Permutations sortable by two stacks in parallel and quarter plane walks

Michael Albert

Department of Computer Science, University of Otago, PO Box 56, Dunedin 9054, New Zealand

Mireille Bousquet-Mélou

CNRS, LaBRI, Université de Bordeaux, 351 cours de la Libération, 33405 Talence Cedex, France

Abstract

At the end of the 1960s, Knuth characterised the permutations that can be sorted using a stack in terms of forbidden patterns. He also showed that they are in bijection with Dyck paths and thus counted by the Catalan numbers. Subsequently, Even & Itai, Pratt and Tarjan studied permutations that can be sorted using two stacks in parallel. This problem is significantly harder. In particular, a sortable permutation can now be sorted by several distinct sequences of stack operations. Moreover, in order to be sortable, a permutation must avoid infinitely many patterns. The associated counting question has remained open for 40 years. We solve it by giving a pair of functional equations that characterise the generating function of permutations that can be sorted with two parallel stacks.

The first component of this system describes the generating function $Q(a, u)$ of square lattice loops confined to the positive quadrant, counted by the length and the number of North-West and East-South factors. Our analysis of the asymptotic number of sortable permutations relies at the moment on two intriguing conjectures dealing with the series $Q(a, u)$. We prove that they hold for loops confined to the upper half plane, or not confined at all. They remain open for quarter plane loops. Given the recent activity on walks confined to cones, we believe them to be attractive *per se*.

Keywords: permutations, stacks, exact and asymptotic enumeration, quarter plane walks
2010 MSC: 05A05, 05A15, 05A99

1. Introduction

If we have a device whose only ability is to rearrange certain sequences of objects, it is natural to ask "What rearrangements can my device produce?" When the device is an abstract one that can operate on sequences of any size,

Email addresses: michael.albert@cs.otago.ac.nz (Michael Albert), bousquet@labri.fr (Mireille Bousquet-Mélou)

this becomes a combinatorial question. Such questions were apparently first considered by Knuth [19] who dealt with the case where the device was a stack, i.e. a storage mechanism operating in a last in, first out manner (Figure 1).

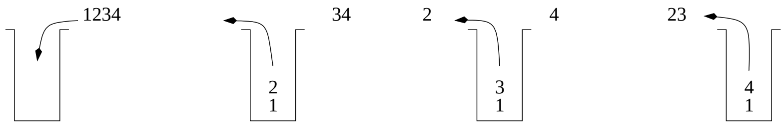

Figure 1: Four steps in the sequence of operations that outputs 2341 from 1234 using a stack. Each arrow shows an operation that is about to be performed.

Using a stack it is clear that the input sequence abc cannot produce the output sequence cab as, in order for c to be the first element output, both a and b must be in the stack together but then they will be output as ba and not as ab. This is in fact the only restriction: a permutation of an input sequence is achievable if and only if it never moves a later item (c) before two earlier items (a and b) while leaving the earlier items in order. In modern language, if the input is $12\cdots n$, the output permutations are those that *avoid the pattern* 312. The stack operations that will produce an output sequence from a given input sequence are easily seen to be uniquely determined. So, it is also routine to count such permutations and to discover that they are enumerated by the Catalan numbers. This is described in Section 2.2.1 of *The Art of Computer Programming* [19]. Knuth also establishes there similar results for *input-restricted deques* (double-ended queues).

Knuth's investigations, nicely described in terms of "railway yard switching networks", were extended by Even & Itai [12], Pratt [25] and Tarjan [29] who considered more general networks of stacks and queues, including the small network consisting of two parallel stacks that we study in this paper (Figure 2). This work was foundational for the study of *permutation classes* which can loosely be described as those collections of permutations that are closed by taking sub-permutations[1]. In our case, we observe indeed that any sub-permutation of a permutation that can be produced using two parallel stacks can itself be produced by this device simply by ignoring any operations that affect elements not in the sub-permutation. The study of permutation classes has been an active and growing field, often concentrating on enumeration, but also dealing with structural properties of these classes. For some general discussions and background see the books [4, 18, 22], and [3] for a survey on stack-sorting.

Despite this activity, most problems related to the rearranging power of Knuth's switchyard networks have turned out to be very hard. For networks consisting of two stacks, the case of parallel stacks seems a bit more manageable than that of two stacks in series. For instance, the list of minimal permutations that cannot be produced by two parallel stacks has been known since 1973 [25], but for two stacks in series it is only known to be infinite [23]. Similarly, it

[1]To be clear, given a permutation of $\{1, 2, \ldots, n\}$ written as a word in one line notation, we form a sub-permutation by taking any subword (of length k say) and then replacing the symbols of the subword by $\{1, 2, \ldots, k\}$ while maintaining their relative values. For instance taking the subword of 15324 occurring in the second, fourth and fifth positions (524) illustrates that 312 is a sub-permutation of 15324.

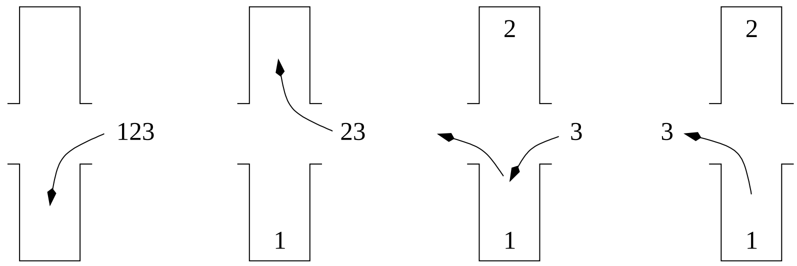

Figure 2: The permutation 312 cannot be produced with a single stack, but can be produced with two parallel stacks as shown here. Note that several distinct sequences of operations produce it.

has just been proved this year that one can decide in polynomial time if a permutation can be sorted by two stacks in series [24], while the corresponding result follows from a 1971 paper for two parallel stacks [12] (see also [26]). However, the questions "How many permutations of length n can be produced by two stacks in series, or by two stacks in parallel?" have remained equally open for forty years.

We answer the latter question in this paper, by giving a system of two functional equations that defines the generating function $\sum_n s_n t^n$, where s_n is the number of permutations of length n that can be produced with two parallel stacks. Denton [11] has presented an algorithm for this problem whose complexity is $O(n^5 2^n)$ (for enumerating the sortable permutations of length n). The form of the functional equations we obtain is such that we have, in principle, a polynomial time algorithm though we have not tried to estimate its precise complexity.

We also determine the exponential growth of the numbers s_n, modulo some conjectures that deal with square lattice walks confined to the quarter plane. These walks naturally encode the admissible sequences of stack operations, in the same way as Dyck paths do in the case of a single stack. Our conjectures deal with the enumeration of quarter plane walks counted by the length and by the number of *corners* of certain types. Walks confined to a quadrant have attracted a lot of attention in the past decade (see e.g. [5, 6, 7, 8, 21, 16]), and we think that our conjectures are interesting quite independently of the original stack sorting question.

Finally we remark that in this metaphor of "devices rearranging input" there are two common viewpoints. As described above, Knuth tended to view the input as arriving in fixed order $12 \cdots n$ and then the question is "How many permutations can be produced?". Tarjan on the other hand tended to think of the objective being to sort the input permutation, so the enumerative question becomes "How many permutations can be sorted?". Of course, passing to inverses, the two viewpoints are equivalent to one another: if a sequence of operations produces π from the identity, then the same sequence, applied to π^{-1}, produces the identity. We will be adopting the first viewpoint.

The outline of the paper is as follows. In Section 2, we describe a set of *canonical operation sequences* such that each permutation that can be produced using two parallel stacks is obtained by exactly one canonical operation sequence. In Section 3, we establish a system of functional equations that characterises the generating function of canonical sequences, and thus, of permutations that can be produced by two parallel stacks. The first equation in this system defines the generating function of quarter plane walks, weighted by their length and the number of North-West and East-South factors (also called *corners*). In Section 4, we state two conjectures about this generating function, and provide evidence for them by proving that they hold if we only impose on walks a half plane restriction, or no restriction at all. In Section 5 we derive from our system of equations the exponential growth of the number of permutations of length n produced by two parallel stacks, assuming the conjectures of Section 4. We conclude with a few comments on our results and conjectures in Section 6.

2. Canonical operation sequences

Throughout this paper we consider the action of two stacks in parallel, and attempt to count permutations of length n that such a machine can produce. These permutations are said to be *achievable*. The primary issue in this question, as opposed to the case of a single stack, is that there is no one-to-one correspondence between sequences of operations of the machine and achievable permutations. That is, several sequences of operations may produce the same permutation: we then say that they are *equivalent*. The most obvious case is that of the identity permutation of length n: there are at least 2^n ways to produce it using two stacks (alternate input and output operations, allowing the freedom of choice as to which stack to use — in fact there are more ways, since we can delay some output steps if we choose the following input to be to the other stack).

In this section we define a family of operation sequences, called *canonical*, such that each operation sequence is equivalent to exactly one canonical sequence. Canonical sequences are thus in one-to-one correspondence with achievable permutations.

In order to proceed further, we present three equivalent descriptions of operation sequences. Recall what the basic scenario is: input items numbered consecutively from 1 through n are processed by two stacks, each of which is capable of containing an arbitrarily large amount of data, but whose operations are limited to input (I) and output (O); an output operation produces the most recently entered item (i.e. items are processed in a last-in first-out fashion). Items are output as a sequence, and after all the input has been processed and the stacks emptied, the result is a permutation of the original input (Figure 2).

Operation sequences are encoded as words over the alphabet $\{I_1, I_2, O_1, O_2\}$, the subscripts determining which stack is referred to. Note that both stacks must be empty at the end, and that one cannot output from an empty stack. This means that a word over $\{I_1, I_2, O_1, O_2\}$ is an operation sequence if and only if it contains the same number of I_i as O_i letters for $i = 1, 2$, and, in each prefix, the number of I_i letters is at least as great as the number of O_i letters for $i = 1, 2$. Equivalently, it is a *shuffle* of two Dyck words, one on the letters I_1 and O_1, and the other on the letters I_2 and O_2. The *type* of an operation sequence is the word on $\{I, O\}$ obtained by deleting the subscripts on its letters.

We consider square lattice walks which begin at $(0, 0)$ and use steps $\mathsf{E} = (1, 0)$, $\mathsf{N} = (0, 1)$, $\mathsf{W} = (-1, 0)$ and $\mathsf{S} = (0, -1)$. Such a walk is a *loop* if it ends at $(0, 0)$. It is a *quarter plane walk* if it remains in the quadrant $\{(x, y) : x \geq 0, y \geq 0\}$. There is an obvious one-to-one correspondence between operation sequences and quarter plane loops (replace I_1 by E, I_2 by N, O_1 by W and O_2 by S). Under this correspondence, the (x, y) coordinate reached after processing a prefix of an operation sequence simply records the number of items in each stack at that point. The number of quarter plane loops consisting of $2n$ steps is well known to be $C_n C_{n+1}$, where $C_n = \binom{2n}{n}/(n+1)$is the n^{th} Catalan number [15, 2]. Observe that the type of an operation sequence corresponds to the projection of the associated loop on the diagonal $x = y$.

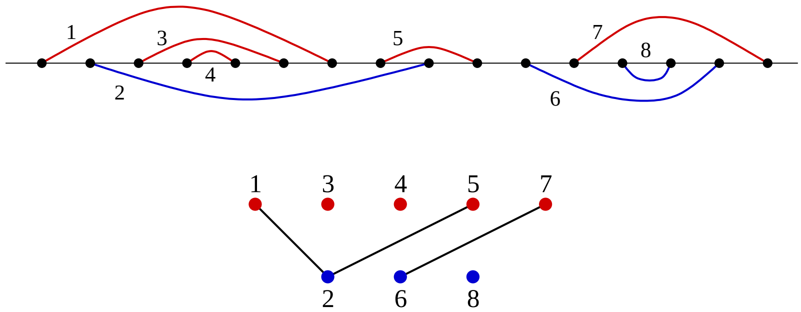

Figure 3: An illustration of the arch system associated with the operation sequence $I_1I_2I_1I_1O_1O_1O_1I_1O_2O_1I_2I_1I_2O_2O_2O_1$, and its associated graph. The arches are labelled using the left-to-right order of their left endpoint. This arch system has five connected components, and one left-right pair (between arches 2 and 5). The output permutation is 43125867.

A third perspective on these objects arises from considering them as bi-coloured *arch systems* (Figure 3). This is the two-dimensional counterpart of the standard bijection between Dyck paths and (one-coloured) arch systems [28, Exercise 6.19o]. For an operation sequence of length $2n$, take $2n$ points arranged along a line, labelled from 1 to $2n$. These points represent time, that is, the $2n$ steps of the operation sequence. For each item k in $\{1, \dots, n\}$, build an arch joining i to j where i (resp. j) is the time at which k is input to (resp. output from) a stack. If k is processed by the first stack, the arch will be above the line (and will be thought of as *red*), and otherwise below the line (and thought of as *blue*). Observe that the arches above the line do not cross, nor do the ones below the line — but there are no further restrictions on such systems. The operation sequence is easily recovered by scanning from left to right the $2n$ points of the arch system, writing I (resp. O) if an arch opens (resp. closes) at this point, and 1 (resp. 2) if this arch is above (resp. below) the line. Upon closing the supporting line into a cycle, an arch system can also be seen as a *rooted planar cubic map with a distinguished Hamiltonian cycle*. In this disguise, they were already considered by Tutte [30].

We use the following simple terminology:

- the *first* arch is the one which has least left endpoint; more generally, the k^{th} arch is the one with k^{th} smallest left endpoint;
- an arch joining i to j *moves* the element k of $\{1, \dots, n\}$ that is input to a stack at time i and output from at time j.

Observe that the k^{th} arch moves item k.

Our aim in this section is to describe a set of operation sequences in bijection with achievable permutations. A first observation is that two sequences obtained from one another by commuting pairs of adjacent letters I_1O_2 or I_2O_1 are equivalent. An operation sequence *outputs eagerly* if it contains neither I_1O_2 nor I_2O_1 as a factor. In other words, if the next item of the permutation which it is producing is already present in one of the two stacks (necessarily at the top of the stack), then it is output immediately, before any other input (necessarily to the other stack) is carried out. Such sequences correspond to walks in the plane containing no ES or NW factor and to arch systems in which the left endpoint of an arch of one colour is never followed immediately by the right endpoint of an arch of the opposite colour — a configuration that we call a *left-right pair* (see Figure 3).

The following lemma is due to Pratt [25] who stated it in a somewhat more general context and with different terminology.

Lemma 1. *If a permutation can be produced by some operation sequence, then it can be produced by one that outputs eagerly.*

Proof. Assign the ordering $O_1 < O_2 < I_1 < I_2$ to the operation letters. If an operation sequence $v = sI_1O_2t$ (respectively sI_2O_1t) produces a permutation π, then $v' = sO_2I_1t$ (respectively sO_1I_2t) is also an operation sequence and produces π. The sequence v' is in each case lexicographically smaller than v so after a finite number of transformations of this type, an operation sequence generating π is obtained that contains none of the forbidden factors. ∎

More simply we could just say that "it can't hurt to output an element as soon as it is possible to do so", which is essentially the content of Pratt's observation.

A second source of ambiguity in operation sequences is the possibility of reflecting one or several (well chosen) arches in the horizontal line. For instance, reflecting all arches gives an equivalent arch system. The same holds if we reflect one arch joining two consecutive points of the line. Which groups of arches can one thus reflect?

We say that two arches of different colours *cross* if they cross once the one below the line is reflected. We sometimes consider the arches as vertices of a graph, two arches being adjacent if they cross (Figure 3, bottom). This graph is then bipartite. We refer to its connected components as the *(connected) components* of the arch system, and call a non-empty arch system *connected* if its corresponding graph is. In terms of operation sequences, or equivalently quarter plane loops, this means that no proper factor is an operation sequence (this may be already clear to the reader, but will be proved when enumerating connected arch systems in Section 3). Connected components were also considered by Tutte in a planar map context [30, Sec. 8].

