

MONOTONE SUBSETS IN LATTICES AND THE SCHENSTED SHAPE OF A SÓS PERMUTATION

Karl Liechty^{*1} and T. Kyle Petersen^{†2}

^{1,2}*Department of Mathematical Sciences, DePaul University, Chicago, IL, U.S.A.*
klichthy@depaul.edu, t.kyle.petersen@depaul.edu

Submitted: Aug 4, 2021; Accepted: May 12, 2022; Published: Jun 30, 2022

© The authors. Released under the CC BY license (International 4.0).

Abstract. For a fixed irrational number α and $n \in \mathbb{N}$, we look at the shape of the sequence $(f(1), \dots, f(n))$ after Schensted insertion, where $f(i) = \alpha i \bmod 1$. Our primary result is that the boundary of the Schensted shape is approximated by a piecewise linear function with at most two slopes. This piecewise linear function is explicitly described in terms of the continued fraction expansion for α . Our results generalize those of Boyd and Steele, who studied longest monotone subsequences. Our proofs are based on a careful analysis of monotone sets in two-dimensional lattices.

Keywords. Longest increasing subsequence, Schensted shape, geometry of numbers, Sós permutations

Mathematics Subject Classifications. 05A05, 11H06, 11B57, 11K06

1. Introduction

Fix a real number α , and let $f = f_\alpha(i) = \alpha i \bmod 1 = \alpha i - \lfloor \alpha i \rfloor$ denote the function that returns the fractional part of multiples of α . It suffices to restrict attention to $\alpha \in (0, 1)$, since $f_\alpha = f_{\alpha+k}$ for any integer k . Now, for any positive integer n we can consider the sequence $(f(1), \dots, f(n))$. These sequences have been studied quite a bit since the 1950's when Steinhaus made the following conjecture: The points $\{e^{2\pi i f(j)}\}_{j=1}^n$ divide the unit circle into $(n-1)$ pieces of at most 3 distinct lengths. This conjecture was proved shortly thereafter by Sós [S58], Surányi, [Sur58] and Świerczkowski [S59], independently of one another. This result, known as the Three Gaps Theorem, has since been proven many times by many different methods. It has found applications in and connections to such disparate places as quantum mechanics [Ble90, Ble91], plant growth [vR86], combinatorics [AB98], music theory [Nar17], Riemannian geometry [BS08], and of course number theory [Har52, S58, vR88].

^{*}Supported by Simons Foundation Collaboration Grant #357872.

[†]Supported by Simons Foundation Collaboration Grant #353772.

In this paper we study a somewhat coarser object: the permutation induced by the list $(f(1), \dots, f(n))$. Let $w = w(n, \alpha)$ denote the sorting permutation for this list of numbers, i.e., the lexicographically first permutation such that $f(w(1)) \leq \dots \leq f(w(n))$. We call such a permutation a *Sós permutation*, following [BCKPT21]. (Actually, [BCKPT21] studies the slightly larger set of permutations generated by fractional parts of lines $f_{\alpha, \beta}(i) = \alpha i + \beta \pmod{1}$, depending on two independent real parameters. But, as discussed in Remark 3.12 at the end of Section 3.4, it will be enough for our purposes to focus on the $\beta = 0$ case.) Let Sós_n denote the set of all such permutations. That is, for fixed n , let

$$\text{Sós}_n = \{\pi \in S_n : \pi = w(n, \alpha) \text{ for some } \alpha \in (0, 1)\}.$$

These permutations satisfy their own Three Gap Property, as was noticed by Sós [S58]. See Equation (1.1) below.

Our goal is to characterize the *Schensted shape* of these permutations.

1.1. The shape of a permutation

The phrase “Schensted shape” comes from a bijection known as *Schensted insertion* or more generally, the *Robinson–Schensted–Knuth correspondence* (RSK for short). This is a bijection between the set of permutations S_n and ordered pairs of *standard Young tableaux*. Drawn in the so-called French style, these tableaux are lower-left justified arrays of n boxes filled with the integers $1, \dots, n$ such that the numbers increase across rows and up columns. If we read the lengths of the rows of boxes, we get a partition of n , which we call the *shape* of the tableaux. The precise definition of the correspondence between permutations and pairs of tableaux is not important for now. The important thing to know is that the correspondence gives each permutation a well-defined shape in the form of an integer partition. For the permutation w , we write $\text{sh}(w)$ for its shape. A feature of Schensted insertion is that if $w \mapsto (P, Q)$, then $w^{-1} \mapsto (Q, P)$. In particular, $\text{sh}(w) = \text{sh}(w^{-1})$, and when we come to prove our main results it will be convenient to work with w^{-1} rather than w .

For example, if $\alpha = .3$ and $n = 7$, we have

$$(f(1), f(2), f(3), f(4), f(5), f(6), f(7)) = (.3, .6, .9, .2, .5, .8, .1),$$

which when sorted is

$$(f(7), f(4), f(1), f(5), f(2), f(6), f(3)) = (.1, .2, .3, .5, .6, .8, .9).$$

Thus $w = 7415263$ is a permutation in Sós_7 . The permutation $w^{-1} = 3572461$ is order-isomorphic to the list $(.3, .6, .9, .2, .5, .8, .1)$ prior to sorting. The image of w under Schensted insertion is shown in Figure 1.1. This pair of tableaux has shape $\text{sh}(w) = (3, 3, 1)$, corresponding to the row lengths of the tableaux.

Let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$ be an integer partition with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m > 0$ and $\sum \lambda_i = n$. Following Romik [Rom15], we define the *planar set* of λ to be the collection of n boxes given by

$$A_\lambda = \bigcup_{\substack{1 \leq k \leq m, \\ 1 \leq l \leq \lambda_k}} [l-1, l] \times [k-1, k],$$

$$7415263 \leftrightarrow \left(\begin{array}{|c|c|c|} \hline 7 & & \\ \hline 4 & 5 & 6 \\ \hline 1 & 2 & 3 \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 3 & & \\ \hline 2 & 5 & 7 \\ \hline 1 & 4 & 6 \\ \hline \end{array} \right)$$

Figure 1.1: A permutation w corresponds to a pair of Young tableaux under the Schensted insertion map. In this example, the shape of w is $\text{sh}(w) = (3, 3, 1)$.

which sits in the first quadrant. We will use the notation ∂A_λ to denote the piece of the boundary of A_λ which is strictly in the first quadrant, i.e., the part of the boundary which is not on one of the coordinate axes. The length of the i th row of A_λ , read from bottom to top, is λ_i . We also let $\lambda' = (\lambda'_1, \dots, \lambda'_{m'})$ denote the conjugate partition, i.e., λ'_i is the height of the i th column of A_λ , read from left to right. This planar set is our concrete realization of the partition $\lambda = \text{sh}(w)$.

Schensted defined his map to study the longest monotone subsequences of permutations (or data strings generally). In particular, [Sch61] shows the *arm*, $\text{arm}(w) = \lambda_1$, is the length of the longest increasing subsequence of w , and the *leg*, $\text{leg}(w) = \lambda'_1$, is the length of the longest decreasing subsequence of w .

As later shown by Greene [Gre74], Schensted’s map allows the following characterization of collections of monotone subsequences. Let I_k denote the size of the largest subsequence formed by the union of k increasing subsequences of w . Equivalently, I_k is the size of the largest subsequence containing no decreasing subsequence of length $k + 1$. In an analogous fashion, let D_k denote the size of the largest subsequence obtained as a union of k decreasing subsequences of w .

Theorem 1.1 (Greene). *For each $k \leq n$,*

$$I_k = \lambda_1 + \dots + \lambda_k \quad \text{and} \quad D_k = \lambda'_1 + \dots + \lambda'_k.$$

Thus, the partition $\text{sh}(w)$ contains all the information one needs about monotone subsequences in w , and a “limit shape theorem” for planar sets A_λ can yield detailed information about the expected number and size of monotone subsequences in random permutations.

The literature abounds with examples of shape theorems for random permutations and related objects, many of which are catalogued and explained in Romik’s book [Rom15]. The most well-known example is the now famous result obtained independently by Vershik and Kerov [VK77] and Logan and Shepp [LS77] that describes a precise limit shape for a uniformly random permutation $w \in S_n$ under the Schensted insertion map. See equation (1.7) below. Our original motivation was the following related question.

Big Question. Given a uniformly random Sós permutation, what is its expected shape?

Unfortunately, at this time we lack the understanding to adequately address this question, though this paper provides some first steps. One issue is that “uniform” could mean two different things here. It could mean:

- *combinatorially uniform*, i.e., for each n we select an element of Sós_n with probability $1/|\text{Sós}_n|$, or
- *geometrically uniform*, i.e., for each n we select α uniformly in $(0, 1)$ and generate $w(n, \alpha)$.

As explained in [BCKPT21], a result of Surányi shows that Sós permutations are in bijection with *Farey intervals*. For fixed n , a Farey interval is of the form $(a/b, c/d)$, $0 \leq a/b < c/d \leq 1$, where a/b and c/d are consecutive reduced fractions of denominator at most n . Surányi’s result says the following.

Theorem 1.2 (Surányi’s bijection). *There is a bijection between Farey intervals and Sós permutations. That is, $w(n, \alpha) = w(n, \alpha')$ if and only if α and α' lie in the same Farey interval.*

Thus, by counting Farey intervals we see the number of Sós permutations is

$$|\text{Sós}_n| = \sum_{k \leq n} \varphi(k) \underset{n \rightarrow \infty}{\sim} \frac{3n^2}{\pi^2},$$

where φ is the Euler totient function (see, e.g., [HW08, Section 18.5]). We see the set Sós_n represents a very small subset of full symmetric group S_n , and it is not surprising that the permutations in Sós_n are highly structured. Indeed, it was shown by Sós [S58, Theorem 1] that for any $\pi \in \text{Sós}_n$ corresponding to the Farey interval $(a/b, c/d)$, we have $\pi(1) = b$ and

$$\pi(i+1) - \pi(i) = \begin{cases} b & \text{if } \pi(i) \leq n - b \\ b - d & \text{if } n - b < \pi(i) < d \\ -d & \text{if } d \leq \pi(i). \end{cases} \quad (1.1)$$

Further, Surányi’s result shows how to compute the geometrically uniform distribution on Sós_n . If $\pi = w(n, \alpha)$ with $a/b \leq \alpha < c/d$, then the geometric probability of π is $c/d - a/b = 1/bd$. In particular, the identity permutation has probability $1/n$ in the geometrically uniform distribution.

As the set Sós_n is so highly structured, we suspect that the distribution of shapes over this set is quite different from the case of the full symmetric group. But without any clear tools to analyze the full distribution, we shift to a related question.

Narrow Question. Does the shape of a typical Sós permutation resemble the shape of a uniformly random permutation?

The answer to the Narrow Question is an emphatic “**no**” because, as we will see stated precisely in Theorem 1.7:

The Schensted shape of a Sós permutation has a boundary that is approximately piecewise linear, with at most two slopes.

Unpacking this claim is the main objective of the current paper. We hope someone can produce more insight on the Big Question in future work.

1.2. Arms and legs for $w(n, \alpha)$

Results concerning monotone subsequences from finite strings of real numbers date back at least to the work of Erdős and Szekeres [ES35], who proved that any sequence of $n^2 + 1$ distinct real

numbers has a monotone subsequence of length $n + 1$. For uniformly random permutations in all of S_n , the Logan–Shepp/Vershik–Kerov limit shape theorem implies that both the longest increasing subsequence, $\text{arm}(w)$, and longest decreasing subsequence, $\text{leg}(w)$, are asymptotic to $2\sqrt{n}$ as $n \rightarrow \infty$. Actually the problem of the asymptotic length of the longest increasing subsequence in a random permutation, known as Ulam’s problem, is a fascinating story which predates [LS77, VK77] by about 15 years. We refer interested readers to the book [Rom15] and the survey article [Ste95].

In contrast with the behavior of uniformly random permutations in S_n , the papers by Boyd and Steele [BS79] and del Junco and Steele [dJS78] found an interesting non-convergent behavior for the Sós permutations $w(n, \alpha)$. That is, while the paper [dJS78] found

$$\frac{\log(\text{arm}(w(n, \alpha)))}{\log(n)} \rightarrow \frac{1}{2} \quad \text{and} \quad \frac{\log(\text{leg}(w(n, \alpha)))}{\log(n)} \rightarrow \frac{1}{2}$$

for almost all α , the paper [BS79] found there is no α for which either $\text{arm}(w(n, \alpha))/\sqrt{n}$ or $\text{leg}(w(n, \alpha))/\sqrt{n}$ tends to a limit!

The results of [BS79] imply that as a function of n , the quantity $\text{arm}(w(n, \alpha))/\sqrt{n}$ oscillates between local maxima and minima that can be precisely described in terms of the principal convergents in the continued fraction expansion of α , and as a consequence $\limsup \text{arm}(w(n, \alpha))/\sqrt{n}$ is finite if and only if the terms of the continued fraction are bounded. In this paper, we will reprove the results of Boyd and Steele in the process of extending them to a tight approximation for the Schensted shape of $w(n, \alpha)$. Before we present the precise version of this approximation, it is helpful to understand Boyd and Steele’s results in more detail.

We first introduce some notation for continued fractions. (See [HW08] for a general reference on the topic.) For a real number x , its (simple) continued fraction expansion is defined recursively via

$$a_0 = \lfloor x \rfloor, \quad x_0^* = \frac{1}{x - a_0},$$

and

$$a_k = \lfloor x_{k-1}^* \rfloor, \quad x_k^* = \frac{1}{x_{k-1}^* - a_{k-1}},$$

provided x_{k-1}^* is not an integer. We write

$$x = [a_0; a_1, a_2, \dots] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}}.$$

If for some k , x_{k-1}^* is an integer, then a_k is the final term in the sequence, and x is rational.

For a fixed real number $\alpha = [a_0; a_1, a_2, \dots]$, we have a sequence of rational numbers

$$\frac{p_k}{q_k} = [a_0; a_1, a_2, \dots, a_k], \tag{1.2}$$

known as the *principal convergents* to α . We denote by $\delta_k = |\alpha - p_k/q_k|$ the distance between α and its k th principal convergent, and let $\beta_k = \delta_k^{-1}$ denote the reciprocal of this distance. A fundamental result about approximating with convergents is the approximation (see, e.g., [HW08,

Theorem 164])

$$\delta_k < \frac{1}{q_{k+1}q_k} < \frac{1}{q_k^2}. \quad (1.3)$$

We can now give precise statements of the main results of Boyd and Steele.

Theorem 1.3 (Arm and leg approximations, see Theorem 1 and Corollary 1 of [BS79]). *For $\beta_{2h} \leq n < \beta_{2h+2}$,*

$$q_{2h+1}(1 + n\delta_{2h+1}) - 2 < \text{arm}(w(n, \alpha)) \leq q_{2h+1}(1 + n\delta_{2h+1})$$

while for $\beta_{2k-1} \leq n < \beta_{2k+1}$,

$$q_{2h}(1 + n\delta_{2h}) - 2 < \text{leg}(w(n, \alpha)) \leq q_{2h}(1 + n\delta_{2h}).$$

These bounds follow naturally from the approach we take in this paper, and a proof is presented in Sections 3.2 (see Proposition 3.2) and 3.6. Our proof is rather similar to the one in [BS79], with the main difference being that we describe the integer programming problem presented in [BS79, Lemma 3] as a problem of longest increasing paths on lattices. Our description lends itself nicely to a simultaneous treatment of increasing and decreasing subsequences which is necessary to understand the full Schensted shape.

Dividing the expressions in Theorem 1.3 by \sqrt{n} and computing critical points with respect to n , it is easy to obtain the following result.

Theorem 1.4 (Normalized arm and leg extrema, see Theorem 2 of [BS79]). *We have the following local extremes for the normalized arm length $\text{arm}(w(n, \alpha))/\sqrt{n}$ and the normalized leg length $\text{leg}(w(n, \alpha))/\sqrt{n}$.*

1. *When n is the nearest integer to β_{2h} , the normalized arm length achieves a local maximum of approximately*

$$M_{2h}^+ = \frac{1}{q_{2h}\delta_{2h}^{1/2}}$$

and the normalized leg length achieves a local minimum of approximately

$$m_{2h}^- = 2q_{2h}\delta_{2h}^{1/2}.$$

2. *When n is the nearest integer to β_{2h+1} , the normalized arm length achieves a local minimum of approximately*

$$m_{2h+1}^+ = 2q_{2h+1}\delta_{2h+1}^{1/2}$$

and the normalized leg length achieves a local maximum of approximately

$$M_{2h+1}^- = \frac{1}{q_{2h+1}\delta_{2h+1}^{1/2}}.$$

We use the example of $\alpha = e = [2; 1, 2, 1, 1, 4, 1, 1, 6, \dots]$ to illustrate here. In Figure 1.2 we plot $\text{arm}(w(n, e))/\sqrt{n}$ and $\text{leg}(w(n, e))/\sqrt{n}$ to give the reader a feel for the behavior of these sequences. We can see the resemblance of the continued fraction itself reflected in the heights of the peaks and valleys. Note the horizontal scale is logarithmic.

