Journal of Combinatorial Theory, Series A 117 (2010) 272–288

Contents lists available at ScienceDirect

Journal of Combinatorial Theory, Series A

www.elsevier.com/locate/jcta

Annular embeddings of permutations for arbitrary genus

I.P. Goulden [a], William Slofstra [b]

[a] Department of Combinatorics and Optimization, University of Waterloo, Canada
[b] Department of Mathematics, University of California, Berkeley, United States

ARTICLE INFO

Article history:
Received 5 March 2008
Available online 8 December 2009

Keywords:
Annular embedding
Combinatorial bijection
Cycle distribution
Dipole
Map enumeration
Ordered tree
Permutation counting
Rooted forest

ABSTRACT

In the symmetric group on a set of size $2n$, let $\mathcal{P}_{2n}$ denote the conjugacy class of involutions with no fixed points (equivalently, we refer to these as "pairings", since each disjoint cycle has length 2). Harer and Zagier explicitly determined the distribution of the number of disjoint cycles in the product of a fixed cycle of length $2n$ and the elements of $\mathcal{P}_{2n}$. Their famous result has been reproved many times, primarily because it can be interpreted as the genus distribution for 2-cell embeddings in an orientable surface, of a graph with a single vertex attached to n loops. In this paper we give a new formula for the cycle distribution when a fixed permutation with two cycles (say the lengths are p, q, where $p+q=2n$) is multiplied by the elements of $\mathcal{P}_{2n}$. It can be interpreted as the genus distribution for 2-cell embeddings in an orientable surface, of a graph with two vertices, of degrees p and q. In terms of these graphs, the formula involves a parameter that allows us to specify, separately, the number of edges between the two vertices and the number of loops at each of the vertices. The proof is combinatorial, and uses a new algorithm that we introduce to create all rooted forests containing a given rooted forest.

1. Introduction

Let $[p]=\{1,\dots,p\}$, and $\mathcal{S}_p$ be the set of permutations of $[p]$, for $p \geqslant 0$. When $p \geqslant 0$ is even, let $\mathcal{P}_p$ be the set of *pairings* on $[p]$, which are partitions of the set $[p]$ into disjoint pairs (subsets of size 2). We refer to the single element of $\mathcal{P}_0$ as the *empty* pairing. Where the context is appropriate, we shall also regard $\mathcal{P}_p$ as the conjugacy class of involutions with no fixed points in $\mathcal{S}_p$. In this latter

E-mail addresses: ipgoulden@uwaterloo.ca (I.P. Goulden), slofstra@math.berkeley.edu (W. Slofstra).

0097-3165/$ – see front matter
doi:10.1016/j.jcta.2009.11.009

context, each pair becomes a disjoint cycle consisting of that pair of elements. Of course, the number of pairings in $\mathcal{P}_p$ is $(p-1)!! = \prod_{j=1}^{\frac{1}{2}p}(2j-1)$, with the empty product convention that $(-1)!! = 1$.

Now, for $p > 0$ and even, let $\gamma_p = (1\,2\ldots p)$, in disjoint cycle notation, and let $\mathcal{A}_p = \{\mu\gamma_p^{-1}\colon\ \mu \in \mathcal{P}_p\}$. Let $a_{p,k}$ be the number of permutations in $\mathcal{A}_p$ with k cycles in the disjoint cycle representation, for $k \geqslant 1$. The generating series for these numbers is given by $A_p(x) = \sum_{k\geqslant 1} a_{p,k}x^k$. Harer and Zagier [4] obtained the following result.

Theorem 1.1. *(See Harer and Zagier [4].) For a positive, even integer p, with $n = \frac{1}{2}p$,*

$$A_p(x) = (2n-1)!!\sum_{k\geqslant 1} 2^{k-1}\binom{n}{k-1}\binom{x}{k}.$$

Other proofs of Theorem 1.1 have been given by Itzykson and Zuber [5], Jackson [6], Kerov [7], Kontsevich [8], Lass [10], Penner [12] and Zagier [15] (see also the survey by Zvonkin [16], Section 3.2.7 of Lando and Zvonkin [9] and the discussion in Section 4 of the paper by Haagerup and Thorbjornsen [2]). Recently, Goulden and Nica [3] gave a direct bijective proof of Theorem 1.1. In the present paper, we consider a similar bijective approach to extend this important result of Harer and Zagier to the case in which the permutation γ_p is replaced by a fixed permutation with two cycles in its disjoint cycle representation. Some additional notation is required.

Let $[q]' = \{1', \ldots, q'\}$, and let $\mathcal{S}_{p,q}$ be the set of permutations of $[p] \cup [q]'$, for $p, q \geqslant 0$. Let $\mathcal{P}_{p,q}$ be the set of pairings on $[p] \cup [q]'$, for $p, q \geqslant 0$, where $p+q$ is even (we refer to the single element of $\mathcal{P}_{0,0}$ as the *empty* pairing). A pair in a pairing is called *mixed* if it consists of one element from $[p]$ and one element from $[q]'$. Where the context is appropriate, we shall also regard $\mathcal{P}_{p,q}$ as the conjugacy class of involutions with no fixed points in $\mathcal{S}_{p,q}$. For $p, q \geqslant 1$, we consider the permutation $\gamma_{p,q} = (1\,2\ldots p)(1'\,2'\ldots q')$, and let $\mathcal{A}_{p,q}^{(s)} = \{\mu\gamma_{p,q}^{-1}\colon\ \mu \in \mathcal{P}_{p,q}$ has s mixed pairs$\}$, and $a_{p,q,k}^{(s)}$ be the number of permutations in $\mathcal{A}_{p,q}^{(s)}$ with k cycles in the disjoint cycle representation, for $k \geqslant 1$. Consider the generating series

$$A_{p,q}^{(s)}(x) = \sum_{k\geqslant 1} a_{p,q,k}^{(s)} x^k.$$

The main result of this paper is the following expression for $A_{p,q}^{(s)}(x)$.

Theorem 1.2. *For $p, q, s \geqslant 1$, with p, q, s of the same odd–even parity and $n = \frac{1}{2}(p+q)$, we have*

$$A_{p,q}^{(s)}(x) = p!q!\sum_{k=1}^{n+1}\sum_{i=0}^{\lfloor\frac{1}{2}p\rfloor}\sum_{j=0}^{\lfloor\frac{1}{2}q\rfloor}\frac{1}{2^{i+j}i!j!(n-i-j)!}\binom{x}{k}\binom{n-i-j}{k-1}\Delta_{k,p,q}^{(s)},$$

where

$$\Delta_{k,p,q}^{(s)} = \binom{k-1}{\frac{1}{2}(p-s)-i}\binom{k-1}{\frac{1}{2}(q-s)-j} - \binom{k-1}{\frac{1}{2}(p+s)-i}\binom{k-1}{\frac{1}{2}(q+s)-j}.$$

Note that Theorem 1.2 gives a summation of nonnegative terms, since for all choices of summation indices k, i, j with $k-1 \leqslant n-i-j$ (so that $\binom{n-i-j}{k-1}$ is nonzero), the difference $\Delta_{k,p,q}^{(s)}$ is nonnegative.

