A bijection proving the Aztec diamond theorem by combing lattice paths

Frédéric Bosio   Marc A. A. van Leeuwen

Laboratoire de Mathématiques et Applications
Université de Poitiers
Poitiers, France

{Frederic.Bosio,Marc.van-Leeuwen}@math.univ-poitiers.fr

Abstract

We give a bijective proof of the Aztec diamond theorem, stating that there are $2^{n(n+1)/2}$ domino tilings of the Aztec diamond of order n. The proof in fact establishes a similar result for non-intersecting families of $n + 1$ Schröder paths, with horizontal, diagonal or vertical steps, linking the grid points of two adjacent sides of an $n \times n$ square grid; these families are well known to be in bijection with tilings of the Aztec diamond. Our bijection is produced by an invertible "combing" algorithm, operating on families of paths without non-intersection condition, but instead with the requirement that any vertical steps come at the end of a path, and which are clearly $2^{n(n+1)/2}$ in number; it transforms them into non-intersecting families.

1 Introduction

The term "Aztec diamond" introduced by Elkies, Kuperberg, Larsen and Propp [EKLP92] refers to a diamond-shaped set of squares in the plane, obtained by taking a triangular array of squares aligned against two perpendicular axes, and completing it with its mirror images in those two axes; the *order* of the diamond is the number of squares along each of the sides of the triangular array. Their main result concerns counting the number of domino tilings (i.e., partitions into subsets of two adjacent squares) of the Aztec diamond.

Theorem 1 (Aztec diamond theorem). *There are exactly $2^{\binom{n+1}{2}}$ domino tilings of the Aztec diamond of order n.*

This result has been proved in various manners; the original article alone gives four different proofs, all closely related to a correspondence that it establishes between the domino tilings and certain pairs of alternating sign matrices. A first proof that involves

associating a collection of disjoint paths to a domino tiling was given by Brualdi and Kirkland [BrKi05]; they use such an association (where the paths are in fact closed oriented cycles) to translate the enumeration problem into one of computing the determinant of a $\{-1, 0, 1\}$-matrix of order $n^2 + n$, which they reduce to the determinant of a Hankel matrix of Schröder numbers of order n, and then finally evaluate using some more algebraic manipulations. Their method is simplified by Eu and Fu [EuFu05] who, using a somewhat different correspondence illustrated in Figure 1, bring domino tilings of an order n Aztec diamond into bijection with non-intersecting families of n lattice paths linking two adjacent sides of an $n \times n$ square grid (the common corner is unused), the paths having horizontal, diagonal or vertical steps. With aid of the Lindström-Gessel-Viennot method, they express the number of such families as the determinant of a Hankel matrix of Schröder numbers (somewhat different from the one in [BrKi05]), and evaluate it to the proper power of 2 through a clever interplay between algebraic and combinatorial viewpoints.

In this paper we propose another proof of the Aztec diamond theorem in terms of non-intersecting families of lattice paths. We start by a further simplification of the argument of [EuFu05], giving the enumeration of disjoint path families as a determinant of a matrix of Delannoy rather than Schröder numbers (the Delannoy paths they count are less restricted than Schröder paths are), which can be evaluated recursively using just row and column operations. Our main result however is a *bijective* proof of this enumeration, by giving a reversible procedure that constructs such non-intersecting families from an array of $\binom{n+1}{2}$ independent values taken from $\{0, 1\}$ (that we shall call bits). Indeed we use these values to first construct a family of $n + 1$ (possibly intersecting) paths P_i with $0 \leqslant i \leqslant n$, where there are 2^i possibilities for P_i; then we modify the family by a succession of operations that may interchange steps among its paths, so as to ensure they all become disjoint. These modifications are invertible step-by-step; to make this precise we specify at each intermediate point of the transformation precise conditions on the family that ensure that continuing both in the forward direction and in the backward direction can be completed successfully. As a consequence we obtain the descriptions of a number of collections of intermediate families of paths, which are all equinumerous to the set of disjoint path families.

Ours is not strictly speaking the first bijective proof of the Aztec diamond theorem. Indeed the fourth proof of the original paper, though not formulated as a bijective proof, does give a "domino-shuffling" procedure (which is more explicitly described in [JPS98, section 2]), with the aid of which one can build domino tilings of Aztec diamonds of increasing order, in a manner that uses a net influx of $n(n+1)/2$ bits of external information (each passage from a tiling of order $i - 1$ to order i uses i bits), and such that all these bits can be recovered from the final tiling produced. However, in spite of some superficial similarities, the procedure we present is quite different in nature. The main differences are that our procedure operates not on tilings but on families of (possibly intersecting) paths; that it proceeds in a regular forward progression rather than alternating deconstruction, shuffling, and construction steps; and that this progression involves parts of the final configuration, successively attaining their final state, rather than a passage through complete
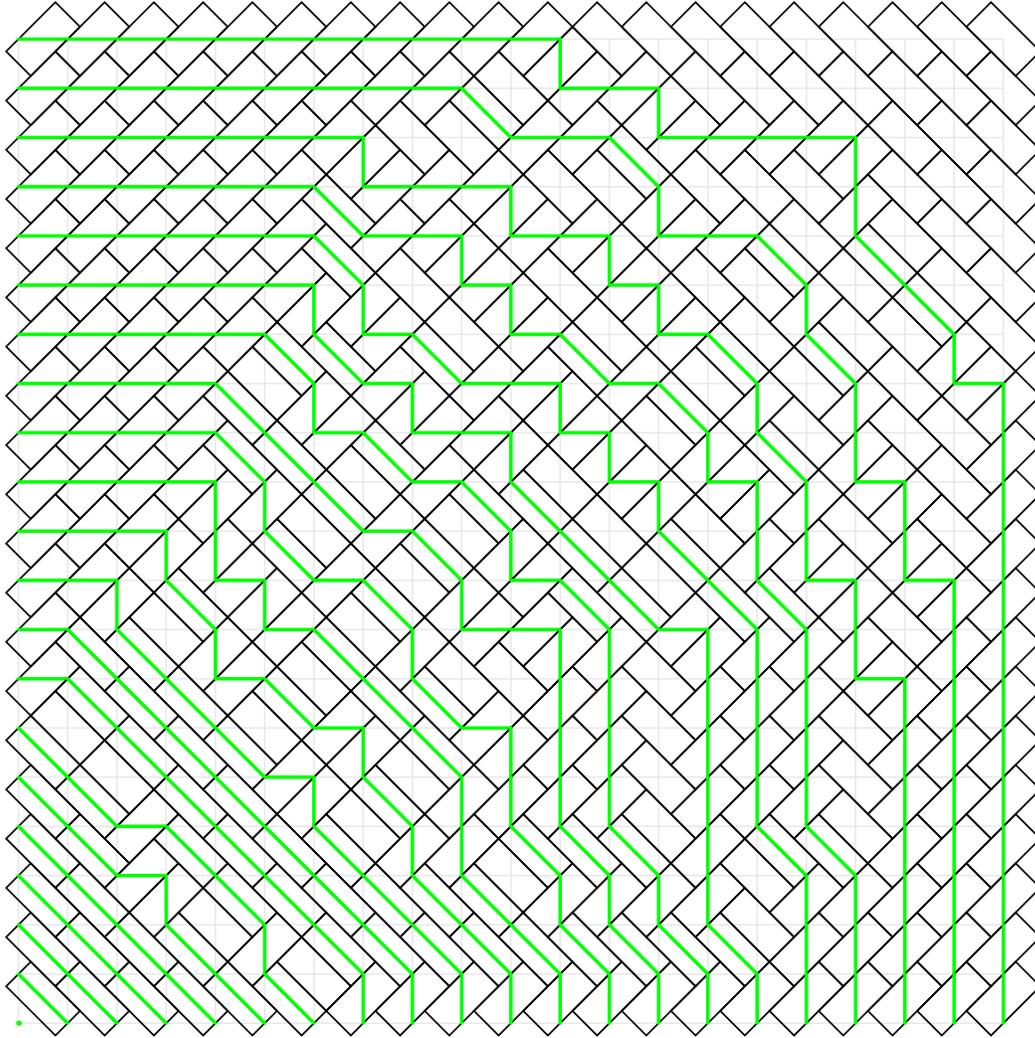
Figure 1: A domino tiling of the Aztec diamond of order 20, and (in green) the corresponding family of 21 disjoint paths

tilings of increasing order. We shall say a bit more about domino shuffling towards the end of our paper. Like domino shuffling, our algorithm provides an efficient method to produce large "random" examples of disjoint families of lattice paths as in Figure 2 (or of domino tilings), which illustrate the "arctic circle" phenomenon of [JPS98].

The domino tiling point of view in fact plays no role at all in our construction; indeed we discovered the known connection with tilings of Aztec diamonds only after the first author had found the bijective proof, formulated as a proof of an enumeration formula for families of lattice paths. In this paper we shall more or less follow the route by which we approached the problem, leaving the connection with Aztec diamonds aside until the final section. Henceforth n will be the number of paths in a family, which is *one more*
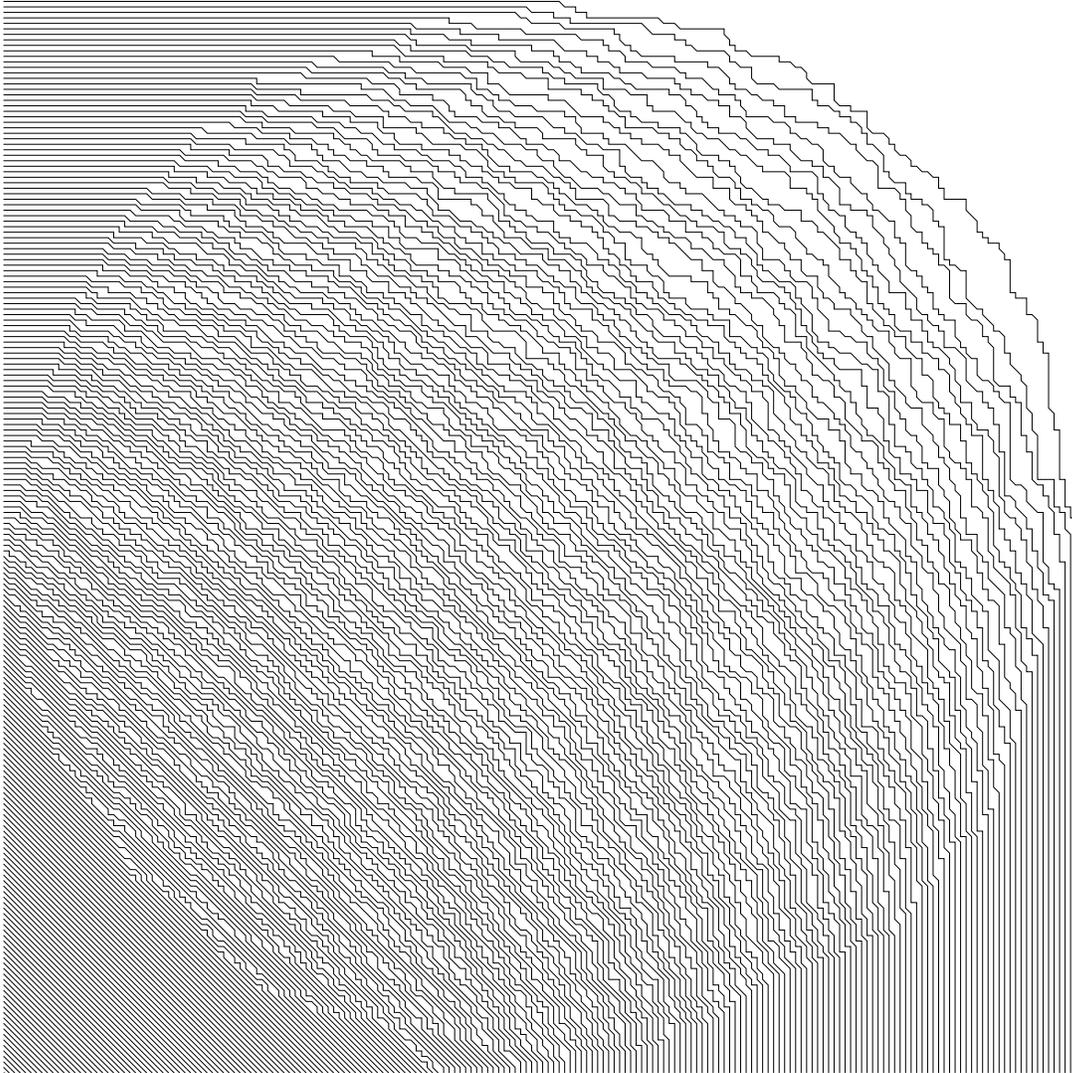
Figure 2: A random disjoint family of 196 paths

than the order of the corresponding Aztec diamond.