Definition 2. *An arch system or its corresponding operation sequence is* standard *if the first arch of each component is red (that is, above the line). It is* canonical *if, in addition, it outputs eagerly.*

The following lemma is illustrated by Figure 4.

Lemma 3. *If a permutation π is achievable, then it can be produced by a canonical operation sequence.*

Proof. By Lemma 1, π can be produced by a sequence that outputs eagerly. Let us take such a sequence v, and reflect the components that do not begin with

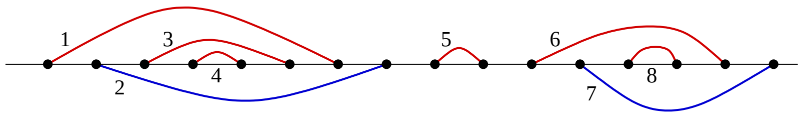

Figure 4: The canonical arch system that is equivalent to the arch system of Figure 3. Note that the left-right pair created by edges 2 and 5 in Figure 3 has disappeared (these edges do not cross any more). Also, the colours of the two rightmost components (edges 6, 7, 8) have changed. The output permutation is still 43125867.

a red arch. By definition of components, this does not create crossings between arches lying on the same side of the line, so that one obtains another operation sequence w. This sequence outputs eagerly since v does.

It remains to prove that w produces π. But this is clear, because the k^{th} arch of w moves item k in and out exactly at the same time as the k^{th} arch of v does. In particular, items are output in the same order. ∎

Let us now address the uniqueness of a canonical operation sequence for each achievable permutation.

Lemma 4. *If $v = v_1 \cdots v_{2n}$ and $w = w_1 \cdots w_{2n}$ are two equivalent operation sequences, both of which output eagerly, then they have the same type.*

Proof. Suppose that the letters v_i and w_i are of the same type, for $1 \leq i \leq j$, and let us prove that this is also true for v_{j+1} and w_{j+1}. After the j^{th} operation, v and w have performed the same number of input and output operations, and since they are equivalent, the items that are currently in the stacks according to the v sequence, are the same as those currently in the stacks according to the w sequence (though their disposition between the stacks may differ). The items that have not been moved yet are also the same for both sequences. Since v and w output eagerly, if the next item to be output is already in the stacks (for v and w) it will be output immediately by both operation sequences. If not, both must perform an input operation at this point. In either case, the types of the next operation in v and w agree. ∎

Lemma 5. *If v and w are two equivalent operation sequences having the same type, then for each i, the i^{th} operation in v moves the same item as the i^{th} operation in w.*

Proof. Let $2n$ be the length of v and w. Recall that the input of the stack is the identity permutation $12 \cdots n$, and let us denote by $\pi_1 \cdots \pi_n$ the permutation produced by v (and w). If the i^{th} operation has type I, and k inputs have taken place before, then the item moved by the i^{th} operation is $k+1$. Similarly, if the i^{th} operation has type O, and k outputs have taken place before, then the item moved by the i^{th} operation is π_{k+1}. Hence v and w move the same item at each time. ∎

We can now conclude the discussion of this section.

Proposition 6. *Every achievable permutation π is produced by a unique canonical operation sequence.*

Proof. The existence of a canonical sequence producing π is guaranteed by Lemma 3. Now suppose that two canonical operation sequences v and w produce π. By Lemma 4, they have the same type, and by Lemma 5, they move the same element at time i, for each i. These two properties mean that v and w only differ by the colouring of some arches. However, once we colour the first arch in a component, the colours of all the other arches of that component are fixed (because two arches that cross must have different colours). But v and w are standard, so that the first arch of each component is red in v and w. This implies that v and w coincide. ∎

It will be useful to define *primitive* objects. First, note that the concatenation of two arch systems (or two operation sequences) w_1 and w_2 is an arch system w. Moreover, w is canonical if and only if w_1 and w_2 are canonical. We say that a non-empty arch system (or operation sequence) is *primitive* if it cannot be written as a non-trivial concatenation. This means that the corresponding quarter plane walk only visits the origin of the lattice at the beginning and at the end. Clearly, a connected arch system is primitive. An arch system is an arbitrary sequence of primitive arch systems, and a similar statement holds for canonical arch systems. The permutations produced by a primitive canonical arch system are also said to be primitive.

Connection with results of Even & Itai [12]. In 1971, Even and Itai gave the following characterization of permutations achievable with two parallel stacks. To a permutation $\pi = \pi_1 \cdots \pi_n$, associate a graph $G(\pi)$ with vertices $1, 2, \dots, n$ and an edge from i to j (with $i < j$) if there exists $k > j$ such that kij is a subsequence of π. Then π is achievable if and only if this graph is bicolourable. Moreover, Even & Itai proved that in this case, one can produce π by putting items out as soon as possible (eager output) and otherwise putting the first available item from the input into the stack corresponding to its colour. This is related to our results as follows: if one colours $G(\pi)$ in such a way that the smallest element in each connected component is red, then the operation sequence described by Even and Itai is exactly the canonical operation sequence associated with π. Moreover, the graph associated with this operation sequence (as in Figure 3) coincides with $G(\pi)$. Since an edge of $G(\pi)$ gives rise to a pair of crossing arches in *any* operation sequence that produces π, this means that canonical operation sequences minimise the number of arch crossings.

3. Exact enumeration

In this section, we derive a system of functional equations that characterises the length generating function $S(t)$ of achievable permutations by two stacks in parallel:

$$S(t) = 1 + t + 2t^2 + 6t^3 + 23t^4 + 103t^5 + 513t^6 + 2760t^7 + 15741t^8 + O(t^9).$$

The first equation in this system characterises the generating function $\mathcal{Q}(a, u; x, y)$ of quarter plane walks, when counted by the length (variable u), the number of NW or ES corners (variable a), and the coordinates of their endpoint (variables x and y):

$$\begin{aligned}\mathcal{Q}(a, u; x, y) =& 1 + (x + y)u + (2 + 2xy + x^2 + y^2)u^2 \\ &+ \left((a + 4)(x + y) + 3x^2y + 3xy^2 + x^3 + y^3\right) u^3 + O(u^4).\end{aligned}$$

By setting $x = y = 0$, and replacing u by $\sqrt{u}$, one obtains the generating function $Q(a,u)$ of quarter plane *loops*, counted by half-length (u) and NW or ES corners (a):

$$Q(a,u) = 1 + 2u + (8+2a)u^2 + (44+24a+2a^2)u^3 + O(u^4).$$

Equivalently, $Q(a,u)$ counts arch systems by the number of arches (u) and the number of left-right pairs (a). The last series involved in our system is the generating function of standard connected arch systems, counted by the number of arches (v) and the number of left-right pairs (b):

$$C(b,v) = v + bv^2 + b(b+2)v^3 + O(v^4).$$

The reason why we have three different length variables (t, u and v) and two different corner variables (a and b) will be made clear below.

For a ring $\mathbb{K}$, we denote by $\mathbb{K}[u]$ (resp. $\mathbb{K}[[u]]$) the ring of polynomials (resp. formal power series) in u with coefficients in $\mathbb{K}$. This notation is generalised to several variables. For instance, $\mathcal{Q}(a,u;x,y) \in \mathbb{N}[a,x,y][[u]]$.

Theorem 7. *The generating function $\mathcal{Q}(a,u;x,y) \equiv \mathcal{Q}(x,y)$ of quarter plane walks is characterised by the following equation:*

$$\begin{aligned}(1 - u(x+\bar{x}+y+\bar{y}) - u^2(a-1)(x\bar{y}+y\bar{x}))\mathcal{Q}(x,y) = \\ 1 - u\bar{y}(1+ux(a-1))\mathcal{Q}(x,0) - u\bar{x}(1+uy(a-1))\mathcal{Q}(0,y), \quad (1)\end{aligned}$$

where $\bar{x} = 1/x$ and $\bar{y} = 1/y$. The generating function for quarter plane loops is

$$Q(a,u) = \mathcal{Q}(a,\sqrt{u};0,0).$$

The generating function $C(b,v)$ for connected standard arch systems is characterised by

$$Q(a,u) = 1 + 2C\left(1 - \frac{1-a}{Q},\, uQ^2\right), \quad (2)$$

where Q stands for $Q(a,u)$. Finally, the generating function $S(t) \equiv S$ of permutations that can be produced by two parallel stacks is characterised by

$$S(t) = 1 + C\left(1 - \frac{1}{S},\, tS^2\right). \quad (3)$$

Proof. The equation defining $\mathcal{Q}(x,y)$ translates a simple recursive description of quarter plane walks, according to which a walk is:

- either empty,
- or obtained by adding an E (resp. N) step at the end of another quarter plane walk,
- or obtained by adding an ES (resp. NW) corner to a walk that does not end on the x- (resp. y-) axis,
- or obtained by adding a S (resp. W) step to a walk that does not end on the x- (resp. y-) axis and whose final step is not E (resp. N).

Moreover, these four cases are disjoint. We now write the contribution to $\mathcal{Q}(x, y)$ of each case, using the following basic remarks:

- the generating function of walks ending with an E (resp. N) step is $ux\mathcal{Q}(x, y)$ (resp. $uy\mathcal{Q}(x, y)$),
- the generating function of walks ending on the x- (resp. y-) axis is $\mathcal{Q}(x, 0)$ (resp. $\mathcal{Q}(0, y)$).

These two observations allow us to express $\mathcal{Q}(x, y)$ as follows:

$$\begin{aligned} \mathcal{Q}(x,y) &= 1 + u(x+y)\mathcal{Q}(x,y) \\ &+ au^2x\bar{y}\left(\mathcal{Q}(x,y) - \mathcal{Q}(x,0)\right) + au^2\bar{x}y\left(\mathcal{Q}(x,y) - \mathcal{Q}(0,y)\right) \\ &+ u\bar{y}\left(\mathcal{Q}(x,y) - \mathcal{Q}(x,0) - ux\mathcal{Q}(x,y) + ux\mathcal{Q}(x,0)\right) \\ &\qquad + u\bar{x}\left(\mathcal{Q}(x,y) - \mathcal{Q}(0,y) - uy\mathcal{Q}(x,y) + uy\mathcal{Q}(0,y)\right). \end{aligned}$$

This gives the first equation of the proposition. It is equivalent to a recurrence relation defining the coefficient of u^n in $Q(a, u)$, and thus characterises this series.

Let us now relate the series $Q(a, u) = \mathcal{Q}(a, \sqrt{u}; 0, 0)$ and $C(b, v)$. Let w be a non-empty quarter plane loop, or equivalently an arch system. The first arch of w belongs to some connected component c, which may be standard or not. The arches of w that do not belong to c do not cross the edges of c. So the whole system w is obtained by inserting an arch system between each pair of adjacent points of c, and after the last point of c (Figure 5). If c has n arches then there are $2n$ positions to make such insertions. Ignoring the corner parameter for the moment we obtain:

$$Q(1, u) = 1 + 2C(1, uQ^2),$$

where Q stands for $Q(1, u)$. On the right-hand side, the factor 2 corresponds to the choice of colour for the first arch (since the series C only counts *standard* connected arch systems), the u enumerates the arches of c, and the Q^2 allows for the inserted arch systems. It remains to account for the number of left-right pairs. If an arch system w_1 is inserted between the endpoints of two arches of c that do not form a left-right pair in c, then the only left-right pairs it creates are those that are already present in w_1. If w_1 is non-empty and inserted in a left-right pair of c, then it destroys that left-right pair, but adds any that it might contain itself. Hence, a connected arch system c with n arches and k left-right pairs contributes $v^n b^k$ in $C(b, v)$ and gives rise, by insertion of $2n$ arch systems, to a set of arch systems counted by

$$u^n Q(a,u)^{2n-k} \left(a + (Q(a,u) - 1)\right)^k.$$

The term $Q(a, u)^{2n-k}$ corresponds to insertions in places that are not left-right pairs, while each left-right pair gives rise to a term a (insertion of an empty system w_1) and a term $Q(a, u) - 1$ (insertion of a non-empty w_1). This gives (2) by summing over all possible values of n and k and multiplying by 2 (since c is not necessarily standard).

In order to prove that this equation uniquely defines $C(b, v) = \sum_{k,n} c_{k,n} b^k v^n$, it suffices to extract from (2) the coefficient of $a^k u^n$: this gives an expression of $c_{k,n}$ in terms of the coefficients $c_{\ell,m}$ for $m < n$ and of the coefficients of Q.

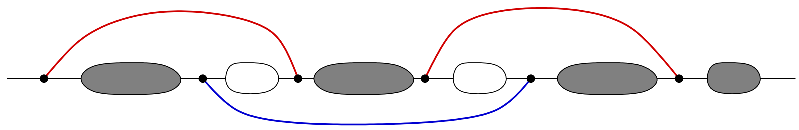

Figure 5: The structure of an arch system: a connected system c with n arches (here, $n = 3$), in which $2n$ arbitrary arch systems are inserted. Here, c has two left-right pairs. The arch systems that are inserted there (shown in white) destroy these left-right pairs, unless they are empty.

Using the same argument we can finally derive a functional equation for the generating function $S(t)$ of achievable permutations. Indeed, Proposition 6 tells us that they are in bijection with canonical arch systems, that is, with standard arch systems having no left-right pairs. Such systems w are obtained from a (standard) connected system c as before, but all inserted arch systems must be canonical; moreover, one cannot insert an empty system in a left-right pair of c. Hence a connected arch system c with n arches and k left-right pairs gives rise to a set of canonical arch systems counted by

$$t^n S(t)^{2n-k} \left(S(t) - 1\right)^k .$$

This gives (3) by summing over all possible values of n and k. This equation is equivalent to a recurrence relation defining the coefficient of t^n in $S(t)$, and thus characterises the series $S(t)$. ∎

It will be convenient to relate the functional equation (2) to a compositional inversion in the ring $\mathbb{Q}[[a, t]]$ of bivariate power series with rational coefficients.

Proposition 8. *Let $Q(a, u)$ and $C(b, v)$ be defined as above, and define the bivariate series A, U, B and V as follows:*

$$\left\{\begin{array}{rcl} A(b, v) & = & 1 + (1 + 2C(b, v))(b - 1), \\ U(b, v) & = & \dfrac{v}{(1 + 2C(b, v))^2}, \end{array}\right. \qquad \left\{\begin{array}{rcl} B(a, u) & = & 1 - \dfrac{1 - a}{Q(a, u)}. \\ V(a, u) & = & uQ(a, u)^2. \end{array}\right. \tag{4}$$

Then it follows from (2) that

$$A(B(a, u), V(a, u)) = a \quad \text{and} \quad U(B(a, u), V(a, u)) = u,$$

so that, by inversion in $\mathbb{Q}[[a, u]]$,

$$B(A(b, v), U(b, v)) = b \quad \text{and} \quad V(A(b, v), U(b, v)) = v. \tag{5}$$

The identity $B(A(b, v), U(b, v)) = b$ can be rewritten as

$$Q(A(b, v), U(b, v)) = 1 + 2C(b, v). \tag{6}$$

The proof is an elementary calculation. A consequence is that we can eliminate the series $C(b, v)$ from the system of Theorem 7, and thus obtain an equation defining $S(t)$ in terms of $Q(a, u)$. This relation looks nicer when we introduce the generating function $S^\bullet(t)$ that counts *primitive* canonical operation sequences, defined at the end of Section 2.

Corollary 9. *The series $Q(a,u)$, $S(t) \equiv S$ and $S^\bullet(t) \equiv S^\bullet$ that count quarter plane loops, achievable permutations and primitive achievable permutations respectively, are related by*

$$S = \frac{1}{1 - S^\bullet}$$

and

$$Q\left(-S^\bullet, \frac{t}{(1+S^\bullet)^2}\right) = \frac{1+S^\bullet}{1-S^\bullet}.$$

The second equation characterises $S^\bullet$ in terms of Q.

Proof. The first identity is a direct consequence of the definition of primitive achievable permutations. For the second one, specialise (6) to $b = 1 - 1/S$ and $v = tS^2$, and use (3). ∎

The equation above is the most efficient way we have found to compute the coefficients of $S^\bullet$ and S.

4. Corners in square lattice walks

Our analysis of the asymptotic behaviour of the number of achievable permutations of length n, performed in the next section, relies on three conjectures which have intrinsic combinatorial interest.