Before we discuss the proof of Theorem 1.2, we shall consider its significance. A major reason that Harer and Zagier's result (Theorem 1.1) is important (as evidenced by so many published proofs) is that it can be restated as an equivalent geometric problem in terms of maps. A *map* is an embedding of a connected graph (with loops and multiple edges allowed) in an orientable surface, partitioning the surface into disjoint regions (called the *faces* of the map) that are homeomorphic to discs (this is called a two-cell embedding). A *rooted* map is a map with a distinguished edge and incident vertex (so, the map is "rooted" at that end of the distinguished edge). The well-known embedding theorem allows us to consider this as equivalent to a pair of permutations and their product (see, *e.g.*, Tutte [14], where the terminology "rotation system" is used to describe this triple of permutations). From this point of view, the kth coefficient $a_{p,k}$ in the generating series $A_p(x)$ evaluated in Theorem 1.1 is equal to the number of rooted maps with 1 vertex, n edges and k faces (where $n = \frac{1}{2}p$, as in Theorem 1.1). Denoting the genus of the surface in which such a map is embedded by g, then the Euler–Poincaré Theorem implies that $1 - n + k = 2 - 2g$, or that $k = n - 2g + 1$.

Similarly, Theorem 1.2 has a geometric interpretation. Let $\mathsf{C}_{p,q}$ be the conjugacy class of $\mathcal{S}_{p+q}$ in which there are two disjoint cycles, of lengths p and q. Then the coefficient $a^{(s)}_{p,q,k}$ in the generating series $A^{(s)}_{p,q}(x)$ is equal to $(2n-1)!/|\mathsf{C}_{p,q}|$ times the number of rooted maps with 2 vertices (of degrees p and q), n edges (exactly s of which join the two vertices together, plus $\frac{1}{2}(p-s)$ that are loops at the vertex of degree p, plus $\frac{1}{2}(q-s)$ that are loops at the vertex of degree q), and k faces (where $n = \frac{1}{2}(p+q)$, as in Theorem 1.2). In this case, if we denote the genus of the surface in which such a map is embedded by g, then we obtain $k = n - 2g$.

Of course, since genus is a nonnegative integer, we must have $a^{(s)}_{p,q,n+1} = 0$, and indeed the coefficient of x^{n+1} in the summation for $A^{(s)}_{p,q}(x)$ given in Theorem 1.2 is zero, since the summand corresponding to $k = n+1$, $i = j = 0$ (which has $\binom{x}{n+1}$ as a factor) is itself equal to zero. For the planar case, which corresponds to $g = 0$, the only nonzero summand that contributes to the coefficient of x^n in the summation of Theorem 1.2 corresponds to $k = n$, $i = j = 0$, and this gives immediately that

$$a^{(s)}_{p,q,n} = s\binom{p}{\frac{1}{2}(p-s)}\binom{q}{\frac{1}{2}(q-s)}.$$

This checks with the straightforward computation that one can make to determine this value by elementary means – there are s edges between the two vertices; between the ends of these edges at each vertex is an even number of vertices, joined by loops without crossings (and there is Catalan number of such arrangements for each such even interval).

This explains the term "genus" in the title; the term "annular" is adapted from its usage in Mingo and Nica [11]. It refers to an equivalent embedding for a map with two vertices, in an annulus. The ends of the edges incident with one of the vertices (say the one of degree p) are identified with p points arranged around the disc on the exterior of the annulus, and the ends incident with the other vertex are identified with q points arranged around the disc on the interior of the annulus. The points corresponding to the two ends of an edge are joined by an arc in the interior of the annulus.

We have been able to find one relevant enumerative result (Jackson [6]) in the literature about such maps, in which the total number of edges is specified, but not the exact number joining the two vertices together. To compare this result to our main result, we must sum over $s \geqslant 1$ (since the underlying graph must be connected, then s, the number of edges joining the two vertices together, must be positive), and thus define

$$A_{p,q}(x) = \sum_{s \geqslant 1} A^{(s)}_{p,q}(x).$$

Then Jackson [6] has considered the case $p = q = n$, and obtained the following result, restated in terms of our notation (by applying the proportionality constant $(2n-1)!/|\mathsf{C}_{n,n}| = n$).

Theorem 1.3. *(See Jackson [6].) For $n \geqslant 1$,*

$$A_{n,n}(x) = n! \sum_{j=0}^{\lfloor \frac{1}{2}(n-1) \rfloor} \sum_{i=0}^{n-2j-1} \sum_{k=0}^{\lfloor \frac{1}{2}(n-2j-1) \rfloor} 4^{-k}\binom{2k}{k}\binom{n}{2k}\binom{2j}{j}\binom{n-2j-1}{i}\binom{x+j+i}{n}.$$

By slightly modifying Jackson's [6] integration argument we are able to obtain the following expression for $A_{p,q}(x)$, with arbitrary p, q of the same parity.

Theorem 1.4. *For $1 \leqslant p \leqslant q$, with $p+q$ even, and $n = \frac{1}{2}(p+q)$,*

$$A_{p,q}(x) = p!q! \sum_{j=0}^{\lfloor \frac{1}{2}(p-1) \rfloor} \sum_{i=0}^{n-2j-1} \sum_{k=0}^{\lfloor \frac{1}{2}(p-2j-1) \rfloor} \frac{1}{2^{n-p+2k} k!(p-2k)!(n-p+k)!} \times \binom{2j}{j} \binom{n-2j-1}{i} \binom{x+j+i}{n}.$$

We have been unable to adapt the integration method to determine $A^{(s)}_{p,q}(x)$ itself.

The proof of Theorem 1.2 is based on a combinatorial model that is developed in Section 3. As a consequence, it is sufficient to enumerate a particular graphical object that we call a *canonical paired array*. In Section 4, we give two combinatorial bijections that, together, reduce the enumeration of canonical paired arrays to a simpler class of objects called *vertical* paired arrays. The latter are explicitly enumerated in Section 5, thus completing the proof of Theorem 1.2. One of the combinatorial conditions on paired arrays is that two graphs associated with them must be acyclic. Because of this, a key component of Sections 4 and 5 is the enumeration of rooted forests which contain a given forest as a subgraph. Thus, as a preliminary, in Section 2 we give a new bijection for this fundamental combinatorial problem.