We give some definitions in section 2, and in section 3 enumerate disjoint families of lattice paths using a determinant evaluation. In section 4 we give some illustrations and considerations leading to an informal approach to our algorithm, followed by a more formal statement and proof in section 5. Finally we detail in section 6 the bijection between disjoint families of Delannoy paths and domino tilings (even though fairly clear in a pictorial example, it is worth while to state and prove it carefully), which in fact is not specific to the Aztec diamond at all; we also discuss some complementary matters.

2 Definitions

We shall consider paths through points in a square lattice whose basic steps are either by a unit vector in one of two directions along the axes, or a diagonal step by the sum of those two vectors. We shall call these Delannoy paths. Concretely, since we shall want our paths to connect points on the two borders of the positive quadrant $\mathbf{N} \times \mathbf{N} \subset \mathbf{Z} \times \mathbf{Z}$, we take our basic steps to be by one of the vectors $(0, +1)$, $(-1, +1)$ and $(-1, 0)$, and the step will then respectively be called horizontal, diagonal or vertical. This terminology implies that we think of the first index (or coordinate) as varying vertically, and the second index as varying horizontally, as in matrices. We shall frequently refer to a set of vertically aligned points as a "column"; in column k the constant second index is equal to k. However for visualisation it will be slightly more convenient to have the first index increase upwards rather than (as in matrices) downwards, so this is what we shall do. This amounts to using the convention of Cartesian coordinates, but with the order of these coordinates interchanged.

Definition 2.1. A Delannoy path from p to q, for points $p, q \in \mathbf{Z} \times \mathbf{Z}$, is a sequence $P = (p_0, p_1, \ldots, p_k)$ with $k \in \mathbf{N}$, $p_i \in \mathbf{Z} \times \mathbf{Z}$ for $0 \leqslant i \leqslant k$, $p_0 = p$, $p_k = q$, and $p_{i+1} - p_i \in \{(0, +1), (-1, +1), (-1, 0)\}$ for $0 \leqslant i < k$. The *support* of P is $\mathrm{supp}(P) = \{p_0, p_1, \ldots, p_k\}$.

We denote by $a_{i,j}$ be the number of Delannoy paths from $(i, 0)$ to $(0, j)$ (a number also known as the Delannoy number $D(i, j)$). Then

$$a_{i,0} = 1 = a_{0,j} \quad \text{and} \quad a_{i+1,j+1} = a_{i,j+1} + a_{i+1,j} + a_{i,j} \quad \text{for all } i, j \in \mathbf{N}. \tag{1}$$

Definition 2.2. The infinite matrix of these numbers is $A = (a_{i,j})_{i,j \in \mathbf{N}}$; its upper-left $n \times n$ sub-matrix is $A_{[n]} = (a_{i,j})_{0 \leqslant i,j < n}$, for any $n \in \mathbf{N}$.

Applying the Lindström-Gessel-Viennot method to the determinant of $A_{[n]}$ leads to the following class of families consisting of n Delannoy paths each.

Definition 2.3. If $\pi \in \mathcal{S}_n$ is a permutation of $[n] = \{0, 1, \ldots, n-1\}$, then we shall call "$\pi$-family" any n-tuple $(P_0, P_1, \ldots, P_{n-1})$ where P_i is a Delannoy path from $(i, 0)$ to $(0, \pi_i)$ for $i \in [n]$. In the special case where π is the identity permutation of $[n]$, a π-family will be simply called an "n-family". A π-family is called *disjoint* if $\mathrm{supp}(P_0)$, $\mathrm{supp}(P_1)$, ... and $\mathrm{supp}(P_{n-1})$ are all disjoint.

A π-family cannot be disjoint unless π is the identity permutation. We shall use general π-families only in the initial interpretation of $\det(A_{[n]})$: after reducing its evaluation to counting disjoint families, we shall only deal with n-families.

Definition 2.4. A Schröder n-family is an n-family $(P_0, \ldots, P_{n-1})$ with the property that for each i the path P_i does not pass to the side of the origin of the anti-diagonal line joining its initial and final points: in formula, for each point $(k, l) \in \mathrm{supp}(P_i)$ one has $k + l \geqslant i$.

Paths in a Schröder n-family are (similar to) actual Schröder paths. A simple induction argument shows that any disjoint n-family is a Schröder n-family. In formulating our bijective proof for the enumeration of disjoint n-families, we shall employ only Schröder n-families, but which are not necessarily disjoint. A type of Schröder paths of particular interest is the following.

Definition 2.5. A Delannoy path $(p_0, p_1, \ldots, p_k)$ from $(i, 0)$ to $(0, i)$ is called cliff-shaped if $p_i = (k - i, i)$, in other words if the first i steps of the path are either horizontal or diagonal, and any remaining steps are vertical. A cliff-shaped (Schröder) n-family is an n-family whose paths are cliff-shaped paths.

Clearly any cliff-shaped Delannoy path is a Schröder path; therefore the qualification "Schröder" when used after "cliff-shaped" is redundant. For a cliff-shaped path from $(i, 0)$ to $(0, i)$, the first i steps can be chosen independently to be horizontal or diagonal, after which the rest of the path is determined; therefore there are 2^i such paths, and in all $2^{\binom{n}{2}}$ cliff-shaped n-families.

3   Enumeration of disjoint Schröder n-families

If we denote the set of π-families by $F(\pi)$ then we have $\#F(\pi) = \prod_{i \in [n]} a_{i, \pi_i}$ by definition of the numbers $a_{i,j}$, and we can therefore evaluate

$$\det(A_{[n]}) = \sum_{\pi \in \mathcal{S}_n} \operatorname{sg}(\pi) \prod_{i \in [n]} a_{i, \pi_i} = \sum_{\pi \in \mathcal{S}_n} \operatorname{sg}(\pi) \# F(\pi). \tag{2}$$

Now the Lindström-Gessel-Viennot method says we can replace the latter summation by its contribution from disjoint families only, since all other contributions cancel out. Indeed if a π-family $(P_0, \ldots, P_{n-1})$ has any pair of distinct paths P_i, P_j whose supports have non-empty intersection, one can modify P_i and P_j by interchanging their parts beyond (in the obvious sense) some point of that intersection to obtain a π'-family, with $\pi' = \pi \circ (i\ j)$ and hence $\operatorname{sg}(\pi') = -\operatorname{sg}(\pi)$, which therefore gives an opposite contribution to the summation. It remains to make this cancellation systematic, which can be done by fixing a rule that chooses for every non-disjoint family a pair $\{i, j\}$ and a point of intersection of the supports of P_i and P_j, in such a way that the same choices will be produced for the family obtained after modifying P_i and P_j by the ensuing interchange; this will ensure one obtains a sign-reversing *involution* of the set of non-disjoint families. This rule can be chosen in a multitude of ways (although it is not *entirely* trivial to do so, since the modification may change the set of candidate pairs $\{i, j\}$ of indices), and leave it to the reader to choose one.

Since a π-family can only be disjoint if π is the identity permutation, we find that $\det(A_{[n]})$ is equal to the number of disjoint Schröder n-families. On the other hand this determinant can be easily evaluated recursively using algebraic manipulations. If $n > 0$ and $E_{[n]} = (\delta_{i,j} - \delta_{i+1,j})_{i,j \in [n]}$ is the upper unitriangular $n \times n$ matrix with entries -1 directly above the diagonal and zeroes elsewhere above the diagonal, then the product

$A'_{[n]} = E^\top_{[n]} A_{[n]} E_{[n]} = (a'_{i,j})_{i,j \in [n]}$ has entries $a'_{i,j}$ that are $\delta_{i,j}$ if $i = 0$ or $j = 0$, and are otherwise given by

$$a'_{i,j} = a_{i,j} - a_{i,j-1} - a_{i-1,j} + a_{i-1,j-1} = 2a_{i-1,j-1} \qquad \text{if } i,j > 0, \tag{3}$$

where the latter equality is a consequence of the recursion relation (1). This means that $A'_{[n]}$ can be written in $(1, n-1) \times (1, n-1)$ block matrix form

$$E^\top_{[n]} A_{[n]} E_{[n]} = \begin{pmatrix} 1 & 0 \\ 0 & 2A_{[n-1]} \end{pmatrix} \qquad \text{if } n > 0, \tag{4}$$

from which, since $\det(E_{[n]}) = 1$, it follows that $\det(A_{[n]}) = 2^{n-1} \det(A_{[n-1]})$ when $n > 0$, so

$$\det(A_{[n]}) = 2^{\binom{n}{2}} \qquad \text{for all } n \in \mathbf{N}. \tag{5}$$

This proves

Theorem 2. *For $n \in \mathbf{N}$, the number of disjoint Schröder n-families is $2^{\binom{n}{2}}$.* $\qquad\square$

4   Some illustrations, and informal approach to a bijection

In this section we give some illustrations of the problem at hand, and some considerations and examples that might help appreciate the bijective proof of theorem 2 that we shall give. Impatient readers may skip to the next section where this proof is given, and which is independent of the current one. There the bijection will be formalised in the form of pseudo-code; a computer program that implements this algorithm, and which was used to prepare the illustrations in this paper, is available from the website [vLee12] of the second author.

We shall start by listing all $2^{\binom{4}{2}} = 64$ disjoint Schröder 4-families, to give an impression of the variety these present. They are displayed in Figure 3, ordered by increasing number of non-diagonal steps from bottom(-left) to top(-right).

A first fact that is apparent in this figure is that the number of horizontal steps (which always equals the number of vertical steps), or by complementation the number of diagonal steps, follows a (symmetric) binomial distribution for $m = 6 = \binom{4}{2}$ independent trials, as the frequencies are $1, 6, 15, 20, 15, 6, 1$ respectively for $0, 1, \ldots, 6$ such steps. Even more remarkably (if less obviously), the joint distribution of the number of vertical steps in each of the four columns (vertical lines of the grid), which we shall call the column counts, is the product of *independent* binomial distributions for $m = 0, 1, 2, 3$ respectively. The corresponding statements remain true for the collection of all disjoint Schröder n-families for any $n \in \mathbf{N}$ (this is in fact a corollary to our bijective proof). By an obvious symmetry one also has the corresponding statement for the joint distribution of the number of horizontal steps on each of the four horizontal lines of the grid (row counts), with m increasing from bottom to top.
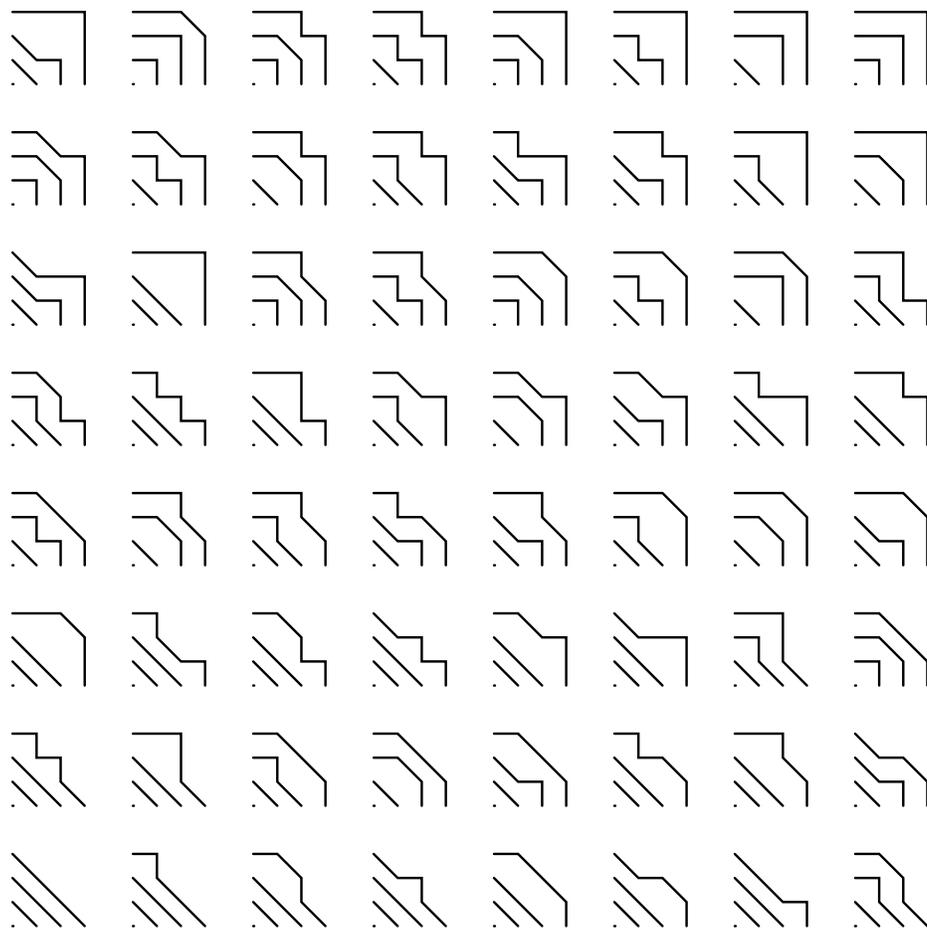
Figure 3: The collection of all disjoint Schröder 4-families