Conjecture 10. *The series $Q(a,u)$ is $(a+1)$*-positive. *That is, it can be expanded as*

$$Q(a,u) = \sum_{n \geq 0} u^n P_n(a+1),$$

where $P_n(x) \in \mathbb{N}[x]$.

Of course, it is combinatorially clear that $Q(a,u)$ is a power series in u with coefficients in $\mathbb{N}[a]$, and hence in $\mathbb{Z}[a+1]$. What is not clear is why the coefficient of $(a+1)^k$ should be non-negative. This has been checked on a computer up to half-length $n = 100$, using the functional equation (1).

Much of our analysis depends on being able to estimate the radius of convergence of various bivariate series as a function of one of the variables. In this context the name of the other variable is not important (and indeed we will most often be subsituting more or less complicated expressions for it) and so we generally suppress it, using a $\cdot$ instead, as in the conjecture below.

Conjecture 11. *For $a \geq -1$, the radius of convergence of $Q(a,\cdot)$ is*

$$\rho_Q(a) = \begin{cases} \dfrac{1}{(2+\sqrt{2+2a})^2} & \text{if } a \geq -1/2, \\[2ex] -\dfrac{a}{2(a-1)^2} & \text{if } a \in [-1,-1/2]. \end{cases} \tag{7}$$

Conjecture 12. *The series $Q_2'(a,u) := \frac{\partial Q}{\partial u}(a,u)$ is convergent at $u = \rho_Q(a)$ for $a \geq -1/3$.*

We shall only use the first part of Conjecture 11 (in fact, for $a \geq -1/3$ only). The analytic techniques of [21, 13] may open a way to its proof. In fact, Kilian Raschel was able to predict these values for the radius from a (not yet rigorous) application of these techniques. We have also checked this conjecture numerically, using the ratio test (Figure 6, left). Regarding Conjecture 12, if we assume that

$$q_n(a) := [u^n]Q(a,u) \sim \kappa(u)\rho_Q(a)^{-n} n^{\gamma(a)}$$

we would need $\gamma(a) < -2$. This is in good agreement with the estimates of $\gamma(a)$ shown in Figure 6, right. It is likely that $Q_2'(a, \rho_Q(a))$ does not converge at $a = -1/2$.

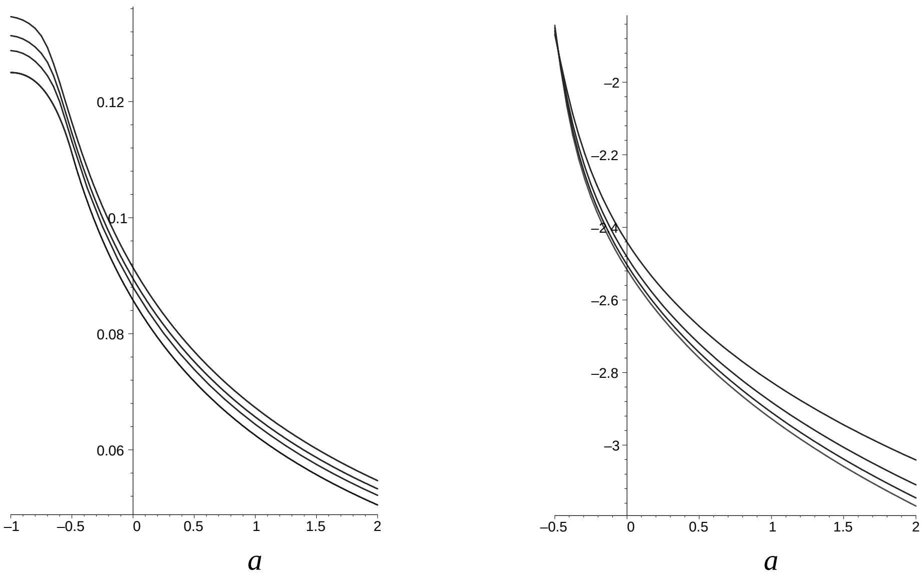

Figure 6: Left: The three top curves show the ratio $q_{n-1}(a)/q_n(a)$, where $q_n(a) = [u^n]Q(a,u)$, for $n = 40$, 60 and 100 and $a \in [-1, 2]$. The curves seem to accumulate, as n grows, on the conjectured radius (bottom curve). Right: the curves $n^2(1 - q_{n-1}(a)q_{n+1}(a)/q_n(a)^2)$, shown for $n = 40$, 60, 80 and 100, provide estimates for the exponent $\gamma(a)$.

In the following subsections, we gather more evidence for these conjectures. In particular, we prove Conjectures 10 and 11 for general loops (Section 4.3), and for loops confined to the upper half plane (Section 4.4). Note that Conjecture 12 does not hold for these more general loops. The fact that Conjecture 11 holds for general loops and half plane loops is reminiscent of a recent result according to which the growth constant of (unweighted) loops confined to a wedge is independent of this wedge [10, Sec. 1.5].

We also prove the conjectures for $a = 1$ and $a = -1$. In the latter case Conjecture 10 then simply means that $Q(-1,u)$ has non-negative coefficients; these coefficients are in fact very nice, see Proposition 15. It is interesting to refine the enumeration by taking into account the number of E steps with a new variable s. This gives rise to a generating function denoted $Q(a,s,u)$. In fact, we have the following refined conjecture.

Conjecture 13. *The series $Q(a,s,u)$ is $(a+1)$-positive, as well as the series $Q^\bullet(a,s,u)$ counting primitive quarter plane loops.*

This has been checked on a computer up to half-length $n = 40$, using the following refinement of the functional equation (1):

$$(1 - u(sx + s\bar{x} + y + \bar{y}) - u^2 s(a-1)(x\bar{y} + y\bar{x}))\mathcal{Q}(x,y) = \\ 1 - u\bar{y}(1 + usx(a-1))\mathcal{Q}(x,0) - us\bar{x}(1 + uy(a-1))\mathcal{Q}(0,y). \quad (8)$$

This equation characterises the series $\mathcal{Q}(x,y) \equiv \mathcal{Q}(a,s,u;x,y)$ that counts quarter plane walks by NW and ES corners (a), horizontal steps (s), total length (u) and coordinates of the endpoint (x,y). In particular, the above defined series $Q(a,s,u)$ is $\mathcal{Q}(a,\sqrt{s},\sqrt{u};0,0)$. Of course, $Q(a,s,u) = 1/(1-Q^\bullet(a,s,u))$, and the second part of Conjecture 13 implies the first one. We discuss in Section 6 further investigations on these conjectures.

Before we embark on our results, we want to report an observation, due (independently) to Olivier Bernardi and Julien Courtiel, which might be useful to prove the above conjectures. It tells that the pair $(\mathsf{NW},\mathsf{ES})$ can be replaced by other pairs of corners. Let us say that two words on the alphabet $\{\mathsf{N},\mathsf{S},\mathsf{E},\mathsf{W}\}$ are *shuffle-equivalent* if they have the same projections on $\{\mathsf{N},\mathsf{S}\}$, and also on $\{\mathsf{E},\mathsf{W}\}$. For instance, the words $\mathsf{NEWSSWWNES}$ and $\mathsf{ENWWSSNWES}$ are shuffle-equivalent. A *shuffle class* is an equivalence class for this relation.

Proposition 14. *There exists an involution Φ on square lattice walks that exchanges the number of NW and WN factors, fixes the number of ES and SE factors and acts inside shuffle classes.*

Consequently, in every shuffle class, the following bi-statistics of corners are equidistributed: $(\mathsf{NW},\mathsf{ES})$*,* $(\mathsf{WN},\mathsf{ES})$*,* $(\mathsf{WN},\mathsf{SE})$ *and* $(\mathsf{NW},\mathsf{SE})$*.*

Proof. To construct $\Phi(w)$, read backwards every maximal factor of w consisting of N and W steps: this transforms every NW factor into a WN factor, and vice-versa. For instance, the word $\mathsf{ENWWSSNWES}$ becomes $\mathsf{EWWNSSWNES}$. This is clearly an involution, which satisfies the announced properties.

The equidistribution of $(\mathsf{NW},\mathsf{ES})$ and $(\mathsf{WN},\mathsf{ES})$ follows. The equivalence with the other pairs follows from simple variants of the involution Φ. ∎

4.1. Some results on quarter plane loops

Proposition 15. *The series $Q(1,u)$ counting quarter plane loops by the half-length is*

$$Q(1,u) = \sum_{i,j\geq 0} \binom{2i+2j}{2i} C_i C_j u^{i+j} = \sum_{n\geq 0} C_n C_{n+1} u^n,$$

where $C_i = \binom{2i}{i}/(i+1)$ is the i^{th} Catalan number. This can be refined by taking into account the number of E steps (with a variable s):

$$Q(1,s,u) = \sum_{i,j\geq 0} \binom{2i+2j}{2i} C_i C_j s^i u^{i+j}. \quad (9)$$

The value of $Q(a,s,u)$ at $a=-1$ is just as remarkable:

$$Q(-1,s,u) = \sum_{i,j\geq 0} \binom{i+j}{i} C_i C_j s^i u^{i+j}. \quad (10)$$

In particular, the coefficients of $Q(-1, s, u)$ are non-negative, which is a very weak form of Conjecture 13.

Proof. When $a = 1$, we do not take corners into account. The results dealing with $Q(1, u)$ and $Q(1, s, u)$ are well-known and easy to prove: it suffices to observe that a quarter plane loop is obtained by shuffling two Dyck paths, one on the alphabet $\{\mathsf{N}, \mathsf{S}\}$ and the other on the alphabet $\{\mathsf{E}, \mathsf{W}\}$. Since there are C_i Dyck paths of length $2i$, this gives directly the expression of $Q(1, s, u)$, and hence the first expression of $Q(1, u)$. The second one follows using the Chu-Vandermonde summation. See also [9] for a (recursive) bijective proof, and [2] for a non-recursive one.

For the case $a = -1$, we work from the functional equation (8). The proof, inspired by recent progress of general quarter plane walks [8], is a bit long. It is given in Appendix Appendix A. ∎

Remark. The above expressions imply that $Q(a, u)$ is *D-finite* for $a = 1$ and $a = -1$. That is, it satisfies a linear differential equation (LDE) in u with polynomials in $\mathbb{Q}[u]$. We suspect that $Q(a, u)$ does not satisfy any LDE with coefficients in $\mathbb{Q}[a, u]$. Using the Maple package gfun [27], we have tried in vain to guess an LDE for $Q(0, u)$ from the first 300 coefficients.

Let us now discuss the radius of convergence of $Q(a, \cdot)$. We begin with a simple lemma, which is often used in a statistical physics context.

Lemma 16. *Let $F(a, u) = \sum_{n\geq 0} f_n(a)u^n$ be a formal power series in u with coefficients in $\mathbb{N}[a]$, such that $f_n(a)$ has degree at most n. Assume that F is not a polynomial.*

For $a \geq 0$, let $\rho(a)$ be the radius of convergence of the series $F(a, \cdot)$. Then ρ is a non-increasing function on $[0, +\infty)$, which is finite and continuous on $(0, +\infty)$.

Proof. Since F is not a polynomial, there exist infinitely many n such that $f_n(a) \neq 0$. In this case, we have, for $a > 0$:

$$f_n(a) \geq \min(1, a^n).$$

This shows that $\rho(a) \leq \max(1, a^{-1})$, and is, in particular, finite for $a > 0$.

That $\rho(a)$ is non-increasing comes from the fact that $f_n(a)$ is non-decreasing. Finally, if $0 < a \leq a'$, we have

$$f_n(a') \leq \left(\frac{a'}{a}\right)^n f_n(a)$$

(since f_n has degree at most n), which gives

$$\rho(a') \geq \frac{a}{a'}\,\rho(a),$$

and, together with $\rho(a') \leq \rho(a)$, establishes the continuity of ρ in $(0, +\infty)$. ∎

Proposition 17. *For fixed a, let $\rho_Q(a)$ be the radius of convergence of $Q(a, \cdot)$. Then*

$$\rho_Q(-1) = \frac{1}{8}, \qquad \rho_Q(1) = \frac{1}{16},$$

and ρ is a non-increasing function on $[0, +\infty)$, continuous on $(0, +\infty)$. Moreover, for $a \geq 0$,

$$\rho_Q(a) \geq \frac{1}{(2+\sqrt{2+2a})^2}. \tag{11}$$

The series $Q_2'(a, \rho_Q(a))$ converges for $a = -1$ and $a = 1$.

If $Q(a, u)$ is $(a+1)$-positive, then ρ_Q is non-increasing on $[-1, +\infty)$ and continuous on $(-1, +\infty)$.

Note that Conjecture 11 says that the bound (11) is tight.

Proof. The first two results follow from the explicit expressions of Proposition 15. At $a = 1$, we simply apply Stirling's formula to the second expression of $Q(1, u)$ to obtain the radius. More precisely, we find

$$[u^n]Q(1, u) \sim \frac{4}{\pi} 16^n n^{-3}.$$

At $a = -1$, we have to determine the asymptotic behaviour of a sum of positive terms. We use the approach described in [1, Section 3], and find

$$[u^n]Q(-1, u) \sim \frac{1}{\pi} 8^{n+1} n^{-3}.$$

These estimates imply the convergence of $Q_2'(a, \rho_Q(a))$ at $a = 1$ and $a = -1$.

Since a walk of half-length n has at most $(n-1)$ NW or ES factors, the properties of ρ_Q on $[0, +\infty)$ are a direct application of Lemma 16.

The lower bound of $\rho_Q(a)$ for $a \geq 0$ follows from the fact that $Q(a, u^2)$ is dominated by the series counting *all* square lattice walks by the length and the number of NW and ES corners. This series is easily seen to be

$$\frac{1}{1-4u-2u^2(a-1)},$$

and its radius is $1/(2+\sqrt{2+2a})$. See the proofs of Propositions 18 and 21 for details. This shows that the radius of $Q(a, u^2)$ is at least $1/(2+\sqrt{2+2a})$, which is equivalent to (11). ∎

4.2. General loops: Generating functions

We now address the enumeration of general loops according to the number of ES and NW corners. Their generating function can be obtained by two successive coefficient extractions in a rational generating function. This will allow us to prove that Conjectures 10 ($(a+1)$-positivity) and 11 (radius of convergence) hold for general loops.

Proposition 18. *The generating function $\mathcal{W}(a, s, t; x, y)$ counting square lattice walks by the number of horizontal steps (s), the number of vertical steps (t), the number of ES and NW corners (a) and the coordinates of the endpoint (x, y) is rational, and given by*

$$\mathcal{W}(a, s, t; x, y) = \frac{1}{1 - s\,(x+\bar{x}) - t(y+\bar{y}) - st\,(a-1)\,(x\bar{y}+\bar{x}y)}. \tag{12}$$

The generating function that only counts walks ending on the x-axis is algebraic, and given by

$$\begin{aligned}\mathcal{W}_{-,0}(a,s,t;x) &:= [y^0]\mathcal{W}(a,s,t;x,y)\\ &= \frac{1}{\sqrt{(1-s(x+\bar{x}))^2-4t^2(1+sx(a-1))(1+s\bar{x}(a-1))}}.\end{aligned} \tag{13}$$

The generating function that only counts loops is D-finite, and given by

$$\mathcal{W}_{0,0}(a,s,t) := [x^0y^0]\mathcal{W}(a,s,t;x,y) = \sum_{j\geq 0}\binom{2j}{j}t^{2j}\mathcal{W}_{0,0,j}(a,s) \tag{14}$$

where

$$\mathcal{W}_{0,0,j}(a,s) = [x^0]\frac{(1+sx(a-1))^j(1+s\bar{x}(a-1))^j}{(1-s(x+\bar{x}))^{2j+1}}. \tag{15}$$

The generating function

$$A(a,s,t) := \sum_{j\geq 0} t^{2j}\mathcal{W}_{0,0,j}(a,s) \tag{16}$$

is biquadratic, and can be written as

$$A(a,s,t) = \frac{1+t^2(1-a^2)T}{\left(1-t^2(a^2-1)T\right)^2-t^2(1+2(a+1)T)^2}\sqrt{\frac{1+4T-t^2(a-1)^2T}{1-t^2(a-1)^2T}}, \tag{17}$$

where $T\equiv T(a,s,t)$ is the unique series in $\mathbb{Q}[a][[s,t]]$ that satisfies $T(a,0,t)=0$ and

$$T = s^2\,\frac{1+4T-t^2(a-1)^2T}{1-t^2-t^2(a+1)^2T}. \tag{18}$$

As explained in the proof below, finding the expression of $\mathcal{W}(a,s,t;x,y)$ is simple, and then the rest of the proposition follows using two consecutive coefficient extractions. We will however give at the end of this subsection an alternative, more combinatorial proof of (12-15), which explains in particular the factor $\binom{2j}{j}$ occurring in (14).