2. A bijection for rooted forests

In this section, we consider the basic combinatorial question of how many rooted forests with a given set of root vertices contain a given rooted forest. We give a bijection for this that differs from the standard ones in the literature, like the Cycle Lemma (see, *e.g.*, [13, p. 67]) or the Prüfer Code (see, *e.g.*, [13, p. 25]), because it is more convenient for our constructions involving paired arrays.

Suppose we have a rooted forest F (all edges directed towards a root vertex in each component) on vertex-set $[k]$, whose components are the rooted trees $T_1, \dots, T_{m+n}$, $m, n \geqslant 1$. Suppose the root vertex of T_j is r_j, $j = 1, \dots, m$, and the root vertex of T_{m+j} is s_j, $j = 1, \dots, n$. For convenience, we order the trees so that $r_1 < \dots < r_m$, $s_1 < \dots < s_n$, and we let S denote the union of the sets of vertices in the trees $T_{m+1}, \dots, T_{m+n}$.

Theorem 2.1 *(Forest Completion Theorem). There is a bijection between $[k]^{m-1} \times S$ and the set of rooted forests on vertex-set $[k]$ with root vertices $s_1, \dots, s_n$ that contain F as a subforest.*

Proof. We describe such a mapping, which we call the "Forest Completion Algorithm" (FCA). Consider the m-tuple $a = (a_1, \dots, a_m) \in [k]^{m-1} \times S$. We construct the forest corresponding to a iteratively, in $m+1$ stages $0, 1, \dots, m$. At every stage, we have a forest G containing F as a subforest, a permutation π of $[m]$, and a sequence $b = (b_1, \dots, b_m)$ in $[k]^m$. Initially, at stage 0, we have $G = F$, π is the identity permutation and $b = a$. Then, for $i = 1, \dots, m$:

- if b_i is in a different component of G from r_i, then add an arc directed from r_i to b_i in G, and leave π and b unchanged;
- otherwise (so b_i is in the same component of G as r_i), add an arc directed from r_i to b_m in G (to obtain the new G), switch $\pi(i)$ and $\pi(m)$ in π, and switch b_i and b_m in b.

The forest corresponding to the m-tuple a is the terminating forest G. We call the terminating permutation π the "Forest Completion Permutation" (FCP). The significance of the FCP is that it identifies precisely the arcs that are added to F – they are $(r_i, a_{\pi(i)})$, $i = 1, \dots, m$. In our examples throughout the paper, we shall specify the second line in the two line representation of π – the list of images $(\pi(1), \dots, \pi(m))$.

In Fig. 1 we give an example of the FCA with $k = 9$, $m = 3$, $n = 2$. The trees T_1, T_2, T_3, with $r_1 = 2$, $r_2 = 4$, $r_3 = 7$, are given in the box at the top left; the trees T_4, T_5, with $s_1 = 6$, $s_2 = 8$, are given in

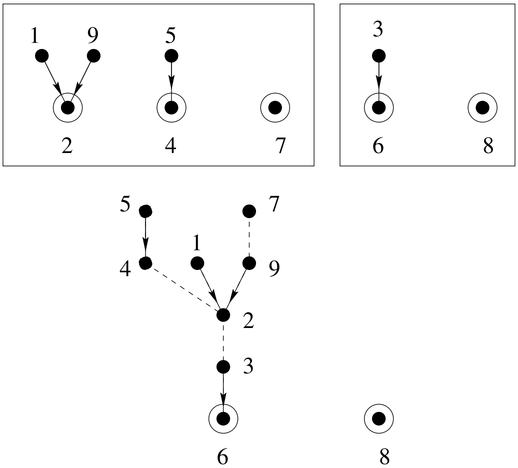

Fig. 1. A rooted forest and subforest.

the box at the top right. Then, corresponding to the triple $a=(9,2,3)$, we construct the forest at the bottom of Fig. 1. The corresponding FCP is $(3,2,1)$.

In analyzing this mapping, it is convenient to use the term “safe” to describe a vertex in a component of a forest rooted at one of the vertices $s_1,\dots,s_n$. Thus, initially, b_m is safe. It is trivial to prove by induction that, after stage i, for $i=1,\dots,m-1$, G is a forest with root vertices $r_{i+1},\dots,r_m,s_1,\dots,s_n$, and that b_m is safe for G (which implies that $r_{i+1},\dots,r_m$ are in different components of G from b_m). Thus, at stage m, b_m is indeed in a different component of G from r_m, so we successfully add the final arc from r_m to b_m, to obtain a terminating forest G rooted at $s_1,\dots,s_n$ (this explains our use of “safe” – for extending our forest, it is always safe to add the new arc directed to b_m). This proves that the FCA does indeed produce a rooted forest on vertex-set $[k]$ with root vertices $s_1,\dots,s_n$ that contains F as a subforest.

To prove that the FCA is a bijection, we describe its inverse. Suppose that we are given a rooted forest H on vertex-set $[k]$, with root vertices $s_1,\dots,s_n$, and that we wish to remove the arcs directed from r_i to c_i, $i=1,\dots,m$, where $r_1,\dots,r_m$ are distinct non-root vertices, with the convention that $r_1<\dots<r_m$. We proceed iteratively, through stages $m,\dots,1,0$. At every stage we have a subforest G of H, a permutation σ of $[m]$, and a sequence $b=(b_1,\dots,b_m)$ in $[k]^m$. Initially, at stage m, we have $G=H$, σ is the identity permutation, and $b=c$. Then, for $i=m-1,\dots,0$, remove the arc from r_{i+1} to b_{i+1} in G (to get the new G), and:

- if b_m is safe for (the new) G, leave σ and b unchanged;
- otherwise (so b_m is not safe for G), switch σ_{i+1} and σ_m in σ, and switch b_{i+1} and b_m in b.

We claim that the m-tuple corresponding to the forest H is the terminating m-tuple b, so that this mapping uniquely reverses the FCA. In fact, it is easy to establish that this mapping uniquely reverses the FCA stage by stage, since it is trivial to prove by induction that the values of b and G after stage i of the above mapping are exactly the same as b and G after stage i of the FCA. It is also easy to prove that the FCP is given by σ^{-1} for the terminating σ.