One also has a similar statement for joint distribution of what we shall call inter-column counts, the number of horizontal steps connecting each pair of successive columns; now m decreases, from $n-1$ between the leftmost pair of columns to 1 between the rightmost pair. This statement can be seen to be equivalent to the one about column counts, if one uses the duality illustrated in Figure 4; this is a bijection between the set of disjoint Schröder n-families and the set of such families transformed by a central reflection sending the origin to $(n-\frac{1}{2}, n-\frac{1}{2})$ (grid points are mapped to centres of squares of the original grid). This correspondence is such that halfway on each horizontal or vertical step of a disjoint n family, the step crosses a vertical respectively horizontal step of the dual family. By contrast to these facts, the joint distribution of the number of horizontal (or equivalently vertical) steps in each of the individual paths that make up a disjoint family does not satisfy any such independence.

Given these observations, one may hope to find a bijection between disjoint Schröder n-families and triangular arrays of $\binom{n}{2}$ "bits" (values in $\{0, 1\}$) in such a way that, for a
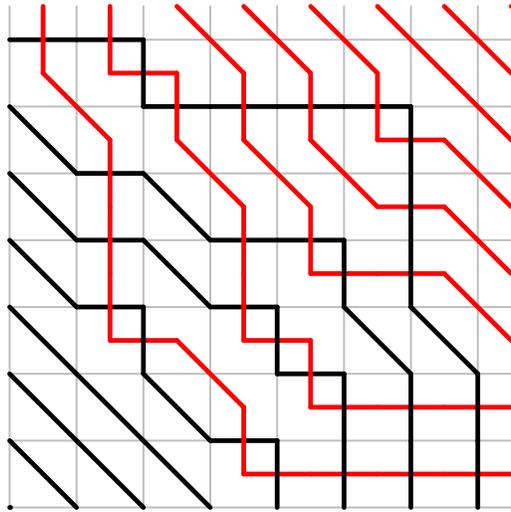
Figure 4: A disjoint 8-family and (in red) its dual family

certain arrangement of the triangle into columns of length i for $i = 0, 1, \ldots, n-1$, the sum of the bits in column j will give the column count for column j of the corresponding disjoint n-family.

Looking at just the 6 paths with a single horizontal and vertical step, in the bottom line of Figure 3, one sees that the point of intersection of the lines containing these steps are all different, and form the triangle of all grid points that are not visited by the paths of the unique 4-family with diagonal steps only (at the bottom left). This might suggest associating the triangular array of bits to those grid points, in the hope to find a bijection with disjoint n-families such that the column counts *and* the row counts of an n-family are given by the sums of these bits along the corresponding column or row. This is easily seen to be impossible however: the joint distribution of the column counts and row counts of disjoint n-families is different from the joint distribution of column sums and row sums in such triangular array of bits. For instance for any $c \leqslant n$ there exist disjoint n-families with c horizontal and c vertical steps, all of them contributing to the *same* column count respectively row count; when $c \geqslant 2$ the corresponding situation cannot occur for the column and row sums of a triangular array of bits. On the other hand it may be checked in the example that the joint distribution of the column counts and the *inter-column* counts over all disjoint 4-families is precisely that of column sums and row sums in such triangular arrays of $\binom{4}{2} = 6$ bits. This suggests that in formulating a bijection one should prefer to abandon the transposition symmetry, and instead focus on (say) vertical alignment only. Indeed our bijection will be such that column counts (of vertical steps) and inter-column counts (of horizontal steps) can be immediately read off from the triangular bit-array. However, the way these steps are distributed within their column respectively inter-column space will not be so easy to read off.

The starting point of our bijection will then be to translate a triangular array of $\binom{n}{2}$ bits into a cliff-shaped n-family, where line i of the triangle (viewed in some appropriate direction) determines the cliff-shaped path P_i. For such families of paths column counts and inter-column counts are defined, just like for disjoint families; this time one has the particular circumstance that only path P_i contributes to the column count for column i. So each bit has two associated indices: that of a path P_i (which also gives the column to which it may contribute a vertical step), and that of a an inter-column space, from column j to $j+1$, to which it may contribute a horizontal step; the triangle runs through values $0 \leqslant j < i < n$. There does not seem to be a particularly good way to view our triangle as positioned geometrically in some way relative to the n-family. A *somewhat* suggestive choice would be to position the bits at the (midpoints of) the horizontal steps in the special "no diagonal steps" n-family (the one at the top right of Figure 3); then column sums of bits give the corresponding inter-column counts for the path family, and the rows sum the corresponding (after transposition) column counts.

Our main task will then be to find a systematic and reversible way to take any cliff-shaped n-family and redistribute its horizontal and vertical steps among the different paths, keeping each of these steps within its inter-column space respectively within its column, so as to obtain a disjoint family. We can give some heuristic arguments to explain the form that our algorithm will take. For the redistribution of steps, the vertical steps will play a passive role, since the fact that within column k they are originally all concentrated in the path P_k makes that they initially carry very little information. So we shall operate primarily on the initial parts of cliff-shaped paths, which contain a mix of horizontal and diagonal steps; whenever we move a horizontal step from one path to another (exchanging it with a diagonal step), a corresponding vertical step will also be moved between the paths so as to keep the ending point of each path unchanged.

An important aspect of our "untangling" procedure will be that it operates essentially on parts of the paths that contain only horizontal and diagonal steps. Since redistributing vertical steps in column k may move them from path P_k into paths P_i with $i > k$, it is practical to so treat columns sequentially by *decreasing* value of k, and to leave column k as it is once the vertical steps it contains are redistributed. In this way we avoid having "polluted" paths with vertical steps in the columns under consideration, and their parts beyond column k (the column where redistribution currently takes place) can be ignored by the procedure.

If the initial cliff-shaped n-family happens to be already disjoint, our procedure will just leave it as it is. Although this condition is satisfied by a vanishingly small fraction of the families as n increases (notwithstanding the 26 such cases out of 64 for $n = 4$), the idea of acting only when necessary is an important principle for the procedure. The following setting is one where action will be required: we have two successive paths P_i, P_{i+1} with $i \geqslant k$, whose parts up to the point where they enter column k (which parts do not contain any vertical steps) intersect. At the point in time where we start considering column k (vertical steps having been redistributed in all columns beyond it), this situation may occur for $i = k$: since P_k cannot have been involved in any of the previous operations, it is in its initial state, and could be any cliff-shaped path. In particular there is no

reason to suppose anything about its position relative to P_{k+1}. And even though P_{k+1} can have been operated upon, and therefore may be more likely to have certain forms than others, it certainly *can* also involve any sequence of horizontal and diagonal steps before entering column k. Indeed P_{k+1} could also be in its initial state, as would happen if no action at all was required before considering column k, and as is certainly the case at the very beginning, when for $k = n - 2$ we consider the paths P_{n-2}, P_{n-1}. So apart from the absence of vertical steps we cannot assume anything about the first k steps of P_i and P_{i+1}. On the other hand we shall assume that in column k only P_i may have vertical steps initially, and also that beyond column k the paths are already disjoint.
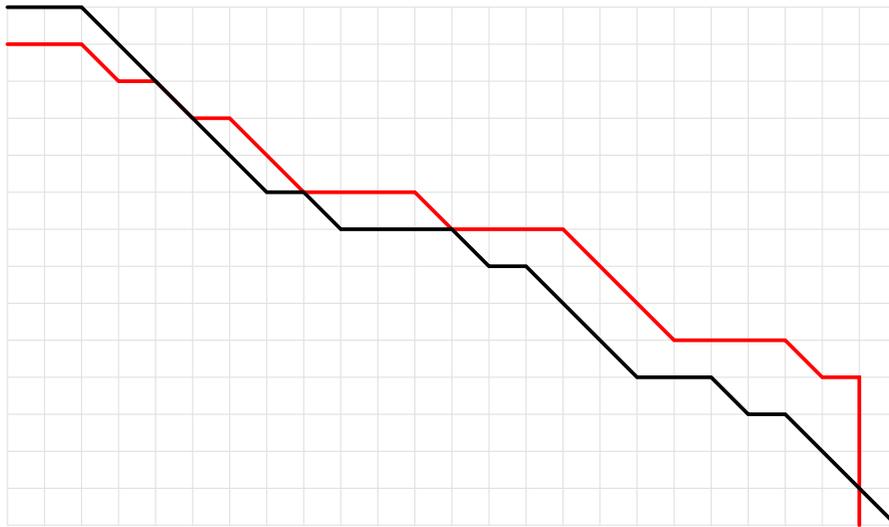

Figure 5: Initial parts of a pair of paths in need of untangling

A typical situation is depicted in Figure 5; the red path is P_i and the black one P_{i+1}. The paths have been truncated to their initial parts relevant to the task of untangling: path P_{i+1} has no vertical steps in column $k = 23$ and passes to column $k + 1$, while path P_i does have at least the vertical steps in column k shown. It may be that P_i continues further downwards (as it will when $i = k$, since then P_i is cliff-shaped), or it may pass to column $k + 1$ as well; but if it does, it must do so while staying below P_{i+1}.

Since the paths depicted first meet in column 4, the principle to act only when needed suggests leaving everything up to column 3 intact. We might then avoid the collision in column 4 either by taking a diagonal step in P_i or by taking a horizontal step in P_{i+1}, but if we want to keep the number of horizontal steps unchanged, and more precisely the number of horizontal steps from column 3 to column 4, the only (easy) way to achieve this is by making *both* these changes. As this transfers a horizontal step from P_i to P_{i+1}, we shall also need to transfer a vertical step, in column k. As we shall see below, the latter transfer combined with the initial absence of vertical steps in P_{i+1} is a key point in

being able to reverse the modification(s) made, as it serves as witness for the effort that was required to make the pair of paths disjoint.