Proof. To establish the expression of $\mathcal{W}(a,s,t;x,y)$, we use the step-by-step construction of walks that was used in Theorem 7 to establish an equation for $\mathcal{Q}(x,y)$. The argument is simplified by the fact that we have no boundary. That is, the terms $\mathcal{Q}(x,0)$ and $\mathcal{Q}(0,y)$ that occur in (1) disappear, and one obtains (12).

There are several ways to obtain the expression (13) for walks ending on the x-axis. One can for instance write a system of algebraic equations by decomposing walks at their first visit on the x-axis, in the spirit of what one usually does to count Dyck paths (see e.g. [14, Sec. V.4]). We can also expand (12) directly in y:

$$[y^0]\mathcal{W}(a,s,t;x,y) = [y^0]\frac{1}{(1-s(x+\bar{x}))\left(1-\frac{ty(1+s(a-1)\bar{x})}{1-s(x+\bar{x})}-\frac{t\bar{y}(1+s(a-1)x)}{1-s(x+\bar{x})}\right)}$$
$$= [y^0]\sum_{i,j\geq 0}\binom{i+j}{i}\frac{t^{i+j}y^{j-i}(1+s(a-1)\bar{x})^j(1+s(a-1)x)^i}{(1-s(x+\bar{x}))^{i+j+1}}$$
$$= \sum_{j\geq 0}\binom{2j}{j}\frac{t^{2j}(1+s(a-1)\bar{x})^j(1+s(a-1)x)^j}{(1-s(x+\bar{x}))^{2j+1}}, \quad (19)$$

which is equivalent to (13) since $\sum_{j\geq 0}\binom{2j}{j}v^j = (1-4v)^{-1/2}$.

Let us now count loops. Equations (14) and (15) are easily obtained by extracting the coefficient of x^0 in (19). We now want to obtain an expression for

$$A(a,s,t) := \sum_{j\geq 0} t^{2j}\mathcal{W}_{0,0,j} = [x^0]R(a,s,t;x),$$

where

$$R(a,s,t;x) = \frac{1-s(x+\bar{x})}{(1-s(x+\bar{x}))^2 - t^2(1+sx(a-1))(1+s\bar{x}(a-1))}.$$

As in Appendix Appendix A, we want to extract a coefficient in a rational fraction (the above series R, specialised to $a=-1$, is in fact related to the series (A.7) considered in the appendix). Even though this is not vital, we find it convenient to have one main length variable u, that is, to replace s by su and t by u. The denominator of $R(a,su,u;x)$ is a Laurent polynomial in x, symmetric in x and $\bar{x}$, of degree 2. It has four roots, which are Laurent series in u with coefficients in $\mathbb{Q}(\sqrt{a},s)$ (we refer to [28, Chapter 6] for generalities on solutions of polynomial equations with coefficients in $\mathbb{K}(u)$, for a field $\mathbb{K}$ of characteristic 0). Two of the roots, denoted X_1 and X_2, are actually power series in u, and they vanish at $u=0$:

$$X_{1,2} = su \pm \sqrt{a}su^2 + \frac{s}{2}\left(a+1+2s^2\right)u^3 + O\left(u^4\right).$$

The other two are $\bar{X}_1 := 1/X_1$ and $\bar{X}_2 := 1/X_2$. We now perform a partial fraction expansion of $R(a,su,u;x)$ with respect to x:

$$\begin{aligned} R(a,su,u;x) &= \frac{X_1X_2(1-su(x+\bar{x}))}{s^2u^2(1-xX_1)(1-xX_2)(1-\bar{x}X_1)(1-\bar{x}X_2)} \\ &= \frac{\alpha_1}{1-xX_1}+\frac{\alpha_2}{1-xX_2}+\frac{\alpha_1\bar{x}X_1}{1-\bar{x}X_1}+\frac{\alpha_2\bar{x}X_2}{1-\bar{x}X_2} \qquad (20)\end{aligned}$$

where

$$\alpha_1 = \frac{1-su(X_1+\bar{X}_1)}{s^2u^2(X_1-\bar{X}_1)(1-\bar{X}_1X_2)(X_1-\bar{X}_2)}$$

and symmetrically for α_2. Since X_1 and X_2 are multiples of u, we can read off from (20) the coefficient of x^0 in R:

$$\begin{aligned} A(a,su,u) &= [x^0]R(a,su,u;x) = \alpha_1+\alpha_2 \\ &= \frac{1-2su(X_1+X_2)+X_1X_2}{s^2u^2(X_1-\bar{X}_1)(X_2-\bar{X}_2)(1-X_1X_2)}.\end{aligned}$$

We finally eliminate X_1 and X_2 using the algebraic equations they satisfy (recall that they cancel the denominator of $R(a, su, u; x)$). This gives an algebraic equation for $A(a, su, u)$. This equation has two distinct factors, both of degree 2 in A^2. Only one of these factors has some roots in $\mathbb{Q}[a, s][[u]]$ (where we expect $A(a, su, u)$ to be): this factor is the minimal algebraic equation satisfied by $A(a, su, u)$. After replacing s by s/u and then u by t, we obtain the minimal algebraic equation satisfied by $A(a, s, t)$.

This equation only involves even powers of s and t. Let us now replace s^2 by its expression in terms of a, t and T derived from (18). The resulting equation factors into two terms, each of degree one in $A(a, s, t)^2$. Only one of these terms has a solution in $\mathbb{Q}[a][[s, t]]$ with constant term 1, and solving it for $A(a, s, t)$ gives (17). ∎

Another proof of (12-15) can be given by considering another pair of corners, as allowed to us by Proposition 14. The following proposition explains in particular the factor $\binom{2j}{j}$ occurring in (14).

Proposition 19. *Let v be a word on $\{\mathsf{N}, \mathsf{S}\}$. The generating function of walks whose vertical projection is v, counted by the number of horizontal steps (s), the abscissa of the endpoint (x) and the number of* NW *and* SE *factors (or any equivalent statistic from Proposition 14; variable a) only depends on $|v|_{\mathsf{N}}$ and $|v|_{\mathsf{S}}$. Its value is*

$$\mathcal{A}\mathcal{B}^{|v|_{\mathsf{N}}}\mathcal{C}^{|v|_{\mathsf{S}}}, \tag{21}$$

where

$$\mathcal{A} = \frac{1}{1 - s(x + \bar{x})}, \quad \mathcal{B} = \frac{1 + s\bar{x}(a-1)}{1 - s(x + \bar{x})} \quad \text{and} \quad \mathcal{C} = \frac{1 + sx(a-1)}{1 - s(x + \bar{x})}.$$

Proof. Write $v = v_1 \cdots v_n$. The walks we want to count read $w_0 v_1 w_1 \cdots v_n w_n$, where the w_i's are words on the alphabet $\{\mathsf{E}, \mathsf{W}\}$. Observe that NW and SE factors can only be created just after a N or S step. In particular, the contribution of w_0 is $\mathcal{A}$. After a N step v_i, a NW factor is created if and only if w_i begins with the letter W; this shows that the contribution of w_i is

$$1 + s\,\frac{a\bar{x} + x}{1 - s(x + \bar{x})},$$

which is precisely $\mathcal{B}$. Similarly, the factors w_i following a step $v_i = \mathsf{S}$ contribute the series $\mathcal{C}$. ∎

Application. One can now rederive the expressions (12) and (13) of $\mathcal{W}(a, s, t; x, y)$ and $\mathcal{W}_{-,0}(a, s, t; x, y)$ by summing (21), respectively over all walks on $\{\mathsf{N}, \mathsf{S}\}$ and over walks on $\{\mathsf{N}, \mathsf{S}\}$ ending at ordinate 0. Moreover, the series $A(a, s, t)$ given by (17) can now be understood as the generating function of loops which have vertical projection NSNSNS...

4.3. General loops: $(a+1)$*-positivity and radius of convergence*

Proposition 20. *The series $\mathcal{W}_{0,0}(a, s, t)$ that counts general loops is $(a + 1)$-positive: for $i, j \geq 0$, the coefficient of $s^i t^j$ in this series is a polynomial in $(a + 1)$ with non-negative coefficients.*

Moreover,

$$\mathcal{W}_{0,0}(1,s,t) = \sum_{i,j\geq 0} s^{2i}t^{2j}\binom{2i+2j}{2i}\binom{2i}{i}\binom{2j}{j},$$

while

$$\mathcal{W}_{0,0}(-1,s,t) = \sum_{i,j\geq 0} s^{2i}t^{2j}\binom{i+j}{i}\binom{2i}{i}\binom{2j}{j}. \tag{22}$$

When $s=t=u$,

$$\mathcal{W}_{0,0}(1,u,u) = \sum_{n\geq 0} u^{2n}\binom{2n}{n}^2.$$

Proof. By (14), the first statement means that for all j, the series $\mathcal{W}_{0,0,j}(a,s)$ is $(a+1)$-positive, or, equivalently, that the quartic series given in (17) is $(a+1)$-positive. Let us first prove that T, given by (18), is $(a+1)$-positive. This follows by observing that (18) can be written as

$$T = s^2\left(\frac{4t^2(a+1)T}{1-t^2-t^2(a+1)^2T} + \frac{4T}{1-t^2(a+1)^2\frac{T}{1-t^2}} + \frac{1}{1-\frac{t^2}{1-t^2(a+1)^2T}}\right).$$

Indeed, this equation is equivalent to a recurrence relation defining the coefficient of $s^{2i}t^{2j}$ in T by recurrence on $i+j$. This recurrence expresses this coefficient, denoted $T_{i,j}$, as a polynomial in $(a+1)$ and the $T_{k,\ell}$ for $k+\ell<i+j$ *with non-negative coefficients*, and thus proves $(a+1)$-positivity of T.

Now the rational factor in (17), one converted in partial fractions of T, reads

$$\frac{1}{2\left(1-t-2t(a+1)T-t^2(a^2-1)T\right)} + \frac{1}{2\left(1+t+2t(a+1)T-t^2(a^2-1)T\right)}.$$

As a rational series in t and T (with coefficients in $\mathbb{Q}[a]$), it is thus the even part in t of the series

$$\frac{1}{1-t-2t(a+1)T-t^2(a^2-1)T} = \frac{1}{1-t}\times\frac{1}{1-t(a+1)T\left(2+\frac{t(a+1)}{1-t}\right)},$$

which can be expanded in t, T and $(a+1)$ with non-negative coefficients. This proves the $(a+1)$-positivity of the rational factor in (17). Finally, using the equation (18) satisfied by T, the square root factor in (17) can be written

$$\frac{1}{\sqrt{1-\frac{4s^2}{1-t^2-t^2(a+1)^2T}}},$$

which is an $(a+1)$-positive series in s, t and T, and thus an $(a+1)$-positive series in s and t. This proves finally that the series $A(a,s,t)$ is $(a+1)$-positive, as well as the generating function $\mathcal{W}_{0,0}(a,s,t)$ of general loops.

The rest of the proof is now easier. The expression of $\mathcal{W}_{0,0}(1,s,t)$ follows from the fact that square lattice loops are just shuffles of one-dimensional loops. The expression of $\mathcal{W}_{0,0}(1,u,u)$ follows by the Chu-Vandermonde identity. Alternatively, it can be proved combinatorially by projecting loops on the diagonals

$x = \pm y$. The expression of $\mathcal{W}_{0,0}(-1, s, t)$ is of course more surprising, but it comes out easily from the work we have already done. By comparing the expressions (14) of $\mathcal{W}_{0,0}$ and (16) of $A(a, s, t)$, we see that what we have to prove reads

$$\begin{aligned} A(-1, s, t) &= \sum_{i,j \geq 0} s^{2i} t^{2j} \binom{i+j}{i} \binom{2i}{i} \\ &= \sum_{i \geq 0} \frac{s^{2i}}{(1-t^2)^{i+1}} \binom{2i}{i} \\ &= \frac{1}{\sqrt{(1-t^2)(1-4s^2-t^2)}}. \end{aligned}$$

This is readily proved by specializing (18) and (17) to $a = -1$. In particular, T becomes rational for this value of a. ■

Proposition 21. *Let $a \geq -1$. The series $\mathcal{W}(a, u, u; 1, 1)$ that counts square lattice walks by the length (variable u) and the number of* NW *and* ES *corners (a) has radius of convergence*

$$\frac{1}{2 + \sqrt{2+2a}}.$$

The same holds for the series $\mathcal{W}_{-,0}(a, u, u; 1)$ that counts walks ending on the x-axis.

The series $W(a, u) := \mathcal{W}_{0,0}(a, \sqrt{u}, \sqrt{u}; 1)$ that counts loops by their half-length and number of NW *and* ES *corners radius of convergence given by* (7).

Proof. It follows from Proposition 18 that

$$\mathcal{W}(a, u, u; 1, 1) = \frac{1}{1 - 4u - 2u^2(a-1)}.$$

This rational series has two poles,

$$\rho_1 = \frac{1}{2 + \sqrt{2+2a}} \quad \text{and} \quad \rho_2 = \frac{1}{2 - \sqrt{2+2a}},$$

the latter being only defined for $a \neq 1$ (recall that we already assume that $a \geq -1$). For $a \in [-1, 1)$, both poles are real and positive, with $\rho_1 \leq \rho_2$, and thus the radius is ρ_1. For $a \geq 1$, ρ_1 is the only positive singularity, and hence must be the radius by Pringsheim's theorem [14, Thm. IV.6, p. 240] (we could alternatively invoke the continuity Lemma 16).

Let us now consider walks ending on the x-axis, with generating function

$$\mathcal{W}_{-,0}(a, u, u; 1) = \frac{1}{\sqrt{(1 - 4u - 2u^2(a-1))(1 + 2u^2(a-1))}}.$$

This series has four singularities, namely ρ_1 and ρ_2 given above, as well as

$$\rho_{3,4} = \pm \frac{1}{\sqrt{2-2a}}$$

which are undefined if $a = 1$. For $a \in [-1, 1)$, all singularities are real, and ρ_1 has minimal modulus, and hence is the radius. For $a \geq 1$, the only real positive singularity is ρ_1, which must be the radius by Pringsheim's theorem.