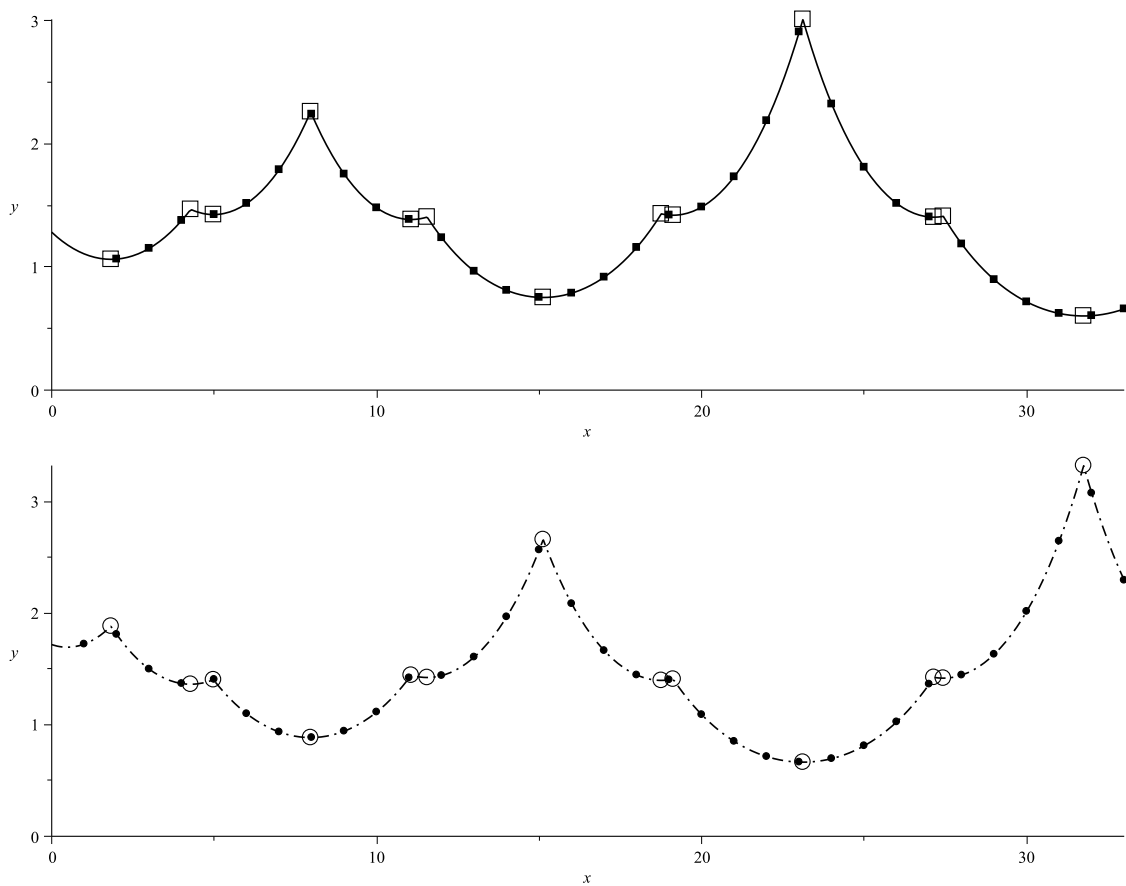


Figure 1.2: Plot of normalized arm and leg lengths for $w(2^n, e)$ for $n \leq 33$. Solid squares mark the points $(n, \text{arm}(w(2^n, e))/\sqrt{2^n})$, while open squares highlight local maxima/minima on the curve. The solid curve predicts the arm length. Solid circles mark the points $(n, \text{leg}(w(2^n, e))/\sqrt{2^n})$, while open circles highlight local maxima/minima on the curve. The dashed curve predicts the leg length.

1.3. Shape of $w(n, \alpha)$

When the arm and leg lengths are close to their local extreme values, one can predict the Schensted shape of $w(n, \alpha)$ using Theorem 1.4 above. A simple situation is when n is approximately equal to β_{2h} . Here, Theorem 1.4 tells us the normalized arm is approximately M_{2h}^+ and the normalized leg is approximately m_{2h}^- . Since, according to Theorem 1.4, their product is $M_{2h}^+ m_{2h}^- = 2$, it is reasonable to predict that the normalized shape is approximated by a triangle with x -intercept M_{2h}^+ and y -intercept m_{2h}^- . The boundary of the shape is thus approximated by the line

$$y = m_{2h}^- - \frac{m_{2h}^-}{M_{2h}^+} x = 2q_{2h}\delta_{2h}^{1/2} - 2q_{2h}^2\delta_{2h}x. \quad (1.4)$$

The case when n is approximately β_{2h+1} is similar. In this case we would guess the shape is a triangle whose boundary is described by

$$y = M_{2h+1}^- - \frac{M_{2h+1}^-}{m_{2h+1}^+} x = \frac{1}{q_{2h+1}\delta_{2h+1}^{1/2}} - \frac{1}{2q_{2h+1}^2\delta_{2h+1}} x. \quad (1.5)$$

These predictions are verified in Theorem 1.7 below.

When n is between β_{2h} and β_{2h+1} , we find the shape of $w(n, \alpha)$ is only slightly more complicated: the boundary is approximated by a piecewise linear function whose two linear components are parallel to the lines (1.4) and (1.5), and with x - and y -intercepts given by the arm and leg measurements predicted by Theorem 1.3. This is our main result, which we have split into two parts.

Theorem 1.5 (Estimating λ and λ'). *Suppose $\alpha \in (0, 1)$ is an irrational number with principal convergents $\{p_i/q_i\}$. Fix $n \in \mathbb{N}$ large enough so that $n > 1/\alpha$ and $n > 1/(1 - \alpha)$. Denote $\lambda = \text{sh}(w(n, \alpha))$, the shape of the permutation $w(n, \alpha)$. With h such that $\beta_{2h} \leq n < \beta_{2h+2}$, define the values x_0 and y_0 to be*

$$x_0 := \frac{1 - n\delta_{2h+1}}{q_{2h+*}(\delta_{2h+1} - \delta_{2h+*})}, \quad y_0 := \frac{1 - n\delta_{2h+*}}{q_{2h+1}(\delta_{2h+*} - \delta_{2h+1})},$$

where

$$* = \begin{cases} 0 & \text{if } \beta_{2h} \leq n < \beta_{2h+1}, \\ 2 & \text{if } \beta_{2h+1} \leq n < \beta_{2h+2}. \end{cases}$$

For all $1 \leq k \leq y_0 - 1$, we have

$$|\lambda_k - q_{2h+1}(1 + n\delta_{2h+1}) + 2kq_{2h+1}^2\delta_{2h+1}| < 4 + 2q_{2h+1}^2\delta_{2h+1} < 6,$$

and for all $1 \leq k \leq x_0 - 1$, we have

$$|\lambda'_k - q_{2h+*}(1 + n\delta_{2h+*}) + 2kq_{2h+*}^2\delta_{2h+*}| < 4 + 2q_{2h+*}^2\delta_{2h+*} < 6.$$

Remark 1.6. In Theorem 1.5 above and Theorem 1.7 below, the technical assumption that $n > 1/\alpha$ and $n > 1/(1 - \alpha)$ simply ensures that $w(n, \alpha)$ is not the identity permutation or its reverse. In these cases, the Schensted shape of $w(n, \alpha)$ is trivial, consisting of a single row or column of length n .

The above theorem indicates that the boundary of the planar set A_λ is very close to one of the two negatively sloped lines

$$y = q_{2h+*}(1 + n\delta_{2h+*}) - 2q_{2h+*}^2\delta_{2h+*}x \quad \text{and} \quad x = q_{2h+1}(1 + n\delta_{2h+1}) - 2q_{2h+1}^2\delta_{2h+1}y.$$

It turns out that these two lines intersect exactly at the point (x_0, y_0) (see equations (3.20) and (3.21) in Section 3.5) and so the boundary of A_λ is uniformly close to the piecewise linear function

$$L(n, \alpha; x) = \begin{cases} q_{2h+*}(1 + n\delta_{2h+*}) - 2q_{2h+*}^2\delta_{2h+*}x & \text{for } 0 \leq x \leq x_0, \\ \frac{q_{2h+1}(1+n\delta_{2h+1})-x}{2q_{2h+1}^2\delta_{2h+1}} & \text{for } x_0 \leq x \leq q_{2h+1}(1 + n\delta_{2h+1}). \end{cases} \quad (1.6)$$

We capture this idea in our next main result, where for fixed x , $\text{dist}(L(n, \alpha; x), \partial A_\lambda)$ denotes the minimum Euclidean distance from $(x, L(x))$ to a point on ∂A_λ .

Theorem 1.7 (Two Slope Theorem). *For a fixed irrational number $\alpha \in (0, 1)$ and $n \in \mathbb{N}$ satisfying $n > 1/\alpha$ and $n > 1/(1 - \alpha)$, let λ and h be as in Theorem 1.5, and let $L(n, \alpha; x)$ be the piecewise linear function defined in (1.6). We have the uniform estimate*

$$\text{dist}(L(n, \alpha; x), \partial A_\lambda) < 8, \quad \text{for all } 0 \leq x \leq q_{2h+1}(1 + n\delta_{2h+1}).$$

For example if $\alpha = e$, let us take $n = 4700$. We have

$$\delta_6 = e - \frac{p_6}{q_6} = .0003331 \dots > \frac{1}{4700} = .0002127 \dots > .00002803 \dots = \frac{p_7}{q_7} - e = \delta_7,$$

so $\beta_6 < 4700 < \beta_7$. We compute the slopes as

$$-2q_6^2\delta_6 = -1.01332 \dots \quad \text{and} \quad -\frac{1}{2q_7^2\delta_7} = -3.53850 \dots$$

It is useful to contrast the shape described in Theorem 1.7 with the Logan–Shepp/Vershik–Kerov limit shape for a uniformly random permutation in S_n . They showed that with high probability, after rescaling by a factor of $1/\sqrt{n}$, the boundary of the corresponding Young diagram for a uniformly random permutation is approximated very well by the curve

$$x + y = \frac{2}{\pi} \left((x - y) \sin^{-1} \left(\frac{x - y}{2} \right) + \sqrt{4 - (x - y)^2} \right), \quad 0 \leq x \leq 2, 0 \leq y \leq 2. \quad (1.7)$$

This curve lies in stark contrast to the piecewise linear boundary for a Sós permutation described in Theorem 1.7. See Figure 1.3.

Remark 1.8. The uniform bounds presented in Theorems 1.5 and 1.7 are not optimal. They can be improved, even by the methods of this paper. Numerical examples such as Figure 1.3 indicate a very tight bound which seems to be less than 2. Since we were not able to prove a bound close to this, we focused on readability rather than obtaining the sharpest bounds possible, reasoning that not much is gained by a marginally better bound which is still not optimal.

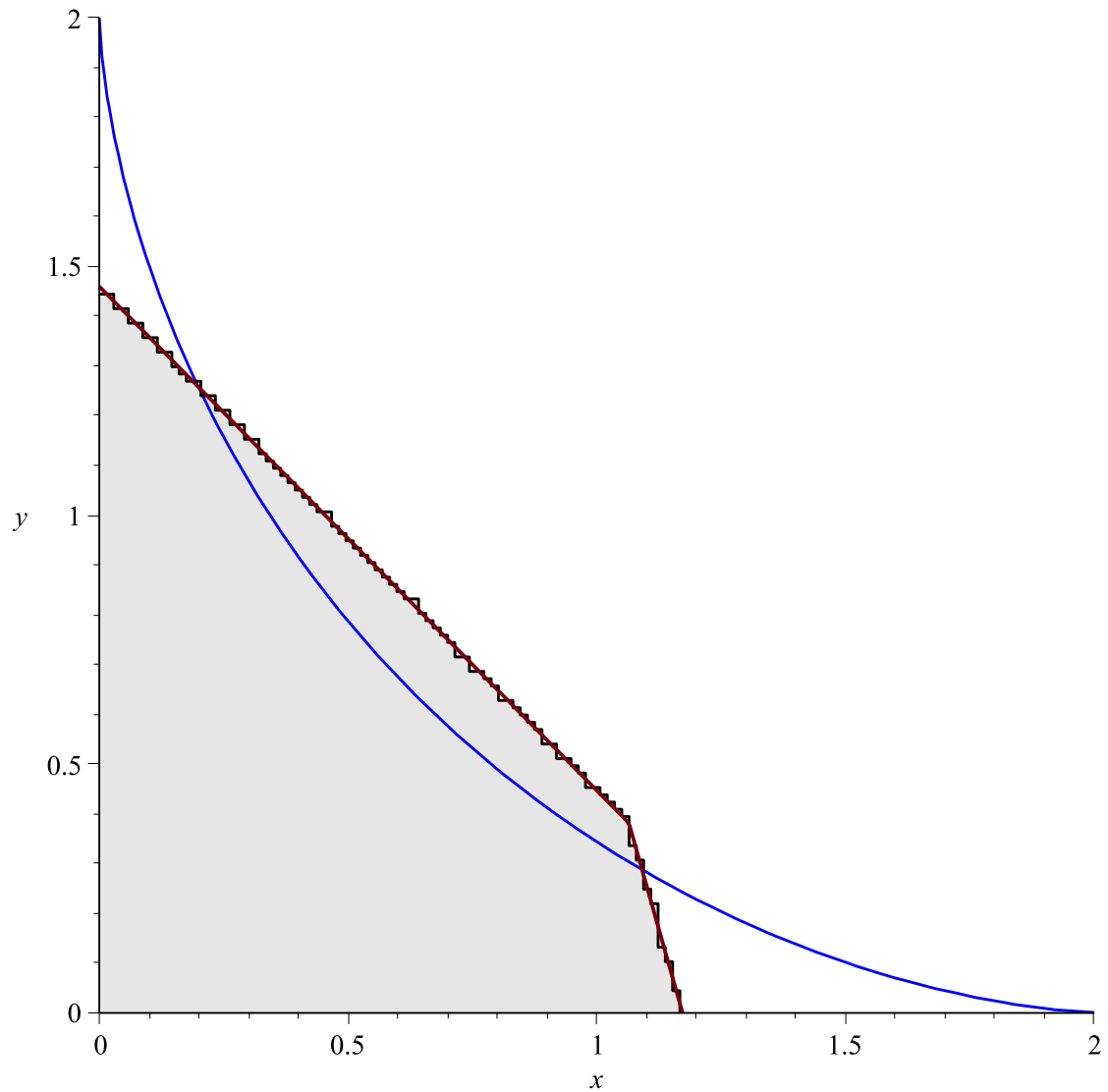


Figure 1.3: The planar set of $w(4700, e)$, with boundary approximated by $L(4700, e; x)$ in red. Here the axes are scaled by a factor of $1/\sqrt{n}$, and for contrast we have superimposed the Logan–Shepp/Vershik–Kerov limit shape for a uniformly random permutation in blue.

1.4. Plan for the rest of the paper

We will prove our main theorems (Theorems 1.3, 1.5, and 1.7) as follows. In Section 2, we will translate our problem of monotone subsequences in a Sós permutation to the problem of monotone paths in an integer lattice. Next, in Section 3, we will characterize the shape of permutations given by the lattice $\{(i, a \cdot i \bmod N) : i = 1, \dots, n\}$, for any integers $0 < a < N$ and $0 < n < N$ with a and N relatively prime. We show in Section 3.6 how this case of rational numbers implies our results for any irrational α , which will complete the proofs.

Finally, in Section 4, we describe a few further directions this research could take.

2. Increasing and decreasing lattice paths

In this section we investigate the geometric interpretation of monotone subsequences of a permutation w via *increasing subsets* of the points $\{(i, w(i))\}$. This is an old idea, which (as explained in [Rom15]) goes back to Hammersley [Ham72]. For Sós permutations we have the special advantage that the set $\{(i, f_\alpha(i))\}$ is a lattice, which, up to vertical scaling, is the same as $\{(i, w^{-1}(i))\}$. As mentioned, $\text{sh}(w) = \text{sh}(w^{-1})$, so it suffices to study increasing subsets in this set of points. The remainder of this section lays out all the details of our approach to increasing subsets in the context of lattices.

2.1. Lattices and monotone lattice walks

In general, a *two-dimensional integer lattice* L is the set of all integer linear combinations of two linearly independent vectors \mathbf{a} and \mathbf{b} in \mathbb{R}^2 , denoted $L = \mathbb{Z}[\mathbf{a}, \mathbf{b}] = \{k \mathbf{a} + l \mathbf{b} : k, l \in \mathbb{Z}\}$. A *lattice vector* refers to the displacement vector between any two points in L . We write points in L as ordered pairs (a, b) , while vectors are given in angle brackets, e.g., $\langle a, b \rangle$. By abuse of notation we use boldface letters for both points and vectors. The meaning will usually be clear from context, and we will alert the reader when care must be taken.

Given a vector $\mathbf{x} = \langle x_1, x_2 \rangle$, we denote the slope of the vector by

$$m_{\mathbf{x}} := x_2/x_1,$$

provided $x_1 \neq 0$. (If $x_1 = 0$ we write $m_{\mathbf{x}} = \pm\infty$ according to whether the second component x_2 is positive or negative.) We will call a non-zero vector \mathbf{x} an *increasing vector* if both components of \mathbf{x} are non-negative; we will call \mathbf{x} a *decreasing vector* if $x_1 \geq 0$ and $x_2 \leq 0$. Note that under this definition, the vector $\langle x_1, 0 \rangle$ is both increasing and decreasing whenever $x_1 > 0$. A sequence of increasing lattice vectors $\mathbf{x}_1, \dots, \mathbf{x}_k$ is called an *increasing walk*, and a sequence of decreasing lattice vectors $\mathbf{y}_1, \dots, \mathbf{y}_j$ is called a *decreasing walk*. If we think of each vector in the sequence as a step on the lattice beginning at the origin, then the increasing walk $\mathbf{x}_1, \dots, \mathbf{x}_k$ begins at $(0, 0)$ and ends at $\mathbf{x} = \sum_{i=1}^k \mathbf{x}_i$ in the first quadrant, while the decreasing walk $\mathbf{y}_1, \dots, \mathbf{y}_j$ begins at $(0, 0)$ and ends at $\mathbf{y} = \sum_{i=1}^j \mathbf{y}_i$ in the fourth quadrant.

Conversely, given any lattice point \mathbf{x} in the first quadrant, there are a number of increasing

walks from $(0, 0)$ to \mathbf{x} . Denote the set of such walks by

$$\text{walk}^+(\mathbf{x}) = \left\{ (\mathbf{x}_1, \dots, \mathbf{x}_k) : \mathbf{x}_i \text{ are all increasing vectors and } \sum_{i=1}^k \mathbf{x}_i = \mathbf{x} \right\}.$$

Similarly, for any lattice point \mathbf{y} in the fourth quadrant we define the collection of decreasing walks from $(0, 0)$ to \mathbf{y} taking negatively sloped steps:

$$\text{walk}^-(\mathbf{y}) = \left\{ (\mathbf{y}_1, \dots, \mathbf{y}_j) : \mathbf{y}_i \text{ are all decreasing vectors and } \sum_{i=1}^j \mathbf{y}_i = \mathbf{y} \right\}.$$

The *increasing lattice length* of a vector \mathbf{x} in L is the greatest number of steps in an increasing walk to \mathbf{x} :

$$\ell^+(\mathbf{x}) = \max\{k : (\mathbf{x}_1, \dots, \mathbf{x}_k) \in \text{walk}^+(\mathbf{x})\}.$$

We similarly define the *decreasing lattice length* of \mathbf{y} :

$$\ell^-(\mathbf{y}) = \max\{k : (\mathbf{y}_1, \dots, \mathbf{y}_k) \in \text{walk}^-(\mathbf{y})\}.$$

Globally, for a finite subset S of a lattice, we denote the maxima of these quantities by

$$\ell^+ = \ell^+(S) = \max\{\ell^+(\mathbf{x}) : \mathbf{x} \in S\} \quad \text{and} \quad \ell^- = \ell^-(S) = \max\{\ell^-(\mathbf{y}) : \mathbf{y} \in S\}.$$

It will be important to understand lattice length in terms of vectors with lattice length one. A helpful way to characterize such vectors is in terms of their *shadow*. For a lattice vector $\mathbf{x} = \langle x, y \rangle$, define the set $\text{shadow}(\mathbf{x})$ to be the set of nonzero lattice points in the rectangle spanned by $\langle x, 0 \rangle$ and $\langle 0, y \rangle$, not including \mathbf{x} itself. The key feature of shadows is the following.