Having "switched step directions" between columns 3 and 4, the remainders of P_i and P_{i+1} are shifted down respectively up by one unit. It might seem that the next (and only) remaining problem that needs resolving occurs in the passage to column 15, where the original path P_i rises *two* units above P_{i+1} for the first time, so that the mentioned remainders meet in spite of the shifts. However, while switching step directions in the passages to columns 4 and 15 only (and moving two vertical steps to P_{i+1}) would succeed in making the paths disjoint, the result leaves insufficient information to reconstruct the set of steps that were adjusted, and hence the initial paths. The modified steps cause the new paths to move apart at a point where they are as close together as they may get, but so do the passages to columns 8 and 12 (in the modified paths), with nothing to distinguish the situation in these columns from that in columns 4 and 15. Therefore, we shall instead switch directions *every* time that the height of P_i above P_{i+1} first reaches a new nonnegative value, which in the example happens for the values $0, 1, 2, 3$ when passing respectively to columns 4, 6, 15 and 23. The result of those four interchanges, and of moving 4 vertical steps from P_i to P_{i+1} in column 23, is shown in Figure 6.
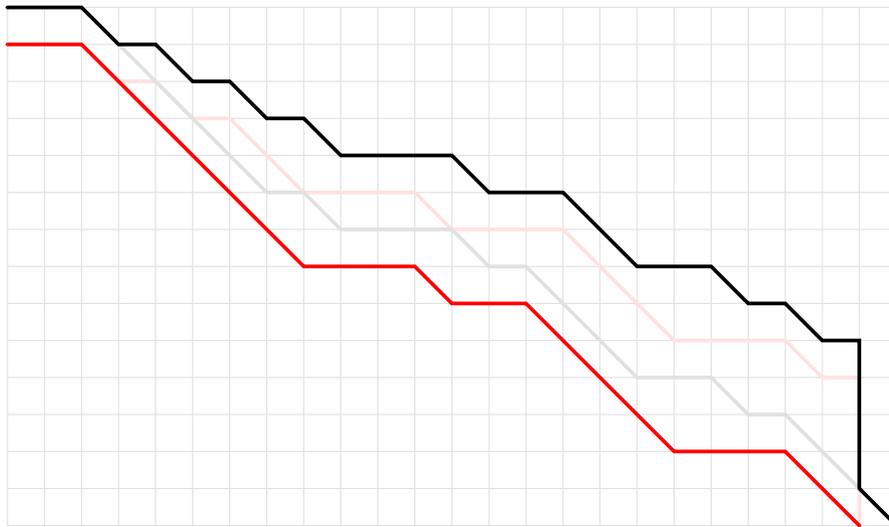

Figure 6: Initial parts of the pair of paths untangled

One can view this transformation in terms of a single path Δ of a new kind, defined by the "difference" of P_{i+1} and P_i: one that makes a down-step whenever between a pair of columns P_{i+1} has a diagonal step and P_i a horizontal one, an up-step when P_{i+1} has a horizontal step and P_i a diagonal one, and a neutral step when P_{i+1} and P_i have the same type of step (in Δ the two kinds of neutral steps are distinguished, so that no information is lost). Then P_i and P_{i+1} are disjoint if and only if the maximal depth d beneath its

starting level to which Δ descends is 0, and we have described a procedure to transform any Δ into a path with $d = 0$, by reversing all those down-steps that lead to a new left-to-right minimum. The procedure is well known in this setting, and in various equivalent guises; see for instance [vLee10] and references therein. The mapping it defines becomes injective when restricted to paths with a given initial value of d. This can be seen by viewing the transformation as obtained by iterating as long as possible the operation of reversing the first down-step that leads to the *globally* minimal level (which is initially d); this iteration produces the reversals from right to left. Each such operation is invertible by reversing the last up-step starting at the globally minimal level, so given d one can undo the entire transformation by repeating this inverse operation d times.

The procedure described allows making a pair of successive paths P_i, P_{i+1} disjoint up to column k, and is reversible provided that P_{i+1} initially had no vertical steps in that column. Assuming that the paths $P_{k+1}, \ldots, P_{n-1}$ have previously been made disjoint, we can use this procedure to make P_k disjoint from P_{k+1}. But since this in general involves moving parts of both paths away from each other, it may cause P_{k+1} to intersect P_{k+2} even though they were disjoint before. In fact one could not expect being able to make $P_k, \ldots, P_{n-1}$ disjoint so easily: one needs to potentially introduce vertical steps in column k for *all* these paths. After all, once this disjointness is obtained, further transformations will no longer change column k, and for each $i \geqslant k$ there certainly exist disjoint n-families in which P_i has one or more vertical steps in column k.

An obvious idea is then to continue applying the untangling procedure as long as there are pairs of adjacent paths that intersect. But unless this process proceeds in a very orderly fashion, it will be problematic to invert, and could even fail to terminate. Fortunately it turns out that the process is indeed very orderly: if after untangling P_i and P_{i+1} we need to untangle P_{i+1} and P_{i+2}, then this may cause P_{i+1} to "bounce back" towards P_i, but when this happens the extra space that their initial untangling had produced between P_i and P_{i+1} is always sufficient to absorb the displacement of P_{i+1}, thus avoiding any new intersection between them. Given this fact (proved in the next section), a single sweep of untangling of paths by increasing value of i, starting at $i = k$, will suffice. The sweep will end when no new intersections are produced, which at the very last is bound to happen after untangling the final paths P_{n-2} and P_{n-1}, if one ever gets to that point.

The succession of intermediate paths families during such a sweep of executions of the untangling procedure for increasing values of i is illustrated in Figure 7, with the path P_i for the next such execution in red. In the very last such execution, the paths are found to be disjoint already and nothing is changed.

To put everything together, it remains to start with a cliff-shaped n-family determined by a triangular array of $\binom{n}{2}$ bits, and apply the above "sweeps" distributing vertical steps in column k among the paths, for $k = n - 2, \ldots, 2, 1$. This process is illustrated in Figure 8, showing the transformation of a cliff-shaped 49-family into a disjoint 49-family in several stages, including the initial and final ones. To avoid distraction, not yet treated cliff-shaped paths, which intersect each other and the already "combed" ones, are in light blue.
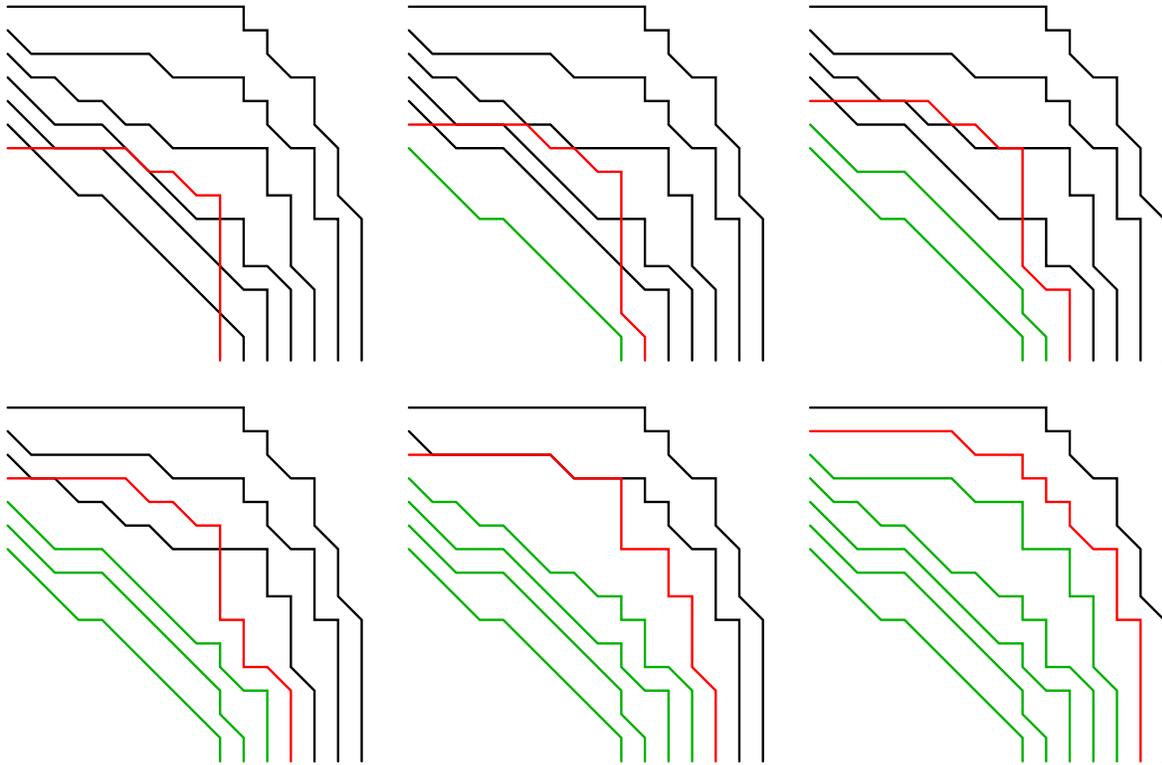
Figure 7: One sweep of untangling until disjointness is obtained

5 A bijective proof

We shall now formulate a bijective proof of theorem 2, by giving an algorithmically defined bijection between the set of disjoint Schröder n-families and the set of cliff-shaped Schröder n-families, the latter set having the number of elements mentioned in the theorem. We shall focus first on the direction from cliff-shaped to disjoint n-families, where the goal is to remove intersections between pairs of paths (this is what "combing" in our title refers to). However our claim that the map so defined is a bijection depends the existence of an inverse transformation defined for any disjoint n-families, and which defines an inverse mapping; in this direction the "goal" is to move, for all k, all vertical steps in column k towards path P_k, where they will appear at the end, so that the paths become cliff-shaped. Whenever one algorithm transforms a family of one type into another, the other algorithm applied to the family produced will realise a step-by-step inverse of the initial transformation.

Our basic operations, forward and backward, operate on a pair of successive paths (P_i, P_{i+1}) in a Schröder n-family (the others are ignored), and depend on an additional parameter $k \leqslant i$. These paths should have no vertical steps in columns $j < k$; the operations will not introduce such steps either. Moreover they leave each of the paths
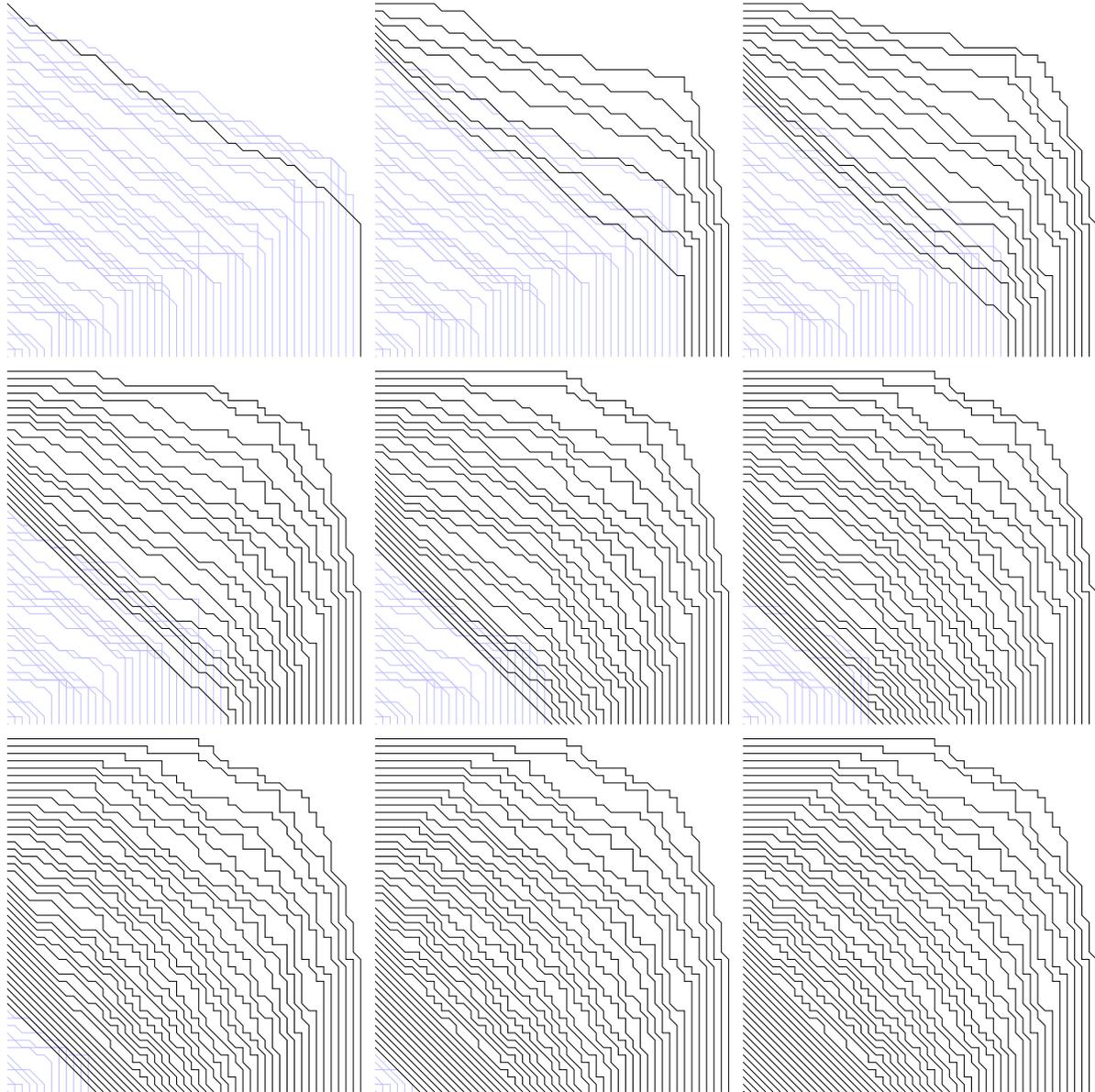
Figure 8: Several intermediate phases of combing 49 paths