The result follows, since the FCA is a bijection between the required sets. □

Of course, it is an immediate enumerative consequence of Theorem 2.1 that there are $k^{m-1}|S|$ rooted forests on vertex-set $[k]$ with root vertices $s_1,\dots,s_n$ that contain F as a subforest. In the special case $m+n=k$ (so that F has no edges) this gives the classical result that there are $k^{k-n-1}n$ rooted forests on vertex-set $[k]$, with a prescribed set of n root vertices (see, *e.g.*, [13, p. 25]).

3. The combinatorial model

3.1. Paired surjections

The combinatorial model for our proof of Theorem 1.2 is based on a *paired surjection*, which has the following definition.

Definition 3.1. For $p, q, s, k \geqslant 1$, with p, q, s of the same odd–even parity, let $\mathcal{B}^{(s)}_{p,q,k}$ be the set of ordered pairs (μ, ϕ), where $\mu \in \mathcal{P}_{p,q}$ has s mixed pairs, and ϕ is a surjection from $[p] \cup [q]'$ onto $[k]$, satisfying the condition

$$\phi\big(\mu(i)\big) = \phi\big(\gamma_{p,q}(i)\big) \quad \text{for all } i \in [p] \cup [q]'. \tag{1}$$

Such an ordered pair (μ, ϕ) is called a *paired surjection*. Let $b^{(s)}_{p,q,k} = |\mathcal{B}^{(s)}_{p,q,k}|$.

In the following result, the generating series $A^{(s)}_{p,q}(x)$ evaluated in Theorem 1.2 is expressed in terms of the numbers $b^{(s)}_{p,q,k}$ of paired surjections. Paired surjections are closely related to shift-symmetric partitions, that arose in Goulden and Nica [3]. Indeed, the proof of the following result is identical to the proof of Proposition 1.3 in Goulden and Nica [3], and is hence omitted.

Proposition 3.2. *For $p, q, s \geqslant 1$, with p, q, s of the same odd–even parity, we have*

$$A^{(s)}_{p,q}(x) = \sum_{k \geqslant 1} b^{(s)}_{p,q,k} \binom{x}{k}.$$

We consider $(\mu, \phi) \in \mathcal{B}^{(s)}_{p,q,k}$, and construct various objects associated with (μ, ϕ). First let $C_i = \phi^{-1}(i) \cap [p]$ and $C'_i = \phi^{-1}(i) \cap [q]'$, for $i \in [k]$. Let $D = \{i\colon\ |C_i| \geqslant 1\}$, and $D' = \{i\colon\ |C'_i| \geqslant 1\}$, and let $m_i = \max C_i$, $i \in D$, and $m'_i = \max C'_i$, $i \in D'$. Suppose that 1 is contained in C_a, and that $1'$ is contained in C'_b. Define $\psi : D \setminus \{a\} \to D$ by $\psi(i) = j$ when $\phi(\mu(m_i)) = j$, and $\psi' : D' \setminus \{b\} \to D'$ by $\psi'(i) = j$ when $\phi(\mu(m'_i)) = j$.

Now, if $\psi(i) = j$, then (interpreting 1 as $p + 1$) condition (1) means that $m_i + 1 \in C_j$, so we have $m_i < m_j$. This implies that the functional digraph of ψ (the directed graph on vertex-set D with an arc directed from i to $\psi(i)$ for each $i \in D \setminus \{a\}$) is actually a tree, in which all arcs are directed towards vertex a (which we consider as the root of this tree). We denote this rooted tree by T. Similarly, the functional digraph of ψ', on vertex-set D', is also a tree, with all arcs directed towards vertex b (which we consider as the root of this tree). We denote this rooted tree by T'.

One condition that the paired surjection (μ, ϕ) satisfies is that the number of mixed pairs containing an element of C_i is equal to the number of mixed pairs containing an element of C'_i for all $i \in [k]$. (For the reason that this necessary condition arises, see the discussion of “unique label recovery” in the next section.) We call this the *balance* condition for (μ, ϕ). The fact that ϕ is a surjection is equivalent to $|C_i| + |C'_i| \geqslant 1$, for $i \in [k]$, and we call this the *nonempty* condition for (μ, ϕ). The fact, established above, that the graphs of ψ and ψ' are trees is called the *tree* condition for (μ, ϕ).

3.2. A graphical model

Now we consider a graphical representation for the paired surjection (μ, ϕ), called its *labelled paired array*. This is an array of cells, arranged in k columns, indexed $1, \ldots, k$ from left to right, and two rows. In column i of row 1, place an ordered list of $|C_i|$ vertices, labelled by the elements of C_i from left to right; in column i of row 2, place an ordered list of $|C'_i|$ vertices, labelled by the elements of C'_i from left to right. For each pair of μ draw an edge between the vertices whose labels are given by the pair.

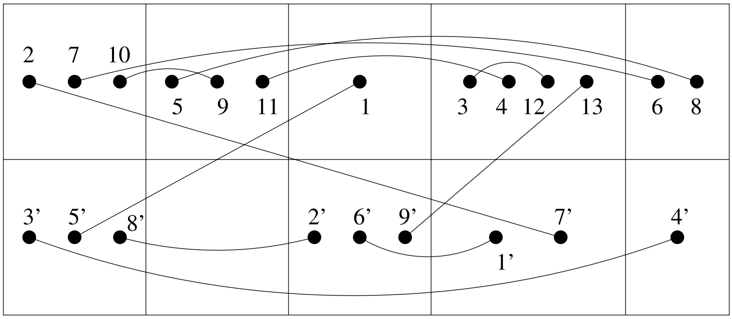

Fig. 2. A labelled paired array.

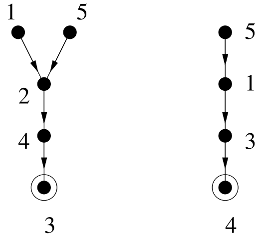

Fig. 3. Two rooted trees.

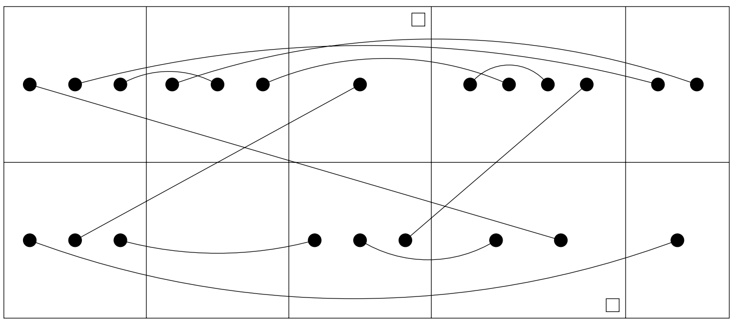

Fig. 4. A paired array.