Observation 2.1 (Empty shadows). *Let $\mathbf{x} \in L$. We have $\text{shadow}(\mathbf{x}) = \emptyset$ if and only if \mathbf{x} has lattice length one.*

Proof. For simplicity, assume \mathbf{x} is an increasing lattice vector. Every sequence in $\text{walk}^+(\mathbf{x})$ necessarily stays in the shadow of \mathbf{x} . Therefore, if no lattice points are inside the shadow, the lattice length must be one.

Conversely, if there is a lattice point in the shadow, call the vector from the origin to that point \mathbf{x}' . Note also that $\mathbf{x} - \mathbf{x}'$ is an increasing lattice vector, and thus the path $(\mathbf{x}', \mathbf{x} - \mathbf{x}')$ shows $\ell^+(\mathbf{x}) > 1$. \square

From now on, we will assume that L is a finite lattice on a torus, given by all multiples of an integer a modulo b , with $a < b$ and $\gcd(a, b) = 1$. That is, we now declare

$$L = L_{a,b} = \{(i, a \cdot i) \pmod{b} : i \in \mathbb{Z}\} = L[\langle 1, a \rangle, \langle -1, b - a \rangle] \pmod{b}$$

We can picture this lattice as lying in the square $[0, b] \times [0, b]$. We can interpret any point (x, y) in the square as an increasing vector $\mathbf{x} = \langle x, y \rangle$ or as a decreasing vector $\mathbf{y} = \mathbf{x} - \langle 0, b \rangle = \langle x, y - b \rangle$. Hence to each point we can assign an increasing length, $\ell^+(\mathbf{x})$, and a decreasing length, $\ell^-(\mathbf{y})$. This is illustrated in Figure 2.1 for $a = 51$ and $b = 71$. The marked point is $(30, 39)$, so that $\mathbf{x} = \langle 30, 39 \rangle$ and $\mathbf{y} = \langle 30, -32 \rangle$.

We will soon show that vectors of lattice length one possess many nice properties and will aid in characterizing lattice length for any point in L . The first thing to show is how these vectors come from the Euclidean algorithm.

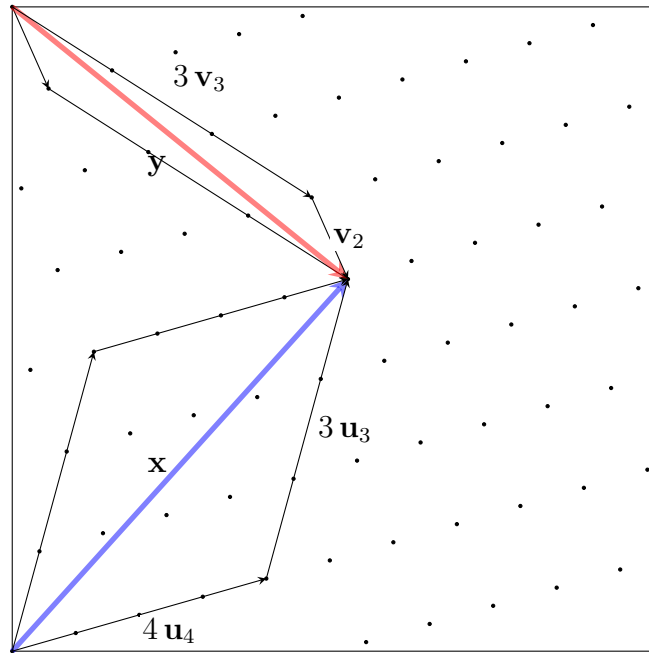


Figure 2.1: The lattice $L_{51,71}$. The vector $\mathbf{x} = \langle 30, 39 \rangle$ has $\ell^+(\mathbf{x}) = 7$, while the vector $\mathbf{y} = \mathbf{x} - \langle 0, 71 \rangle$ has $\ell^-(\mathbf{y}) = 4$.

2.2. The slow Euclidean algorithm and continued fractions

For input integers a and b , the usual Euclidean algorithm produces a sequence of remainders r_i and integers s_i and t_i such that $r_i = s_i b + t_i a$. We now describe a version of the extended Euclidean algorithm that records not only every stage of the usual Euclidean algorithm, which invokes the division algorithm in each step, but also the steps of subtraction that the division summarizes. We will refer to this algorithm as the “slow” Euclidean algorithm.

Initialize the algorithm with two integers a and b such that $1 \leq a \leq b$. We will index our steps with pairs (i, j) , where at each step, (i, j) , we have an integer combination

$$r_{i,j} = s_{i,j}b + t_{i,j}a. \tag{2.1}$$

We initialize

$$\begin{aligned} r_{-1,1} = r_{-1} = b, & \quad s_{-1,1} = s_{-1} = 1, & \quad t_{-1,1} = t_{-1} = 0 \\ r_{0,1} = r_0 = a, & \quad s_{0,1} = s_0 = 0, & \quad t_{0,1} = t_0 = 1. \end{aligned}$$

For each $i \geq 1$, rather than divide r_{i-1} by r_{i-2} (as in the common version of the algorithm), we successively subtract r_{i-1} from r_{i-2} until we can no longer subtract and remain nonnegative. That is, we set $r_{i,1} = r_{i-2} - r_{i-1}$, and while $r_{i,j-1} \geq r_{i-1}$ we set

$$r_{i,j} = r_{i,j-1} - r_{i-1} = r_{i-2} - jr_{i-1}.$$

We similarly define the coefficients $s_{i,j}$ and $t_{i,j}$ by setting

$$s_{i,j} = s_{i-2} - js_{i-1} \quad \text{and} \quad t_{i,j} = t_{i-2} - jt_{i-1}, \tag{2.2}$$

where j ranges over the same values as the $r_{i,j}$. We iterate i when we reach a value of j' for which $r_{i-1} > r_{i,j'} \geq 0$. At this point we set $r_i = r_{i,j'}$, $s_i = s_{i,j'}$, and $t_i = t_{i,j'}$. We will call these terms with a single subscript the ‘‘simple’’ remainders and s, t -coefficients, since they are the ones that appear in the usual (fast) Euclidean algorithm (with division rather than subtraction at each step). The terms with two subscripts ($r_{i,j}, s_{i,j}, t_{i,j}$) we refer to as ‘‘slow’’ remainders and coefficients. We also remark that this j' is the usual integer quotient, since $r_{i-2} = r_{i-1}j' + r_i$ with $0 \leq r_i < r_{i-1}$. The algorithm terminates when we find $r_{i,j} = 0$.

The Euclidean algorithm has a well known connection to continued fraction expansions. Recalling the language of continued fractions and their convergents from (1.2), we write the rational number a/b as

$$\frac{a}{b} = [a_0; a_1, a_2, \dots, a_k] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_k}}}}$$

For each $i = 0, \dots, k$, we thus have sequences of relatively prime pairs (p_i, q_i) such that

$$\frac{p_i}{q_i} = [a_0; a_1, \dots, a_i],$$

with these fractions giving the principal convergents to a/b . It follows from the definition that

$$p_i = p_{i-2} + a_i p_{i-1} \quad \text{and} \quad q_i = q_{i-2} + a_i q_{i-1}.$$

Moreover, we can define intermediate convergents, for $j = 1, \dots, a_i$, via

$$p_{i,j} = p_{i-2} + j p_{i-1} \quad \text{and} \quad q_{i,j} = q_{i-2} + j q_{i-1}, \quad (2.3)$$

with the property that

$$\frac{p_{i,j}}{q_{i,j}} = [a_0; a_1, \dots, a_{i-1}, j].$$

We can see the recurrences in (2.3) for the $p_{i,j}$ and $q_{i,j}$ are identical, up to sign, to the recurrences in (2.2) for the $s_{i,j}$ and $t_{i,j}$. Furthermore, we have

$$\delta_i = |a/b - p_i/q_i| \quad \text{and} \quad \delta_{i,j} = |a/b - p_{i,j}/q_{i,j}|,$$

or equivalently,

$$bq_i \delta_i = |q_i a - p_i b| \quad \text{and} \quad bq_{i,j} \delta_{i,j} = |q_{i,j} a - p_{i,j} b|.$$

Up to sign, these identities are equivalent to that of Equation (2.1) expressing the $r_{i,j}$ in terms of $s_{i,j}$ and $t_{i,j}$.

As we have the same linear recurrence relations (up to sign) with the same initial values, we make the following observation relating the data in the slow Euclidean algorithm for the pair (a, b) with the data in the sequence of convergents of the continued fraction for a/b .

Observation 2.2 (Slow Euclidean algorithm and continued fraction convergents). *Let (a, b) be a relatively prime pair of integers. Let $r_{i,j}, s_{i,j}, t_{i,j}$ be as defined in the slow Euclidean algorithm, and $p_{i,j}, q_{i,j}, \delta_{i,j}$ as defined for the sequence of convergents for the continued fraction expansion $a/b = [a_0; a_1, \dots, a_k]$. Then the following statements hold for $i = 0, \dots, k$ and $j = 1, \dots, a_i$.*

i	j	a_i	$r_{i,j}$	$s_{i,j}$	$t_{i,j}$	\mathbf{u}	\mathbf{v}
-1	1		71	1	0		
0	1		51	0	1	$\mathbf{u}_1 = (1, 51)$	
1	1	1	20	1	-1		$\mathbf{v}_1 = (1, -20)$
2	1		31	-1	2	$\mathbf{u}_2 = (2, 31)$	
	2	2	11	-2	3	$\mathbf{u}_3 = (3, 11)$	
3	1	1	9	3	-4		$\mathbf{v}_2 = (4, -9)$
4	1	1	2	-5	7	$\mathbf{u}_4 = (7, 2)$	
5	1		7	8	-11		$\mathbf{v}_3 = (11, -7)$
	2		5	13	-18		$\mathbf{v}_4 = (18, -5)$
	3		3	18	-25		$\mathbf{v}_5 = (25, -3)$
	4	4	1	23	-32		$\mathbf{v}_6 = (32, -1)$
6	1		1	-28	39	$\mathbf{u}_5 = (39, 1)$	
	2	2	0	-51	71		

Table 2.1: The steps of the slow Euclidean algorithm give unit lattice vectors.

- We have $a/b \geq p_{i,j}/q_{i,j}$ when i is odd, $a/b \leq p_{i,j}/q_{i,j}$ when i is even.
- With

$$\pm(a/b - p_{i,j}/q_{i,j}) < 0,$$

we have

$$r_{i,j} = bq_{i,j}\delta_{i,j}, \quad s_{i,j} = \mp p_{i,j}, \quad t_{i,j} = \pm q_{i,j}. \tag{2.4}$$

- Moreover, the a_i are the sizes of the blocks in the slow Euclidean algorithm, i.e., $r_i = r_{i-2} - a_i r_{i-1}$.

Another helpful way to express these relationships is to rethink (2.1) for the slow remainders as

$$r_{i,j} = bt_{i,j}(a/b + s_{i,j}/t_{i,j}) = b|t_{i,j}| |a/b - |s_{i,j}/t_{i,j}|| = b|t_{i,j}|\delta_{i,j}, \tag{2.5}$$

keeping in mind that $s_{i,j}$ and $t_{i,j}$ have opposite sign.

2.3. Unit lattice vectors

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d$ denote the vectors $\langle t_{i,j}, r_{i,j} \rangle$, such that $t_{i,j} > 0$ are listed in order of appearance in the slow Euclidean algorithm. Similarly, denote the vectors $-\langle t_{k,l}, r_{k,l} \rangle$, with $t_{k,l} < 0$, by $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_e$. The set $U = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d\}$ is the set of positively sloped pairs of this type, while, $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_e\}$ is the collection of such negatively sloped vectors. We call the vectors in these sets *unit lattice vectors*, since we will prove they are precisely those vectors with lattice length 1.

By construction (since generally speaking the $r_{i,j}$ decrease while the $t_{i,j}$ increase) we have their slopes in decreasing order of magnitude:

$$m_{\mathbf{u}_1} > m_{\mathbf{u}_2} > \dots > m_{\mathbf{u}_d} > 0 > m_{\mathbf{v}_e} > \dots > m_{\mathbf{v}_2} > m_{\mathbf{v}_1}.$$

So long as $a < b$, the two steepest vectors are $\mathbf{u}_1 = \langle 1, a \rangle$ and $\mathbf{v}_1 = \langle -1, b - a \rangle$. Also, notice that each consecutive difference is $\mathbf{u}_{i+1} - \mathbf{u}_i = -\langle t_{i'}, r_{i'} \rangle$ for some i' . That is, each difference is a simple (t, r) -pair, appearing at the bottom of the block of the algorithm just above \mathbf{u}_{i+1} . This fact can be restated as $\mathbf{u}_{i+1} - \mathbf{u}_i = \mathbf{v}_j$ for some j . Likewise, the differences between consecutive entries in V are simple (t, r) -pairs from set U : $\mathbf{v}_{k+1} - \mathbf{v}_k = \langle t_{k'}, r_{k'} \rangle = \mathbf{u}_l$.

In Table 2.1 we have the steps of the slow Euclidean algorithm for $a = 51$ and $b = 71$, with the vectors in sets U and V identified.

2.4. Basis pairs from the Euclidean algorithm

It is well known that vectors $\mathbf{x} = \langle x_1, x_2 \rangle$ and $\mathbf{y} = \langle y_1, y_2 \rangle$ form an integer basis for $L_{a,b}$ if and only if the area of the parallelogram spanned by \mathbf{x} and \mathbf{y} is b , i.e., if:

$$\det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} = x_1 y_2 - y_1 x_2 = \pm b.$$

It so happens that many pairs of vectors $\langle t_{i,j}, r_{i,j} \rangle$ in the extended Euclidean algorithm have this property.

We begin with a helpful observation, which allows us to focus on the coefficients $s_{i,j}$ and $t_{i,j}$.

Observation 2.3. *If $r = sb + ta$ and $r' = s'b + t'a$, then*

$$\det \begin{pmatrix} t & t' \\ r & r' \end{pmatrix} = (ts' - t's)b,$$

and thus $\{\langle t, r \rangle, \langle t', r' \rangle\}$ is a basis for $L_{a,b}$ if and only if

$$\det \begin{pmatrix} t & t' \\ s & s' \end{pmatrix} = \pm 1.$$

We now provide a collection of bases for $L_{a,b}$ that come from the Euclidean algorithm.

Proposition 2.4 (Nice basis pairs). *Fix integers a and b with $1 \leq a < b$. Then for each pair of indices i and j appearing in the slow Euclidean algorithm for a and b , we have bases of the form*

$$\{\langle t_i, r_i \rangle, \langle t_{i+1,j}, r_{i+1,j} \rangle\}, \{\langle t_i, r_i \rangle, \langle t_{i+2,1}, r_{i+2,1} \rangle\}, \text{ and } \{\langle t_{i,j}, r_{i,j} \rangle, \langle t_{i,j+1}, r_{i,j+1} \rangle\}.$$

In other words, the vector at the bottom of one block of the Euclidean algorithm forms a basis with each of the vectors in the next block, as well as the first vector of the block after that. Moreover, adjacent vectors within each block give bases.

As especially nice cases, consecutive elements of U and V give bases. That is, $\{\mathbf{u}_i, \mathbf{u}_{i+1}\}$ is a basis for each $i = 1, \dots, d - 1$, and $\{\mathbf{v}_j, \mathbf{v}_{j+1}\}$ is a basis for each $j = 1, \dots, e - 1$.

Proof. By Observation 2.3, it will suffice to show the relevant determinants of (s, t) -pairs have absolute value 1.

We make the elementary observation that for any two pairs $(s, t), (s', t')$ such that

$$\det \begin{pmatrix} t & t' \\ s & s' \end{pmatrix} = ts' - t's = \pm 1,$$

then

$$\det \begin{pmatrix} t & t' + jt \\ s & s' + js \end{pmatrix} = ts' + jts - t's - jts = ts' - t's = \pm 1.$$

In other words, if $\{\mathbf{x}, \mathbf{y}\}$ form a basis, then $\{\mathbf{x}, \mathbf{y} + j\mathbf{x}\}$ also forms a basis.

The result now follows by induction, with base cases of $\{\langle 0, b \rangle, \langle 1, a \rangle\}$ and $\{\langle 1, a \rangle, \langle -1, b - a \rangle\}$, since

$$\{\langle t_i, r_i \rangle, \langle t_{i+1, j}, r_{i+1, j} \rangle\} = \{\langle t_i, r_i \rangle, \langle t_{i-1}, r_{i-1} \rangle - j\langle t_i, r_i \rangle\},$$

$$\{\langle t_i, r_i \rangle, \langle t_{i+2, 1}, r_{i+2, 1} \rangle\} = \{\langle t_i, r_i \rangle, \langle t_{i+1}, r_{i+1} \rangle - \langle t_i, r_i \rangle\},$$

and

$$\{\langle t_{i, j}, r_{i, j} \rangle, \langle t_{i, j+1}, r_{i, j+1} \rangle\} = \{\langle t_{i, j}, r_{i, j} \rangle, \langle t_{i, j}, r_{i, j} \rangle - \langle t_{i-1}, r_{i-1} \rangle\}. \quad \square$$

Our next goal is to prove that the (t, r) -pairs appearing in the Euclidean algorithm give not only nice basis pairs for L , but that they are also the vectors of lattice length one.

2.5. Characterizing unit lattice vectors

We will now show that vectors in $U \cup V$, i.e., those $\langle t_{i, j}, r_{i, j} \rangle$ that come from the slow Euclidean algorithm, are the vectors of lattice length one in $L_{a, b}$.

Proposition 2.5 (Unit lattice length vectors). *Let \mathbf{x} be an increasing lattice vector and let \mathbf{y} be a decreasing lattice vector. Then $\mathbf{x} \in U$ if and only if $\ell^+(\mathbf{x}) = 1$. Likewise, $\mathbf{y} \in V$ if and only if $\ell^-(\mathbf{y}) = 1$.*

Proof. We will prove only the statement for the increasing vectors, since the argument is identical in the decreasing case.