unchanged beyond column k, so the only way vertical steps play a role is by a possible transfer between the paths of vertical steps in column k. The forward operation defines a bijection from the pairs of such paths for which P_{i+1} does not have any vertical steps in column k while P_i has enough of such steps in a sense to be made precise, to the pairs of such paths with disjoint supports. In the context where we shall apply the operations, these conditions will be satisfied, and P_i, P_{i+1} will also be disjoint beyond column k.

In what follows the following assumptions are tacitly made: all paths will be assumed to without vertical steps in columns $j < k$, the paths P_i and P'_i are Schröder paths from

$(i, 0)$ to $(0, i)$, and the paths P_{i+1} and P'_{i+1} are Schröder paths from $(i+1, 0)$ to $(0, i+1)$. The absence of vertical steps allows the parts of such paths up to column k to be viewed as graphs of functions: for $\delta \in \{0, 1\}$ and $0 \leqslant j \leqslant k$, let $h_\delta(j)$ be the greatest (and unique, unless $j = k$) value v with $(v, j) \in \operatorname{supp}(P_{i+\delta})$. These functions h_0 and h_1 are weakly decreasing, and their values decrease by at most 1 at each step.

For defining the forward operation (with the mentioned assumptions on P_i, P_{i+1}), put

$$d_j = \max\{\, h_0(j') + 1 - h_1(j') \mid 0 \leqslant j' \leqslant j \,\} \qquad \text{for } 0 \leqslant j \leqslant k. \tag{6}$$

The sequence $(d_0, d_1, \ldots, d_k)$ is weakly increasing, and by at most 1 at each step; it starts with $d_0 = 0$. One will (still) have $d_k = 0$ if and only if the paths P_i and P_{i+1} have disjoint supports up to column k. Now define h'_0, h'_1 by

$$h'_0(j) = h_0(j) - d_j \quad \text{and} \quad h'_1(j) = h_1(j) + d_j \qquad \text{for } 0 \leqslant j \leqslant k. \tag{7}$$

There is at most one pair of paths (P'_i, P'_{i+1}), unchanged from their final points in column k on with respect to (P_i, P_{i+1}), that gives rise to (h'_0, h'_1) in the same way as (P_i, P_{i+1}) gives rise to (h_0, h_1). Our operation is defined only when such (P'_i, P'_{i+1}) exists, and then replaces P_i by P'_i and P_{i+1} by P'_{i+1}.

For any $j < k$, the steps in P'_i, P'_{i+1} from column j to column $j + 1$ will be of the same type as the corresponding steps in P_i, P_{i+1} respectively, unless $d_j < d_{j+1}$. By (6), the latter case occurs only in situations where $h_0(j) = h_0(j+1)$ and $h_1(j) > h_1(j+1)$, in other words when the step from column j to column $j+1$ is horizontal in P_i and diagonal in P_{i+1}. When indeed $d_j < d_{j+1}$, these directions are interchanged in P'_i, P'_{i+1}: the step from column j to column $j + 1$ is diagonal in P'_i and horizontal in P'_{i+1}. This situation arises d_k times in all. As a result, P'_{i+1} has $(h_1(k) + d_k, k)$ as first point in column k, after which it has d_k vertical steps to reach the point $(h_1(k), k)$ where the original path P_{i+1} enters column k.

The path P'_i on the other hand will have d_k vertical steps *less* in column k than P_i has. The (unique) condition for the existence of (P'_i, P'_{i+1}) then is that P_i has at least that many such steps to begin with. So we can detail the requirement alluded to above that P_i have enough vertical steps in column k: we must assume that it has at least d_k such steps, as defined in (6). An equivalent, maybe more natural, way of stating this requirement is that if we would modify P_i by removing all its vertical steps from column k and insert them into column 0 instead (shifting all intermediate steps), then the resulting Delannoy path would have its support disjoint from that of P_{i+1}.

It is clear that $h'_0(j) < h'_1(j)$ for all $0 \leqslant j \leqslant k$, since

$$h'_1(j) - h'_0(j) - 1 = h_1(j) - h_0(j) - 1 + 2d_j \geqslant d_j, \tag{8}$$

and $d_j \geqslant 0$; moreover for $j = k$ one gets that $h'_0(k) < h'_1(k) - d_k = h_1(k)$, which is the first coordinate of the point where P_{i+1} enters into column k, and by the assumption that P_{i+1} has no vertical steps in column k, this point is also the last one of P'_{i+1} in that column. This shows that the supports of P'_i and P'_{i+1} are disjoint up to column k inclusive. In fact this inequality shows that these paths leave at least d_j empty places between them

in any column $j < k$, so whenever d_j increases with j, the modified paths are forced to remain further and further apart. Thus an increase $d_j < d_{j+1}$ not only implies one has a diagonal step in P'_i and a horizontal step in P'_{i+1} between columns j and $j + 1$, but also that the paths then continue to leave this increased number d_{j+1} of spaces (or more) between them, until they enter column k.

The backward operation uses this property to detect the points of increase of d_j from the shape of the paths P'_i, P'_{i+1} alone (so that it can then reconstruct $(d_0, \ldots, d_k)$), but needs to distinguish this situation from one where the *original* difference $h_1(j) - h_0(j)$ increases at $j + 1$ without ever falling back subsequently. But in the latter case one has $d_k = d_j < h'_1(j) - h'_0(j)$ (the equality follows from "not falling back", and the inequality from (8)), whereas in the case $d_j < d_{j+1}$ one has instead $d_k \geqslant d_{j+1} = d_j + 1 = h'_1(j) - h'_0(j)$ (the final equality holds because the maximum in (6) must be attained for $j' = j$). Therefore one can tell the two cases apart provided that d_k is known. But that is the case: P_{i+1} has no vertical steps in column k, so one can read off d_k as the number of vertical steps of P'_{i+1} in column k.

So we can now formulate the backward operation, which can be applied to a pair of paths (P'_i, P'_{i+1}) with supports disjoint up to column k inclusive. We start by defining functions h'_0, h'_1 in terms of respectively P'_i, P'_{i+1}, as before, and in addition let d be the number of vertical steps of P'_{i+1} in column k; then define the sequence $(d_0, d_1, \ldots, d_k)$ by

$$d_j = \min(\{d\} \cup \{ h'_1(j') - h'_0(j') - 1 \mid j \leqslant j' \leqslant k \}) \qquad \text{for } 0 \leqslant j \leqslant k. \qquad (9)$$

We then find h_0, h_1 by using equation (7) in the opposite direction:

$$h_0(j) = h'_0(j) + d_j \quad \text{and} \quad h_1(j) = h'_1(j) - d_j \qquad \text{for } 0 \leqslant j \leqslant k, \qquad (10)$$

and finally take (P_i, P_{i+1}) to be the unique pair of paths, unchanged with respect to (P'_i, P'_{i+1}) from their final points in column k onwards, giving rise to (h_0, h_1).

Several easy verifications suffice to see that this backward operation is well defined. The sequence $(d_0, d_1, \ldots, d_k)$ is weakly increasing by at most one at each step, and satisfies $d_k = d$ (since the supports of P'_i and P'_{i+1} are disjoint in column k) and $d_0 = 0$ (since the disjointness of the supports of P'_i and P'_{i+1} in column j gives $h'_1(j) - h'_0(j) - 1 \geqslant 0$, while $h'_1(0) - h'_0(0) - 1 = 0$). All d vertical steps in column k of P'_{i+1} are absent from P_{i+1} but transferred to P_i, and the steps in P_i and P_{i+1} from column j to $j + 1$ stay of the same kind as respectively in P_i and P_{i+1} when $d_j = d_{j+1}$, while the steps interchange directions when $d_j < d_{j+1}$; this establishes the existence of (P_i, P_{i+1}).

When the pair (P'_i, P'_{i+1}) to which the backward operation is applied was itself obtained by the forward operation from (P_i, P_{i+1}), it can be checked that in the backward operation $d = d_k$, and that the sequence $(d_0, \ldots, d_k)$ is the same as it was in the forward operation (the condition causing $d_j < d_{j+1}$ in the backward operation is equivalent to the one for which we argued that it characterises $d_j < d_{j+1}$ in the forward operation); in this case the pair obtained in the backward operation is therefore the original pair (P_i, P_{i+1}). Conversely, if the backward operation is applied to any applicable pair (P'_i, P'_{i+1}), then the forward operation can be applied to the resulting pair (P_i, P_{i+1}), and it will reconstruct

(P'_i, P'_{i+1}). Again this follows by showing that the forward operation reproduces the same sequence $(d_0, \ldots, d_k)$ as the backward operation, as follows. For a maximal interval of consecutive indices j for which during the backward operation d_j has a constant value, say c, one has the relation $h_0(j) + 1 - h_1(j) = 2c - (h'_1(j) - h'_0(j) - 1)$ throughout. Also the maximal value of this expression is attained for the minimum such j (as well as for the maximum such j, provided it is less than k). Therefore during the forward operation, the value of d_j from (6) will be constant on such intervals as well. On the other hand, when $d_j < d_{j+1}$ during the backward operation, one has $h_0(j) + 1 - h_1(j) = h_0(j+1) - h_1(j+1)$, and together with the constancy result we just gave this shows that $d_j < d_{j+1}$ during the forward operation as well, and therefore that $(d_0, \ldots, d_k)$ is reconstructed identically.

Let us resume the description of these basic operations as somewhat more formalised computational procedures. To that end we need a concrete representation of the n-families of paths operated upon. We choose a representation that facilitates handling paths with a varying number of steps, and allows making evident the simple structure of our operations. An n-family of paths is encoded by a pair of lower triangular matrices (B, D) indexed by $[n] \times [n]$ (recall that $[n] = \{0, 1, \ldots, n-1\}$). The matrix B is strictly lower triangular with entries in $\{0, 1\}$, while D is weakly lower triangular with entries in $\mathbf{N}$. the entry $B_{i,j}$ indicates the direction of the step in P_i between column j and $j + 1$ (a value 0 for horizontal, or 1 for diagonal), and the entry $D_{i,j}$ counts the number of vertical steps of P_i in column j. A cliff-shaped n-family is determined by B alone, and the forward "combing" algorithm will gradually compute D for the corresponding disjoint n-family from it while updating B to match it. The reverse "uncombing" algorithm takes a disjoint n-family encoded by B, D and computes B for the corresponding cliff-shaped n-family from it.