For example, when $p=13, q=9, s=3, k=5$, consider $(\mu,\phi)\in\mathcal{B}^{(s)}_{p,q,k}$, given by $\mu=\{\{1,5'\},\{2,7'\},\{3,12\},\{4,11\},\{5,8\},\{6,7\},\{9,10\},\{13,9'\},\{1',6'\},\{2',8'\},\{3',4'\}\}$, and $\phi^{-1}(1)=\{2,7,10,3',5',8'\}$, $\phi^{-1}(2)=\{5,9,11\}$, $\phi^{-1}(3)=\{1,2',6',9'\}$, $\phi^{-1}(4)=\{3,4,12,13,1',7'\}$, $\phi^{-1}(5)=\{6,8,4'\}$. The corresponding labelled paired array is given in Fig. 2, and the trees T and T' are given in Fig. 3.

Now suppose that we mark the cells in column a of row 1 and in column b of row 2 (by placing a small box in the top right-hand corner of the marked cell in row 1, and in the bottom right-hand corner of the marked cell in row 2), and remove the labels from all vertices – call the resulting object the *paired array* of (μ,ϕ). The ordered list of vertices in each cell is now to be interpreted as a generic totally ordered set, with the given left to right order, and the pairing μ now acts on these ordered sets in the obvious way. For example, the paired array determined from the labelled paired array displayed in Fig. 2 is given in Fig. 4.

What information have we lost when the labels are removed? The answer, perhaps surprisingly, is that no information is lost, since we have *unique label recovery* by applying condition (1) iteratively, as follows: for the first row, place label 1 on the leftmost vertex in the marked cell of row 1; for each i from 2 to p, place label i on the leftmost unlabelled vertex in column $\phi(\mu(i-1))$ of row 1. The same process applied to the second row will place labels $1'$ to q' on the vertices in row 2. (The reader can apply this to the paired array in Fig. 4, to check that indeed the labelled paired array in Fig. 2 is recovered in this way.) The proof that this process always works for a paired array satisfying the balance, nonempty and tree conditions (and the proof that these conditions are necessary for this process to work) requires only a slight modification of the results in Section 3 of [3], and is not given here (note that neither the functions ψ and ψ', nor the trees T and T', depend on the labels of the vertices, and the number of mixed pairs incident with the vertices in each cell of the paired array also does not depend on the labels, so the balance, nonempty and tree conditions can be checked on the paired array alone).

3.3. *Paired arrays*

This motivates us to define a *paired array* in the abstract (and not as obtained by removing the labels from a labelled paired array), and in fact to extend it to a more general class of objects, in the following definition.

Definition 3.3. For $p, q, s, k \geqslant 1$, with p, q, s of the same odd–even parity, we define $\mathcal{PA}^{(s)}_{p,q,k}$ to be the set of arrays of cells, arranged in k columns and 2 rows, subject to the following conditions:

- Each cell contains an ordered list of vertices, so that there is a total of p vertices in the first row, and q vertices in the second row. The vertices are paired (in the language of graph theory, there is a perfect matching on the vertices), so that s pairs join a vertex in the first row to a vertex in the second row (these are the *mixed pairs*). The number of mixed pairs containing a vertex in column i of row 1 is equal to the number of mixed pairs containing a vertex in column i of row 2, for all $i = 1, \dots, k$ (this is called the *balance* condition).
- There is at least one marked (with a small box) cell in row 1, and we denote the set of such columns by R. There is at least one marked (with a small box) cell in row 2, and we denote the set of such columns by R'. There is at least one vertex in every column that is not contained in $R \cup R'$ (this is called the *nonempty* condition).
- Denote the set of columns in which there is at least one vertex in row 1 by D, and the set of columns in which there is at least one vertex in row 2 by D'. Define the function $\psi : D \setminus R \to D$ as follows: if the rightmost vertex in column i of row 1 is paired with a vertex in column j, then $\psi(i) = j$. Similarly, define $\psi' : D' \setminus R' \to D'$ as follows: if the rightmost vertex in column i of row 2 is paired with a vertex in column j, then $\psi'(i) = j$. The functional digraph of ψ is a forest with $|R|$ components (called the *rightmost forest for row* 1); each component is a tree in which all edges are directed towards an element of R (and this is called the *root* of that tree). The functional digraph of ψ' is a forest with $|R'|$ components (called the *rightmost forest for row* 2); each component is a tree in which all edges are directed towards an element of R' (and this is called the *root* of that tree). Together, these specify the *forest* condition.

The elements of $\mathcal{PA}^{(s)}_{p,q,k}$ are called *paired arrays*. A paired array is defined to be *canonical* if $|R| = |R'| = 1$. Define $\mathcal{C}^{(s)}_{p,q,k}$ to be the set of canonical paired arrays in $\mathcal{PA}^{(s)}_{p,q,k}$, and $c^{(s)}_{p,q,k} = |\mathcal{C}^{(s)}_{p,q,k}|$.

The uniqueness of label recovery described in the previous section proves that there is a bijection (via labelled paired arrays) between the set $\mathcal{B}^{(s)}_{p,q,k}$ of paired surjections and the set $\mathcal{C}^{(s)}_{p,q,k}$ of canonical paired arrays, so we have

$$b^{(s)}_{p,q,k} = c^{(s)}_{p,q,k}. \tag{2}$$

(It is straightforward to verify that the conditions for canonical paired arrays imply that *every* column is nonempty.) In this paper we shall determine $b^{(s)}_{p,q,k}$, and hence the generating series $A^{(s)}_{p,q}(x)$ via

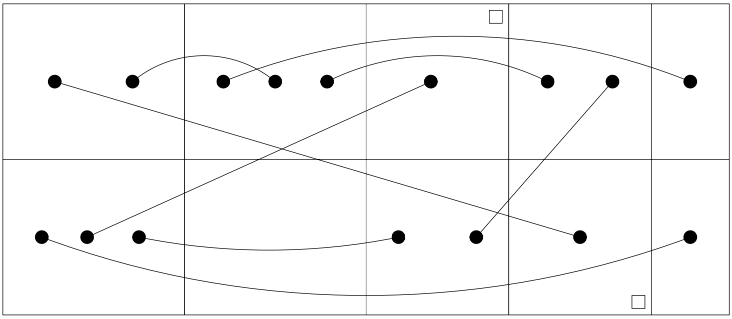

Fig. 5. A minimal paired array.