Finally, the length generating function of loops satisfies

$$W(a, u^2) = \mathcal{W}_{0,0}(a, u, u) = [x^0 y^0]\mathcal{W}(a, u, u; x, y).$$

The Mathematica package HolonomicFunctions [20] allows one to construct a linear differential equation (DE) in u satisfied by this series, starting from a system of DEs satisfied by the rational series $\mathcal{W}(a, u, u; x, y)$ (one with respect to u, one with respect to x, one with respect to y). The DE that we obtain for $W(a, u^2)$ has order two. It translates into a DE of order 2 for $W(a, u)$, in which the coefficient of the second derivative is

$$u\,(1 + 2u(a-1))\left(a + 2u(a-1)^2\right)\left(1 - 4u(a+3) + 4(a-1)^2 u^2\right)\times \\ \left(a + u(a-1)(a-3) + 2u^2(a-1)^3\right).$$

The general theory of linear DEs [14, p. 519] tells us that the singularities of $W(a, u)$ are found among the seven roots of this polynomial, namely, with the above notation:

$$0, \quad \rho_{1,2}^2 = \frac{1}{(2 \pm \sqrt{2a+2})^2}, \quad \rho_3^2 = \frac{1}{2(1-a)}, \quad -\frac{a}{2(a-1)^2}, \\ \frac{3 - a \pm \sqrt{(9-7a)(1+a)}}{4(a-1)^2}. \tag{23}$$

It follows from the expressions of $\mathcal{W}_{0,0}(a, u, u)$ at $a = 1$ and $a = -1$ given in Proposition 20 that the radius of $W(a, u)$ is $1/16$ at $a = 1$ and $1/8$ at $a = -1$ (the proof is similar to the proof of the first part of Proposition 17). Moreover, since $W(a, u)$ is $(a+1)$-positive, the radius is a continuous function of a on $(-1, +\infty)$, non-increasing on $[-1, +\infty)$. And by Pringsheim's theorem, the radius is one of the singularities for $a \geq -1$. It then follows from an elementary study of the functions (23) that the radius of $W(a, u)$ is ρ_1^2 at $a = 1$, and then, by continuity, for $a \in [-1/2, +\infty)$ (since ρ_1^2 does not meet any other root in $(-1/2, +\infty)$). For $a \in [-1, -1/2]$, we have three candidates that would satisfy continuity at $-1/2$ (see Figure 7), but only $-\frac{a}{2(a-1)^2}$ remains below the value $1/8$, and there are no further intersection points with the other two candidates in the interval $[-1, -1/2)$. ■

4.4. Half plane walks

We obtain similar results for loops confined to the upper half plane $\{(x, y) : y \geq 0\}$.

Proposition 22. *The generating function of half plane loops, counted by horizontal steps (s), vertical steps (t), and* NW *and* ES *factors (a), is*

$$\mathcal{H}_{0,0}(a, s, t) = \sum_{j \geq 0} \frac{1}{j+1} \binom{2j}{j} t^{2j} \mathcal{W}_{0,0,j}(a, s),$$

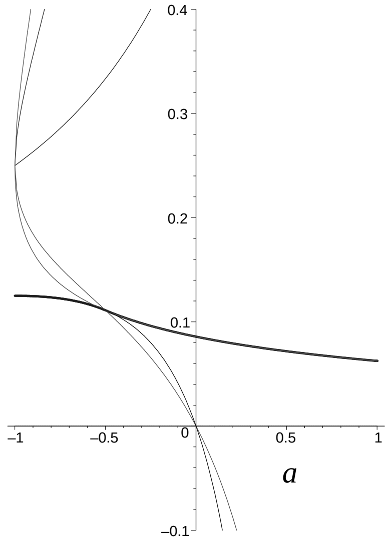

Figure 7: The seven candidates for the radius of $W(a, u)$, shown on the interval $[-1, 1]$. The thick line is the radius.

where $\mathcal{W}_{0,0,j}(a, s)$ *is given by* (15)*. This series is* $(a + 1)$*-positive. Moreover,*

$$\mathcal{H}_{0,0}(1, s, t) = \sum_{i,j\geq 0} s^{2i}t^{2j}\frac{1}{j+1}\binom{2i+2j}{2i}\binom{2i}{i}\binom{2j}{j},$$

while

$$\mathcal{H}_{0,0}(-1, s, t) = \sum_{i,j\geq 0} s^{2i}t^{2j}\frac{1}{j+1}\binom{i+j}{i}\binom{2i}{i}\binom{2j}{j}. \tag{24}$$

For $a \geq -1$*, the series* $\mathcal{H}_{0,0}(a, \sqrt{u}, \sqrt{u})$ *that counts half plane loops by half-length and corners has radius of convergence given by* (7)*.*

Proof. The expression of $\mathcal{H}_{0,0}(a, s, t)$ follows from the analogous expression (14) obtained for $\mathcal{W}_{0,0}(a, s, t)$ by applying Proposition 19. The same proposition allows us to derive the expressions of $\mathcal{H}_{0,0}(1, s, t)$ and $\mathcal{H}_{0,0}(-1, s, t)$ from their counterparts of Proposition 20.

The $(a + 1)$-positivity of $\mathcal{W}_{0,0,j}(a, s, t)$ implies the $(a + 1)$-positivity of $\mathcal{H}_{0,0}(a, s, t)$.

As far as the radius of convergence is concerned, we have for half plane loops

$$\mathcal{H}_{0,0}(a, u, u) = \sum_{j\geq 0} \frac{1}{j+1}\binom{2j}{j} u^{2j}\mathcal{W}_{0,0,j}(a, u),$$

while for general loops,

$$\mathcal{W}_{0,0}(a, u, u) = \sum_{j\geq 0} \binom{2j}{j} u^{2j}\mathcal{W}_{0,0,j}(a, u).$$

We have proved in Proposition 20 that $\mathcal{W}_{0,0,j}(a,u)$ is $(a+1)$-positive. Hence for $a \geq -1$,

$$\frac{1}{n/2+1}[u^n]\mathcal{W}(a,u,u) \leq [u^n]\mathcal{H}(a,u,u) \leq [u^n]\mathcal{W}(a,u,u),$$

and this proves that $\mathcal{H}(a,u,u)$ has the same radius of convergence as $\mathcal{W}(a,u,u)$. ■

Remark. One can also construct the (algebraic) generating functions $\mathcal{H}(a,s,t;x,y)$ (resp. $\mathcal{H}_{-,0}(a,s,t;x)$) that count walks confined to the upper half plane (resp. and ending on the x-axis). When $x=y=1$ and $s=t=u$, the radius of each of these series is found to be $\frac{1}{2+\sqrt{2+2a}}$, as in the unconfined case (Proposition 21). This confirms that the transition found at $a=-1/2$ is really a property of loops.

5. Asymptotic Analysis

5.1. Statement of the results

Recall the relationship between the series $Q(a,u)$ and $S \equiv S(t)$ established in Corollary 9:

$$Q\left(-S^\bullet, \frac{t}{(1+S^\bullet)^2}\right) = \frac{1+S^\bullet}{1-S^\bullet} \tag{25}$$

with $S = 1/(1-S^\bullet)$. Our main theorem below tells us that $S(t)$ reaches its radius of convergence when the pair $(-S^\bullet, t(1+S^\bullet)^{-2})$ reaches the *critical curve* $\{(a,\rho_Q(a)), a \geq -1\}$, where $\rho_Q(a)$ denotes the radius of the series $Q(a,\cdot)$. See Figure 8 for an illustration. However, this theorem relies on the conjectures studied in the previous section.

Theorem 23. *Assume that the series $Q(a,u)$ is $(a+1)$-positive, and that $Q_2'(a,\rho_Q(a)) < \infty$ for $-1/3 \leq a \leq 0$. Let t_c be the radius of convergence of $S = 1/(1-S^\bullet)$. Then $t/(1+S^\bullet(t))^2$ increases on the interval $[0,t_c]$, and on this interval,*

$$\frac{t}{(1+S^\bullet(t))^2} \leq \rho_Q(-S^\bullet(t)), \tag{26}$$

with equality if and only if $t = t_c$. Moreover, $S^\bullet(t_c) \leq 1/3$.

This section is devoted to the proof of this theorem. Before we begin with the proof, let us make the value of t_c more explicit thanks to the conjectured expression of $\rho_Q(a)$ (Conjecture 11).

Corollary 24. *Assume that the assumptions of the above theorem hold, as well as (the first part of) Conjecture 11. Then the radius of convergence of S is*

$$t_c = \left(1 - \frac{\sqrt{2+2a}}{2}\right)^2, \tag{27}$$

where $a = -S^\bullet(t_c)$ satisfies

$$Q\left(a, \frac{1}{(2+\sqrt{2+2a})^2}\right) = \frac{1-a}{1+a}.$$

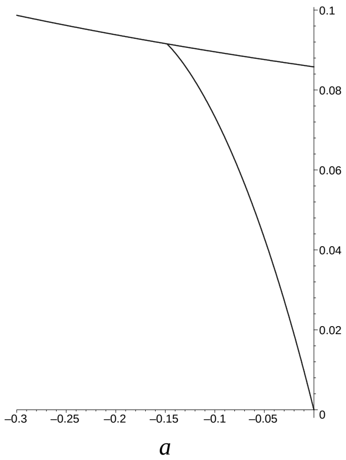

Figure 8: The top curve shows the conjectured radius of $Q(a,\cdot)$. The bottom curve shows the points $\left(-S^\bullet(t), t/(1+S^\bullet(t))^2\right)$ (estimated from the first 70 coefficients of $S^\bullet$) as t grows from 0 to t_c.

Proof. Write $a = -S^\bullet(t_c)$. By Theorem 23, we have $S^\bullet(t_c) \le 1/3 < 1/2$, and so by Conjecture 11,

$$\rho_Q(a) = \frac{1}{(2+\sqrt{2+2a})^2}.$$

Since (26) is an equality at $t = t_c$,

$$t_c = (1-a)^2 \rho_Q(a)$$

and this gives (27). The second identity of the corollary is obtained by setting $t = t_c$ in (25). ■

Using the first 100 terms of the expansion of $Q(a,u)$ in u, we estimate a between -0.15 and -0.148, which would give

$$1/t_c = \limsup s_n^{1/n} \in [8.25, 8.29].$$

This should be compared with two natural upper bounds on s_n: the number of operation sequences of length $2n$ that output eagerly (that is, have no NW nor ES corner), and the number of standard operation sequences of length $2n$. According to Conjecture 11, the growth constant for operation sequences that output eagerly would be $1/\rho_Q(0) = (2+\sqrt{2})^2 \simeq 11.6$. Now the arguments of Theorem 7 imply that the generating function $\tilde{S}(t)$ of standard operation sequences satisfies

$$\tilde{S}(t) = 1 + C(1, t\tilde{S}(t)^2) = 1 + C(t\tilde{S}(t)^2) \tag{28}$$

if we abbreviate $C(1,v)$ by $C(v)$. With the same convention,

$$Q(u) = 1 + 2C(uQ(u)^2).$$

Recall that $Q(u) \equiv Q(1,u)$ has radius $1/16$. Moreover, $Q_c := Q(1/16) = 8 - 64/(3\pi)$. One derives from this that the radius of $C(v)$ is $Q_c^2/16$, and that at this point C takes the value $(Q_c - 1)/2$. Returning to (28), this implies that $\tilde{S}$ equals $(Q_c + 1)/2$ at its radius, and that this radius is

$$\tilde{t}_c = \frac{{Q_c}^2}{4(Q_c + 1)^2}.$$

Taking the reciprocal, this gives the estimate 13.3 for the growth constant of standard operation sequences, which is larger than the growth constant obtained for sequences that output eagerly.

We can also obtain lower bounds on $1/t_c$ directly using the fact that $(s_n)_{n\geq 0}$ is a super-multiplicative sequence, so $s_n^{1/n}$ is increasing. At $n = 100$ this gives a bound $7.2 < t_c$. On the other hand we can do a bit better using $S = 1/(1-S^\bullet)$. Specifically, on the right hand side we can replace $S^\bullet$ by a polynomial truncation of its Taylor series to obtain a power series dominated term by term by S whose radius of convergence therefore is not smaller than t_c. This approximation gives $7.38 < 1/t_c$, using the truncation of $S^\bullet$ of degree 100.

5.2. Relating the singularities of S and C

We begin with a simple lemma.

Lemma 25. *The series $S^\bullet(t)$ and $S(t) = 1/(1 - S^\bullet(t))$ have the same radius of convergence t_c. Moreover, $S^\bullet(t_c) < 1$, so that $S(t_c) < \infty$.*

Proof. Let s_n (resp. $s_n^\bullet$) denote the coefficient of t^n in $S(t)$ (resp. $S^\bullet(t)$). Then $s_n^\bullet \leq s_n$, since $s_n^\bullet$ counts primitive achievable permutations (of size n), while s_n counts all achievable permutations. Recall that s_n also counts canonical operation sequences. If w is a canonical operation sequence (seen as a word on $\{\mathsf{N}, \mathsf{S}, \mathsf{E}, \mathsf{W}\}$), then $\mathsf{E}w\mathsf{W}$ is a primitive canonical operation sequence. This shows that $s_n \leq s_{n+1}^\bullet$.

It follows from these inequalities that S and $S^\bullet$ have the same radius of convergence t_c. The identity $S = 1/(1 - S^\bullet)$ then gives $S^\bullet(t_c) \leq 1$ (otherwise the radius of S would be smaller than that of $S^\bullet$). In particular, the series $S^\bullet(t)$ is convergent at $t = t_c$. The inequality $s_n \leq s_{n+1}^\bullet$ then implies that also $S(t)$ is convergent at $t = t_c$. This implies in turn that $S^\bullet(t_c) < 1$. ∎

Our next lemma exploits the connection between the series $C(b, v)$ and $S(t)$ established in Theorem 7, which can be written as:

$$S(t) = 1 + C\left(S^\bullet, tS^2\right). \tag{29}$$

Lemma 26. *Let t_c be the radius of convergence of $S = 1/(1 - S^\bullet)$, and $\rho_C(b)$ the radius of convergence of $C(b, \cdot)$. Then $S(t)$ increases on the interval $[0, t_c)$, and on this interval,*

$$tS(t)^2 < \rho_C(S^\bullet(t)).$$

Proof. That $S(t)$ increases is obvious since the series S has non-negative coefficients. We now argue *ad absurdum.* Assume that there exists $t_1 < t_c$ such that $t_1 S(t_1)^2 \geq \rho_C(S^\bullet(t_1))$. Let $t_2 \in (t_1, t_c)$. Since $S(t)$ increases strictly with t while $\rho_C(S^\bullet(t))$ decreases weakly, $t_2 S(t_2)^2 > \rho_C(S^\bullet(t_2))$. Let us write $C(b, v) = \sum_{k\geq 0, m\geq 1} c_{k,m} b^k v^m$. The identity (29) gives, for $n \geq 1$,

$$\begin{aligned} s_n &:= [t^n]S(t) = \sum_{k\geq 0, m\geq 1} c_{k,m} a_{n,k,m} \quad \text{where} \\ a_{n,k,m} &:= [t^n]\left(S^\bullet(t)^k t^m S(t)^{2m}\right) \geq 0. \end{aligned} \tag{30}$$

Let us now evaluate the series $C(b,v)$ at $b = S^\bullet(t_2)$ and $v = t_2 S(t_2)^2$. Since $t_2 S(t_2)^2 > \rho_C(S^\bullet(t_2))$ and S has non-negative coefficients, this series should be infinite. However,

$$\begin{aligned}
C(b,v) &= \sum_{k\geq 0, m\geq 1} c_{k,m} b^k v^m &&= \sum_{k\geq 0, m\geq 1} c_{k,m} S^\bullet(t_2)^k t_2^m S(t_2)^{2m} \\
& &&= \sum_{k\geq 0, m\geq 1} c_{k,m} \sum_{n\geq 1} t_2^n a_{n,k,m} \\
& &&= \sum_{n\geq 1} t_2^n \sum_{k\geq 0, m\geq 1} c_{k,m} a_{n,k,m} \\
& &&= \sum_{n\geq 1} t_2^n s_n \qquad \text{by (30)} \\
& &&= S(t_2) - 1 < \infty \qquad \text{since } t_2 < t_c.
\end{aligned}$$

In the third line, we have used the fact that all terms in the sum are non-negative, so that the value of the series is unchanged if we perform any rearrangement of terms.

We have thus obtained a contradiction, and the lemma is proved. ■

The next lemma deals with the series $C(b,v)$ and its radius $\rho_C(b)$. The proof is given in Appendix Appendix B. It is purely combinatorial and in particular, does not use the equations of Section 3.

Lemma 27. *Let $b > 0$. Then $\rho_C(b) \leq 1/4$ and for $v \in [0, \rho_C(b))$,*

$$v < C(b,v) < \frac{1}{2}.$$

The series $A(b,\cdot)$ and $U(b,\cdot)$ defined by (4) have radius of convergence at least $\rho_C(b)$.
The series $C(0,\cdot)$, $A(0,\cdot)$ and $U(0,\cdot)$ have respectively radius $+\infty$, $+\infty$ and $1/2$.