To begin we prove that each element $\mathbf{u} \in U$ has $\ell^+(\mathbf{u}) = 1$. By Observation 2.1, it is enough to show that $\text{shadow}(\mathbf{u}_i) = \emptyset$ for all i . Since $\mathbf{u}_1 = \langle 1, a \rangle$, it is trivially true that $\text{shadow}(\mathbf{u}_1) = \emptyset$. Now we will show that if \mathbf{u}_i has an empty shadow, then \mathbf{u}_{i+1} has an empty shadow as well.

Suppose for a contradiction that $\text{shadow}(\mathbf{u}_i) = \emptyset$ but there is a point $(x, y) \in \text{shadow}(\mathbf{u}_{i+1})$. Then either the vector $\mathbf{x} = \langle x, y \rangle$ or the vector $\mathbf{u}_{i+1} - \mathbf{x} = \langle t - x, r - y \rangle$ is a vector with slope at least $m_{\mathbf{u}_{i+1}}$. Without loss of generality, suppose $m_{\mathbf{x}} \geq m_{\mathbf{u}_{i+1}}$. But if this is the case, then the lattice point (x, y) lies in the parallelogram spanned by \mathbf{u}_i and \mathbf{u}_{i+1} , contradicting the fact that, by Proposition 2.4, $\{\mathbf{u}_i, \mathbf{u}_{i+1}\}$ is a basis.

Hence, we must conclude $\text{shadow}(\mathbf{u}_{i+1}) = \emptyset$, as desired. This proves the first implication.

We now argue for the converse, that if $\mathbf{x} \notin U$, then $\ell^+(\mathbf{x}) > 1$. Suppose \mathbf{x} is an increasing lattice vector. Then for some i , $m_{\mathbf{u}_i} \geq m_{\mathbf{x}} \geq m_{\mathbf{u}_{i+1}}$. Since $\{\mathbf{u}_i, \mathbf{u}_{i+1}\}$ is an integer basis, we

know $\mathbf{x} = r \mathbf{u}_i + s \mathbf{u}_{i+1}$ for integers r and s where at least one of r or s must be positive since \mathbf{x} is an increasing lattice vector.

If $s < 0 < r$, then $m_{\mathbf{x}} > m_{\mathbf{u}_i}$, a contradiction. Similarly, if $r < 0 < s$, then $m_{\mathbf{u}_{i+1}} > m_{\mathbf{x}}$. Thus both r and s are nonnegative, and $\ell^+(\mathbf{x}) \geq r + s \geq 1$. If $r = 0$ and $s = 1$ or vice-versa, then $\mathbf{x} \in \{\mathbf{u}_i, \mathbf{u}_{i+1}\} \subset U$. Otherwise, $\ell^+(\mathbf{x}) > 1$, as desired. \square

2.6. Length in terms of unit vectors

Having established that the vectors $U \cup V$ are precisely those vectors in $L_{a,b}$ with lattice length one, we now describe the length of any lattice vector in terms of these. To facilitate discussion, throughout the remainder of the section we will focus on increasing vectors only. Similar ideas will yield analogous results for decreasing vectors.

Let $P = \bigcup_{i=1}^d P_i$ denote the union of the triangles $P_i = \{\lambda \mathbf{u}_i + \mu \mathbf{u}_{i+1} : \lambda, \mu \in [0, 1]\}$. (Note λ, μ are real numbers in this construction.) The boundary of P is the union of segments in the first quadrant given by

$$\partial P = \{\lambda \mathbf{u}_i + (1 - \lambda) \mathbf{u}_{i+1} : 0 \leq i \leq d, \lambda \in [0, 1]\},$$

where $\mathbf{u}_0 = \mathbf{u}_{d+1} = \mathbf{0}$. Apart from the zero vector, every lattice vector in P has lattice length 1 by Proposition 2.5. Let kP denote the dilation of P by k units. Then

$$\partial(kP) = \{\lambda \mathbf{u}_i + (k - \lambda) \mathbf{u}_{i+1} : 0 \leq i \leq d, \lambda \in [0, k]\}.$$

In the proof of Proposition 2.5, we showed that if \mathbf{x} is an increasing vector, then there is an integer i such that $m_{\mathbf{u}_i} \geq m_{\mathbf{x}} \geq m_{\mathbf{u}_{i+1}}$, and moreover we can write $\mathbf{x} = r \mathbf{u}_i + s \mathbf{u}_{i+1}$ for nonnegative integers r and s . In other words, this means $\mathbf{x} \in \partial(kP_i)$, where $k = r + s$.

Observation 2.6. *Every increasing vector \mathbf{x} lies in $\partial(kP_i)$ for some i and k .*

We now show that the dilation of P gives us a way to understand lattice length.

Proposition 2.7. *For any integer k , $\ell^+(k \mathbf{u}_1) = \ell^+(k \mathbf{u}_d) = k$. If $\mathbf{x} \in \partial(kP) \cap L_{a,b}$, and if \mathbf{x} is not a multiple of \mathbf{u}_1 or \mathbf{u}_d , then $\ell^+(\mathbf{x}) = k$. In particular, if $\mathbf{x} \in \partial(kP_i)$, then there exist nonnegative integers c and d such that $\mathbf{x} = c \mathbf{u}_i + d \mathbf{u}_{i+1}$ and*

$$\ell^+(\mathbf{x}) = c + d.$$

See Figure 2.2 for an illustration of this result for $a = 51$ and $b = 71$.

Proof. Since \mathbf{u}_1 has the greatest slope among the increasing unit vectors, multiples of \mathbf{u}_1 have only one increasing path with unit lattice vectors, namely, $\mathbf{u}_1, \mathbf{u}_1, \dots, \mathbf{u}_1$. Similarly, there is only one increasing path to any multiple of \mathbf{u}_d since it has the smallest slope. This proves $\ell^+(k \mathbf{u}_1) = \ell^+(k \mathbf{u}_d) = k$ as claimed.

We now turn to the points on $\partial(kP)$ that are not multiples of \mathbf{u}_1 or \mathbf{u}_d . We proceed by induction, with the base case of $k = 1$ true by Proposition 2.5.

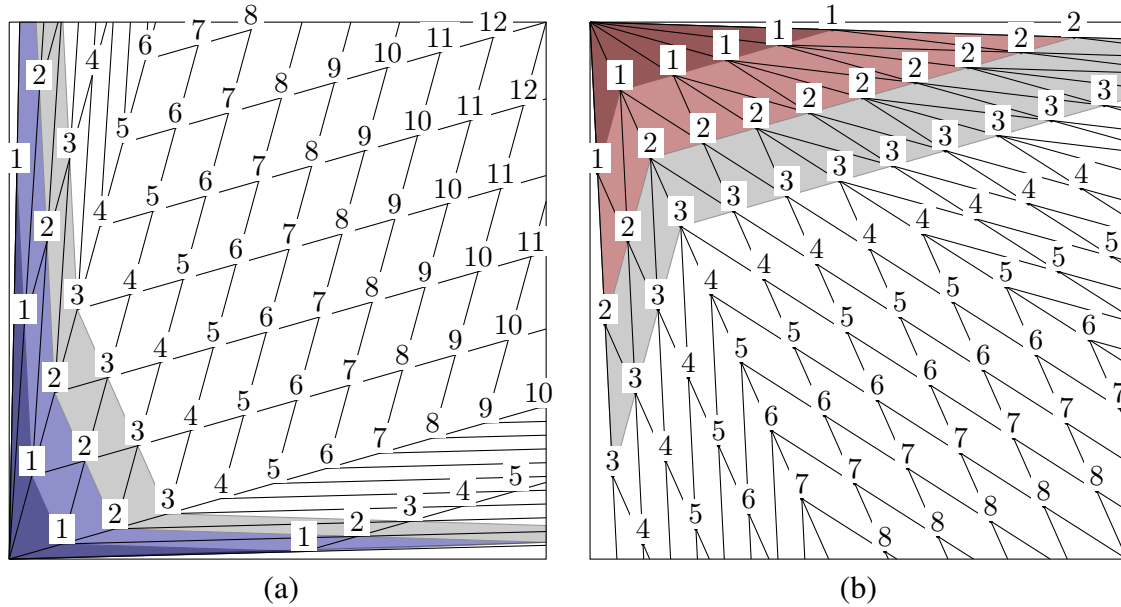


Figure 2.2: The lattice $L_{51,71}$, with points labeled according to (a) increasing lattice length, and (b) decreasing lattice length. Dilations of the polygon P are shaded in blue in (a), and red in (b).

Now suppose the claim is true for some integer $k \geq 1$. Let $\mathbf{x} \in \partial((k + 1)P) \cap L_{a,b}$. By Observation 2.6, there is an i such that $\mathbf{x} \in \partial((k + 1)P_i) \cap L_{a,b}$. Then for some integer c with $0 \leq c \leq k + 1$,

$$\mathbf{x} = c \mathbf{u}_i + (k + 1 - c) \mathbf{u}_{i+1}.$$

This shows that $\ell^+(\mathbf{x}) \geq k + 1$.

Now, in general, any path in $\text{walk}^+(\mathbf{x})$ can be refined into a path consisting only of unit lattice vectors, and we can see

$$\ell^+(\mathbf{x}) = 1 + \max\{\ell^+(\mathbf{x} - \mathbf{u}) : \mathbf{u} \in U \text{ and } \mathbf{x} - \mathbf{u} \in L_{a,b}\}.$$

Now for any \mathbf{u} such that $\mathbf{x} - \mathbf{u} \in L_{a,b}$, the vector $\mathbf{x} - \mathbf{u}$ must lie in the interior of $(k + 1)P$ since it is in the shadow of \mathbf{x} . Again by Observation 2.6, this means $\mathbf{x} - \mathbf{u} \in \partial(jP)$ for some $j \leq k$, and hence by induction $\ell^+(\mathbf{x} - \mathbf{u}) = j \leq k$. Therefore

$$\ell^+(\mathbf{x}) \leq 1 + k.$$

Together with our earlier inequality, we have proved the desired result:

$$\ell^+(\mathbf{x}) = k + 1. \quad \square$$

In later sections, we will want to have a different description of the location of a vector \mathbf{x} that also determines its length. Let us denote three consecutive unit vectors from simple (t, r) -pairs as

$$\mathbf{u} = \langle t_{2i-2}, r_{2i-2} \rangle, \mathbf{v} = \langle t_{2i-1}, -r_{2i-1} \rangle, \text{ and } \mathbf{w} = \langle t_{2i}, r_{2i} \rangle,$$

so that $\mathbf{w} = \mathbf{u} + a_{2i} \mathbf{v}$, where a_{2i} is the corresponding continued fraction term. Then every vector of the form $\mathbf{u} + a \mathbf{v}$, with $0 \leq a \leq a_{2i}$, is a unit lattice vector corresponding to a slow (t, r) -pair. If $\mathbf{u} = \mathbf{u}_j$, then $\mathbf{u} + \mathbf{v} = \mathbf{u}_{j+1}$, $\mathbf{u} + 2\mathbf{v} = \mathbf{u}_{j+2}$, and so on, with $\mathbf{w} = \mathbf{u}_{j+a_{2i}}$. Thus, by Proposition 2.7, any lattice vector in the k -fold dilation of the line segment $(1 - \lambda) \mathbf{u} + \lambda \mathbf{w}$, ($0 \leq \lambda \leq 1$) will also have length k , since this segment is just a union of dilations of segments between consecutive unit lattice vectors. We capture this idea in the following corollary to Proposition 2.7.

Corollary 2.8. *Let \mathbf{u} , \mathbf{v} , and \mathbf{w} , as above, with $\mathbf{w} = \mathbf{u} + a_{2i} \mathbf{v}$. Then the following are equivalent for an increasing vector \mathbf{x} :*

- *Vector \mathbf{x} has $\ell^+(\mathbf{x}) = k$ and slope bounded by that of \mathbf{u} and \mathbf{w} : $m_{\mathbf{w}} \leq m_{\mathbf{x}} \leq m_{\mathbf{u}}$.*
- *There exists an integer $0 \leq a \leq ka_{2i}$ such that $\mathbf{x} = k \mathbf{u} + a \mathbf{v} = k \mathbf{w} - (ka_{2i} - a) \mathbf{v}$.*

2.7. Best unit vectors to approximate a given slope

To use our notions of lattice length to find longest increasing/decreasing subsequences it is important for us to consider the unit vectors with slopes closest to ± 1 . In fact, we do better, and find the analogous unit lattice vectors for any slope $\pm \tau$.

From this point forward, fix a real slope $\tau \geq 1$ (our illustrations will use $\tau = 1$ for ease of viewing) and let \mathbf{a} , \mathbf{b} , \mathbf{c} , and \mathbf{d} denote those consecutive unit lattice vectors such that

$$m_{\mathbf{a}} \geq \tau > m_{\mathbf{b}} \quad \text{and} \quad m_{\mathbf{c}} > -\tau \geq m_{\mathbf{d}}.$$

Note $\mathbf{a} = \mathbf{u}_i$ and $\mathbf{b} = \mathbf{u}_{i+1}$ for some i , while $\mathbf{c} = \mathbf{v}_{j+1}$ and $\mathbf{d} = \mathbf{v}_j$ for some j . We will show that \mathbf{c} and \mathbf{d} are nice linear combinations of \mathbf{a} and \mathbf{b} .

For example, in Table 2.1, we see that for $a = 51$ and $b = 71$ and $\tau = 1$, we have $\mathbf{a} = \mathbf{u}_3 = \langle 3, 11 \rangle$, $\mathbf{b} = \mathbf{u}_4 = \langle 7, 2 \rangle$, $\mathbf{c} = \mathbf{v}_3 = \langle 11, -7 \rangle$, and $\mathbf{d} = \mathbf{v}_2 = \langle 4, -9 \rangle$.

First we make some observations about the slow Euclidean algorithm in order to describe where \mathbf{a} , \mathbf{b} , \mathbf{c} , and \mathbf{d} occur. The idea behind the following observations is merely that ‘‘simple remainders decrease’’ and ‘‘remainders decrease within subtraction blocks’’ while ‘‘ t -coefficients increase in absolute value and alternate signs from block to block.’’

Observation 2.9. *We have the following characterizations.*

- *Let i and j be the lexicographically first pair such that $\tau > r_{i,j}/t_{i,j} > 0$. Then $\mathbf{b} = \langle t_{i,j}, r_{i,j} \rangle$ is the unit vector with greatest slope less than τ and*

$$\mathbf{a} = \begin{cases} \langle t_{i,j-1}, r_{i,j-1} \rangle & \text{if } j > 1, \\ \langle t_{i-2}, r_{i-2} \rangle & \text{if } j = 1. \end{cases}$$

- *Let k and l be the lexicographically first pair such that $\tau > -r_{k,l}/t_{k,j} > 0$. Then $\mathbf{c} = \langle -t_{k,l}, -r_{k,l} \rangle$ is the unit vector with least slope greater than $-\tau$ and*

$$\mathbf{d} = \begin{cases} \langle -t_{k,l-1}, -r_{k,l-1} \rangle & \text{if } l > 1, \\ \langle -t_{k-2}, -r_{k-2} \rangle & \text{if } l = 1. \end{cases}$$

From now on, let us write $\mathbf{x} = \mathbf{c} - \mathbf{d}$ and $\mathbf{y} = \mathbf{b} - \mathbf{a}$. As discussed Section 2.2, it follows from the definition of the slow Euclidean algorithm that $\mathbf{x} \in U$ and $\mathbf{y} \in V$ are unit lattice vectors and moreover that they are simple remainder pairs (of the form $\langle t_i, r_i \rangle$). These vectors will provide a convenient way for us to frame our results.

Observation 2.10. *Let i be minimal such that $\tau > r_i/|t_i| = b\delta_i > 0$.*

1. *If $t_i > 0$, i.e., if i is even, then \mathbf{b} is in the i th block: $\mathbf{b} = \langle t_{i,j}, r_{i,j} \rangle$ for some j , and \mathbf{c} is in the $i + 1$ st block: $\mathbf{c} = -\langle t_{i+1,k}, r_{i+1,k} \rangle$ for some k . Moreover:*

- $\mathbf{y} = -\langle t_{i-1}, r_{i-1} \rangle$, $t_{i-1} < 0$, and by minimality of i , $\tau \leq r_{i-1}/|t_{i-1}| = -m_{\mathbf{y}}$,
- $\mathbf{x} = \langle t_i, r_i \rangle$, with $m_{\mathbf{x}} \leq m_{\mathbf{b}} < \tau$.

2. *If $t_i < 0$, i.e., if i is odd, then \mathbf{c} is in the i th block: $\mathbf{c} = -\langle t_{i,j}, r_{i,j} \rangle$ for some j , and \mathbf{b} is in the $i + 1$ st block: $\mathbf{b} = \langle t_{i+1,k}, r_{i+1,k} \rangle$ for some k . Moreover:*

- $\mathbf{y} = -\langle t_i, r_i \rangle$, with $0 < -m_{\mathbf{y}} = r_i/|t_i| < \tau$,
- $\mathbf{x} = \langle t_{i-1}, r_{i-1} \rangle$, $t_{i-1} > 0$, and by minimality of i , $\tau \leq r_{i-1}/t_{i-1} = m_{\mathbf{x}}$.

In either case, precisely one of \mathbf{x} and \mathbf{y} has slope of absolute value less than τ . Moreover, as \mathbf{x} and \mathbf{y} come from simple (t, r) -pairs in adjacent blocks, $\{\mathbf{x}, \mathbf{y}\}$ is a basis for $L_{a,b}$ by Proposition 2.4.

Now that we have very tightly identified where vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{x}$, and \mathbf{y} appear in the slow Euclidean algorithm, we are ready to characterize the linear relationships between them. There are four cases, which are equivalent up to transformations of the lattice.

