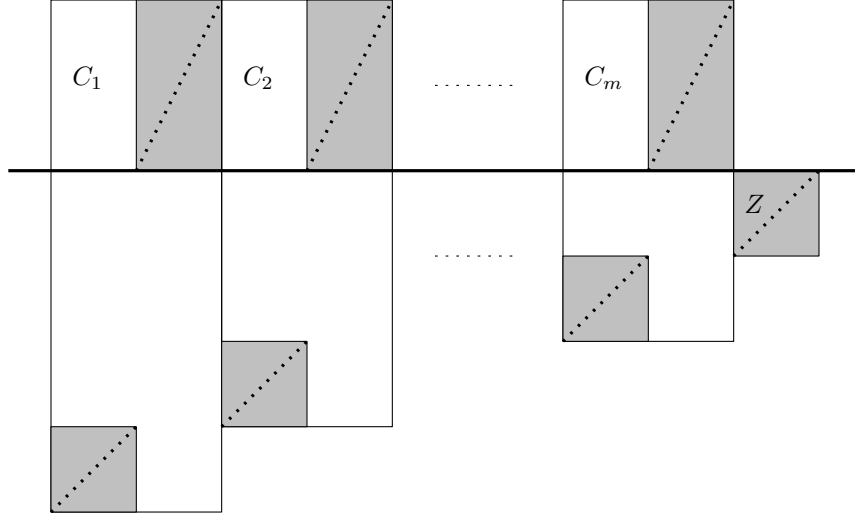


Figure 22: Transforming B_k .

for $i \geq 1$ and

$$\beta_{i+1} \geq d\beta_i + \frac{d^{i+1}}{3T},$$

so that $\alpha_k \leq 3Td^k + 3kTd^{k-1}$ and $\beta_k \geq \frac{k}{3T}d^k$. Hence, $r \geq \frac{k}{3T(3T+3kT/d)}$. Setting $k = d$, we see that $r \geq \frac{d}{18T^2}$ can be taken arbitrarily large. In practice, we will set $d = 100T^3$, so that $r \geq 3T + 1$.

By Lemma 3.4, we can go from B_k to $C_1 \cdots C_{m-1}C'_m$ where each block C_i and C'_m is increasing and $|C_i^-|, |C'_m| \geq \lfloor r \rfloor \geq 3T$. We can decompose C'_m into $C''_m Z C'''_m$, where $C''_m \prec Z \prec 0 \prec C'''_m$, $|Z| \geq T$ and $|C''_m| \geq 2T$. We can go from $C''_m Z C'''_m$ to $C''_m C'''_m Z$. Now set $C_m = C'_m C'''_m$. Thus, we have gone from B_k to $C_1 \cdots C_m Z$, where $C_1^- \prec C_2^- \prec \cdots \prec C_m^- \prec Z \prec 0$ and $|C_i^-| \geq 2T$ for each i (see Figure 22).

For each i , decompose C_i into $D_i E_i F_i$, where $D_i \prec E_i \prec 0 \prec F_i$ and $|E_i| = 2t + 1$, so that $|D_i| \geq T + 2t + 1$. Decompose X' as $X' = X_m \cdots X_1$ with $|X_i| = |F_i|$, thus fixing the value of a .

Note that $\mathbf{A}_k \prec C_1$. By bringing X_1 towards the centre and reflecting F_1 , we can go from $X_1 \mathbf{A}_k D_1 E_1 F_1$ to

$$\overline{F_1} \overline{\mathbf{E}_1} P_1,$$

for some decreasing $P_1 \prec E_1$ (see Figure 23).

Note that P_1 consists of elements from X_1 , A_k and D_1 , so $P_1 \prec ZJ$. Set P_1 aside by moving it all the way to the right, just before Z . Repeating this process for all the C_i , we can go from

$$X' L_k W_k \mathbf{A}_k B_k R_k J$$

to

$$L_k W_k \overline{F_1} \overline{F_2} \cdots \overline{F_m} \overline{\mathbf{E}_m} P_m \cdots P_1 Z R_k J.$$

Bringing ZJ to the left, we can go to

$$L_k W_k \overline{F_1} \cdots \overline{F_m} \overline{\mathbf{E}_m} Z J P_m \cdots P_1 R_k$$