The forward basic operation, which will make paths P_i, P_{i+1} disjoint up to column $k \leqslant i$ inclusive, assumes $D_{i,k}$ is already determined, and at the end of its execution transfers part of its value to $D_{i+1,k}$ (taken to be 0 initially). Its description in procedure 1 uses local variables $cur \in \mathbf{Z}$ recording the current value of $h_0(j) + 1 - h_1(j)$, and $d \in \mathbf{N}$ recording the maximum of cur so far. In this pseudo-code '$\leftarrow$' denotes assignment of a new value, and we write indices in square brackets to remind that matrices are treated as arrays with individually assignable entries.

$untangle(i, k)$:
$cur \leftarrow 0, \quad d \leftarrow 0$
for j **from** 0 **to** $k - 1$ **do**
    $cur \leftarrow cur + B[i+1, j] - B[i, j]$
    if $cur > d$ **then**
        $d \leftarrow cur$
        $B[i, j] \leftarrow 1, \quad B[i+1, j] \leftarrow 0$ {interchange directions of steps}
    end if
end for
$D[i, k] \leftarrow D[i, k] - d, \quad D[i+1, k] \leftarrow d$ {transfer d vertical steps to P_{i+1}}

Procedure 1: Forward operation on paths $i, i+1$ up to column k inclusive

The backward operation in procedure 2 retraces the steps of procedure 1 using the same local variables cur and d. While the sequence of values of d retraces those in procedure 1 in reverse order, the values of cur are different: they record the current value of $h'_1(j) - h'_0(j) - 1$ for the functions h'_0, h'_1 corresponding to the disjoint paths described by the initial values for procedure 2; in particular $cur \geqslant 0$ throughout the execution. In order to set cur correctly, it assumes that the values $h'_0(k)$ and $h'_1(k)$, where paths P_i and P_{i+1} respectively enter column k (values that are not directly available in our encoding) have been stored beforehand as elements h_i, h_{i+1} of an auxiliary array (not local to $cliffify$); these values are also updated to reflect the effect of the operation.

$cliffify(i, k)$:
$d \leftarrow D[i+1, k], \quad cur \leftarrow h[i+1] - h[i] - 1 \quad \{0 \leqslant d \leqslant cur\}$
$D[i+1, k] \leftarrow 0, \quad D[i, k] \leftarrow D[i, k] + d \quad \{\text{transfer } d \text{ vertical steps to } P_i\}$
$h[i+1] \leftarrow h[i+1] - d, \quad h[i] \leftarrow h[i] + d \quad \{\text{adapt entry point into column}\}$
for j **from** $k - 1$ **down to** 0 **do**
   $cur \leftarrow cur + B[i+1, j] - B[i, j]$
   if $cur < d$ **then**
      $d \leftarrow cur$
      $B[i, j] \leftarrow 0, \quad B[i+1, j] \leftarrow 1$
   end if
end for

Procedure 2: Backward operation on paths $i, i+1$ up to column k inclusive

We can now formulate somewhat more formally what was proved above about the forward and backward operations, as statement about the given procedures. For conciseness we denote by $\mathrm{Pathfam}(n)$ the set of pairs of matrices (B, D) where B is strictly lower triangular $[n] \times [n]$ matrix with entries in $\{0, 1\}$, while D is weakly lower triangular $[n] \times [n]$ matrix with entries in $\mathbf{N}$.

The procedures obviously only inspect and alter a small part of these matrices, but there is no need to make explicit mention of that fact. As we proceed along the path P_i, the values of $B_{i,j}$ and $D_{i,j}$ encountered are subtracted from the level; therefore the inequalities below are in the opposite direction as the corresponding comparison of the levels of two paths. Also we have chosen to leave out the respective initial levels i and $i+1$ of the paths P_i and P_{i+1} from the expressions, so when interpreting the inequalities as comparisons of levels, one should take into account the difference in offset.

Proposition 5.1. *For $0 \leqslant k \leqslant i < n - 1$, procedure 1 defines a bijection, and procedure 2 defines the inverse bijection, between on one hand the set of pairs $(B, D) \in \mathrm{Pathfam}(n)$ satisfying*

$$D_{i+1,k} = 0, \quad and$$
$$\sum_{j'=0}^{j-1} B_{i+1,j'} \leqslant D_{i,k} + \sum_{j'=0}^{j-1} B_{i,j'}, \qquad for \; 0 \leqslant j \leqslant k,$$

and on the other hand the set of pairs $(B, D) \in \mathrm{Pathfam}(n)$ satisfying

$$\sum_{j'=0}^{j-1} B_{i+1,j'} \leqslant \sum_{j'=0}^{j-1} B_{i,j'}, \qquad \text{for } 0 \leqslant j < k, \text{ and}$$

$$\sum_{j=0}^{k-1} B_{i+1,j} + D_{i+1,k} \leqslant \sum_{j=0}^{k-1} B_{i,j}.$$

The relations $h_{i'} = i' - \sum_{j=0}^{k-1} B_{i',j}$ for $i' = i, i+1$ are assumed to hold initially in procedure 2, and continue to hold after its execution. $\qquad\square$

We now build an algorithmic bijection corresponding to theorem 2 by repeated application of basic operations. The iteration itself is straightforward, although a bit of work will remain to show that the goal is attained. For a given value of k, we shall start calling $untangle(k, k)$ to make P_k and P_{k+1} disjoint (recall that P_k does not extend beyond column k), then $untangle(k + 1, k)$ to make P_{k+1} and P_{k+2} disjoint up to column k, and so forth up to $untangle(n − 2, k)$ to make the last two paths P_{n-1} and P_{n-2} disjoint up to column k. We shall show that the disjointness obtained in a step is not lost in the following step, so this iteration will result in paths $P_k, \ldots, P_{n-1}$ being disjoint up to column k. Placing the iteration within another iteration, in which k decreases from $n − 2$ to 0, we ensure that all paths that extend beyond column k are already disjoint when this inner iteration starts. Since the parts beyond column k are unaffected by it, the inner iteration will in fact achieve that $P_k, \ldots, P_{n-1}$ are entirely disjoint, and at the end of the outer iteration the whole n-family will be disjoint. Note that in general applying $untangle(i, k)$ *will destroy* the disjointness of P_{i+1} and P_{i+2} up to column k, which explains why the inner iteration is needed.

Since each call $untangle(i, k)$ takes care of setting the value of $D_{i+1,k}$ for use in the subsequent call $untangle(i + 1, k)$, all that remains to do is to ensure that $D_{k,k}$ is set correctly before the inner iteration at k starts; this is easy since the number of final vertical steps in the cliff-shaped path P_k is equal to its number of horizontal steps. We obtain the combing algorithm described in procedure 3.

```
for k from n − 1 down to 0 do
    D[k, k] ← k − ∑_{0⩽j<k} B[k, j]   {initialise diagonal entry}
    for i from k to n − 2 do
        untangle(i, k)
    end for
end for
```

Procedure 3: Combing algorithm from cliff-shaped to disjoint n-families

A first verification to be made is that the condition of proposition 5.1 is satisfied whenever $untangle(i, k)$ is invoked. This is clear initially when $i = k$, since the initialisation of $D_{k,k}$ gives that $D_{k,k} + \sum_{j'=0}^{j-1} B_{i,j'} = k − \sum_{j'=j}^{k-1} B_{i,j'} \geqslant j$. To prove that the inequality is satisfied when $i > k$, we need the hypothesis that the paths P_i and P_{i+1} were disjoint just before $untangle(i−1, k)$ was executed. This means that one has $\sum_{j'=0}^{j-1} B_{i+1,j'} \leqslant \sum_{j'=0}^{j-1} B_{i,j'}$ for $0 \leqslant j \leqslant k$ at the start of $untangle(i − 1, k)$. If $d = D_{i,k}$ is the final value obtained

by this variable during that execution, then for any such j the value of $\sum_{j'=0}^{j-1} B_{i,j'}$ is decreased by at most d by the procedure, and since the values $B_{i+1,j'}$ are unaffected, one obtains $\sum_{j'=0}^{j-1} B_{i+1,j'} \leqslant d + \sum_{j'=0}^{j-1} B_{i,j'}$ at the end of $untangle(i-1,k)$, and therefore at the beginning of $untangle(i,k)$; this is the condition required.

A reverse (uncombing) algorithm is also easy to formulate. Here both B and D have well defined values initially, and the only initialisation required is that of the vector h, which should give the levels at which paths P_i and P_{i+1} enter column k at the point where $cliffify(i,k)$ is invoked, as mentioned in proposition 5.1. Since procedure 2 takes care of updating the vector h according to the changes to B it produces, these initialisations are easily integrated into the uncombing algorithm, which only needs to take care of the passage from column $k-1$ to k. We obtain the algorithm described in procedure 4.

---

for k **from** 0 **to** $n-1$ **do**
    for i **from** $n-1$ **down to** k **do**
        if $k=0$ **then**
            $h[i] \leftarrow i$ {initialise height function for column 0}
        else
            $h[i] \leftarrow h[i] - B[i, k-1]$ {adapt height function to column k}
        end if
        if $i < n-1$ **then**
            $cliffify(i,k)$
        end if
    end for
end for

Procedure 4: Uncombing algorithm from disjoint to cliff-shaped n-families

---

For this algorithm it is easy to see that in the inner loop for k, the condition of proposition 5.1 is satisfied, provided that the paths $P_k, \ldots, P_{n-1}$ are disjoint at the start of the loop. Indeed the condition when calling $cliffify(i,k)$ precisely requires the disjointness of P_i and P_{i+1}, and although a preceding $cliffify(i+1,k)$ may have changed the entries $B_{i+1,j}$ that describe P_{i+1}, this can only have made them smaller, moving P_{i+1} away from P_i. In column k the vertical steps introduced by $cliffify(i+1,k)$ come *before* the unchanging point where P_{i+1} leaves that column, so this does not endanger disjointness with P_i either. It is on the other hand not obvious that the paths $P_{k+1}, \ldots, P_{n-1}$ will again be disjoint at the end of the inner loop (and of course P_k in general *will not* be disjoint from them). This brings us to the main technical verification that needs to be done in order to conclude that we have described well defined combing and uncombing bijections.

Proposition 5.2. *Let* Pathfam(n,k) *denote the subset of* Pathfam(n) *of pairs* (B,D) *encoding n-families without any vertical steps in any non-final column before column k (so $D_{i,j} = 0$ whenever $0 \leqslant j < k$ and $j < i < n$) and for which the supports of the paths $P_k, \ldots, P_{n-1}$ are all disjoint. Then for each $k < n$, the inner loop at k of procedure 3 defines a bijection, and the one of procedure 4 defines the inverse bijection, between* Pathfam$(n, k+1)$ *and* Pathfam(n,k)

Proof. We have already seen that, when starting in the forward direction from an element of Pathfam$(n, k+1)$, the calls $untangle(i, k)$ in the inner loop of procedure 3 are invoked under the proper conditions: the number of units of $D_{i,k}$ (vertical steps) that such a call transfers to $D_{i+1,k}$ does not exceed the value of $D_{i,k}$ at that point. Starting in the backward direction from an element of Pathfam(n, k), the inner loop of procedure 4 will also invoke the calls of $cliffify(i, k)$ under the proper conditions, and they will transfer all units from $D_{i+1,k}$ to $D_{i,k}$, so that in the end all units of column k of D have been combined into $D_{k,k}$. The only point left to prove is the disjointness of the supports of the indicated set of paths at the completion of the inner loop, in both directions. This was assumed and remains unchanged beyond column k, and for column k the verifications were done in proposition 5.1 (the disjointness in that column obtained by $untangle(i, k)$ is not endangered by a following $untangle(i + 1, k)$). So only the parts of the paths in columns $j < k$ need to be considered.