Proposition 2.11. *We have the following expressions relating $\mathbf{a}, \mathbf{b}, \mathbf{c}$, and \mathbf{d} .*

1. *If $m_{\mathbf{y}} \leq -\tau$, then $0 < m_{\mathbf{x}} < \tau$. Furthermore,*

(a) *if $|\mathbf{y}| \leq |\mathbf{x}|$, then there exists an integer $s \geq 1$ such that:*

$$\begin{pmatrix} \mathbf{a} & \mathbf{c} \\ \mathbf{b} & \mathbf{d} \end{pmatrix} = \begin{pmatrix} \mathbf{x} - s\mathbf{y} & \mathbf{y} + \mathbf{x} \\ \mathbf{x} - (s-1)\mathbf{y} & \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{c} - (s+1)\mathbf{d} & (s+1)\mathbf{b} - s\mathbf{a} \\ \mathbf{c} - s\mathbf{d} & \mathbf{b} - \mathbf{a} \end{pmatrix}.$$

(b) *if $|\mathbf{y}| > |\mathbf{x}|$, then there exists an integer $s \geq 1$ such that:*

$$\begin{pmatrix} \mathbf{a} & \mathbf{c} \\ \mathbf{b} & \mathbf{d} \end{pmatrix} = \begin{pmatrix} -\mathbf{y} + \mathbf{x} & \mathbf{y} + s\mathbf{x} \\ \mathbf{x} & \mathbf{y} + (s-1)\mathbf{x} \end{pmatrix} = \begin{pmatrix} s\mathbf{c} - (s+1)\mathbf{d} & (s+1)\mathbf{b} - \mathbf{a} \\ \mathbf{c} - \mathbf{d} & s\mathbf{b} - \mathbf{a} \end{pmatrix}.$$

2. *If $0 > m_{\mathbf{y}} > -\tau$, then $m_{\mathbf{x}} \geq \tau$. Furthermore,*

(a) *if $|\mathbf{y}| \leq |\mathbf{x}|$, then there exists an integer $s \geq 1$ such that:*

$$\begin{pmatrix} \mathbf{a} & \mathbf{c} \\ \mathbf{b} & \mathbf{d} \end{pmatrix} = \begin{pmatrix} \mathbf{x} + (s-1)\mathbf{y} & \mathbf{y} \\ \mathbf{x} + s\mathbf{y} & \mathbf{y} - \mathbf{x} \end{pmatrix} = \begin{pmatrix} s\mathbf{c} - \mathbf{d} & \mathbf{b} - \mathbf{a} \\ (s+1)\mathbf{c} - \mathbf{d} & s\mathbf{b} - (s+1)\mathbf{a} \end{pmatrix}.$$

(b) *if $|\mathbf{y}| > |\mathbf{x}|$, then there exists an integer $s \geq 1$ such that:*

$$\begin{pmatrix} \mathbf{a} & \mathbf{c} \\ \mathbf{b} & \mathbf{d} \end{pmatrix} = \begin{pmatrix} \mathbf{x} & \mathbf{y} - (s-1)\mathbf{x} \\ \mathbf{y} + \mathbf{x} & \mathbf{y} - s\mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{c} - \mathbf{d} & \mathbf{b} - s\mathbf{a} \\ (s+1)\mathbf{c} - s\mathbf{d} & \mathbf{b} - (s+1)\mathbf{a} \end{pmatrix}.$$

Proof. We will prove case (1a) in detail. The other cases follow from symmetries, as indicated in Figure 2.3.

Suppose that $-\mathbf{y} = \mathbf{a} - \mathbf{b} = \langle t_{i-1}, r_{i-1} \rangle$ with $t_{i-1} < 0$ and $\tau \leq r_{i-1}/|t_{i-1}| = -m_{\mathbf{y}}$, whereas $\mathbf{x} = \mathbf{c} - \mathbf{d} = \langle t_i, r_i \rangle$ with $\tau > r_i/t_i = m_{\mathbf{x}}$. From Observation 2.10, \mathbf{x} and \mathbf{y} are of unit lattice length and form an integer basis for the lattice. Moreover, by considering their slopes, we know that \mathbf{d} and \mathbf{c} are in the nonnegative span of \mathbf{x} and \mathbf{y} . Further, we suppose that $|\mathbf{x}| \geq |\mathbf{y}|$. Since \mathbf{x} is positively sloped and \mathbf{y} is negatively sloped with $|\mathbf{y}| \leq |\mathbf{x}|$, we have

$$m_{\mathbf{x}+\mathbf{y}} > -1 \geq -\tau.$$

We claim that $\mathbf{y} = \mathbf{d}$. Indeed, suppose \mathbf{d}' is a lattice vector whose slope satisfies $m_{\mathbf{d}} < m_{\mathbf{d}'} \leq -\tau$. Then $\mathbf{d}' = k\mathbf{y} + l\mathbf{x} = (k-l)\mathbf{y} + l(\mathbf{x}+\mathbf{y})$ for some nonnegative integers k and l . Since $m_{\mathbf{x}+\mathbf{y}} > -\tau$ and \mathbf{y} and \mathbf{d}' have slopes less than or equal to $-\tau$, we see that $k > l > 0$. But also the vector $(k-1)\mathbf{y} + l\mathbf{x}$ must have negative slope (else, it is positively sloped and $|\mathbf{y}| > |(k-1)\mathbf{y} + l\mathbf{x}| > |\mathbf{x}|$), so we can write $\mathbf{d}' = ((k-1)\mathbf{y} + l\mathbf{x}) + \mathbf{y}$, the sum of two negatively sloped lattice vectors. This shows the lattice length of any such \mathbf{d}' has $\ell^-(\mathbf{d}') > 1$.

Now, given that $\mathbf{d} = \mathbf{y}$, we have $\mathbf{c} = \mathbf{y} + \mathbf{x}$ by the definition of \mathbf{x} .

Observation 2.10 tells us that in the context of the slow Euclidean algorithm, $\mathbf{x} = \langle t_i, r_i \rangle$ is the simple pair at the bottom of the block containing \mathbf{b} . Thus adding some number of copies of $-\mathbf{y}$ to \mathbf{x} will lead to \mathbf{b} (say $s-1$ of them) and by definition, one more copy gives \mathbf{a} . This proves

$$\mathbf{b} = \mathbf{x} - (s-1)\mathbf{y} \quad \text{and} \quad \mathbf{a} = \mathbf{x} - s\mathbf{y}.$$

The expressions for \mathbf{a} and \mathbf{b} in terms of \mathbf{c} and \mathbf{d} (and vice-versa) now follow from easy algebra. \square

2.8. The vectors \mathbf{x} and \mathbf{y} in terms of the convergents for a/b .

The vectors \mathbf{x} and \mathbf{y} will play an important role in what follows, and they admit a nice interpretation in terms of the continued fraction expansion of a/b . From Observation 2.2, and specifically (2.4) and (2.5), we have generally that

$$\langle t_{i,j}, r_{i,j} \rangle = \langle q_{i,j}, \pm bq_{i,j}\delta_{i,j} \rangle = \langle q_{i,j}, \pm bq_{i,j}|a/b - p_{i,j}/q_{i,j}| \rangle.$$

Since \mathbf{y} is a decreasing vector whose entries are a simple (t, r) -pair, we have $\mathbf{y} = -\langle t_i, r_i \rangle$, with $t_i = q_{2h+1}$ for some h , i.e.,

$$\mathbf{y} = \langle q_{2h+1}, -bq_{2h+1}\delta_{2h+1} \rangle = \langle q_{2h+1}, -bq_{2h+1}|a/b - p_{2h+1}/q_{2h+1}| \rangle. \quad (2.6)$$

Since \mathbf{x} and \mathbf{y} come from adjacent blocks of the slow Euclidean algorithm, we find the similar formula for \mathbf{x} :

$$\mathbf{x} = \langle q_j, bq_j\delta_j \rangle = \langle q_j, bq_j|a/b - p_j/q_j| \rangle, \quad (2.7)$$

where $j = 2h$ or $j = 2h + 2$.

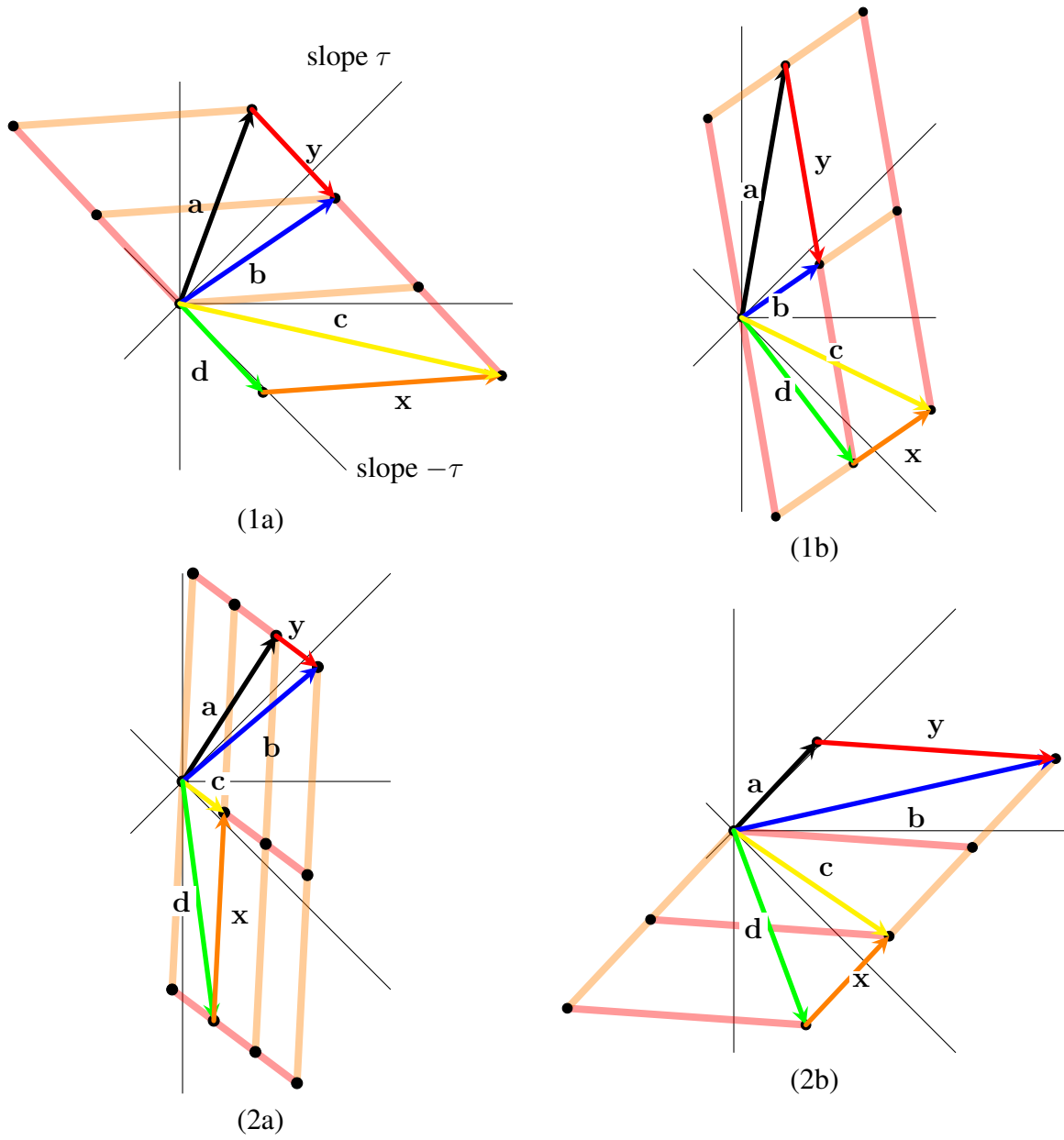


Figure 2.3: The four cases in Proposition 2.11.

Though it has not yet been emphasized, we now notice that the slope of any unit lattice vector $\langle t_{i,j}, r_{i,j} \rangle$ is $m = \pm b\delta_{i,j}$. Thus, following Observation 2.10, we let i be minimal such that

$$r_i/|t_i| = b\delta_i \leq \tau.$$

If i is even, then we have $i = j = 2h+2$ in (2.7). Then we are in the Case 1 from Proposition 2.11, and

$$\begin{aligned} \mathbf{x} &= \langle q_{2h+2}, bq_{2h+2}\delta_{2h+2} \rangle = \left\langle q_{2h+2}, bq_{2h+2} \left| \frac{a}{b} - \frac{p_{2h+2}}{q_{2h+2}} \right| \right\rangle, \\ \mathbf{y} &= \langle q_{2h+1}, -bq_{2h+1}\delta_{2h+1} \rangle = \left\langle q_{2h+1}, -bq_{2h+1} \left| \frac{a}{b} - \frac{p_{2h+1}}{q_{2h+1}} \right| \right\rangle. \end{aligned} \quad (2.8)$$

If i is odd, then we are in the Case 2 from Proposition 2.11, and we have

$$\begin{aligned} \mathbf{x} &= \langle q_{2h}, bq_{2h}\delta_{2h} \rangle = \left\langle q_{2h}, bq_{2h} \left| \frac{a}{b} - \frac{p_{2h}}{q_{2h}} \right| \right\rangle, \\ \mathbf{y} &= \langle q_{2h+1}, -bq_{2h+1}\delta_{2h+1} \rangle = \left\langle q_{2h+1}, -bq_{2h+1} \left| \frac{a}{b} - \frac{p_{2h+1}}{q_{2h+1}} \right| \right\rangle. \end{aligned} \quad (2.9)$$

3. The shape of the permutation $w(n, a/N)$

In this section we consider the shape of Young diagrams for Sós permutations $w = w(n, a/N)$, where a, n and N are non-negative integers. As mentioned, $\text{sh}(w) = \text{sh}(w^{-1})$, and upon vertical rescaling, we see the set

$$\{(i, w^{-1}(i))\}_{i=1}^n$$

(a subset of integer points in $[1, n] \times [1, n]$), the set

$$\{(i, f_{a/N}(i))\}_{i=1}^n$$

(a subset of points in $[1, n] \times [0, 1)$), and the set

$$\{(i, a \cdot i \pmod{N})\}_{i=1}^n$$

(a subset of points in $[1, n] \times [1, N]$), all have the same order structure. Thus we can leverage all the tools developed in Section 2 to study $\text{sh}(w)$. It is convenient to identify the permutation with the third lattice, scaled vertically to fit in a square of size $n \times n$. Throughout this section, define the lattices

$$L := L_{a,N} = \{(i, a \cdot i \pmod{N}) : i \in \mathbb{Z}\} \quad \text{and} \quad L_n := \left\{ \left(x, \frac{n}{N}y \right) : (x, y) \in L_{a,N} \right\}.$$

These lattices differ only by a scaling factor of n/N in the vertical direction. Notice then that lines of slope $\pm\tau = \pm N/n$ in the lattice L are mapped to lines of slope ± 1 in the lattice L_n . We define also

$$L_n^{\square} := L_n \cap [0, n]^2 \quad \text{and} \quad L_n^{\square} := L_n \cap ([0, n] \times [-n, 0]).$$

In summary, the shape $\lambda = \text{sh}(w)$ is characterized by increasing paths in L_n^{\square} and decreasing paths in L_n^{\square} .

In Proposition 2.11, it is shown that there are unit lattice vectors \mathbf{x} and \mathbf{y} forming a basis for L whose slopes satisfy one of several inequalities, relative to a fixed slope τ . Take $\tau = N/n$, and let vectors \mathbf{x}_n and \mathbf{y}_n be obtained from the vectors \mathbf{x} and \mathbf{y} in Proposition 2.11 by $\mathbf{x}_n = \langle 1, n/N \rangle \circ \mathbf{x}$, $\mathbf{y}_n = \langle 1, n/N \rangle \circ \mathbf{y}$, where \circ is the Hadamard (entrywise) product. The slopes of these rescaled basis vectors satisfy one of the following strings of inequalities:

$$m_{\mathbf{y}_n} \leq -1 \leq m_{\mathbf{y}_n + \mathbf{x}_n} < 0 < m_{\mathbf{x}_n} \leq 1 \leq m_{\mathbf{x}_n - s\mathbf{y}_n}, \tag{3.1}$$

$$m_{\mathbf{y}_n} \leq -1 \leq m_{\mathbf{y}_n + s\mathbf{x}_n} < 0 < m_{\mathbf{x}_n} \leq 1 \leq m_{\mathbf{x}_n - \mathbf{y}_n}, \tag{3.2}$$

$$m_{\mathbf{y}_n - \mathbf{x}_n} \leq -1 \leq m_{\mathbf{y}_n} < 0 < m_{\mathbf{x}_n + s\mathbf{y}_n} \leq 1 \leq m_{\mathbf{x}_n}, \tag{3.3}$$

$$m_{\mathbf{y}_n - s\mathbf{x}_n} \leq -1 \leq m_{\mathbf{y}_n} < 0 < m_{\mathbf{y}_n + \mathbf{x}_n} \leq 1 \leq m_{\mathbf{x}_n}, \tag{3.4}$$

where, in each case, $s \geq 1$ is the minimal positive integer that makes the inequalities true. The four cases above correspond to the cases 1(a), 1(b), 2(a), and 2(b), respectively, from Proposition 2.11. Recall from Section 2.8 that \mathbf{x} and \mathbf{y} are determined by the minimal value of i such that $N\delta_i \leq \tau$, or

$$\delta_i = \left| \frac{a}{N} - \frac{p_i}{q_i} \right| \leq \frac{\tau}{N} = \frac{1}{n},$$

where p_i/q_i is the i th convergent for a/N .

A rescaling of the formulas (2.8) and (2.9) gives

$$\mathbf{x}_n = \left\langle q_{2h+2}, nq_{2h+2} \left| \frac{a}{N} - \frac{p_{2h+2}}{q_{2h+2}} \right| \right\rangle, \quad \mathbf{y}_n = \left\langle q_{2h+1}, -nq_{2h+1} \left| \frac{a}{N} - \frac{p_{2h+1}}{q_{2h+1}} \right| \right\rangle, \tag{3.5}$$

if $i = 2h + 2$ is even; and

$$\mathbf{x}_n = \left\langle q_{2h}, nq_{2h} \left| \frac{a}{N} - \frac{p_{2h}}{q_{2h}} \right| \right\rangle, \quad \mathbf{y}_n = \left\langle q_{2h+1}, -nq_{2h+1} \left| \frac{a}{N} - \frac{p_{2h+1}}{q_{2h+1}} \right| \right\rangle. \tag{3.6}$$

if $i = 2h + 1$ is odd.

3.1. Symmetries between cases

There are two fundamental lattice transformations that allow us to reduce our analysis of a priori eight cases (increasing/decreasing paths for each of Inequalities (3.1)–(3.4)) down to only two cases. In terms of the basis pair (\mathbf{x}, \mathbf{y}) , these two transformations are:

$$\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix} = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \mapsto \begin{pmatrix} -y_2 & x_2 \\ y_1 & -x_1 \end{pmatrix} = \begin{pmatrix} \mathbf{x}^* & \mathbf{y}^* \end{pmatrix} = \rho \begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix},$$

and:

$$\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix} = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \mapsto \begin{pmatrix} y_1 & x_1 \\ -y_2 & -x_2 \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{x}} & \bar{\mathbf{y}} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}.$$

It is easily verified that $\rho^2 = \omega^2$ is the identity map. Simple calculations yield the following observation.