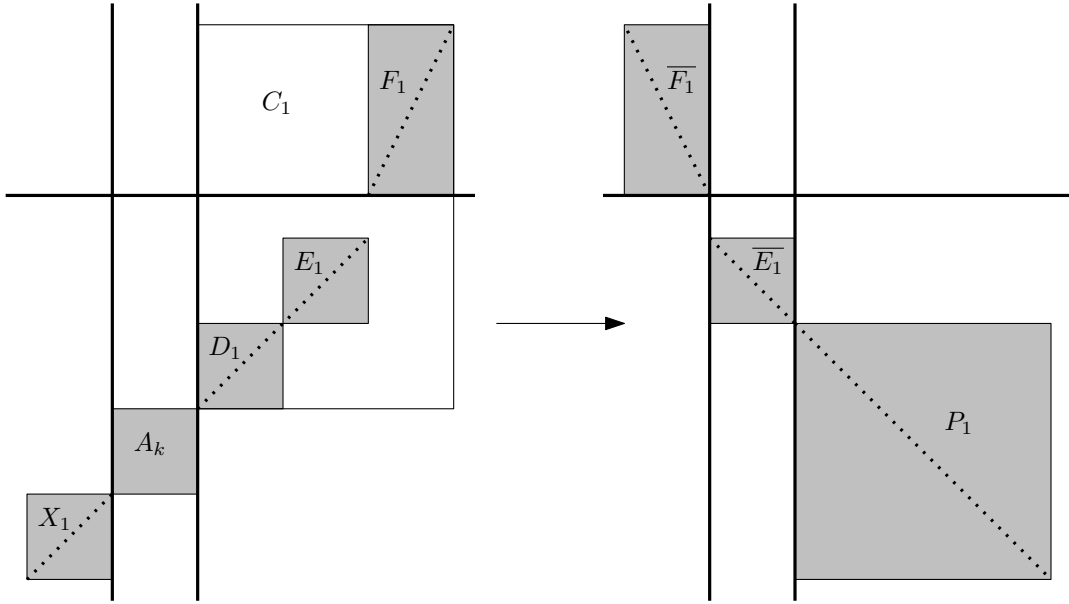


Figure 23: Moving C_1^+ across.

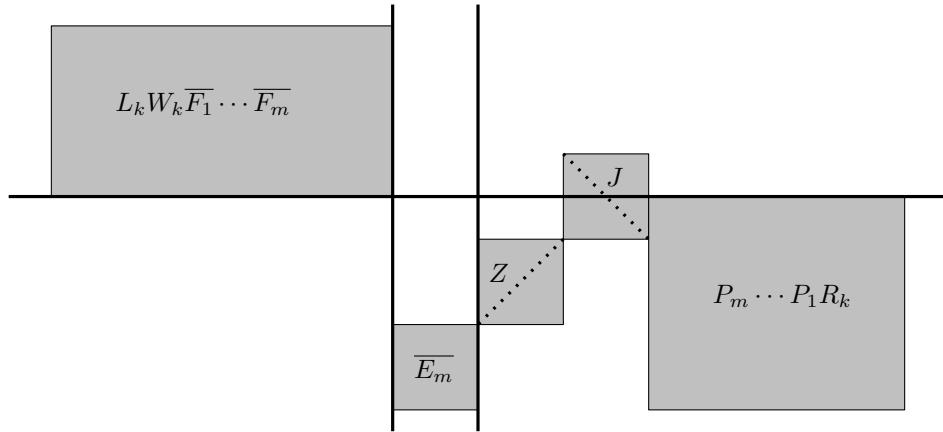


Figure 24: Almost done!

(see Figure 24).

Finally, shift J to go from $\overline{\mathbf{E}_m}ZJ$ to $\mathbf{J}Z'$, where $Z' \prec 0$. Now sort $L_k W_k \overline{F_1} \cdots \overline{F_m}$ and $Z' P_m \cdots P_1 R_k$ in decreasing order to get the reverse of the identity. Since this must be a centred sequence $[a, b] \rightarrow \mathbb{Z}$, we have $a = -b$. Moreover, we have $b = 3t + 1 + |Y|$ and, by Lemma 3.5, $|Y| \leq 10d^{2k+1}n$ with $k = d = 100T^3$ and $n = 1$, giving $b = 2^{2^{O(t)}}$, as required.