Corollary 28. *For $t \in [0, t_c]$ one has*

$$S(t) \leq \frac{3}{2} \quad \text{and} \quad S^\bullet(t) \leq \frac{1}{3}.$$

Proof. The identities are obvious if $t = 0$, so let us assume $t > 0$. Then $S^\bullet(t) > 0$, and by Lemma 26, the pair $(S^\bullet(t), tS(t)^2)$ lies in the domain of convergence of $C(b,v)$ for $t \in [0, t_c)$. Hence (29) holds in this interval, and implies that $S(t) \leq 3/2$ by Lemma 27. Since $S = 1/(1 - S^\bullet)$, this means that $S^\bullet(t) \leq 1/3$ in this interval. These inequalities hold at $t = t_c$ as well by continuity. ■

5.3. Relating the singularities of S and Q

We first establish a weak form of Theorem 23.

Lemma 29. *Assume that the series $Q(a,u)$ is $(a+1)$-positive. There exists $t_1 \in [0, t_c]$ such that*

$$\frac{t_1}{(1 + S^\bullet(t_1))^2} = \rho_Q(-S^\bullet(t_1)).$$

Moreover for any such t_1, the function $t(1 + S^\bullet(t))^{-2}$ is increasing on $[0, t_1]$.

Proof. Recall that $S^\bullet(t) < 1/3$ for $t \in [0, t_c]$ (Corollary 28), and assume that the first part of the lemma is wrong. By continuity of $S^\bullet$ and ρ_Q, this means that (26) holds strictly on $[0, t_c]$. Then for $t \in [0, t_c]$, the pair $(-S^\bullet(t), t(1 + S^\bullet(t))^{-2})$ lies in the (open) domain of convergence of Q, and by Corollary 9,

$$Q\left(-S^\bullet(t), \frac{t}{(1+S^\bullet(t))^2}\right) = \frac{1+S^\bullet(t)}{1-S^\bullet(t)}. \tag{31}$$

This holds in particular at $t = t_c$. We will now use the implicit function theorem to define an analytic continuation of $S^\bullet$ at t_c. Consider the equation

$$Q\left(-S^\circ(t), \frac{t}{(1+S^\circ(t))^2}\right) = \frac{1+S^\circ(t)}{1-S^\circ(t)}$$

as the implicit definition of a function $S^\circ(t)$. The implicit function theorem guarantees the existence of a (unique) analytic solution $S^\circ(t)$ defined in a neighbourhood of t_c and satisfying $S^\circ(t_c) = S^\bullet(t_c)$, provided

$$-Q'_1 - \frac{2t_c}{(1+S^\bullet(t_c))^3} Q'_2 \neq \frac{2}{(1-S^\bullet(t_c))^2},$$

where Q'_1 and Q'_2 denote the derivatives of Q taken at

$$\left(-S^\bullet(t_c), t_c(1+S^\bullet(t_c))^{-2}\right).$$

But the $(a+1)$-positivity of Q, together with the fact that $S^\bullet(t_c) < 1$, implies that the left-hand side is negative, while the right-hand side is positive. So the implicit function theorem applies. By (31), the function S° must coincide with $S^\bullet$ on an interval of the form $(t_c - \varepsilon, t_c)$, for some $\varepsilon > 0$. It thus constitutes an analytic continuation of $S^\bullet$ at t_c, which is impossible by Pringsheim's theorem (see [14, Thm. IV.6, p. 240]). We have thus reached a contradiction, which proves the first part of the lemma.

Now (31) holds for $t \in [0, t_1]$ by analytic continuation and continuity at t_1. The right-hand side increases with t, and thus the left-hand side must also increase. However, due to the $(a+1)$-positivity of $Q(a, u)$, it reads

$$\sum_{k,n \geq 0} q_{k,n}(1 - S^\bullet(t))^k \left(\frac{t}{(1+S^\bullet(t))}\right)^n,$$

with $q_{n,k} \geq 0$, and if $t(1 + S^\bullet(t))^{-2}$ would decrease, even locally (or weakly), so would the whole left-hand side (because $(1 - S^\bullet(t))$ decreases). Hence $t(1 + S^\bullet(t))^{-2}$ increases on $[0, t_1]$. ∎

We are now ready for the

Proof of Theorem 23. The bound (26) holds strictly at $t = 0$, but by Lemma 29, it cannot be strict on $[0, t_c]$. Let t_1 be the smallest value of $[0, t_c]$ where the equality holds. We have to prove that $t_1 = t_c$. We argue *ad absurdum*. The argument is illustrated by Figure 9. Let us denote

$$\begin{aligned} a_1 &= -S^\bullet(t_1), & u_1 &= \frac{t_1}{(1+S^\bullet(t_1))^2}, \\ b_1 &= S^\bullet(t_1), & v_1 &= \frac{t_1}{(1-S^\bullet(t_1))^2} = t_1 S(t_1)^2. \end{aligned} \tag{32}$$

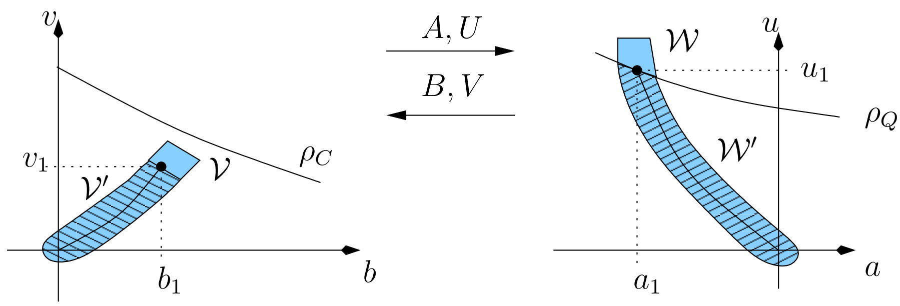

Figure 9: Illustration for the proof of Theorem 23. Left: the (b, v)-plane of the series $C(b, v)$. Right: the (a, u)-plane of the series $Q(a, u)$.

As in the proof of the previous lemma, our objective is to obtain a contradiction by constructing an analytic continuation of the map $u \mapsto Q(a_1, u)$ at $u = u_1$. However, it will take a bit of work before we can establish our starting point, namely that (37) holds on an interval $[u_1 - \varepsilon, u_1]$.

By Lemma 26, the closed curve $\overline{\mathcal{C}}_C := \{(S^\bullet(t), tS(t)^2), t \in [0, t_1]\}$ lies in the region $\mathcal{D}_C = \{(b, v) \in \mathbb{C}^2, |v| < \rho_C(|b|)\}$. We adopt the convention

$$\rho_C(0) = \lim_{b \to 0^+} \rho_C(b), \tag{33}$$

which makes $\mathcal{D}_C$ open and connected (recall that $C(0, v) = v$, so that the radius of $C(0, \cdot)$ is infinite; Lemma 27 implies that the above value of $\rho_C(0)$ is less than $1/4$. Note that C is analytic in $\mathcal{D}_C$, as well as A and U (by Lemma 27). Hence (29) holds for $t \in [0, t_1]$, and the definition of A and U in terms of C (i.e. equation (4)) thus gives

$$A(S^\bullet, tS^2) = -S^\bullet \quad \text{and} \quad U(S^\bullet, tS^2) = \frac{t}{(1 + S^\bullet)^2}. \tag{34}$$

Let $\mathcal{V}$ be an open neighbourhood of $\overline{\mathcal{C}}_C$ contained in $\mathcal{D}_C$. By the open mapping theorem in two variables [17, Thm. 6.3], the image by (A, U) of $\mathcal{V}$ is a neighbourhood $\mathcal{W}$ (in $\mathbb{C}^2$) of $\overline{\mathcal{C}}_Q := (A, U)\left(\overline{\mathcal{C}}_C\right) = \{(-S^\bullet(t), t(1 + S^\bullet(t))^{-2}), t \in [0, t_1]\}$ (see Figure 9). Let $\mathcal{D}_Q = \{(a, u) \in \mathbb{C}^2, |u| < \rho_Q(|a + 1| - 1)\}$ (with the same convention as in (33) for defining $\rho_Q(-1)$). Then by continuity of ρ_Q (Proposition 17), $\mathcal{D}_Q$ is open and connected, and Q is analytic in $\mathcal{D}_Q$. By definition of t_1, the domain $\mathcal{D}_Q$ contains $\mathcal{C}_Q := \{(-S^\bullet(t), t(1 + S^\bullet(t))^{-2}), t \in [0, t_1)\}$. Let $\mathcal{W}'$ be the (necessarily open) connected component of $\mathcal{W} \cap \mathcal{D}_Q$ containing $\mathcal{C}_Q$, and let

$$\mathcal{V}' = \{(b, v) \in \mathcal{V} : (A(b, v), U(b, v)) \in \mathcal{W}'\}.$$

Then $\mathcal{V}'$ is a connected open neighbourhood of $\mathcal{C}_C := \{(S^\bullet(t), tS(t)^2), t \in [0, t_1)\}$. By analytic continuation of (6), we have, for $(b, v) \in \mathcal{V}'$,

$$Q(A(b, v), U(b, v)) = 1 + 2C(b, v), \tag{35}$$

and in particular $Q(A(b, v), U(b, v)) \neq 0$ since $|C(b, v)| < 1/2$ in $\mathcal{D}_C$ (see Lemma 27). Hence $1/Q(A, U)$ has no pole in $\mathcal{V}'$, the series $B(A, U)$ is analytic in $\mathcal{V}'$, and by analytic continuation of (5), we have, for $(b, v) \in \mathcal{V}'$,

$$B(A(b, v), U(b, v)) = b \quad \text{and} \quad V(A(b, v), U(b, v)) = v. \tag{36}$$

Let $(a, u) \in \mathcal{W}'$. By definition of $\mathcal{W}$ and $\mathcal{V}'$, there exists $(b, v) \in \mathcal{V}'$ such that $A(b, v) = a$ and $U(b, v) = u$. The above identities show that b and v are unique, and given by

$$b = B(a, u) \quad \text{and} \quad v = V(a, u).$$

In particular, the identity (35) reads, for $(a, u) \in \mathcal{W}'$,

$$Q(a, u) = 1 + 2C\left(1 - \frac{1-a}{Q(a, u)}, uQ(a, u)^2\right).$$

Recall that $\mathcal{W}'$ is the connected component of $\mathcal{W} \cap \mathcal{D}_Q$ containing $\mathcal{C}_Q$, and that $\mathcal{W}$ contains a ball centered at the point (a_1, u_1). This implies that $\mathcal{W}'$ contains a segment $\{(a_1, u) : u \in [u_1 - \varepsilon, u_1)\}$ with $\varepsilon > 0$. Hence the identity

$$Q(a_1, u) = 1 + 2C\left(1 - \frac{1-a_1}{Q(a_1, u)}, uQ(a_1, u)^2\right) \tag{37}$$

holds in this segment, and by continuity at u_1 as well. Taking the limit $(b, v) \to (b_1, v_1)$ in (36) shows, in combination with (32) and (34), that

$$1 - \frac{1-a_1}{Q(a_1, u_1)} = b_1, \quad \text{and} \quad u_1 Q(a_1, u_1)^2 = v_1.$$

Recall that C is analytic in the neighborhood of (b_1, v_1). We can now mimic the implicit function argument used in the proof of Lemma 29. Consider the equation

$$Q^\circ(u) = 1 + 2C\left(1 - \frac{1-a_1}{Q^\circ(u)}, uQ^\circ(u)^2\right) \tag{38}$$

as the implicit definition of a function $u \mapsto Q^\circ(u)$. The implicit function theorem guarantees the existence of a (unique) analytic solution defined in a neighbourhood of u_1 and satisfying $Q^\circ(u_1) = Q(a_1, u_1)$, provided that

$$1 \neq 2\frac{1-a_1}{Q(a_1, u_1)^2} C_1' + 4u_1 Q(a_1, u_1) C_2', \tag{39}$$

where the derivatives C_1' and C_2' are taken at the point (b_1, v_1). By differentiating (37) with respect to u, we obtain, for $u \in [u_1 - \varepsilon, u_1)$,

$$Q_2'(a_1, u) = Q_2'(a_1, u)\left(2\frac{1-a_1}{Q(a_1, u)^2} C_1' + 4uQ(a_1, u) C_2'\right) + 2Q(a_1, u)^2 C_2',$$

where the derivatives are evaluated at $\left(1 - \frac{1-a_1}{Q(a_1,u)}, uQ(a_1, u)^2\right)$. Recall that by definition of t_1, the point (a_1, u_1) lies on the critical curve of Q. Since we have assumed that $Q_2'(a, u)$ is finite on this curve, There exists in a neighbourhood of u_1 a (unique) analytic function $Q^\circ(u)$ satisfying (38) and $Q^\circ(u_1) = Q(a_1, u_1)$. By (37), it coincides with $Q(a_1, u)$ on the segment $[u_1 - \varepsilon, u_1)$, and thus constitutes an analytic continuation of $u \mapsto Q(a_1, u)$ at $u_1 = \rho_Q(a_1)$. This contradicts Pringsheim's theorem, and we have thus proved that $t_1 = t_c$.

It now follows from Lemma 29 that $t(1 + S^\bullet(t))^{-2}$ is increasing on $[0, t_c]$. Finally, the bound on $S^\bullet(t_c)$ comes from Corollary 28. ∎

6. Some questions and observations

The work we have presented opens up some obvious related or more general questions.

6.1. Some questions raised directly by our work

Of course, the conjectures of Section 4 remain open. For the readers who would be interested in exploring them, we discuss some possible improvements of these conjectures in Section 6.3.

Also, we do not know anything about the nature of the series $S(t)$: is it D-finite, is it differentially algebraic? Nor do we know the nature of $Q(a, u)$ for a generic value of a, nor even for $a = 0$. The right plot of Figure 6 suggests that the exponent in the asymptotic behaviour of $q_n(a, u)$ varies continuously with a (and is not constant), which would rule out D-finiteness for a generic value of a. For comparison, for unconfined loops we predict from the differential equation that the exponent is -1 for $a \geq -1$, except at $a = -1/2$ where it is $-3/4$ (this could almost certainly be made into a rigorous proof). For loops confined to the upper half plane, we find an exponent -2 except at $a = -1/2$ where it should be $-7/4$.

Finally, it would be interesting to obtain an asymptotic estimate of s_n, not only its exponential growth constant. We have submitted the first 70 values of s_n to Tony Guttmann who predicts, using differential approximants, that $s_n \sim \kappa\, t_c^{-n} n^{-\gamma}$, for some positive constant κ, where $\gamma \simeq -2.48$ might be $-5/2$ if we expect it to be rational, and $t_c \simeq 0.12075$, so $1/t_c \simeq 8.28$.

6.2. Other rearranging devices

All of the questions we have asked about stacks in parallel can equally well be asked about other devices whose only purpose is to permute data. Specifically we could consider double ended queues (deques) and multiple stacks in parallel.

The action of a deque was considered by Knuth [19, Sec. 2.2.1]. A deque behaves very much like two stacks in parallel, treating the inputs at either end as corresponding to inputs to two stacks. The difference is that the bottoms of the stacks are effectively connected meaning that an element can be input at one end and output from the other. The arch system diagrams extend naturally to this context viewing the whole picture as a cylinder (by connecting the upper and lower edges), so we are also allowed "arches" that loop around — starting above the line and finishing below it (or vice versa). Of course the non-crossing criterion must still be satisfied. In this case there are further sources of non-uniqueness and one needs to develop a new notion of canonical sequences.

Likewise, one could consider a system of m stacks in parallel for any $m \geq 2$. Operation sequences now correspond to loops in $\mathbb{N}^m$, and the arches of our arch systems are now coloured with m colours instead of 2. Section 2 extends without any difficulty, provided we define for each connected arch system a *standard* colouring. The main difficulty comes later, when one relates loops in $\mathbb{N}^m$ to connected arch systems: one has to determine in how many ways a standard connected arch system can be re-coloured (when $m = 2$, this is the factor 2 in Eq. (2)), and this question requires further investigation.

6.3. More on $(a+1)$-positivity of loops

In our attempts for proving Conjecture 10 (the generating function of quarter plane loops is $(a+1)$-positive), we have tried to see if stronger properties hold. We believe that the following observations may be useful for the readers who would be interested in exploring this conjecture.