Observation 3.1. *Let (\mathbf{x}, \mathbf{y}) be a lattice basis pair. Then:*

1. (\mathbf{x}, \mathbf{y}) satisfies inequalities (3.1) if and only if $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies (3.2) if and only if $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ satisfies (3.4), and
2. (\mathbf{x}, \mathbf{y}) satisfies inequalities (3.3) if and only if $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies (3.4) if and only if $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ satisfies (3.2).

Moreover, both transformations ρ and ω map increasing paths to decreasing paths and vice-versa. Therefore we can translate results for increasing paths in case (3.1) to results about decreasing paths in case (3.2), and so on. A conceptual diagram for our cases is as follows:

$$(\text{Case (3.1)} \nearrow) \xleftrightarrow{\rho} (\text{Case (3.2)} \searrow) \xleftrightarrow{\omega} (\text{Case (3.3)} \nearrow) \xleftrightarrow{\rho} (\text{Case (3.4)} \searrow)$$

and

$$(\text{Case (3.2)} \nearrow) \xleftrightarrow{\rho} (\text{Case (3.1)} \searrow) \xleftrightarrow{\omega} (\text{Case (3.4)} \nearrow) \xleftrightarrow{\rho} (\text{Case (3.3)} \searrow).$$

Thus, it suffices to only verify our results in detail for increasing paths in cases (3.1) and (3.2), with results translated via ρ and ω to all other cases. In fact, the differences in the arguments used for case (3.2) versus those in (3.1) are negligible (all of Subsection 3.2 applies equally to both cases), so we focus on (3.1) for the duration of this section and leave the modifications for (3.2) to the reader.

With this assumption now fixed, and with a slight abuse of notation to streamline notation, we write our basis vectors as

$$\mathbf{x} = \mathbf{x}_n = \langle x_1, x_2 \rangle, \quad \mathbf{y} = \mathbf{y}_n = \langle y_1, y_2 \rangle,$$

and assume the inequalities of case (3.1). In particular, Subsection 3.3 and Subsection 3.4 make repeated use of the fact that $1 \geq m_{\mathbf{x}} > 0 > m_{\mathbf{x}+\mathbf{y}} \geq -1 \geq m_{\mathbf{y}}$.

3.2. Lines of slope $m_{\mathbf{y}}$

Let f_j be the line of slope $m_{\mathbf{y}}$ passing through the point $j\mathbf{x}$. That is, $f_j = \{t\mathbf{y} + j\mathbf{x} : t \in \mathbb{R}\}$. Points $(x, y) \in f_j$ satisfy

$$y = \frac{y_2}{y_1}x + j \left(x_2 - \frac{y_2}{y_1}x_1 \right) = \frac{y_2}{y_1}x + \frac{nj}{y_1}, \quad (3.7)$$

where we have used the identity $n = y_1x_2 - x_1y_2$ for the second equality. Similarly, let g_i be the line of slope $m_{\mathbf{x}}$ passing through the point $i\mathbf{y}$. These lines are described by the equations

$$y = \frac{x_2}{x_1}x - \frac{ni}{x_1}. \quad (3.8)$$

In what follows, our convention will be to treat j as a real-valued parameter, and i will typically be integer-valued. With that in mind, define the *crossing set*

$$\chi_j := \{f_j \cap g_i : i \in \mathbb{Z}\} \cap [0, n]^2.$$

If j is an integer, then f_j and g_i necessarily intersect at lattice points, in which case $\chi_j = f_j \cap L_n^{\square}$. We denote the number of crossings by $l_j = |\chi_j|$.

As a function of j , l_j is a step function which changes only by jumps of magnitude 1 or 2. More concretely, we see l_j jumps when one or more of the following quantities are integers:

$$\frac{x_1}{y_1}j, \quad \frac{x_2}{y_2}j, \quad \frac{x_1}{y_1}(y_2 + j) - x_2, \quad \text{or} \quad \frac{x_2}{y_2}(y_1 - j) - x_1. \tag{3.9}$$

These four cases correspond to a line $f_j, j \geq 0$, meeting a line $g_i, i \in \mathbb{Z}$, along the left edge, bottom edge, right edge, and top edge, respectively, of the box $[0, n]^2$. Since each line f_j intersects the boundary of the box $[0, n]^2$ in at most two points, we see that there are at most two distinct integers among the values in (3.9).

Suppose $j \in \mathbb{N}$. By Corollary 2.8, we know that lattice vectors \mathbf{v} in χ_j have $\ell^+(\mathbf{v}) \leq j$, and that this bound is sharp. In particular, χ_{ℓ^+} contains the lattice vectors of maximal length, where we recall $\ell^+ = \text{arm}(w)$ denotes the longest increasing lattice walk in L_n , and $\ell^- = \text{leg}(w)$ denotes the longest decreasing walk.

We can now find bounds for ℓ^+ and, with a slight reorientation, ℓ^- .

Proposition 3.2 (Longest increasing and decreasing paths). *We have the following bounds on ℓ^+ and ℓ^- :*

$$y_1 - y_2 - 2 < \ell^+ \leq y_1 - y_2, \quad \text{and} \quad x_1 + x_2 - 2 < \ell^- \leq x_1 + x_2. \tag{3.10}$$

By substituting our expressions for x_1, x_2, y_1, y_2 in terms of continued fraction convergents from Equations (3.5) and (3.6), Proposition 3.2 yields Theorem 1.3 for rational numbers $\alpha = a/N$.

Proof. Notice that if we solve for j such that $\langle n, n \rangle \in f_j$, we find, from (3.7),

$$n = \frac{y_2}{y_1}n + \frac{j}{y_1}n,$$

and thus $j = y_1 - y_2$. Since $y_1 - y_2$ is a priori rational, we see the lines with integer index intersecting $[0, n]^2$ are $f_0, f_1, \dots, f_{\lfloor y_1 - y_2 \rfloor}$. Thus $\ell^+ \leq y_1 - y_2$.

It is possible that this last line $f = f_{\lfloor y_1 - y_2 \rfloor}$ has no lattice points inside $[0, n]^2$. If so, let \mathbf{x}' be the lowest point on f above $[0, n]^2$ and let $\mathbf{x}'' = \mathbf{x}' + \mathbf{y}$ be the leftmost point on f to the right of $[0, n]^2$. Note that since f intersects $[0, n]^2$, $x'_2 \geq n \geq x'_1$, while $x''_2 \leq n \leq x''_1$.

The vector $\mathbf{x} - \mathbf{y}$ has positive slope, and so $\mathbf{x}''' = \mathbf{x}' - (\mathbf{x} - \mathbf{y}) = \mathbf{x}'' - \mathbf{x}$ is to the left of \mathbf{x}' (so $x'''_1 < n$) and below \mathbf{x}'' (so $x'''_2 < n$). In other words, $\mathbf{x}''' \in \chi_{\lfloor y_1 - y_2 \rfloor - 1} \neq \emptyset$. This shows ℓ^+ is at least $\lfloor y_1 - y_2 \rfloor - 1$, i.e., $y_1 - y_2 - 2 < \ell^+$, which proves the desired result.

The case of ℓ^- follows the same argument for increasing paths in the lattice with basis $\mathbf{x}^* = \langle -y_2, y_1 \rangle$ and $\mathbf{y}^* = \langle x_2, -x_1 \rangle$, where we find $y_1^* - y_2^* = x_2 + x_1$. □

For example, in Figure 2.2, we have $a = 51, n = N = 71$, and the upper bound is realized. There, $\mathbf{x} = \langle 7, 2 \rangle$ and $\mathbf{y} = \langle 4, -9 \rangle$ and we see $\ell^+ = 4 + 9 = 13$, while $\ell^- = 7 + 2 = 9$.

Remark 3.3. One may compare the bounds in Proposition 3.2 with Corollary 1 in [BS79]. The quantities $y_1 - y_2$ and $x_1 + x_2$, denoted λ_n^+ and λ_n^- in [BS79], are the solutions to the linear programming problem in that paper, which tightly approximate ℓ^+ and ℓ^- (denoted l_n^+ and l_n^- in that paper), the solutions to an integer programming problem.

In what follows, it will be convenient to partition the collection of points in L_n^{\square} according to where they intersect the lines f_j , and according to where those lines intersect the boundary of $[0, n]^2$. Define

$$\mathcal{L} = \bigcup_{0 \leq j \leq y_1} \chi_{[j]}, \quad \mathcal{M} = \bigcup_{y_1 < j \leq -y_2} \chi_{[j]}, \quad \text{and} \quad \mathcal{R} = \bigcup_{-y_2 < j \leq y_1 - y_2} \chi_{[j]},$$

so that \mathcal{L} contains points from lines f_j that intersect $[0, n]^2$ on the left and bottom edges, \mathcal{M} contains points from lines that intersect on the top and bottom, while set \mathcal{R} contains points from lines that intersect on the top and right edges.

Notice that, as a function of j , l_j is weakly increasing for $0 \leq j \leq y_1$ and weakly decreasing for $|y_2| \leq j \leq y_1 - y_2$. In what follows, for a fixed positive integer k , we will need to identify the first and last lines f_j which contain at least k crossings. With that in mind, define for $k = 1, 2, 3, \dots, l_{y_1}$,

$$j_k := \min_{0 \leq j \leq y_1} \{j : l_j \geq k\}, \quad j'_k := \max_{|y_2| \leq j \leq y_1 - y_2} \{j : l_j \geq k\}.$$

We now present some straightforward results for counting both the number of crossings $l_j = |\chi_j|$, as well as estimating the value of j_k and j'_k .

Proposition 3.4 (Counting crossings and lines). *The numbers l_j satisfy the bounds*

$$\frac{jn}{|y_1 y_2|} - 1 \leq l_j \leq \frac{jn}{|y_1 y_2|} + 1, \quad \text{for } 0 \leq j \leq y_1, \quad (3.11)$$

$$\frac{n}{|y_2|} - 1 \leq l_j \leq \frac{n}{|y_2|} + 1, \quad \text{for } y_1 < j \leq |y_2|, \quad (3.12)$$

$$\frac{n(y_1 - y_2 - j)}{|y_1 y_2|} - 1 \leq l_j \leq \frac{n(y_1 - y_2 - j)}{|y_1 y_2|} + 1, \quad \text{for } |y_2| < j \leq y_1 - y_2. \quad (3.13)$$

Thus, for any $k = 1, 2, \dots, l_{y_1}$ we have

$$\frac{(k-1)|y_1 y_2|}{n} \leq j_k \leq \frac{(k+1)|y_1 y_2|}{n}, \quad (3.14)$$

$$y_1 - y_2 - \frac{(k+1)|y_1 y_2|}{n} \leq j'_k \leq y_1 - y_2 - \frac{(k-1)|y_1 y_2|}{n}. \quad (3.15)$$

Proof. The inequalities (3.11)–(3.13) are simply given by the Euclidean length of line segment $f_j \cap [0, n]^2$ divided by $|y|$, plus or minus 1.

The inequality (3.14) follows from (3.11) along with the fact that j_k satisfies $l_{j_k - \epsilon} < k \leq l_{j_k}$ for any $\epsilon > 0$. Similarly (3.15) follows from (3.13) along with the fact that $l_{j'_k + \epsilon} < k \leq l_{j'_k}$ for any $\epsilon > 0$. \square

3.3. Lemmas for lines of slope m_y

In this subsection, we record two useful lemmas regarding the lines of slope m_y . We first compare crossings at a fixed distance from the bottom left and top right corner. Define

$$\bar{j} := y_1 - y_2 - j.$$

Lemma 3.5. For any $j \leq y_1$,

$$|l_j - l_{\bar{j}}| \leq 1.$$

Proof. We can be completely explicit here. The line $g_{x_2-x_1}$ passes through $\langle n, n \rangle$. If $i = x_2 - x_1 \in \mathbb{Z}$, then the lines of slope m_x passing through $[0, n]^2$ have 180 degree rotational symmetry, and we have $l_j = l_{\bar{j}}$ for all j .

In general, however, $i \notin \mathbb{Z}$, and $g_{\lfloor i \rfloor}$ and $g_{\lceil i \rceil}$ are the nearest integer lines to this top corner. Let A denote the horizontal spacing between lines of slope m_x , and let B denote the vertical spacing. Then from Equation (3.8) we deduce

$$A = \frac{n}{x_2} \quad \text{and} \quad B = \frac{n}{x_1}.$$

Let a denote the horizontal distance from $g_{\lfloor i \rfloor}$ to $\langle n, n \rangle$ along the line $y = n$, and let b denote the vertical distance from $g_{\lceil i \rceil}$ to $\langle n, n \rangle$ along the line $x = n$. Then solving for these values concretely (again using (3.8)), we have:

$$a = \frac{n}{x_2} ((x_2 - x_1) - \lfloor x_2 - x_1 \rfloor) \quad \text{and} \quad b = \frac{n}{x_1} (\lceil x_2 - x_1 \rceil - (x_2 - x_1)).$$

For any $j \in \mathbb{Q}$, write $w(j) = -\frac{n}{y_2}j$ and $h(j) = \frac{n}{y_1}j$ to denote the width and height of the triangle formed by the origin and the x - and y -intercepts of the line f_j . Note that by definition, $w(j)$ (resp. $h(j)$) is also the horizontal (resp. vertical) distance between $f_{\bar{j}}$ and the corner $\langle n, n \rangle$.

For $j < y_1$, define l_j^{left} to be the number of integer-indexed lines g_i that intersect the line $x = 0$ with a positive vertical component at most $h(j)$ and define l_j^{bot} to be the number of lines g_i that intersect the line $y = 0$ with a positive horizontal component at most $w(j)$. Then counting the line g_0 that passes through the origin, we have

$$l_j = l_j^{\text{left}} + l_j^{\text{bot}} + 1.$$

Similarly, let l_j^{top} denote the number of lines g_i that intersect the line $y = n$ at a horizontal component $w(j) \leq x < n$, and let l_j^{right} denote the number of lines that intersect the line $x = n$ at height $h(j) \leq y < n$. Then (unless $x_2 - x_1 \in \mathbb{Z}$ as discussed earlier), we have

$$l_{\bar{j}} = l_j^{\text{top}} + l_j^{\text{right}}.$$

Now suppose $w(j) = mA + r$ for some $0 \leq r < A$. Then

$$l_j^{\text{bot}} = m \quad \text{and} \quad l_j^{\text{top}} = m + \begin{cases} 0 & \text{if } r < a, \\ 1 & \text{if } a \leq r < A. \end{cases}$$

Similarly if $h(j) = m'B + r'$,

$$l_j^{\text{left}} = m' \quad \text{and} \quad l_j^{\text{right}} = m' + \begin{cases} 0 & \text{if } r' < b, \\ 1 & \text{if } b \leq r' < B. \end{cases}$$

Summing, the result follows since $l_j^{\text{bot}} + l_j^{\text{left}} = l_j - 1$, and thus,

$$l_{\bar{j}} = l_j^{\text{top}} + l_j^{\text{right}} = l_j + \begin{cases} -1 & \text{if } r < a, r' < b \\ 1 & \text{if } a \leq r < A, b \leq r' < B, \\ 0 & \text{otherwise.} \end{cases} \quad \square$$

We now introduce notation to indicate mapping an arbitrary crossing set χ_j into an integer-indexed crossing set. For $j < y_1$, we define

$$\psi_j = \{\mathbf{w} + (\lceil j \rceil - j) \mathbf{x} : \mathbf{w} \in \chi_j\} \subseteq \chi_{\lceil j \rceil},$$

and for $j > |y_2|$, we define

$$\phi_j = \{\mathbf{w} - (j - \lfloor j \rfloor) \mathbf{x} : \mathbf{w} \in \chi_j\} \subseteq \chi_{\lfloor j \rfloor}.$$

By construction $\psi_j, \phi_j \in L_n^{\square}$.

Lemma 3.6. *Suppose $j < y_1$. If \mathbf{z} is the i th lattice point from the top in ψ_j , then $\mathbf{z} - (i + 1) \mathbf{x} \notin [0, n]^2$. Similarly, if $j > |y_2|$ and \mathbf{z} is the i th lattice point from the bottom in ϕ_j , then $\mathbf{z} + (i + 1) \mathbf{x} \notin [0, n]^2$.*

Proof. We argue only for the case of $j < y_1$, as the case of $j > |y_2|$ follows by 180 degree rotation.

First, we make the simple observation that because $\mathbf{x} - \mathbf{y}$ is in the first quadrant, $x_1 > y_1$. Suppose $\mathbf{z}' = \langle z_1, z_2 \rangle$ is the i th crossing from the top in χ_j . Then $\mathbf{z}' - i \mathbf{y} \notin [0, n]$, and $z_1 - iy_1 < 0$. But then $z_1 - ix_1 < z_1 - iy_1 < 0$ as well, and hence $\mathbf{z}' - i \mathbf{x} \notin [0, n]^2$.

Now, $\mathbf{z} = \mathbf{z}' + \delta \mathbf{x}$, where $\delta = (\lceil j \rceil - j) < 1$. Thus $\mathbf{z} - (i + 1) \mathbf{x} = \mathbf{z}' - i \mathbf{x} - (1 - \delta) \mathbf{x}$, which is clearly outside of $[0, n]^2$. \square

We now define the quantity

$$J_0 = \frac{y_1 y_2 (x_1 - x_2)}{x_1 y_2 + x_2 y_1}. \quad (3.16)$$

which will play a role in much of what follows.

Lemma 3.7. *Fix k such that $j = \max\{j_k, \bar{j}_k\} \leq J_0$. Let \mathbf{w} be the topmost point in χ_j and let $\mathbf{v} = \mathbf{w} + (\lceil j \rceil - j) \mathbf{x}$ be the topmost point in ψ_j . Similarly, let \mathbf{w}' be the topmost point in $\chi_{\bar{j}}$ and let $\mathbf{v}' = \mathbf{w}' - (\bar{j} - \lfloor \bar{j} \rfloor) \mathbf{x}$ be the topmost point in $\phi_{\bar{j}}$. Then there is an increasing path of length $\lfloor \bar{j} \rfloor - \lceil j \rceil$ from \mathbf{v} to \mathbf{v}' .*

Proof. For each j and \mathbf{v}, \mathbf{v}' defined as above, let m_j denote the slope of the line from \mathbf{v} to \mathbf{v}' . Corollary 2.8 says there is an increasing path of length $\lfloor \bar{j} \rfloor - \lceil j \rceil$ from \mathbf{v} to \mathbf{v}' if $m_{\mathbf{x}} \leq m_j \leq m_{\mathbf{u}}$.