Proposition 3.2, by giving a combinatorial reduction for canonical paired arrays, thus directly determining $c^{(s)}_{p,q,k}$.

4. Removing non-mixed pairs and vertical paired arrays

A *redundant* pair in a paired array is a vertex pair that is *not* mixed, and does *not* contain a vertex that is rightmost in an unmarked cell. A *minimal* paired array is a paired array without redundant pairs. We define $\mathcal{M}^{(s)}_{p,q,k}$ to be the set of minimal, canonical paired arrays in $\mathcal{PA}^{(s)}_{p,q,k}$, and $m^{(s)}_{p,q,k} = |\mathcal{M}^{(s)}_{p,q,k}|$. A *vertical* paired array is a paired array in which all pairs are mixed. We define $\mathcal{V}^{(s)}_{k,i,j}$ to be the set of vertical paired arrays in $\mathcal{PA}^{(s)}_{s,s,k}$, in which there are $i+1$ marked cells in row 1 and $j+1$ marked cells in row 2, for $i, j \geqslant 0$, and let $v^{(s)}_{k,i,j} = |\mathcal{V}^{(s)}_{k,i,j}|$.

In this section, we remove non-mixed pairs from a canonical paired array in two stages, thus reducing the enumeration of canonical paired arrays to the enumeration of vertical paired arrays. The next result gives the first stage of this reduction, in which the problem is reduced to the enumeration of minimal, canonical paired arrays.

Theorem 4.1. *For $p, q, s, k \geqslant 1$, with p, q, s of the same odd–even parity, we have*

$$c^{(s)}_{p,q,k} = \sum_{i,j \geqslant 0} \binom{p}{2i}(2i-1)!!\binom{q}{2j}(2j-1)!!\, m^{(s)}_{p-2i,q-2j,k}.$$

Proof. For the proof, it is convenient to introduce some notation. A *partial* pairing on $[p]$ is a pairing on a set $\alpha \subseteq [p]$ of even cardinality. If $|\alpha| = 2i$, then we also call it an i-partial pairing. For each of these partial pairings μ, we call α the *support*, and denote this by $\mathrm{supp}(\mu) = \alpha$. Let $\mathcal{R}_{p,i}$ be the set of i-partial pairings on $[p]$. Similarly, let $\mathcal{R}'_{q,i}$ be the set of i-partial pairings on $[q]'$.

Consider an arbitrary $\alpha \in \mathcal{C}^{(s)}_{p,q,k}$. We now describe a construction for three objects, μ_1, μ_2 and β, obtained from α. We begin by attaching the numbers $1, \dots, p+1$ to the vertices and the small box in row 1 of α, from left to right (under the interpretation that all vertices in column i are to the left of all vertices in column j for $i < j$, and that the small box representing a marking is rightmost in its cell). Let μ_1 be the partial pairing consisting of pairs of numbers attached to the redundant pairs in row 1 of α. We follow the analogous procedure for row 2: we attach primed numbers $1', \dots, (q+1)'$ to the vertices and small box in row 2, and let μ_2 be the partial pairing consisting of the pairs of (primed) numbers attached to the redundant pairs in row 2 of α. Third, we remove all redundant pairs (both vertices and edges) from α, to get the paired array β, with the same marked cells as α. The vertices in each cell of β have the same relative order as they did in α. For example, if α is the paired array in Fig. 4, then we have $\mu_1 = \{\{2, 13\}, \{9, 11\}\}$, $\mu_2 = \{\{5', 7'\}\}$, and β is given in Fig. 5.

Now, the only vertex that can be numbered $p+1$ in row 1 is the rightmost vertex of the rightmost nonempty cell in row 1 (if this cell is not marked), but this vertex cannot appear in a redundant pair since it is rightmost in an unmarked cell. This implies that the numbers on redundant pairs in row 1 all fall in the range $1, \dots, p$, and so μ_1 is a partial pairing on $[p]$. Similarly, μ_2 is a partial pairing on $[q]'$. Also, since the redundant pairs that were removed in the construction do not involve the rightmost vertex in any nonempty cell, β has the same rightmost functions ψ and ψ' as α, and the same mixed pairs as α, so it must satisfy the balance, nonempty and forest conditions, which implies that β is a minimal paired array. Thus we have a mapping

$$\xi : \mathcal{C}^{(s)}_{p,q,k} \to \bigcup_{i,j \geqslant 0} \mathcal{R}_{p,i} \times \mathcal{R}'_{q,j} \times \mathcal{M}^{(s)}_{p-2i,q-2j,k} : \alpha \mapsto (\mu_1, \mu_2, \beta).$$

We now prove that ξ is a bijection. It is sufficient to describe the inverse mapping, so that we can uniquely recover α from an arbitrary triple $(\mu_1, \mu_2, \beta) \in \bigcup_{i,j \geqslant 0} \mathcal{R}_{p,i} \times \mathcal{R}'_{q,j} \times \mathcal{M}^{(s)}_{p-2i,q-2j,k}$. Given (μ_1, μ_2, β), let $\sigma_i = \operatorname{supp}(\mu_i)$, $i = 1, 2$, and $\rho_1 = [p+1] \setminus \sigma_1$, $\rho_2 = [q+1]' \setminus \sigma_2$. Number the vertices and small box in row 1 of β with the elements of ρ_1. Then insert vertices numbered with the elements of σ_1, so that the numbers on all vertices and the small box in row 1 increase from left to right, and so that the vertex numbered with $l \in \sigma_1$ is placed in the same cell as either the vertex or small box numbered with $\min\{i \in \rho_1\colon\ i > l\}$. Inserting vertices numbered from σ_2 in row 2 using an analogous process, pairing the inserted vertices with μ_1 and μ_2, and removing the numbers, we arrive at a paired array α. It is straightforward to check that α satisfies the balance, nonempty, and tree conditions, and that the process described above reverses the numbering scheme used in the mapping ξ. Thus, we have described the mapping ξ^{-1}, and our proof that ξ is a bijection is complete. The result follows from the easily established facts that $|\mathcal{R}_{p,i}| = \binom{p}{2i}(2i-1)!!$, $|\mathcal{R}'_{q,j}| = \binom{q}{2j}(2j-1)!!$. $\square$

The next result gives the second stage of our reduction. We remove non-mixed pairs from a minimal, canonical paired array, and thus show that the enumeration of minimal, canonical paired arrays can be reduced to the enumeration of vertical paired arrays. We use the following notation. For a finite set X, let $\mathcal{L}_{X,i}$ denote the set of i-tuples consisting of i distinct elements of X. Thus $|\mathcal{L}_{X,i}| = (x)_i$, where $x = |X|$ and $(x)_i$ is the *falling factorial*: for positive integers i, $(x)_i = x(x-1)\cdots(x-i+1)$; for $i = 0$, $(x)_i = 1$; otherwise $(x)_i = 0$. In the proof, we use the Forest Completion Algorithm (FCA) and the Forest Completion Permutation (FCP), which were described in Section 2.