It is part of proposition 5.1 that after $untangle(i, k)$ the paths P_i' and P_{i+1}' have disjoint supports up to column k, but (if $i \neq n-2$) the subsequent application of $untangle(i+1, k)$ may move P_{i+1}' in the direction of P_i' again, and we need to show that the resulting path P_{i+1}'' nevertheless stays disjoint from P_i'. Let as before h_0, h_1 be the functions describing the initial paths P_i and P_{i+1}, with h_0', h_1' the ones after modification by $untangle(i, k)$; let h_2 similarly describe the initial path P_{i+2}, and call the functions obtained after $untangle(i + 1, k)$ modifies h_1' and h_2 respectively h_1'' and h_2'. Just as $untangle(i, k)$ determines a sequence $(d_0, \ldots, d_k)$ there is a sequence determined by $untangle(i + 1, k)$ that we call $(e_0, \ldots, e_k)$; then one has equation (7) and similarly $h_1''(j) = h_1'(j) - e_j$, and $h_2'(j) = h_2(j) + e_j$ for $0 \leqslant j \leqslant k$. From (8) we have $h_0'(j) < h_1'(j) - d_j$ and we wish to show $h_0'(j) < h_1''(j) = h_1'(j) - e_j$. It will therefore suffice to show that $e_j \leqslant d_j$ for $0 \leqslant j \leqslant k$. We shall do so by induction on j; the starting case $e_0 = 0 = d_0$ is trivial, so suppose $j > 0$. Then the equivalent of (6) for e_j can be written $e_j = \max(e_{j-1}, h_1'(j) + 1 - h_2(j))$. Now by induction $e_{j-1} \leqslant d_{j-1} \leqslant d_j$, while from the hypothesis $h_1(j) < h_2(j)$ that P_{i+1} and P_{i+2} are initially disjoint we get $h_1'(j) + 1 - h_2(j) = d_j + h_1(j) + 1 - h_2(j) \leqslant d_j$ as well, so indeed $e_j \leqslant d_j$.

Having shown that the inner loop at k of procedure 3 maps Pathfam$(n, k + 1)$ to Pathfam(n, k), we must also prove that conversely the inner loop at k of procedure 4 maps Pathfam(n, k) to Pathfam$(n, k+1)$. The situation is a bit different, in that the disjointness of P_{i+2} and P_{i+1} that we need to show (for $k \leqslant i < n - 2$) is first potentially destroyed by $cliffify(i + 1, k)$, and then must be restored by $cliffify(i, k)$. We can use the same notation as above, but the hypotheses differ: we assume that $cliffify(i + 1, k)$ transforms (h_1'', h_2') into (h_1', h_2) while producing (from right to left) a sequence $(e_0, \ldots, e_k)$, and then $cliffify(i, k)$ transforms (h_0', h_1') into (h_0, h_1) while producing a sequence $(d_0, \ldots, d_k)$. Again the key point is establishing $d_j \geqslant e_j$ for $0 \leqslant j \leqslant k$, since analogously to (8) one has $h_1'(j) = h_1''(j) + e_j < h_2'(j)$, and so the condition $d_j \geqslant e_j$ will imply the desired inequality $h_1(j) = h_1'(j) - d_j < h_2'(j) - e_j = h_2(j)$. This time we use descending induction on j; the initial case $e_k \leqslant d_k$ is a consequence of the fact that $cliffify(i + 1, k)$ transfers all e_k vertical steps of P_{i+1}' to P_i', where they contribute to d_k. In the induction step we use equation (9) in the form $d_j = \min(d_{j+1}, h_1'(j) - h_0'(j) - 1)$, which allows us to prove

$d_j \geqslant e_j$ in two parts, as before: by induction $d_{j+1} \geqslant e_{j+1} \geqslant e_j$, and since $h_1''(j) > h_0'(j)$ (the hypothesis that the original paths P_{i+1}'' and P_i' are disjoint) one also has $h_1'(j) - h_0'(j) - 1 = h_1''(j) + e_j - h_0'(j) - 1 \geqslant e_j$. This completes the proof. $\qquad\square$

We can now state our main result, a bijective version of theorem 2.

Theorem 3. *The algorithm of procedure 3 defines a bijection, and the algorithm of procedure 4 defines the inverse bijection, between on hand the set of cliff-shaped Schröder n-families, encoded by the corresponding strictly lower triangular matrices B with entries in $\{0, 1\}$, and on the other hand the set of disjoint Schröder n-families, encoded by the corresponding pairs (B, D).*

Proof. After pairing each B corresponding to a cliff-shaped n-family with the corresponding diagonal matrix D with diagonal entries $D_{k,k} = k - \sum_{0 \leqslant j < i} B_{i,k}$, the set of cliff-shaped Schröder n-families corresponds to Pathfam(n, n) and the set of disjoint Schröder n-families corresponds to Pathfam$(n, 0)$. Now procedure 3 realises the composite map

$$\text{Pathfam}(n, n) \to \text{Pathfam}(n, n-1) \to \cdots \to \text{Pathfam}(n, 0) \qquad (11)$$

where the individual maps are the bijections of proposition 5.2, and procedure 4 realises the reverse composition of the corresponding inverse bijections. $\qquad\square$

It may be observed that the initial map Pathfam$(n, n) \to$ Pathfam$(n, n-1)$ and the final map Pathfam$(n, 1) \to$ Pathfam$(n, 0)$ are in fact identity maps: the sets of families involved are the same in both cases (with just slightly different descriptions), namely that of the cliff-shaped n-families respectively that of the disjoint n-families, and for $k = n-1$ and $k = 0$ our procedures only perform some administrative actions without any changes to the paths.

6 Some complements and discussion

As we have mentioned in the introduction, and illustrated in Figure 1, there is a bijection between what we have called disjoint n-families and tilings of the Aztec diamond of order $n - 1$. It is not easy to attribute the discovery of this bijection clearly. To our knowledge, a bijection between disjoint families of paths and domino tilings of the Aztec diamond is first described in the literature by Johansson in [Joha02], although unpublished sources witness that the existence of such a correspondence was known in certain circles at least about a decade earlier. (We note that in the cited description, like in the proof of [EuFu05, proposition 2.2], no path of length 0 is present in the disjoint family; curiously an interdiction for the (next) shortest path to pass through the corresponding point, which now becomes necessary in order to obtain the correspondence with tilings of the Aztec diamond, is not formulated either.) However a similar bijection in a slightly different context, and involving "families" of just one path, goes back to Sachs and Zernitz [SaZe94]. That context is originally that of counting dimer coverings (perfect matchings) in a graph describing the adjacency of squares in the *augmented* Aztec diamond, obtained from an

Aztec diamond of order n by replacing the $2 \times 2n$ rectangle it contains by a $3 \times 2n$ rectangle. Each such covering (equivalent to a domino tiling of the augmented Aztec diamond) turns out to be determined (bijectively) by a path from source to sink in a particular orientation of the graph that is illustrated in Figure 9. The observation that those paths can be replaced by paths built up of three types of steps, two of which are not parallel but at a 45°-angle with the corresponding dominoes (therefore leading to what we have called Delannoy paths), is due to Dana Randall [LRaS01], and is mentioned in [Ciu96] and [Stan99, p. 277 (6.49 a)].
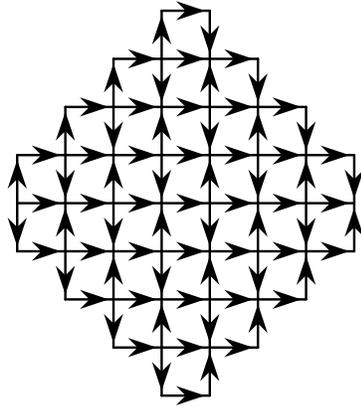

Figure 9: Directed graph for augmented Aztec diamond, order 4

Even though careful inspection of an illustration like our Figure 1 makes it quite obvious that a correspondence between domino tilings and families of disjoint paths can be defined, and that this should be a bijection to an easily described set of such paths, it is worth while to formulate a precise statement and proof. For that we could in principle refer to the arguments given in [Joha02] (and our arguments will in any case be of a very similar nature to those), but we shall here take the occasion to give a statement that tries to isolate the precise properties of the situation at hand that are necessary to ensure that the construction can be carried out. These turn out to be very few, and we therefore obtain a quite general statement about domino tilings, which in particular does not mention the Aztec diamond at all.

Proposition 6.1. *Let S be a finite subset of $\mathbf{Z}^2$, viewed as a set of squares in the plane, with $B = \{ (i,j) \in S \mid i \equiv j \pmod 2 \}$ and $W = S \setminus B$ its subsets of black respectively white squares. Define sets E, I, X of vertical edges with a white square w to their left and a black square b to their right, where E (the "entries") is the set of such edges with $w \notin W$ and $b \in B$, the set X (the "exits") is that of such edges with $w \in W$ and $b \notin B$, and I (the "interior edges") is the set of such edges with $w \in W$ and $b \in B$; formally (identifying*

such an edge with the black square to its right) we set

$$E = \{ b \in B \mid b - (0,1) \notin W \},$$
$$I = \{ b \in B \mid b - (0,1) \in W \},$$
$$X = \{ b \in \mathbf{Z}^2 \setminus B \mid b - (0,1) \in W \}.$$

There is a bijection between on one hand the set of domino tilings of S and on the other hand the set of disjoint families of paths, using steps chosen from $\{(1,1),(0,2),(-1,1)\}$, each path going from a element of E to an element of X while passing through elements of I as intermediate points only, and such that each element of E or of X is the starting respectively ending point of a path.

Finiteness is the only hypothesis made for the set of squares for which domino tilings are considered (and even that could probably be avoided, but this would require some adaptation of the statement). This means of course that very possibly no domino tilings exist at all, and therefore no path families. The most obvious obstruction against the existence of such tiling is a nonzero balance $\#B - \#W$ between black and white squares; this balance is equal to the balance $\#E - \#X$ between entry and exit points for the path which clearly must be zero for path families to exist. But there may be other obstructions, which possibly are more easily recognised in one of the two formulations than in the other.

While the expressions for E, I, X given in the statement of the proposition for convenience identify (as an element of $\mathbf{Z}^2$) any vertical edge that has a black square to its right with that black square, our proof will use a geometric language that distinguishes the two as different kinds of objects.

Proof. First let a domino tiling of S is given. Associate to each domino d of the tiling a pair $(e, e') \in (E \cup I) \times (I \cup X)$, which will serve as a step in one of the paths of the corresponding family whenever $e \neq e'$: take e to be the left edge of the black square of d, and e' the right edge of the white square of d. Every edge in $E \cup I$ occurs as e exactly once, namely for the domino that contains the square $b \in B$ at the right of the edge, and every edge in $I \cup X$ occurs as e' exactly once, namely for the domino the contains the square $w \in W$ at the left of the edge. According to the four possibilities for the orientation and colouring of a domino, each such pair (e, e') obtained either has $e = e'$, or otherwise $e' - e$ is in the set $\{(1,1),(0,2),(-1,1)\}$ of allowed steps. So with I' the subset of I of edges for which the squares w, b to its left and right do not form a domino of the tiling, every $i \in I'$ joins a unique (e_0, i) with a unique (i, e_1'), and $i - e_0$ and $e_1' - i$ are allowed steps. One can then follow unique paths from each entry to each exit and vice versa, and these paths exhaust the set of valid steps obtained from the tiling (monotonicity in the second coordinate makes cycles impossible).

Conversely let a family of paths as described in the proposition be given. For any black square $b \in B$ of S, its left edge e belongs to $E \cup I$. If e is in I but not on any path of the family, then the white square w to the left of e is in W, and we let (w, b) be a domino of the tiling; otherwise we pair up b with the white square to the left of the edge e' reached from e by one forward step on the path passing through it. Similarly the right edge e' of

any square $w \in W$ belongs to $I \cup X$, and w is paired either with the black square the right of e' if $e' \in I$ is not on any path for the family, or otherwise with the black square to the right of the edge e reached by going one step back along the path passing through e'. Clearly this attribution of squares is reciprocal, so one obtains a partition of $S = B \cup W$ into dominoes. The maps from domino tilings to path families and vice versa are inverses of each other, by inspection of the definitions.   □

We note that a similar result can be proved in the same way for lozenge tilings of a subset of triangles in a triangular tiling of the plane, and leads to a bijection with families of disjoint paths in which only two basic steps are allowed. This correspondence is very well known, probably more so than the one for domino tilings, as there is a visually obvious correspondence between lozenge tilings and stacks of cubes (plane partitions), under which the family of paths associated to the tiling traces the edges of "slices" of cubes into which the stack can be cut.