We first show $m_j \geq m_{\mathbf{x}}$. It will be convenient to consider the slope of the line from \mathbf{w} to \mathbf{w}' , which we denote \tilde{m}_j . We will show that $\tilde{m}_j \geq m_{\mathbf{x}}$, which is sufficient to show $m_j \geq \tilde{m}_j$. Indeed, if $\tilde{m}_j \geq m_{\mathbf{x}}$, then the point $\mathbf{v} = \mathbf{w} + (\lceil j \rceil - j) \mathbf{x}$ is weakly below the line segment connecting \mathbf{w} and \mathbf{w}' , and $\mathbf{v}' = \mathbf{w}' - (\bar{j} - \lfloor \bar{j} \rfloor) \mathbf{x}$ is weakly above it. Thus $m_j \geq \tilde{m}_j \geq m_{\mathbf{x}}$. So let us show

that $\tilde{m}_j \geq m_x$. Since $w + (\bar{j} - j)x$ is in $f_{\bar{j}}$, it is enough to show that $w + (\bar{j} - j)x$ has vertical component at most n .

Note that w is very near the point $(0, \frac{nj}{y_1})$, the point at which line f_j intersects the left edge of $[0, n]^2$. In particular, $w_2 \leq \frac{nj}{y_1}$. Computing this vertical component we find:

$$\begin{aligned} w_2 + (\bar{j} - j)x_2 &< \frac{nj}{y_1} + (\bar{j} - j)x_2 = \frac{nj}{y_1} + (\bar{j} + j - 2j)x_2 \\ &\leq j \left(\frac{n}{y_1} - 2x_2 \right) + (y_1 - y_2)x_2. \end{aligned}$$

Using $j \leq \frac{y_1 y_2 (x_1 - x_2)}{x_1 y_2 + x_2 y_1}$ we find

$$\begin{aligned} j \left(\frac{n}{y_1} - 2x_2 \right) + (y_1 - y_2)x_2 &\leq \frac{y_1 y_2 (x_1 - x_2)}{x_1 y_2 + x_2 y_1} \left(\frac{n}{y_1} - 2x_2 \right) + (y_1 - y_2)x_2 \\ &= \frac{y_2 (x_1 - x_2) (n - 2x_2 y_1)}{x_1 y_2 + x_2 y_1} + (y_1 - y_2)x_2 \\ &= -y_2 (x_1 - x_2) + (y_1 - y_2)x_2 \\ &= -y_2 x_1 + y_1 x_2 = n, \end{aligned}$$

where we have used $n = -y_2 x_1 + y_1 x_2$ in two different places. Thus $w + (\bar{j} - j)x$ has vertical component at most n , and $\tilde{m}_j \geq m_x$ as desired.

To show $m_j \leq m_u$, we use a more direct comparison of the two slopes. The line f_j hits the left edge of the box $[0, n]^2$ at height nj/y_1 , so v can be no lower than $nj/y_1 + y_2$. On the other hand, v can be no farther to the right than x_1 , since otherwise $v - x$ would still be in the box. With this in mind, we denote $A = (x_1, nj/y_1 + y_2)$, which is weakly below and to the right of point v .

Similarly, we find a point B which is weakly above and left of v' . The point v' is no higher than n itself. The line $f_{\bar{j}}$ intersects the top edge of the box at horizontal component $-n(\bar{j} - y_1)/y_2$, and so point v' can be no farther left than $-n(\bar{j} - y_1)/y_2 - x_1$. So we can take B to be the point $B = (-n(\bar{j} - y_1)/y_2 - x_1, n)$.

Denoting the slope of the line between the points A and B as m_{AB} , we see that $m_j \leq m_{AB}$, so it suffices to show $m_{AB} \leq m_u$. We can write the slope m_{AB} explicitly:

$$\begin{aligned} m_{AB} &= \frac{n - nj/y_1 - y_2}{-n(\bar{j} - y_1)/y_2 - x_1 - x_1} \\ &= \frac{ny_2(y_1 - j) - y_2^2 y_1}{-ny_1(\bar{j} - y_1) - 2x_1 y_1 y_2} \\ &= \frac{-ny_2(y_1 - j) + y_2^2 y_1}{ny_1(-y_2 - j) + 2x_1 y_1 y_2}. \end{aligned}$$

We now show that $m_{AB} < u_2/u_1 = m_u$ for all $0 < j \leq y_1$. Since $J_0 \leq y_1$, this will complete the proof. Let

$$\begin{aligned} N(j) &= u_1(-ny_2(y_1 - j) + y_2^2 y_1) = j \cdot ny_2 u_1 - ny_2 y_1 u_1 + y_2^2 y_1, \\ D(j) &= u_2(ny_1(-y_2 - j) + 2x_1 y_1 y_2) = -j \cdot ny_1 u_2 - ny_2 y_1 u_2 + 2x_1 y_2 y_1. \end{aligned}$$

Then $m_{AB} < u_2/u_1$ precisely when $N(j) < D(j)$. Solving the inequality $N(j) < D(j)$ for j yields

$$j < \frac{-y_2y_1(u_2 - u_1 + 2x_1/n - y_2/n)}{y_2u_1 + y_1u_2}. \quad (3.17)$$

(Note that we have divided by $y_2u_1 + y_1u_2$, which is necessarily positive since $u_2 > -y_2 > 0$ and $y_1 > u_1 > 0$.) We are left to show that the right hand side of (3.17) is bounded by y_1 .

Since

$$\mathbf{u} = \mathbf{x} - a\mathbf{y}, \quad a = \lfloor x_1/y_1 \rfloor,$$

we can rewrite the right hand side of (3.17) as

$$\frac{-y_2y_1(x_2 - x_1 + (2x_1 - y_2)/n + a(y_1 - y_2))}{x_2y_1 + y_2x_1 - 2ay_1y_2}. \quad (3.18)$$

Consider this expression as a function of a (we know that $a = \lfloor x_1/y_1 \rfloor$, but consider a as free for now). There is a vertical asymptote at $a = x_1/(2y_1) + x_2/(2y_2) < \lfloor x_1/y_1 \rfloor$ (recall that $x_1 > y_1$). To the right of this asymptote the denominator is clearly positive, since it is positive for large enough a . (In fact, at $a = x_1/y_1$, the denominator is $x_2y_1 - y_2x_1 = n$.) We can check that the numerator is positive as well by plugging in $a = x_1/(2y_1) + x_2/(2y_2)$:

$$\begin{aligned} & -y_2y_1(x_2 - x_1 + (2x_1 - y_2)/n + (x_1/(2y_1) + x_2/(2y_2))(y_1 - y_2)) = \\ & -y_2y_1 \left[\frac{x_1}{2} \left(-\frac{y_2}{y_1} - 1 + \frac{4}{n} \right) + \frac{x_2(y_1 + y_2)}{2y_2} - \frac{y_2}{n} \right]. \end{aligned}$$

Since $0 < y_1 \leq -y_2$, each of the three terms inside the brackets is non-negative, and the first and last terms are positive, so the entire expression is as well.

Since both the numerator and denominator of (3.18) are positive for a immediately to the right of the vertical asymptote, the expression is decreasing in $a > x_1/(2y_1) + x_2/(2y_2)$. It follows that the value of (3.18) at $a = \lfloor x_1/y_1 \rfloor > x_1/(2y_1) + x_2/(2y_2)$ is greater than the value in the limit as $a \rightarrow \infty$, which is

$$\lim_{a \rightarrow +\infty} \frac{-y_2y_1a(y_1 - y_2)}{-2ay_1y_2} = \frac{y_1 - y_2}{2} \geq y_1.$$

Thus $N(j) < D(j)$, and therefore $m_{AB} < m_{\mathbf{u}}$, for all $0 < j \leq y_1$. \square

3.4. Constructive use of Greene's Theorem

Recall $I_k = \lambda_1 + \dots + \lambda_k$ denotes the size of the largest subsequence formed by the union of k increasing subsequences. We will prove a nearly exact formula for I_k in terms of counting lines of slope $m_{\mathbf{y}}$, which in turn will give us a uniform estimate for λ_k .

First, we make the easy observation that any union of k increasing subsequences can intersect each crossing set χ_i in at most k points. Otherwise, since \mathbf{y} is negatively sloped, we would have a decreasing run of size $k + 1$. Making this observation for each integer i yields the following upper bound for I_k .

Observation 3.8. For any k ,

$$I_k \leq \sum_{i=1}^{\lfloor y_1 - y_2 \rfloor} \min(l_i, k).$$

We will prove that the above inequality is nearly an equality when k is not too big by constructing a collection of k increasing paths in L_n^{\square} . Here is the general idea of the construction. Fix $j = \max\{j_k, \overline{j'_k}\}$ as in Lemma 3.7, and let P denote the path from the topmost lattice point in ψ_j to the topmost lattice point in $\phi_{\overline{j}}$.

Define set $S = S(j)$ to be the union of the following four sets:

- \mathcal{L}_j , containing of all points in $\chi_1, \chi_2, \dots, \chi_{\lfloor j \rfloor - 1}$,
- \mathcal{M}_j , containing all points in paths $P, P + \mathbf{y}, \dots, P + (k - 1)\mathbf{y}$, (these connect the top k points in ψ_j with the top k points in $\phi_{\overline{j}}$)
- \mathcal{R}_j , containing all points in $\chi_{\lfloor \overline{j} \rfloor + 3}, \chi_{\lfloor \overline{j} \rfloor + 4}, \dots, \chi_{\lfloor y_1 - y_2 \rfloor}$, and
- X , containing all lattice points in $\chi_{\lfloor \overline{j} \rfloor + 1}$ and $\chi_{\lfloor \overline{j} \rfloor + 2}$ of the form $\mathbf{w} + \mathbf{x}$ or $\mathbf{w} + 2\mathbf{x}$, where \mathbf{w} is one of the topmost k points in $\phi_{\overline{j}}$.

See Figure 3.1. We now show S is a union of k increasing paths of cardinality at least 3 less than the upper bound of Observation 3.8.

Proposition 3.9. For any $k \leq l_{J_0}$, let $j = \max\{j_k, \overline{j'_k}\}$. Then the set $S = S(j)$ defined above is a union of k increasing paths. Moreover, $|S| \geq -3 + \sum_{i=1}^{\lfloor y_1 - y_2 \rfloor} \min(l_i, k)$. Thus,

$$-3 + \sum_{i=1}^{\lfloor y_1 - y_2 \rfloor} \min(l_i, k) \leq I_k \leq \sum_{i=1}^{\lfloor y_1 - y_2 \rfloor} \min(l_i, k).$$

Proof. Without loss of generality, we may assume $j_k \leq \overline{j'_k}$, so that $j = j_k$. The case of $\overline{j'_k} < j_k$ follows the same approach under 180 degree rotation.

First, we see each lattice point in \mathcal{L}_j can be connected by some number of \mathbf{x} steps to one of the translates of path P that make up \mathcal{M}_j . This is obvious if $l_j = k$. If $l_j = k + 1$ and $l_{j-\epsilon} < k$, then because l_j is a step function whose steps are of size 1 or 2, it must be that there are crossings at both the points where f_j intersects the boundary of $[0, n]^2$. Hence, $\psi_{j-\epsilon}$ is a subset of the middle $k - 1$ points in ψ_j , which are all in \mathcal{M}_j .

Similarly, any point in sets \mathcal{R}_j or X can be connected by some number of $-\mathbf{x}$ steps to one of these paths at a point in $\phi_{\overline{j}}$. The points of X are defined to be those points of the form $\mathbf{w} + \mathbf{x}$ and $\mathbf{w} + 2\mathbf{x}$, where \mathbf{w} is one of the top k points in $\phi_{\overline{j}}$, so they are obviously part of the same k increasing sets. The argument for \mathcal{R}_j requires some more care.

From Lemma 3.5, we have $|l_j - l_{\overline{j}}| \leq 1$, so because $l_j \leq k + 1$, we have $l_{\overline{j}} \leq k + 2$. From Lemma 3.6 we know that if \mathbf{z} and $\mathbf{z} + \mathbf{y}$ are the two bottommost points in $\phi_{\overline{j}}$, then $\mathbf{z} + 3\mathbf{x} \notin [0, n]^2$, and $(\mathbf{z} + \mathbf{y}) + 2\mathbf{x} \notin [0, n]^2$. Hence crossing set $\chi_{\lfloor \overline{j} \rfloor + 3}$ has at most $l_{\overline{j}} - 2 \leq k$ points, and moreover, each of them can be connected by $-3\mathbf{x}$ to one of the top k points in $\phi_{\overline{j}}$.

Thus, all points in S are connected to one of the k translates of path P and S is the union of k increasing sets.

For the claimed lower bound on the cardinality of S , we observe

$$|S| = |\mathcal{L}_j| + |\mathcal{M}_j| + |\mathcal{R}_j| + |X|,$$

where

$$|\mathcal{L}_j| = \sum_{i=1}^{\lceil j \rceil - 1} l_i, \quad |\mathcal{M}_j| = \sum_{i=\lceil j \rceil}^{\lfloor j \rfloor} k, \quad \text{and} \quad |\mathcal{R}_j| = \sum_{i=\lfloor j \rfloor + 3}^{\lfloor y_1 - y_2 \rfloor} l_i.$$

Finally, we observe that if \mathbf{v} is the topmost lattice point in $\phi_{\bar{j}}$, then

$$\phi_{\bar{j}} \subseteq \{\mathbf{v}, \mathbf{v} + \mathbf{y}, \dots, \mathbf{v} + k\mathbf{y}, \mathbf{v} + (k+1)\mathbf{y}\}.$$

Therefore, $X = \chi_{\lfloor \bar{j} \rfloor + 1} \cup \chi_{\lfloor \bar{j} \rfloor + 2} - \{\mathbf{v} + k\mathbf{y} + \mathbf{x}, \mathbf{v} + (k+1)\mathbf{y} + \mathbf{x}, \mathbf{v} + k\mathbf{y} + 2\mathbf{x}\}$, and so

$$|X| \geq \min(l_{\lfloor \bar{j} \rfloor + 1}, k) - 2 + \min(l_{\lfloor \bar{j} \rfloor + 2}, k) - 1,$$

which completes the proof. \square

3.5. Bounding the shape of w

Let $\text{lines}(k) = \lfloor j'_k \rfloor - \lceil j_k \rceil + 1$ denote the number of integers i such that $l_i \geq k$. Since

$$\lambda_k = I_k - I_{k-1},$$

and since $\sum_i \min(l_i, k) - \sum_i \min(l_i, k-1) = \text{lines}(k)$, we can translate Proposition 3.9 into the following bound on λ_k , provided $k \leq l_{j_0}$:

$$|\lambda_k - \text{lines}(k)| \leq 3. \quad (3.19)$$

The analogous result for λ'_k follows by applying the transformation ρ defined in in Subsection 3.1, which takes the pair (\mathbf{x}, \mathbf{y}) satisfying (3.1) to the pair $(\mathbf{x}^*, \mathbf{y}^*) = (\langle -y_2, y_1 \rangle, \langle x_2, -x_1 \rangle)$ satisfying (3.2). As discussed, each decreasing path in the lattice spanned by (\mathbf{x}, \mathbf{y}) corresponds to an increasing path in the lattice spanned by $(\mathbf{x}^*, \mathbf{y}^*)$.

Define

$$J_0^* = \frac{y_1^* y_2^* (x_1^* - x_2^*)}{x_1^* y_2^* + x_2^* y_1^*} = \frac{x_1 x_2 (y_1 + y_2)}{x_1 y_2 + x_2 y_1},$$

and let $l_j^* = |f_j^* \cap L_n^{\square}|$, where $f_j^* = j\mathbf{x}^* + t\mathbf{y}^*$ is the line of slope $m_{\mathbf{y}^*}$ passing through $j\mathbf{x}^*$.

Now, using the estimates in Proposition 3.4 (3.14) and (3.15), we deduce

$$\left| (y_1 - y_2) - \frac{2|y_1 y_2|}{n} k - \text{lines}(k) \right| < \frac{2|y_1 y_2|}{n} + 1.$$

With the characterization of \mathbf{y} in either (3.5) or (3.6), we have

$$\frac{|y_1 y_2|}{n} = q_{2h+1}^2 \delta_{2h+1}.$$

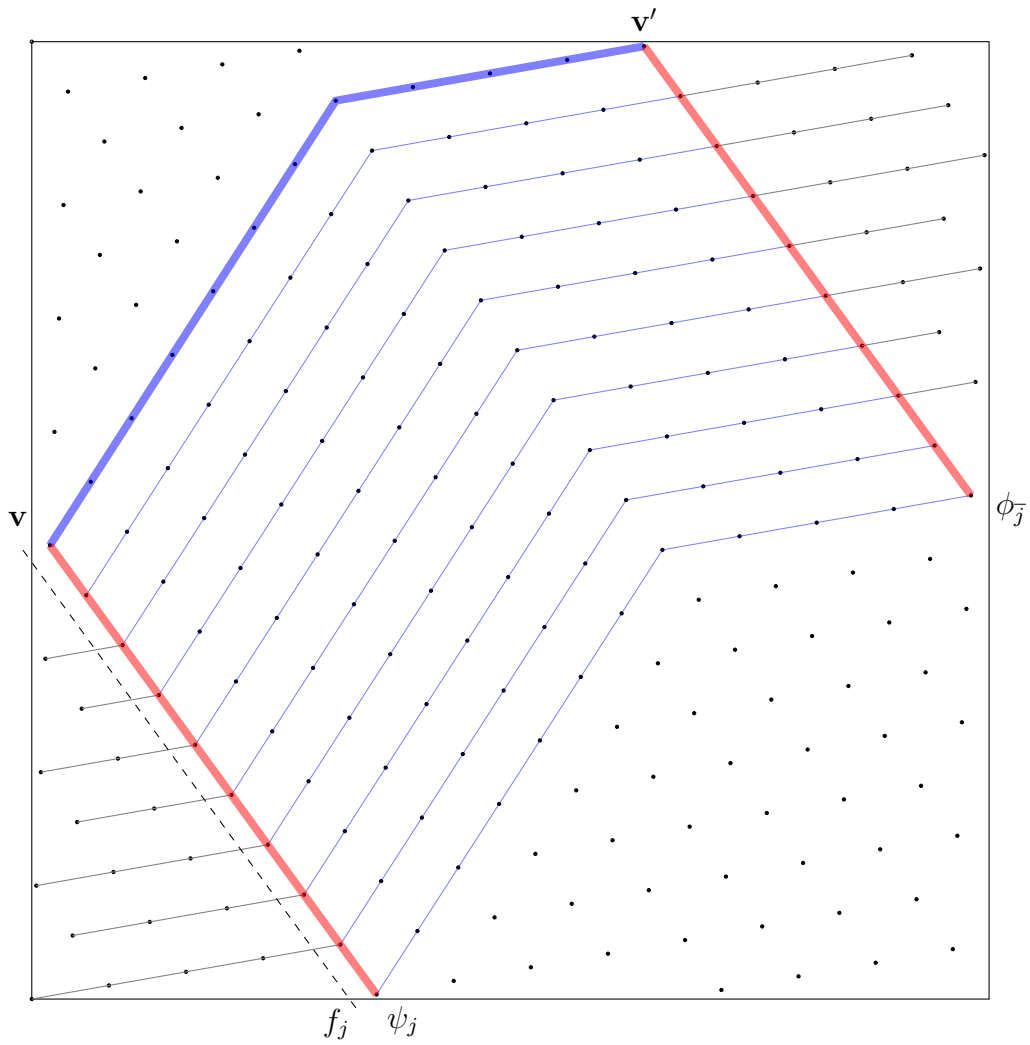


Figure 3.1: Example construction of set S . The blue path illustrates an increasing path P constructed from v , the top point on ψ_j , to v' , the top point on $\phi_{\bar{j}}$. Here $a = 25$, $n = 210$, and $N = 211$. This yields $x = \langle 17, 3 \rangle$ and $y = \langle 8, -11 \rangle$.