Theorem 4.2. *For $p, q, s, k \geqslant 1$, with p, q, s of the same odd–even parity, we have*

$$m^{(s)}_{p,q,k} = (p)_{\frac{1}{2}(p-s)}(q)_{\frac{1}{2}(q-s)}v^{(s)}_{k,\frac{1}{2}(p-s),\frac{1}{2}(q-s)}.$$

Proof. In the proof, we use the notation $i = \frac{1}{2}(p-s)$ and $j = \frac{1}{2}(q-s)$. Note that every element of $\mathcal{M}^{(s)}_{p,q,k}$ has exactly i and j non-mixed pairs in the top and bottom rows, respectively. Taking an arbitrary $\alpha \in \mathcal{M}^{(s)}_{p,q,k}$, we now describe a mapping that is initially identical to that used in the proof of Theorem 4.1. We attach the numbers $1, \dots, p+1$ to the vertices and the small box in row 1, using the same left to right convention as in the proof of Theorem 4.1. Let the pairs of numbers attached to the non-mixed pairs in row 1 be denoted by $(u_1, v_1), \dots, (u_i, v_i)$, where $u_1, \dots, u_i$ are attached to the rightmost vertices in these pairs (with $u_1 < \cdots < u_i$), and $v_1, \dots, v_i$ are attached to the other (not rightmost in their cell) vertices in these pairs.

Suppose that the marked cell in row 1 is in column m, and that the rightmost tree for row 1 of α is T, so T is rooted at vertex m. Now run the inverse of the FCA on T, to remove the arcs directed from $\mathsf{c}(u_\ell)$ to $\mathsf{c}(v_\ell)$, $\ell = 1, \dots, i$ (here, $\mathsf{c}(\ell)$ denotes the column in which the number ℓ appears), and let ρ be the corresponding FCP. Let $\kappa_1 = (v_{\rho^{-1}(1)}, \dots, v_{\rho^{-1}(i)})$.

We follow the analogous procedure for row 2: we attach primed numbers $1', \dots, (q+1)'$ to the vertices and small box in row 2. Let the pairs of numbers attached to the non-mixed pairs in row 2 be denoted by $(x'_1, y'_1), \dots, (x'_j, y'_j)$, where $x'_1, \dots, x'_j$ are attached to the rightmost vertices in these pairs

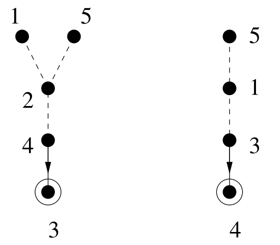

Fig. 6. The rightmost trees for Fig. 5.

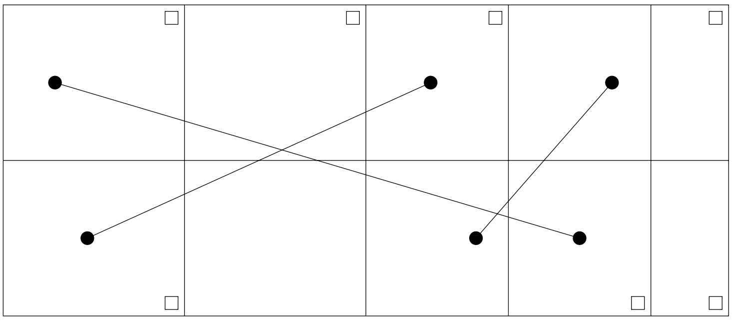

Fig. 7. The vertical paired array for Fig. 5.

(with $x_1 < \cdots < x_j$), and $y'_1, \dots, y'_j$ are attached to the other (not rightmost in their cell) vertices in these pairs.

Suppose that the marked cell in row 2 is in column n, and that the rightmost tree for row 2 of α is T', so T' is rooted at vertex n. Now run the inverse of the FCA on T', to remove the arcs directed from $\mathsf{c}(x'_\ell)$ to $\mathsf{c}(y'_\ell)$, $\ell = 1, \dots, j$, and let τ be the corresponding FCP. Let $\kappa_2 = (y'_{\tau^{-1}(1)}, \dots, y'_{\tau^{-1}(j)})$.

Finally, we mark the cells in row 1 containing $u_1, \dots, u_i$ (in addition to the existing marked cell in column m), and the cells in row 2 containing $x'_1, \dots, x'_j$ (in addition to the existing marked cell in column n), and then remove all non-mixed pairs (both vertices and edges) from α, to get the vertical paired array β. The vertices in each cell of β have the same relative order as they did in α.

For example, suppose that α is the paired array in Fig. 5. Then we have $i = 3$, with $u_1 = 2$, $u_2 = 5$, $u_3 = 10$, $v_1 = 4$, $v_2 = 8$, $v_3 = 3$, and $j = 2$, with $x'_1 = 3'$, $x_2 = 8'$, $y'_1 = 4'$, $y'_2 = 1'$. We have $m = 3$, $n = 4$, and the trees T and T' from Fig. 3 are repeated in Fig. 6. When we run the inverse of the FCA on T to remove the arcs $(1,2)$, $(2,4)$ and $(5,2)$ (for which we use dashed lines in Fig. 6), we obtain $(1,3,2)$ as the FCP, which gives $\kappa_1 = (4,3,8)$. When we run the inverse of the FCA on T' to remove the arcs $(1,3)$ and $(5,1)$ (for which we use dashed lines in Fig. 6), we obtain $(2,1)$ as the FCP, which gives $\kappa_2 = (1', 4')$. The vertical paired array β in this example is given in Fig. 7.