To apply this proposition to obtain the correspondence between domino tilings of the Aztec diamond of order $n-1$ and disjoint Schröder n-families, we may position the Aztec diamond so that all entry points are at its bottom left, and all exists at its bottom right, as it illustrated in the first part of Figure 10. One can then apply a linear transformation with matrix $\frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ to the paths so as to map their basic steps respectively to $(0,1)$, $(-1,1)$ and $(-1,0)$, and then shift them so that path P_i runs from $(i,0)$ to $(0,i)$ as we have assumed. We get such paths for $0 < i < n$, but we do not get P_0 at $(0,0)$; indeed our proposition cannot produce paths with 0 steps, as one has $E \cap X = \emptyset$. However, since the edge corresponding to the origin, marked in yellow in Figure 10, is not in $E \cup I \cup X$, none of the paths P_i for $i > 0$ can pass though it, so we can safely add a path P_0 of 0 steps at the origin to complete the disjoint n-family. The edge in question may be thought of as part of the Aztec diamond configuration, even though neither of the squares it separates belong to the diamond.

The proposition allows us to understand the qualitative difference between the problems of tiling the Aztec diamond and the augmented Aztec diamond: the latter (if properly positioned) gives rise in the path setting to a situation where there is just a single entry point and a single exit, whereas for the Aztec diamond there are as many entry points and exits as the order of the diamond. This difference is illustrated in Figure 10.

From the point of view of domino tilings, the choice to focus on vertical edges between squares with a black square on their right is an arbitrary one among four similar possibilities. This means that with a single domino tiling one can associate four different disjoint path families by making different choices, adapting the direction of the basic steps in the paths, as is illustrated in Figure 11. Note that the duality illustrated in Figure 4 just expresses the relation between two of these disjoint path families associated to the same domino tiling, for classes of edges chosen with the same direction (SW–NE in our figure), but with the opposite relation to the black-white colouring of squares.

There is another way to view this choice of orientation underlying the bijection from tilings to path families. The choice of a collection of edges to focus on (like the vertical edges with a black square to their right) can be seen as the choice of a pairing of the squares of $\mathbf{Z} \times \mathbf{Z}$, the pairs being those separated by the chosen type of edges (in the example any
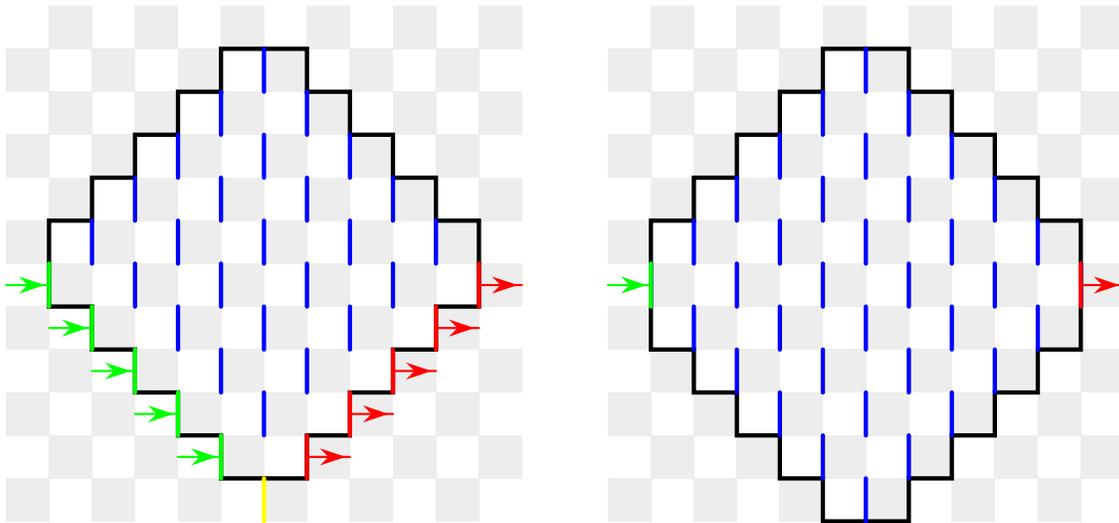
Figure 10: Aztec and augmented diamonds; entry points and exits

black square is paired with the white square to its left). Calling this chosen pairing the "background pairing", the association of paths to a tiling T then is essentially as follows: starting with a black square, move to the white square w paired with it under T, then to the black square paired with w under the background pairing, and repeat; squares that are paired with *the same* square under T as under the background pairing would give short loops, and are excluded. The description of proposition 6.1 combines two successive such steps into one, and places where the background pairing straddles the boundary of S give an entry or exit point for the paths. The fact that the white-to-black move for the background pairing is always in the same direction gives a "forced drift" to paths that makes loops impossible. Yet, one can play the same game even for a background pairing that lacks this "unidirectional" property, although that means that the set of steps a path can choose from varies from one place to another, and loops will in general be possible. If the background matching happens to give, by itself, a tiling of S, then in fact there are no entries and exits, and all paths will be loops; this is what happens in [BrKi05], where an "all horizontal" background matching is chosen, shifted in each row so as to not cross the Aztec diamond boundary.

In the introduction we mentioned that the "domino shuffling" method, like our algorithm, constructs a domino tiling of the Aztec diamond of order n using a sequence of $\frac{n(n+1)}{2}$ bits as input, in an invertible manner. The following more detailed description shows that nevertheless domino shuffling is of an entirely different nature than our procedure.

Domino shuffling constructs tilings of Aztec diamonds of increasing order until the desired order is attained. In passing from one order to the next, first "bad blocks" of two
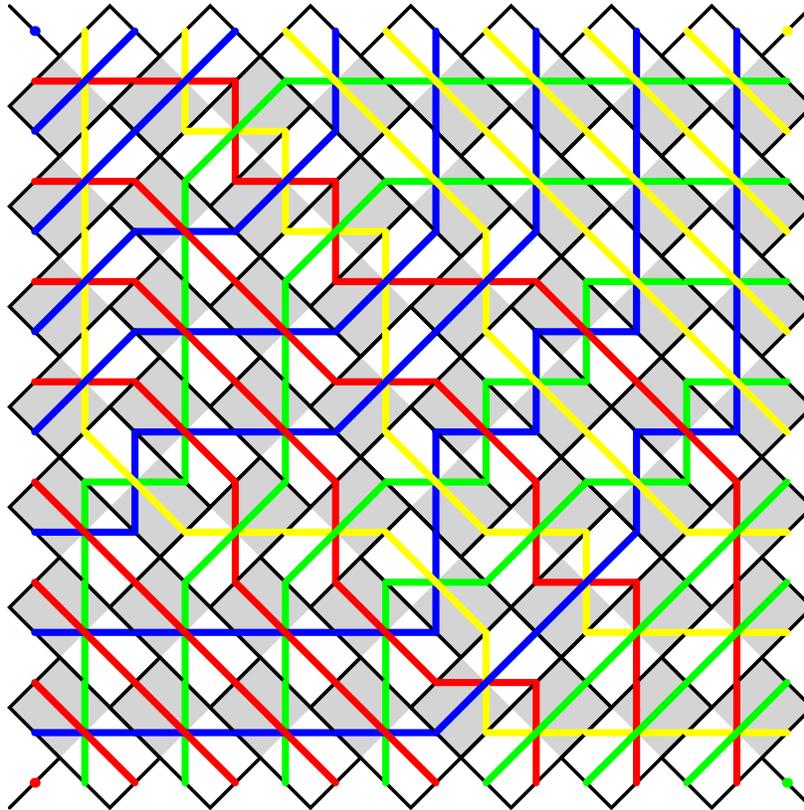
Figure 11: Four families of disjoint paths for a single tiling

dominoes each are removed, then each of the remaining dominoes is shifted one place in one of 4 directions (according to the 4 types that can be found in Figure 11), thus enlarging the diamond, and finally the resulting open space is filled with a choice of "good blocks" of two dominoes each. Each bad or good block has one of two possible tilings, and so carries one bit of information; the net information added is the difference between the numbers of good and bad blocks added, which is always i when passing from order $i - 1$ to order i.

Although the information removed with the bad blocks can be recycled when inserting good blocks, this repeated partial deconstruction/reconstruction gives an essential irregularity of operation to domino shuffling; removal cannot be avoided, as the dominoes in bad blocks would get in the way of the others. The method is based on a representation of tilings by a pair of compatible alternating sign matrices differing by 1 in size, with each matrix restricting the possibilities for choosing the other to a (relatively small) power of 2, those of choosing bad/good blocks. The shuffling procedure can be seen as similar to climbing up a ladder of increasing alternating sign matrices, of which adjacent ones compatible. A complete tiling fixes two such matrices (rungs of the ladder); removal of bad blocks means releasing the smaller matrix of a pair while holding on to the other;

the shuffling proper makes this other matrix switch roles from larger to smaller of a pair; finally insertion of good blocks chooses a new bigger matrix to complete the pair (tiling), and which will serve as support in the next iteration.

We conclude with a brief description of how we found our algorithm without realising at first any connection with the Aztec diamond theorem. It started with a question [M.SE11] posed on the online form Math.StackExchange. It asked for an explanation of the nice evaluation of a the determinant of a matrix with entries defined by a recurrence relation, an insignificant generalisation of the matrix $A_{[n]}$ of Delannoy numbers of section 2. One of the answers given (by "Grigory M") proposed a combinatorial explanation in terms of counting families of non-intersecting lattice paths, but failed to complete the argument by showing that this enumeration was given by the proposed formula. The second author, having came across this question and incomplete answer, was also unable to find a combinatorial argument, but discussed the problem with the first author. Thus discussion did not immediately lead to a bijective proof either, but the first author ultimately found an informal description of the algorithm underlying the bijection presented in this paper. Several months later the second author learned from Christian Krattenthaler that this lattice path enumeration was known to be equivalent to the well known Aztec diamond theorem, but that no quite satisfactory bijective proof of it was known. With interest in the question renewed in this manner, the details of the informal algorithm were worked out, which led to its implementation in a computer program, and the more formal proof presented in the current paper.

References

[BrKi05]   Richard A. Brualdi, Stephen Kirkland "Aztec diamonds and digraphs, and Hankel determinants of Schröder numbers", *J. of Comb. Theory, Series B*, **94** (2005), pp. 334–351

[Ciu96]   Mihai Ciucu, "Perfect matchings of cellular graphs", *J. of Algebraic Combinatorics*, **5** (1996), pp. 87–103.

[EKLP92]   Noam Elkies, Greg Kuperberg, Michael Larsen, James Propp, "Alternating-sign matrices and domino tilings (parts I and II)", *J. of Algebraic Combinatorics*, **1** (1992), pp. 111–132 and 219–234.

[EuFu05]   Sen-Peng Eu, Tung-Shan Fu, "A simple proof of the Aztec diamond theorem", *Electr. J. of Combinatorics*, **12** (2005), #R18.

[JPS98]   William Jockusch, James Propp, Peter Shor, "Random domino tilings and the arctic circle theorem", arXiv:math/9801068v1 [math.CO] (1998)

[Joha02]   Kurt Johansson, "Non-intersecting paths, random tilings and random matrices", *Probab. Theory Related Fields*, **123** (2002),, no. 2, pp. 225–280

[vLee10]   Marc A. A. van Leeuwen, "Some simple bijections involving lattice walks and ballot sequences", arXiv:1010.4847 [math.CO] (2010)

[vLee12]   Marc A. A. van Leeuwen, www-math.univ-poitiers.fr/~maavl/programs/

[LRaS01] Michael Luby, Dana Randall, and Alistair Sinclair, "Markov chain algorithms for planar lattice structures", *SIAM Journal on Computing*, **31** (2001), pp. 167–192.

[Stan99] Richard P. Stanley, "Enumerative Combinatorics, Volume 2", *Cambridge University Press*, 1999, Cambridge Studies in Advanced Mathematics 62

[SaZe94] Horst Sachs, Holger Zernitz, "Remark on the dimer problem", *Discrete Applied Mathematics*, **51** (1994), pp. 171–179.

[M.SE11] http://math.stackexchange.com/questions/86244/