5 Concluding remarks

The main problem left open by our work is whether there are configurations like ours which can be realised by lines. In the literature, such configurations are called *stretchable*. While we suspect that our configurations, as given, are not stretchable, we firmly believe, and conjecture below, that there should be stretchable configurations, perhaps even suitable variants of our construction, where the number of points on either side of each line differs by at least k for any given natural number k .

Conjecture 5.1. *There exists an unbounded function $f : \mathbb{N} \rightarrow \mathbb{N}$ such that for every natural number n there is a set of n points for which the number of points on either side of each line determined by the set differ by at least $f(n)$.*

It may also be that we can take $f(n) = c \log \log n$ for some $c > 0$, matching our Theorem 1.2. However, at present, we do not even know how to find point sets where the number of points on either side of each line determined by the set differ by at least 3.

Our Theorem 1.2 also implies that another result of Pinchasi is essentially best possible, in that any improvement must somehow take into account the fact that one is dealing with lines rather than pseudolines. If we define $f(k)$ to be the maximum size of a finite point set in the plane, not contained in a line, with the property that there is no line through at least two of these points with at least k points on either side of this line, then Pinchasi [11] showed that $f(k) < 2k + C \log \log k$ for some absolute constant C . Once again, this follows from a statement about allowable sequences and therefore holds in the broader context of generalised configurations. Moreover, in this form, the result easily implies Theorem 1.1, since if $n = 2k + C \log \log k$, then either every point is on some pseudoline, in which case we are done, or there exists a pseudoline with at least k points on either side and the number of points on each side differ by at most $C \log \log k$. By the same argument, any improved bound for this result would give an analogous improvement to Theorem 1.1, but our Theorem 1.2 shows that this is already best possible.

One may also study analogues of Kupitz's question in higher dimensions. In three dimensions, the most natural analogue is whether there is a fixed natural number k such that every finite point set has a plane through at least three of its points where the number of points on either side of this plane differ by at most k . Curiously, we do not even know the answer under the far less prescriptive condition that the plane only pass through at least two points. By a simple projection argument, we know that an analogue of Pinchasi's result holds in this context, namely, that if our point set has n points, then there is a plane through at least two points where the number of points on either side of this plane differ by at most $C \log \log n$. However, we do not expect this to be tight.

References

- [1] N. Alon, Research Problems, *Discrete Math.* **257** (2002), 599–624.

- [2] P. Brass, W. Moser and J. Pach, **Research problems in discrete geometry**, Springer, New York, 2005.
- [3] P. Erdős, Some old and new problems in combinatorial geometry, in Convexity and graph theory (Jerusalem, 1981), 129–136, North-Holland Math. Stud., 87, Ann. Discrete Math., 20, North-Holland, Amsterdam, 1984.
- [4] S. Felsner and J. E. Goodman, Pseudoline arrangements, in Handbook of discrete and computational geometry, 3rd edition, 125–157, CRC Press Ser. Discrete Math. Appl., CRC, Boca Raton, FL, 2018.
- [5] J. E. Goodman and R. Pollack, On the combinatorial classification of nondegenerate configurations in the plane, *J. Combin. Theory Ser. A* **29** (1980), 220–235.
- [6] J. E. Goodman and R. Pollack, Semispaces of configurations, cell complexes of arrangements, *J. Combin. Theory Ser. A* **37** (1984), 257–293.
- [7] B. Green, 100 open problems, unpublished manuscript.
- [8] G. Kalai, A discrepancy problem for planar configurations, blog post available at <https://gilkalai.wordpress.com/2010/02/03/a-discrepancy-problem-for-planar-configurations/>.
- [9] Y. S. Kupitz, **Extremal problems in combinatorial geometry**, Lecture Notes Series, 53, Aarhus Universitet, Matematisk Institut, Aarhus, 1979.
- [10] J. Pach, Finite point configurations, in Handbook of discrete and computational geometry, 3rd edition, 3–25, CRC Press Ser. Discrete Math. Appl., CRC, Boca Raton, FL, 2018.
- [11] R. Pinchasi, Lines with many points on both sides, *Discrete Comput. Geom.* **30** (2003), 415–435.