Recalling the Diophantine estimate for δ_{2h+1} in (1.3), we have

$$q_{2h+1}^2 \delta_{2h+1} < \frac{q_{2h+1}^2}{q_{2h+2} q_{2h+1}} = \frac{q_{2h+1}}{q_{2h+2}} < 1.$$

Thus, $\frac{2|y_1 y_2|}{n} + 1 < 3$.

We can now state simple bounds on λ_k , and by symmetry, λ'_k .

Proposition 3.10. *For $k \leq l_{J_0}$,*

$$\left| (y_1 - y_2) - \frac{2|y_1 y_2|}{n} k - \lambda_k \right| < 4 + \frac{2|y_1 y_2|}{n} < 6,$$

and for $k^* \leq l_{J_0}^*$,

$$\left| (x_1 + x_2) - \frac{2x_1 x_2}{n} k^* - \lambda'_{k^*} \right| < 4 + \frac{2x_1 x_2}{n} < 6.$$

Proposition 3.10 establishes that there are two lines that roughly approximate the boundary of shape A_λ :

$$y = (x_1 + x_2) - \frac{2x_1 x_2}{n} x \quad \text{and} \quad x = (y_1 - y_2) - \frac{2|y_1 y_2|}{n} y. \quad (3.20)$$

One can verify (using $x_2 y_1 - y_2 x_1 = n$) that these two lines intersect at the point (x_0, y_0) given by

$$x_0 = \frac{n(y_1 + y_2)}{x_1 y_2 + x_2 y_1} \quad \text{and} \quad y_0 = \frac{n(x_2 - x_1)}{x_1 y_2 + x_2 y_1}. \quad (3.21)$$

Furthermore, from Proposition 3.4 (3.11) we see this point of intersection is very nearly $(l_{J_0}^*, l_{J_0})$ since

$$l_{J_0}^* - 1 \leq x_0 \leq l_{J_0}^* + 1,$$

and

$$l_{J_0} - 1 \leq y_0 \leq l_{J_0} + 1.$$

Let $L(x)$ denote the piecewise linear function

$$L(x) = \begin{cases} (x_1 + x_2) - \frac{2x_1 x_2}{n} x & \text{for } 0 \leq x \leq x_0, \\ \frac{n(y_1 - y_2 - x)}{2|y_1 y_2|} & \text{for } x_0 \leq x \leq y_1 - y_2, \end{cases}$$

and recall that ∂A_λ denotes the boundary of A_λ that does not lie on a coordinate axis. For fixed x , let $\text{dist}(L(x), \partial A_\lambda)$ denote the minimum Euclidean distance from $(x, L(x))$ to a point on ∂A_λ . Proposition 3.10 gives us a uniform bound of 6 on this distance for all points (x, y) with $x \leq x_0 - 1$ and $y \geq y_0 + 2x_1 x_2 / n$, as well as for all (x, y) with $y \leq y_0 - 1$ and $x \geq x_0 + 2|y_1 y_2| / n$.

We now carefully examine the geometry near the point (x_0, y_0) . Proposition 3.10 indicates that at vertical component $y_0 - 1$, the graph of L has horizontal component $x_0 + 2|y_1 y_2| / n$, and the horizontal component of the boundary must satisfy

$$x_0 - 4 < x < x_0 + 4 \left(1 + \frac{|y_1 y_2|}{n} \right) < x_0 + 8.$$

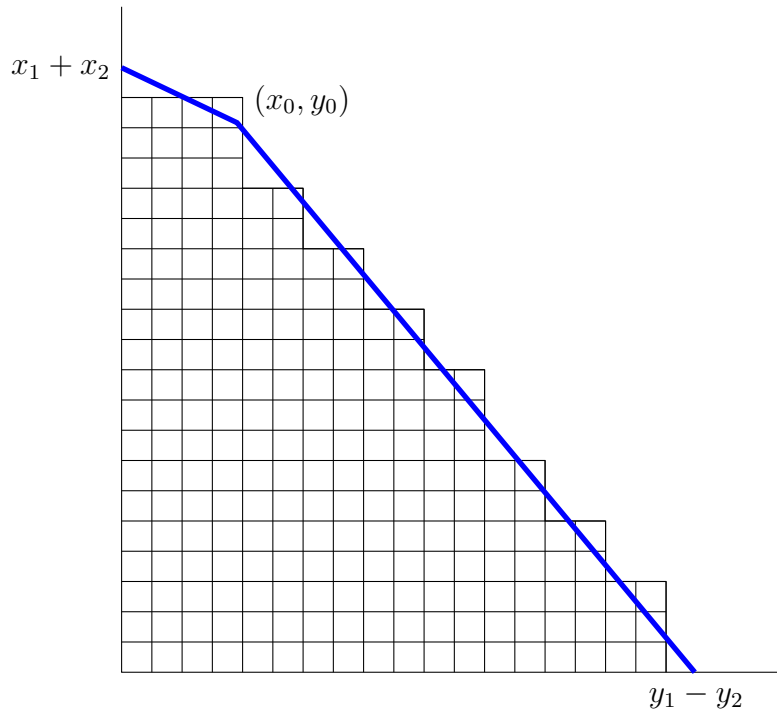


Figure 3.2: The planar set A_λ for $w(210, \frac{25}{211})$, along with the piecewise linear curve $L(x)$.

Similarly, at horizontal component $x_0 - 1$, the vertical component is bounded by

$$y_0 - 4 < y < y_0 + 4 \left(1 + \frac{x_1 x_2}{n} \right) < y_0 + 8.$$

Thus we see the boundary of A_λ must pass through the rectangle given by $(x_0 - 4, x_0 + 8) \times (y_0 - 4, y_0 + 8)$, and so Proposition 3.10 can now be stated in the following form.

Corollary 3.11. For all $0 \leq x \leq y_1 - y_2$,

$$\text{dist}(L(x), \partial A_\lambda) < 8.$$

For example, with $a = 25$, $n = 210$ and $N = 211$ with $w(n, a/N)$ as shown in Figure 3.1, we have $\mathbf{x} = \langle 17, 3\frac{210}{211} \rangle$, $\mathbf{y} = \langle 8, -11\frac{210}{211} \rangle$, and therefore we can compute

$$L(x) = \begin{cases} \frac{4217}{211} - \frac{102}{211}x & \text{for } 0 \leq x \leq \frac{622}{163}, \\ \frac{1999}{88} - \frac{211}{176}x & \text{for } \frac{622}{163} \leq x \leq \frac{3998}{211}. \end{cases}$$

In Figure 3.2 we see $L(x)$ superimposed on A_λ .

Remark 3.12. As mentioned in the introduction, the paper [BCKPT21] actually studied the slightly more general set of permutations generated by fractional parts of lines $f_{\alpha,\beta}(i) = \alpha i + \beta \pmod 1$, depending on two independent real parameters. As discussed in [BCKPT21], it is easily shown that varying β (for fixed n and α) acts by cyclic shifting of the Sós permutation. Intuitively, a cyclic shift has little effect on monotone subsequences. From our perspective, the

vertical shift by β amounts to viewing the lattice $L[\mathbf{x}, \mathbf{y}]$ within the box $[0, n] \times [\beta', n + \beta']$, where $\beta' = n\beta$. All our definitions, lemmas, and constructions are easily adapted to this shift. We leave details to the interested reader.

3.6. Proofs for irrational α

We have already noted that Proposition 3.2 implies Theorem 1.3 for rational numbers. Similarly, by substituting the relevant expressions for x_1, x_2, y_1, y_2 from (3.5) or (3.6) into Proposition 3.10 and Corollary 3.11, we prove Theorems 1.5 and 1.7, respectively, for rational numbers. We now show how to extend those results to all reals.

Fix $n \in \mathbb{N}$ and an irrational number $0 < \alpha < 1$. Denote the sequence of convergents for α as $p_0/q_0, p_1/q_1, \dots$, and let i be the smallest non-negative integer such that

$$\delta_i = \left| \alpha - \frac{p_i}{q_i} \right| \leq \frac{1}{n}. \quad (3.22)$$

If $i = 2h + 2$ is even, then $\frac{p_i}{q_i} < \alpha$, and we have

$$\frac{p_{2h+1}}{q_{2h+1}} - \alpha \geq \frac{1}{n}, \quad 0 < \alpha - \frac{p_{2h+2}}{q_{2h+2}} < \frac{1}{n}.$$

In this case we apply the program of Section 3 with $a = p_m$ and $N = q_m$, where m is an *even* integer satisfying $q_m > n$. Note then that $\frac{p_m}{q_m} < \alpha$, so we have

$$\frac{p_{2h+1}}{q_{2h+1}} - \frac{p_m}{q_m} > \frac{p_{2h+1}}{q_{2h+1}} - \alpha \geq \frac{1}{n}, \quad 0 < \frac{p_m}{q_m} - \frac{p_{2h+2}}{q_{2h+2}} < \alpha - \frac{p_{2h+2}}{q_{2h+2}} < \frac{1}{n},$$

which implies that the convergents appearing in (3.5) are the same as those defined by (3.22).

Similarly, if $i = 2h + 1$ is odd, then we take $a = p_m$ and $N = q_m$, where m is an *odd* integer satisfying $q_m > n$ to yield that the convergents appearing in (3.6) are the same as those defined by (3.22).

With these choices of a and N , all the results of Section 3 can be applied to α . Let $\alpha(m) = p_m/q_m$ denote the convergent we have chosen. Since $\alpha(m)$ has denominator $q_m > n$, both α and $\alpha(m)$ belong to the same Farey interval in $F^{(n)}$. But then Theorem 1.2 tells us the corresponding Sós permutations are equal: $w(n, \alpha) = w(n, \alpha(m))$. Furthermore, these numbers α and $\alpha(m)$ induce same Schensted shape A_λ .

For the rational approximation $\alpha(m)$, let

$$\delta_i(m) = \left| \alpha(m) - \frac{p_i}{q_i} \right|.$$

For fixed n , m can be chosen arbitrarily large, provided m has the desired parity and $q_m > n$. Since our results hold for any such m , we may then take the limit as $m \rightarrow \infty$, in which case $\delta_i(m) \rightarrow \delta_i$. This proves Theorems 1.3, 1.5, and 1.7 for any real α .

4. Further directions

We hope the results in this paper can serve as a starting point for further study. Here are a few directions they could lead.

- As described in the **Big Question** in the introduction, our original motivation was to study the Schensted shape of *random* Sós permutations. We currently cannot describe the probability distribution on partitions which arises from the uniform distributions on $Sós_n$ via Schensted insertion. It would be interesting to see if there is an average shape which has a scaling limit as $n \rightarrow \infty$.
- The set $Sós_n$ represents a rather extreme restriction of the full symmetric group, which is the reason for the markedly different quality of the Schensted shape of a Sós permutation compared to a random one. Numerical experiments indicate that other sequences of the form $(f(1), f(2), \dots, f(n))$ have Schensted shapes similar to that of a uniformly random permutation (in all of S_n) for a fairly general class of nonlinear functions f , including nonlinear polynomials. If f is chosen to interpolate between a nonlinear function and a linear one as $n \rightarrow \infty$, it is very plausible that the Schensted shape would interpolate between the Logan–Shepp/Vershik–Kerov limit shape and the piecewise linear one studied in this paper. For example, one could consider $f(i) = n^{-\gamma}\beta i^2 + \alpha i \pmod 1$ for fixed irrational numbers α and β and an appropriate scaling exponent $\gamma > 0$.
- For pattern avoiding permutations, there are some results on the distribution of the longest increasing subsequence as $n \rightarrow \infty$ [DHW03, MP14]. One could also consider their full Schensted shape. The paper [MP14] considers a different notion of shape, namely the density of points on the graph $(i, w(i))$.

Acknowledgements

Both authors are grateful to Katherine Stange for invaluable discussions. Dan Romik provided helpful references.

References

- [AB98] Pascal Alessandri and Valérie Berthé. Three distance theorems and combinatorics on words. *Enseign. Math. (2)*, 44(1-2):103–132, 1998.
- [BCKPT21] Sarah Bockting-Conrad, Yevgenia Kashina, T. Kyle Petersen, and Bridget Eileen Tenner. Sós Permutations. *Amer. Math. Monthly*, 128(5):407–422, 2021. doi:10.1080/00029890.2021.1889911.
- [Ble90] P. M. Bleher. The energy level spacing for two harmonic oscillators with golden mean ratio of frequencies. *J. Statist. Phys.*, 61(3-4):869–876, 1990. doi:10.1007/BF01027305.

- [Ble91] P. M. Bleher. The energy level spacing for two harmonic oscillators with generic ratio of frequencies. *J. Statist. Phys.*, 63(1-2):261–283, 1991. doi:10.1007/BF01026603.
- [BS79] David W. Boyd and J. Michael Steele. Monotone subsequences in the sequence of fractional parts of multiples of an irrational. *J. Reine Angew. Math.*, 306:49–59, 1979.
- [BS08] Ian Biringer and Benjamin Schmidt. The three gap theorem and Riemannian geometry. *Geom. Dedicata*, 136:175–190, 2008. doi:10.1007/s10711-008-9283-8.
- [DHW03] Emeric Deutsch, A. J. Hildebrand, and Herbert S. Wilf. Longest increasing subsequences in pattern-restricted permutations. *Electron. J. Combin.*, 9(2):Research paper 12, 8, 2002/03. Permutation patterns (Otago, 2003). URL: <https://www.combinatorics.org/ojs/index.php/eljc/article/view/v9i2r12>.
- [dJS78] A. del Junco and J. Michael Steele. Growth rates for monotone subsequences. *Proc. Amer. Math. Soc.*, 71(2):179–182, 1978. doi:10.2307/2042828.
- [ES35] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935. URL: http://www.numdam.org/item?id=CM_1935__2__463_0.
- [Gre74] Curtis Greene. An extension of Schensted’s theorem. *Advances in Math.*, 14:254–265, 1974. doi:10.1016/0001-8708(74)90031-0.
- [Ham72] J. M. Hammersley. A few seedlings of research. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics*, pages 345–394, 1972.
- [Har52] S. Hartman. Über die Abstände von Punkten $n\xi$ auf der Kreisperipherie. *Ann. Soc. Polon. Math.*, 25:110–114 (1953), 1952.
- [HW08] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.
- [LS77] B. F. Logan and L. A. Shepp. A variational problem for random Young tableaux. *Advances in Math.*, 26(2):206–222, 1977. doi:10.1016/0001-8708(77)90030-5.
- [MP14] Sam Miner and Igor Pak. The shape of random pattern-avoiding permutations. *Adv. in Appl. Math.*, 55:86–130, 2014. doi:10.1016/j.aam.2013.12.004.
- [Nar17] Terumi Narushima. *Microtonality and the Tuning Systems of Erv Wilson: Mapping the Harmonic Spectrum*. Routledge, 2017.
- [Rom15] Dan Romik. *The surprising mathematics of longest increasing subsequences*, volume 4 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, New York, 2015.
- [S58] V. Sós. On the distribution mod 1 of the sequence $n\alpha$. *Ann. Univ. Sci. Budapest Eötvös Sect. Math.*, 1:127–134, 1958.

- [Ś59] S. Świerczkowski. On successive settings of an arc on the circumference of a circle. *Fund. Math.*, 46:187–189, 1959. doi:10.4064/fm-46-2-187-189.
- [Sch61] C. Schensted. Longest increasing and decreasing subsequences. *Canadian J. Math.*, 13:179–191, 1961. doi:10.4153/CJM-1961-015-3.
- [Ste95] J. Michael Steele. Variations on the monotone subsequence theme of Erdős and Szekeres. In *Discrete probability and algorithms (Minneapolis, MN, 1993)*, volume 72 of *IMA Vol. Math. Appl.*, pages 111–131. Springer, New York, 1995. doi:10.1007/978-1-4612-0801-3_9.
- [Sur58] J. Surányi. Über die Anordnung der Vielfachen einer reellen Zahl mod 1. *Ann. Univ. Sci. Budapest Eötvös Sect. Math.*, 1:107–111, 1958.
- [VK77] A. M. Veršik and S. V. Kerov. Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux. *Dokl. Akad. Nauk SSSR*, 233(6):1024–1027, 1977.
- [vR86] Tony Peter van Ravenstein. *Number sequences and phyllotaxis*. ProQuest LLC, Ann Arbor, MI, 1986. Thesis (Ph.D.)—University of Wollongong (Australia). URL: <https://ro.uow.edu.au/theses/1555/>.
- [vR88] Tony van Ravenstein. The three gap theorem (Steinhaus conjecture). *J. Austral. Math. Soc. Ser. A*, 45(3):360–370, 1988.