6.3.1. Some properties that may hold

We begin with a strong property dealing with the values found at $a = -1$. If true, it would give a new proof of (10), (22) and (24). Below, we call a *bilateral Dyck path* any one-dimensional walk starting and ending at 0.

(P_1) Let w (resp. v) be a bilateral Dyck path of half-length i (resp. j) on the alphabet $\{\mathsf{E}, \mathsf{W}\}$ (resp. $\{\mathsf{N}, \mathsf{S}\}$). Then the polynomial that counts walks of the shuffle class of vw according to the number of NW and ES corners takes the value $\binom{i+j}{i}$ at $a = -1$.

By Lemma 14, replacing the pair (NW, ES) by (NW, SE), (WN, ES) or (WN, SE) does not change the validity of the statement. This property, observed by Julien Courtiel, has been checked for $i, j \leq 5$ for Dyck paths, and for $i, j \leq 4$ for bilateral Dyck paths.

Our second property deals with $(a+1)$-positivity. We have proved in this paper that for any bilateral Dyck path v on the alphabet $\{\mathsf{N}, \mathsf{S}\}$, the generating function of loops that project vertically on v, counted by the length and the number of NW and ES corners, is $(a+1)$-positive. In fact, this series only depends on the length of v (see Propositions 19 and 20). A similar statement might be true for quarter plane loops.

(P_2) Let v be a Dyck path of half-length j on the alphabet $\{\mathsf{N}, \mathsf{S}\}$. Then the generating function of quarter plane loops that project vertically on v is $(a+1)$-positive (but does not depend on j only).

By Lemma 14, replacing the pair (NW, ES) by (NW, SE), (WN, ES) or (WN, SE) does not change the validity of the statement. This property has been checked for $j \leq 5$ and loops of half length at most 10.

6.3.2. Some properties that do not hold

Our first observation is that $(a+1)$-positivity really appears as a property of loops.

(N_1) There is no $(a+1)$-positivity property for walks ending at a prescribed endpoint (i, j), whether confined to the quarter plane, to the upper half plane or not confined at all.

Examples. For unconfined walks of length 3 ending at $(-1, 2)$, we obtain the polynomial $2a + 1$. Since these walks are confined to the upper half plane, this also provides an example in this case. Finally, for quarter plane walks of length 7 ending at $(5, 0)$, we obtain the polynomial $15a + 12$.

(N_2) There is no $(a+1)$-positivity property inside a shuffle class, even in the quarter plane.

Example. For the shuffle class of (EWEWEW,NNNSSS), we find the polynomial $62a^3 + 292a^2 + 390a + 180$, which is not $(a+1)$-positive.

However, the value at $a = -1$ is conjectured to be very simple (and positive), see Property (P_1) above.

We finally examine a natural extension of (P_2) to bilateral Dyck paths.

(N_3) There is no $(a+1)$-positivity property for loops of the half plane $\{(x, y) : x \geq 0\}$ that project on a fixed bilateral Dyck path v.

Example. For $v = \mathsf{SSNN}$, the series reads

$$u^4 + (4a^2 + 6a + 5)u^6 + O(u^8),$$

and the second coefficient is not $(a+1)$-positive.

Acknowledgements. We are indebted to Cyril Banderier, Olivier Bernardi, Alin Bostan, Julien Courtiel, Tony Guttmann, Pierre Lairez, Kilian Raschel for helpful and interesting discussions. MA thanks LaBRI for its hospitality during visits in 2008 and 2012.

Appendix A. The series $Q(-1, s, u)$

We now prove the second part of Proposition 15, dealing with the case $a = -1$. We start from the functional equation (8). As a warm up, let us give another proof of the case $a = 1$, based on that equation. Our approach is taken from [8]. When $a = 1$, Equation (8) reads:

$$K(x, y)xy\mathcal{Q}(x, y) = xy - ux\mathcal{Q}(x, 0) - usy\mathcal{Q}(0, y), \tag{A.1}$$

with $K(x, y) \equiv K(s, u; x, y) = 1 - u(sx + s\bar{x} + y + \bar{y})$. Observe that the variables x and y are decoupled in the unknown series occurring in right-hand side. Moreover, $K(x, y)$ is left unchanged by the two following involutions:

$$(x, y) \mapsto (\bar{x}, y) \quad \text{and} \quad (x, y) \mapsto (x, \bar{y}).$$

Each involution fixes one coordinate of the pair (x, y): this will play an important role in the solution. Together, these involutions generate a group of order 4, and the orbit of (x, y) is $\{(x, y), (\bar{x}, y), (\bar{x}, \bar{y}), (x, \bar{y})\}$. Let us form the alternating sum of (A.1) over this orbit. Because of the x/y-decoupling, all unknown series on the right-hand side disappear, leaving

$$\begin{aligned} K(x, y)\left(xy\mathcal{Q}(x, y) - \bar{x}y\mathcal{Q}(\bar{x}, y) + \bar{x}\bar{y}\mathcal{Q}(\bar{x}, \bar{y}) - x\bar{y}\mathcal{Q}(x, \bar{y})\right) = \\ xy - \bar{x}y + \bar{x}\bar{y} - x\bar{y} = (x - \bar{x})(y - \bar{y}). \end{aligned}$$

Equivalently,

$$xy\mathcal{Q}(x, y) - \bar{x}y\mathcal{Q}(\bar{x}, y) + \bar{x}\bar{y}\mathcal{Q}(\bar{x}, \bar{y}) - x\bar{y}\mathcal{Q}(x, \bar{y}) = \frac{(x - \bar{x})(y - \bar{y})}{1 - u(sx + s\bar{x} + y + \bar{y})}.$$

To conclude, we observe that, on the left-hand side, the series $xy\mathcal{Q}(x, y)$ consists of monomials in which the exponents of x and y are positive. In the other three series, either the exponent of x, or the exponent of y (or both) is negative. This tells us that $xy\mathcal{Q}(x, y)$ is the *positive part in x and y* of the rational series occurring on the right-hand side. In particular, extracting the coefficient of x^1y^1 in the above equation gives

$$\begin{aligned}\mathcal{Q}(0,0) \equiv \mathcal{Q}(1,s,u;0,0) &= [xy]\frac{(x-\bar{x})(y-\bar{y})}{1-u(sx+s\bar{x}+y+\bar{y})}\\ &= \sum_{n\geq 0} u^n [xy]\big((x-\bar{x})(y-\bar{y})(sx+s\bar{x}+y+\bar{y})^n\big),\end{aligned}$$

which yields

$$\mathcal{Q}(1,s,u;0,0) = \sum_{n\geq 0} u^{2n} \sum_{i=0}^{n} s^{2i}\binom{2n}{2i} C_i C_{n-i}$$

after an elementary calculation. This is equivalent to the expression (9) of $Q(1,s,u) = \mathcal{Q}(1,\sqrt{s},\sqrt{u};0,0)$.

Let us now move to the solution of (8) in the case $a=-1$. The equation reads

$$K(x,y)xy\mathcal{Q}(x,y) = xy - ux(1-2usx)\mathcal{Q}(x,0) - suy(1-2uy)\mathcal{Q}(0,y), \quad \text{(A.2)}$$

where now

$$K(x,y) = 1 - u(sx+s\bar{x}+y+\bar{y}) + 2u^2 s(x\bar{y}+y\bar{x}). \quad \text{(A.3)}$$

The involutions that leave $K(x,y)$ unchanged and fix an element of the pair (x,y) are now

$$(x,y) \mapsto \left(\bar{x}\,\frac{1-2uy}{1-2u\bar{y}}, y\right) \quad \text{and} \quad (x,y) \mapsto \left(x, \bar{y}\,\frac{1-2usx}{1-2us\bar{x}}\right).$$

However, they generate an infinite group, which prevents us from applying the above strategy. But a finite group is still hiding in this equation. Let us introduce new variables X and Y, with

$$x = 2us + X \quad \text{and} \quad y = 2u + Y.$$

The functional equation (A.2) now reads

$$\begin{aligned}&\tilde{K}(X,Y)XY\tilde{\mathcal{Q}}(X,Y) =\\ &\quad (2us+X)(2u+Y) - u(2us+X)(1-2us(2us+X))\mathcal{Q}(2us+X,0)\\ &\quad\quad - us(2u+Y)(1-2u(2u+Y))\mathcal{Q}(0,2u+Y), \quad \text{(A.4)}\end{aligned}$$

with $\tilde{\mathcal{Q}}(X,Y) = \mathcal{Q}(2us+X, 2u+Y)$ and

$$\tilde{K}(X,Y) = \frac{xyK(x,y)}{XY} = 1 - 4u^2(1+s^2) - suX - uY - su\alpha\bar{X} - u\beta\bar{Y},$$

where $\bar{X} = 1/X$, $\bar{Y} = 1/Y$, and $\alpha = (4u^2s^2-1)$ and $\beta = (4u^2-1)$ are independent of X and Y. This Laurent polynomial is now invariant by the (simpler) involutions

$$(X,Y) \mapsto (\alpha\bar{X}, Y) \quad \text{and} \quad (X,Y) \mapsto (X, \beta\bar{Y}).$$

These involutions generate again a group of order four[2], and the orbit of (X,Y) is now

$$\{(X,Y),\left(\alpha\bar X,Y\right),\left(\alpha\bar X,\beta\bar Y\right),\left(X,\beta\bar Y\right)\}.$$

We form the alternating sum of (A.4) over this orbit:

$$\begin{aligned}&\tilde K(X,Y)\times\\&\quad\Big(XY\tilde{\mathcal Q}(X,Y)-\alpha\bar X Y\tilde{\mathcal Q}(\alpha\bar X,Y)+\alpha\beta\bar X\bar Y\tilde{\mathcal Q}(\alpha\bar X,\beta\bar Y)-\beta X\bar Y\tilde{\mathcal Q}(X,\beta\bar Y)\Big)\\&\qquad\qquad=(X-\alpha\bar X)(Y-\beta\bar Y).\end{aligned}$$

Returning to the original variables x and y, this gives, after dividing by $\tilde K(X,Y)$,

$$\begin{aligned}&(x-2su)(y-2u)\mathcal Q(x,y)-\frac{\alpha\bar x(y-2u)}{1-2us\bar x}\mathcal Q\left(\frac{2su-\bar x}{1-2su\bar x},y\right)\\&+\frac{\alpha\beta\bar x\bar y}{(1-2su\bar x)(1-2u\bar y)}\mathcal Q\left(\frac{2su-\bar x}{1-2su\bar x},\frac{2u-\bar y}{1-2u\bar y}\right)-\frac{\beta\bar y(x-2su)}{1-2u\bar y}\mathcal Q\left(x,\frac{2u-\bar y}{1-2u\bar y}\right)\\&\qquad=\frac{(4su-x-\bar x)(4u-y-\bar y)}{K(x,y)},\end{aligned}$$

where $K(x,y)$ is given by (A.3). All the series occurring in this equation are power series in u with coefficients in $\mathbb Q[s,x,\bar x,y,\bar y]$. The series $(x-2su)(y-2u)\mathcal Q(x,y)$ consists of monomials in which the exponents of x and y are always non-negative. In the three other series occurring in the left-hand side, either the exponent of x, or the exponent of y (or both) is negative. This tells us that $(x-2su)(y-2u)\mathcal Q(x,y)$ is the non-negative part in x and y of the rational series occurring in the right-hand side. In particular, extracting from the above equation the coefficient of x^0y^0 gives

$$\begin{aligned}4su^2\mathcal Q(0,0)&\equiv 4su^2\mathcal Q(-1,s,u;0,0)\\&=[x^0y^0]\,\frac{(4su-x-\bar x)(4u-y-\bar y)}{1-u(sx+s\bar x+y+\bar y)+2u^2s(x\bar y+y\bar x)}.\end{aligned}$$

Let us now perform this coefficient extraction, beginning with the constant term in y:

$$\begin{aligned}4su^2\mathcal Q(0,0)&=[x^0y^0]\frac{(4su-x-\bar x)(4u-y-\bar y)}{\left(1-us(x+\bar x)\right)\left(1-\frac{uy(1-2us\bar x)}{1-us(x+\bar x)}-\frac{u\bar y(1-2usx)}{1-us(x+\bar x)}\right)}\\&=[x^0]\frac{4su-x-\bar x}{1-us(x+\bar x)}\times\\&\sum_{i,j\ge 0}\binom{i+j}{i}\frac{u^{i+j}(1-2us\bar x)^j(1-2usx)^i}{(1-us(x+\bar x))^{i+j}}[y^0]\left(4uy^{j-i}-y^{j-i+1}-y^{j-i-1}\right).\end{aligned}$$

[2] In terms of the original variables x and y, these involutions are

$$\Phi:(x,y)\mapsto\left(\frac{2su-\bar x}{1-2su\bar x},y\right)\text{ and }\Psi:(x,y)\mapsto\left(x,\frac{2u-\bar y}{1-2u\bar y}\right).$$

They do not leave $K(x,y)$ invariant, but transform it simply as follows:

$$K(\Phi(x,y))=\frac{1-4s^2u^2}{(1-2sux)(1-2su\bar x)}K(x,y)\text{ and }K(\Psi(x,y))=\frac{1-4u^2}{(1-2uy)(1-2u\bar y)}K(x,y).$$

We thus need to extract from the double sum over (i, j) the summands obtained for $i = j$, for $i = j + 1$ and for $j = i + 1$. This yields three simple sums. Upon exchanging i and j in the third one, this gives

$$\begin{aligned} 4su^2\mathcal{Q}(0,0) &= [x^0]\frac{4su - x - \bar{x}}{1 - us(x + \bar{x})}\sum_{j\geq 0}\frac{u^{2j}(1 - 2us\bar{x})^j(1 - 2usx)^j}{(1 - us(x + \bar{x}))^{2j}} \\ &\quad \times\left(4u\binom{2j}{j} - u\binom{2j+1}{j}\frac{1 - 2usx}{1 - us(x + \bar{x})} - u\binom{2j+1}{j}\frac{1 - 2us\bar{x}}{1 - us(x + \bar{x})}\right) \\ &= 2[x^0](4su - x - \bar{x})\sum_{j\geq 0}C_j u^{2j+1}\frac{(1 - 2us\bar{x})^j(1 - 2usx)^j}{(1 - us(x + \bar{x}))^{2j+1}} \end{aligned} \quad (A.5)$$

with C_j the j^{th} Catalan number. It remains to extract the constant term in x. Let us return for a while to the expression (10) of $Q(-1, s, u) = \mathcal{Q}(-1, \sqrt{s}, \sqrt{u}; 0, 0)$ that we want to establish. It is equivalent to

$$\begin{aligned} &4su^2\mathcal{Q}(-1, s, u; 0, 0) \equiv 4su^2\mathcal{Q}(0, 0) \\ &\qquad = 4\sum_{j\geq 0}C_j u^{2j+1}\sum_{i\geq 0}\binom{i+j}{i}C_i(us)^{2i+1}. \end{aligned} \quad (A.6)$$

Comparing with (A.5) shows that what remains to prove is that for $j \geq 0$,

$$[x^0](4su - x - \bar{x})\frac{(1 - 2su\bar{x})^j(1 - 2sux)^j}{(1 - su(x + \bar{x}))^{2j+1}} = 2\sum_{i\geq 0}\binom{i+j}{i}C_i(us)^{2i+1},$$

or equivalently, by taking the generating function of this collection of identities:

$$\begin{aligned} &[x^0](4su - x - \bar{x})\sum_{j\geq 0}u^{2j}\frac{(1 - 2su\bar{x})^j(1 - 2sux)^j}{(1 - su(x + \bar{x}))^{2j+1}} \\ &\qquad = 2\sum_{j\geq 0}u^{2j}\sum_{i\geq 0}\binom{i+j}{i}C_i(us)^{2i+1}. \end{aligned}$$

This is of course equivalent to prove that (A.5) and (A.6) coincide, but the absence of the factor C_j makes this new task easier. In particular, we are now handling algebraic series. Indeed, all the sums occurring in the above identities can be evaluated in closed form, and what we now need to prove is the following lemma.

Lemma 30. *Let $R(s, u; x)$ be the following rational function:*

$$R(s, u; x) = \frac{(4su - x - \bar{x})(1 - su(x + \bar{x}))}{(1 - su(x + \bar{x}))^2 - u^2\,(1 - 2su\bar{x})(1 - 2sux)}. \quad (A.7)$$

Then its constant term in x is

$$A(s, u) := [x^0]R(s, u; x) = \frac{1}{us}\left(1 - \sqrt{1 - \frac{4u^2s^2}{1 - u^2}}\right). \quad (A.8)$$

Proof. There are several ways of performing this extraction effectively. As in [28, Thm. 6.3.3], we use a partial fraction extraction in x.

The denominator of R is a Laurent polynomial in x, symmetric in x and $\bar{x}$, of degree 2. It has four roots, which are Laurent series in u with coefficients in $\mathbb{C}[s]$. Two of them are actually power series in u, and vanish at $u = 0$:

$$X_{1,2} = us \pm isu^2 + s^3u^3 \pm is(s^2 + 1/2)u^4 + s^3(2s^2 + 1)u^5 + O(u^6)$$

where $i^2 = -1$. The other two are $\bar{X}_1 := 1/X_1$ and $\bar{X}_2 := 1/X_2$. We will now perform a partial fraction expansion of R with respect to x, after writing R as

$$R(s, u; x) = \frac{X_1X_2(4su - x - \bar{x})(1 - su(x + \bar{x}))}{u^2s^2(1 - xX_1)(1 - xX_2)(1 - \bar{x}X_1)(1 - \bar{x}X_2)}. \tag{A.9}$$

In fact, we can also write the factor su in terms of X_1 and X_2, and this will simplify the result of the partial fraction expansion a bit. Indeed, since X_1 and X_2 cancel the denominator of R, we derive from (A.7) that

$$u^2 = \frac{(1 - su(X_1 + \bar{X}_1))^2}{(1 - 2su\bar{X}_1)(1 - 2suX_1)} = \frac{(1 - su(X_2 + \bar{X}_2))^2}{(1 - 2su\bar{X}_2)(1 - 2suX_2)}. \tag{A.10}$$

By solving the second equation for su, we find

$$su = \frac{X_1 + X_2}{2(1 + X_1X_2)}.$$

(There is another solution, $su = \frac{1+X_1X_2}{2(X_1+X_2)}$, but it is excluded since the X_i's are multiples of u.) Returning to (A.9), this gives

$$\begin{aligned} R(s, u; x) &= \\ &\frac{2X_1X_2\left(2(X_1 + X_2) - (x + \bar{x})(1 + X_1X_2)\right)\left(2(1 + X_1X_2) - (x + \bar{x})(X_1 + X_2)\right)}{(X_1 + X_2)^2(1 - xX_1)(1 - xX_2)(1 - \bar{x}X_1)(1 - \bar{x}X_2)} \\ &= \frac{2(1 + X_1X_2)}{X_1 + X_2} + \frac{\alpha_1}{1 - xX_1} + \frac{\alpha_2}{1 - xX_2} + \frac{\alpha_1\bar{x}X_1}{1 - \bar{x}X_1} + \frac{\alpha_2\bar{x}X_2}{1 - \bar{x}X_2} \end{aligned}$$

where

$$\alpha_1 = -\frac{2X_2(1 - X_1^2)}{(X_1 + X_2)^2}$$

and symmetrically for α_2. Since X_1 and X_2 are multiples of u, we can read off the coefficient of x^0 in $R(s, u; x)$:

$$A(s, u) = [x^0]R(s, u; x) = \frac{2(1 + X_1X_2)}{X_1 + X_2} + \alpha_1 + \alpha_2 = \frac{4X_1X_2}{X_1 + X_2}.$$

We finally eliminate X_1 and X_2 using the identities (A.10), and this gives an algebraic equation satisfied by $A(s, u)$. This equation has four distinct factors. One is quartic in A, and the other three are quadratic. Only one factor has a solution that is a power series in u with coefficients in $\mathbb{Q}[s]$, and this solution is precisely (A.8). ∎

This concludes the proof of Proposition 15.

Appendix B. Proof of Lemma 27

Recall that $C(b,v)$ is the generating function for connected arch systems beginning with a red arch, counted by the number of arches (variable v) and the number of left-right pairs (variable b). We denote by $\rho_C(b)$ the radius of convergence of $C(b,\cdot)$. For convenience, we repeat here the lemma we want to prove.

Lemma 27. *Let $b>0$. Then $\rho_C(b)\le 1/4$ and for $v\in[0,\rho_C(b))$,*

$$v < C(b,v) < \frac{1}{2}.$$

The series $A(b,\cdot)$ and $U(b,\cdot)$ defined by (4) have radius of convergence at least $\rho_C(b)$.
The series $C(0,\cdot)$, $A(0,\cdot)$ and $U(0,\cdot)$ have respectively radius $+\infty$, $+\infty$ and $1/2$.

Proof. Let us say that a quarter plane loop is *self-avoiding* if it only visits the point $(0,0)$ at the beginning and at the end, and does not visit any other point twice. It follows from the proof of (1) that, if a quarter plane loop is not connected, it admits a proper factor that is itself a loop. In particular, it is not self-avoiding. Consequently, every quarter plane self-avoiding loop is connected. It is standard if it begins with an E step.

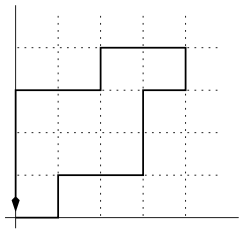

Figure B.10: A staircase polygon.

Let us use this to bound the radius $\rho_C(b)$ from above. A quarter plane self-avoiding loop is a *staircase polygon* if it consists of a sequence of E and N steps, followed by a sequence of W and S steps (Figure B.10). It is well known that the generating function of staircase polygons (according to the half-length) is [28, Exercise 6.19.l]:

$$SP(v)=\sum_{n\ge 1}\frac{1}{n+1}\binom{2n}{n}v^{n+1}=\frac{1-\sqrt{1-4v}}{2},$$

which has radius of convergence 1/4. Since non-degenerate staircase polygons have exactly one NW corner, and no ES corner, the above discussion implies that the series $C(b,v)$ dominates $b(SP(v)-v)$ term by term. Hence $\rho_C(b)$ is at most 1/4.

Let us now prove the inequalities on $C(b,v)$. The lower bound is obtained by counting only the arch system reduced to a single arch. The upper bound follows from another inequality, which is combinatorial in the sense it holds coefficient by coefficient. Namely, we will prove that the series

$$C-v-bv^2-2C(C-v) \tag{B.1}$$

has non-negative coefficients. For $v \in (0, \rho_C(b))$, dividing by $C - v$ gives

$$1 - 2C \geq \frac{bv^2}{C - v} > 0.$$

In particular $C(b, v) < 1/2$.

So let us prove that the series (B.1) has non-negative coefficients. We begin with some terminology. Two arches in an arch system are *parallel* if they are adjacent at both ends, nested, and have the same colour. For instance, the arches 3 and 4 in Figure 3 are parallel. We define the *negative* $-x$ of an arch system x to be the one obtained by interchanging colours — that is, reflecting x in a horizontal line. Let v denote the arch system consisting of a single standard (that is, red) arch (conveniently, the generating function for this arch is also v).

Let $\mathcal{C}$ denote the collection of all standard connected arch systems, counted by $C(b, v)$. We now define two injective maps:

$$\Phi : (\mathcal{C} - v) \times \mathcal{C} \to \mathcal{C} \qquad \text{and} \qquad \Psi : \mathcal{C} \times (\mathcal{C} - v) \to \mathcal{C}$$

whose images are disjoint, and which do not change the total number of arches nor the total number of left-right pairs.

Construction of Φ. Take $x \in \mathcal{C} - v$ and $y \in \mathcal{C}$. If the *last* arch of x (that is, the one with the rightmost right end) is red, let $x' = -x$, otherwise let $x' = x$. Now form $\Phi(x, y)$ as follows (Figure B.11). Place x' to the left of y. Unhook the right end of the last arch of x' and pass it beneath y before reconnecting with the line. Unhook the left end of the first arch of y and pass it above x' before reconnecting with the line.

Let us prove that the resulting arch system $\Phi(x, y)$ is connected. Its graph is obtained by juxtaposing the graphs of x' and y (which are both connected) and adding an edge between them (corresponding to the crossing between the first and last arches of $\Phi(x, y)$). This graph is connected, and so is $\Phi(x, y)$. This graph can also be used to prove that Φ is injective: if we delete the edge connecting the first and last arch, we obtain two connected components. The left one is x' (from which we can find x) and the right one is y.

Note that since $x \neq v$, the first two arches of $\Phi(x, y)$ do not cross. One also checks that the number of left-right pairs behaves additively (this also uses the fact that $x \neq v$).

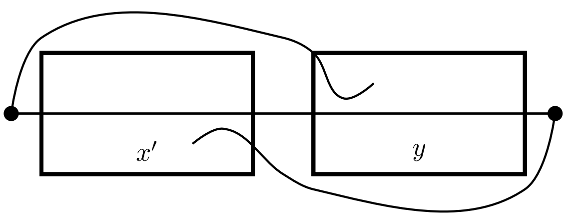

Figure B.11: The construction of $\Phi(x, y)$.

Construction of Ψ. Let $(x, y) \in \mathcal{C} \times (\mathcal{C} - v)$. Place $-x$ on the line, and place a copy of y between its first two points. Now unhook the left end the first arch of y and pass it above the first point of $-x$ before reconnecting with the line (Figure B.12).

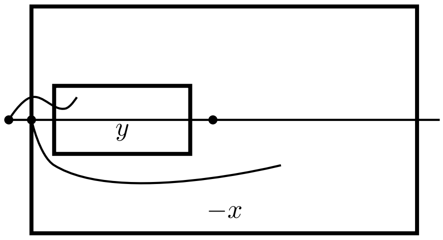

Figure B.12: The construction of $\Psi(x, y)$.

As above, the graph of $\Psi(x, y)$ is obtained by adding an edge joining a vertex of the graph of x to a vertex of the graph of y. This graph is connected, and so is $\Psi(x, y)$. If we delete from the graph of $\Psi(x, y)$ the edge between the first two arches we obtain two connected components from which we can recover x and y. Hence Ψ is injective.

The number of left-right pairs still behaves additively (this uses $y \neq v$). Since the first two arches of $\Psi(x, y)$ *do* cross, the range of Ψ is disjoint from that of Φ. In particular, the union of their images is enumerated by $2C(C - v)$. This series counts a subset of $\mathcal{C}$ which consists of arch systems with at least three arches. We thus conclude that $C - v - bv^2 - 2C(C - v)$ has non-negative coefficients, as claimed.

To finish, it is clear that $A(b, v) = 1 + (b - 1)(1 + 2C(b, v))$ has at least radius ρ_C. Moreover, since $C(b, v)$ has non-negative coefficients, then the first part of the lemma implies that $|C(b, v)| < 1/2$ for $b > 0$ and $|v| < \rho_C(b)$. The function

$$U(b, v) = \frac{v}{(1 + 2C(b, v))^2}$$

is thus analytic in this disk, and thus has radius of convergence at least $\rho_C(b)$. The results stated for $b = 0$ are obvious, since $C(0, v) = v$. ∎

References

References

[1] Bender, E. A. Asymptotic methods in enumeration. *SIAM Rev. 16* (1974), 485–515.

[2] Bernardi, O. Bijective counting of tree-rooted maps and shuffles of parenthesis systems. *Electron. J. Combin. 14*, 1 (2007), Research Paper 9, 36 pp. (electronic).

[3] Bóna, M. A survey of stack-sorting disciplines. *Electron. J. Combin. 9*, 2 (2002/03), Article 1.

[4] Bóna, M. *Combinatorics of Permutations, Second Edition.* Discrete Mathematics and Its Applications. Taylor & Francis, 2012.

[5] Bostan, A., and Kauers, M. The complete generating function for Gessel walks is algebraic. *Proc. Amer. Math. Soc. 138*, 9 (2010), 3063–3078. With an appendix by Mark van Hoeij.

[6] Bostan, A., Raschel, K., and Salvy, B. Non-D-finite excursions in the quarter plane. *J. Combin. Theory Ser. A 121* (2014), 45–63.

[7] Bousquet-Mélou, M. Walks in the quarter plane: Kreweras' algebraic model. *Ann. Appl. Probab. 15*, 2 (2005), 1451–1491.

[8] Bousquet-Mélou, M., and Mishna, M. Walks with small steps in the quarter plane. *Contemp. Math. 520* (2010), 1–40.

[9] Cori, R., Dulucq, S., and Viennot, G. Shuffle of parenthesis systems and Baxter permutations. *J. Combin. Theory Ser. A 43*, 1 (1986), 1–22.

[10] Denisov, D., and Wachtel, V. Random walks in cones. *Ann. Probab.* (to appear). arXiv:1110.1254.

[11] Denton, D. Methods of computing deque sortable permutations given complete and incomplete information. arXiv:1208.1532, 2012.

[12] Even, S., and Itai, A. Queues, stacks, and graphs. In *Theory of machines and computations (Proc. Internat. Sympos., Technion, Haifa, 1971).* Academic Press, New York, 1971, pp. 71–86.

[13] Fayolle, G., and Raschel, K. Some exact asymptotics in the counting of walks in the quarter plane. In *Analysis of Algorithms 2012*, DMTCS Proceedings, pp. 109–124.

[14] Flajolet, P., and Sedgewick, R. *Analytic combinatorics.* Cambridge University Press, Cambridge, 2009.

[15] Guy, R. K., Krattenthaler, C., and Sagan, B. E. Lattice paths, reflections, & dimension-changing bijections. *Ars Combin. 34* (1992), 3–15.

[16] Kauers, M., Koutschan, C., and Zeilberger, D. Proof of Ira Gessel's lattice path conjecture. *Proc. Natl. Acad. Sci. USA 106*, 28 (2009), 11502–11505.

[17] Kaup, L., and Kaup, B. *Holomorphic functions of several variables*, vol. 3 of *de Gruyter Studies in Mathematics.* Walter de Gruyter & Co., Berlin, 1983.

[18] Kitaev, S. *Patterns in Permutations and Words.* Monographs in theoretical computer science. Springer, 2011.

[19] Knuth, D. E. *The art of computer programming. Vol. 1: Fundamental algorithms.* Addison-Wesley Series in Computer Science and Information Processing. London, 1968.

[20] Koutschan, C. A fast approach to creative telescoping. *Math. Comput. Sci. 4*, 2-3 (2010), 259–266.

[21] Kurkova, I., and Raschel, K. On the functions counting walks with small steps in the quarter plane. *Publ. Math. Inst. Hautes Études Sci. 116* (2012), 69–114.

[22] Linton, S., Ruškuc, N., and Vatter, V., Eds. *Permutation patterns*, vol. 376 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2010. Papers from the 5th Conference held at the University of St. Andrews, St. Andrews, June 11–15, 2007.

[23] Murphy, M. M. *Restricted permutations, antichains, atomic classes and stack sorting*. PhD thesis, University of St. Andrews, 2002.

[24] Pierrot, A., and Rossin, D. 2-Stack sorting is polynomial. arXiv:1304.2860, 2013.

[25] Pratt, V. R. Computing permutations with double-ended queues, parallel stacks and parallel queues. In *STOC '73: Proceedings of the fifth annual ACM symposium on Theory of computing* (New York, NY, USA, 1973), ACM Press, pp. 268–277.

[26] Rosenstiehl, P., and Tarjan, R. E. Gauss codes, planar Hamiltonian graphs, and stack-sortable permutations. *J. Algorithms 5*, 3 (1984), 375–390.

[27] Salvy, B., and Zimmermann, P. Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Transactions on Mathematical Software 20*, 2 (1994), 163–177.

[28] Stanley, R. P. *Enumerative combinatorics. Vol. 2*, vol. 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[29] Tarjan, R. Sorting using networks of queues and stacks. *J. Assoc. Comput. Mach. 19* (1972), 341–346.

[30] Tutte, W. T. A census of Hamiltonian polygons. *Canad. J. Math. 14* (1962), 402–417.