Now, for the same reasons as in the proof of Theorem 4.1, we have $1 \leqslant v_\ell \leqslant p$ for $\ell = 1, \dots, i$, so $\kappa_1 \in \mathcal{L}_{[p],i}$, and similarly, $\kappa_2 \in \mathcal{L}_{[q]',j}$. Thus we have a mapping

$$\zeta : \mathcal{M}^{(s)}_{p,q,k} \to \mathcal{L}_{[p],i} \times \mathcal{L}_{[q]',j} \times \mathcal{V}^{(s)}_{k,i,j} : \alpha \mapsto (\kappa_1, \kappa_2, \beta).$$

In fact, ζ is a bijection. Before proving this, for $\zeta(\alpha) = (\kappa_1, \kappa_2, \beta)$, we first note the key dependencies between α and $(\kappa_1, \kappa_2, \beta)$ that follow from the FCA: let F and F' denote the rightmost forests of β for rows 1 and 2, respectively. Then, in α with numbers attached as in the construction above, the column containing the vertex whose number is the last entry of κ_1 is contained in the component of F rooted at m. Similarly, the column containing the vertex whose number is the last entry of κ_2 is contained in the component of F' rooted at n. For example, for $\alpha, \kappa_1, \kappa_2, \beta$ given in Figs. 5–7, the last entries of κ_1 and κ_2 are 8 and $4'$, respectively, corresponding to vertices of α in columns 4 and 3, respectively. Now, the forests F and F' in this case are obtained from the trees T and T', respectively, by removing the dashed edges in Fig. 6. Then, indeed, vertex 4 is contained in the component of F rooted at $m=3$, and vertex 3 is contained in the component of F' rooted at $n=4$.

We now prove that ζ is a bijection by describing the inverse mapping, so that we can uniquely recover α from an arbitrary triple $(\kappa_1,\kappa_2,\beta)\in\mathcal{L}_{[p],i}\times\mathcal{L}_{[q]',j}\times\mathcal{V}^{(s)}_{k,i,j}$. Let F denote the rightmost forest of β for row 1. Let $\{\kappa_1\}$ denote the set consisting of the entries in κ_1, and let $p=s+2i$, $\delta_1=[p+1]\setminus\{\kappa_1\}$. Next (as a generalization of the procedure for the inverse of ξ described in the proof of Theorem 4.1), we number the vertices and small boxes in row 1 of β with the elements of δ_1. Then insert vertices numbered with the elements of $\{\kappa_1\}$, so that the numbers on all vertices and small boxes increase from left to right (with any small box regarded as the rightmost object in its cell) and so that the vertex numbered with $l\in\{\kappa_1\}$ is placed in the same cell as the object numbered with $\min\{t\in\delta_1\colon\ t>l\}$. Now suppose that $\kappa_1=(w_1,\ldots,w_i)$, and use the key dependency noted above: let the column containing the vertex numbered w_i be contained in the component of F rooted at vertex m. Let $u_1<\cdots<u_i$ denote the numbers attached to the small boxes that are not in column m (the columns containing $u_1,\ldots,u_i$ are the root vertices for the components of F not rooted at m). Now, apply the FCA on i-tuple $(\mathsf{c}(w_1),\ldots,\mathsf{c}(w_i))$, to give the tree T rooted at $m=\mathsf{c}(w_i)$, that contains the forest F as a subforest. Let ρ be the FCP. Finally, replace the small boxes numbered $u_1,\ldots,u_i$ by rightmost vertices (in the same cells) numbered $u_1,\ldots,u_i$, pair the vertex numbered u_ℓ with the vertex numbered $w_{\rho(\ell)}$, $\ell=1,\ldots,i$, and remove the numbers from row 1. Repeat the analogous process for row 2, and we arrive at a minimal paired array α. It is straightforward to check that the process described above reverses ζ. Thus, we have described ζ^{-1}, and our proof that ζ is a bijection is complete. The result follows immediately. □

At this stage, we have now reduced the enumeration of canonical paired arrays (and thus the evaluation of $A^{(s)}_{p,q}(x)$) to the enumeration of vertical paired arrays. In the next result, we summarize the results to this stage.

Corollary 4.3. *For $p,q,s,k\geqslant 1$, with p,q,s of the same odd–even parity, we have*

$$A^{(s)}_{p,q}(x)=\sum_{\substack{k\geqslant 1\\ i,j\geqslant 0}}\binom{x}{k}\frac{p!q!}{2^{i+j}i!j!(\frac{1}{2}(p+s)-i)!(\frac{1}{2}(q+s)-j)!}v^{(s)}_{k,\frac{1}{2}(p-s-2i),\frac{1}{2}(q-s-2j)}.$$

Proof. From Proposition 3.2, (2) and Theorems 4.1, 4.2, we obtain

$$\begin{aligned}A^{(s)}_{p,q}(x)&=\sum_{\substack{k\geqslant 1\\ i,j\geqslant 0}}\binom{x}{k}\binom{p}{2i}(2i-1)!!\binom{q}{2j}(2j-1)!!m^{(s)}_{p-2i,q-2j}\\&=\sum_{\substack{k\geqslant 1\\ i,j\geqslant 0}}\binom{x}{k}\binom{p}{2i}(2i-1)!!\binom{q}{2j}(2j-1)!!(p-2i)_{\frac{1}{2}(p-s-2i)}\\&\quad\times(q-2j)_{\frac{1}{2}(q-s-2j)}v^{(s)}_{k,\frac{1}{2}(p-s-2i),\frac{1}{2}(q-s-2j)},\end{aligned}$$

and the result follows by routine simplification. □

5. Enumeration of vertical paired arrays

In this section, we enumerate vertical paired arrays, in Theorem 5.2. First, we enumerate a full, vertical paired array, that arises as follows: Note that for every column of a vertical paired array, the cells in rows 1 and 2 have the same number of vertices, because of the balance condition. A *full*, vertical paired array is a vertical paired array with a positive number of vertices in every column. Let $\mathcal{F}^{(s)}_{k,i,j}$ be the set of full, vertical paired arrays in $\mathcal{V}^{(s)}_{k,i,j}$, and $f^{(s)}_{k,i,j}=|\mathcal{F}^{(s)}_{k,i,j}|$. In Theorem 5.1 we shall give an explicit construction for the elements of $\mathcal{F}^{(s)}_{k,i,j}$, and thus obtain an explicit formula for $f^{(s)}_{k,i,j}$. To help in the proof of this result, we first introduce some terminology and notation associated with an arbitrary $\alpha \in \mathcal{F}^{(s)}_{k,i,j}$. A vertex is said to be *dependent* if it is paired with the rightmost vertex of an unmarked cell (in the other row). If the rightmost vertex of an unmarked cell in row 1, column u is paired with the rightmost vertex of an unmarked cell in row 2, column v, then we call this a *shared pair* of α. In this case, the rightmost forest F for row 1 of α contains arc (u, v) and the rightmost forest F' for row 2 of α contains arc (v, u), and we call each of these a *shared arc*.

Now, canonically number the vertices in row 1 of α $1, \ldots, s$, from left to right, and number the vertices in row 2 of α $1', \ldots, s'$, from left to right. Let E be the subforest of F with only the shared arcs of F. Suppose that F has $n \geqslant 1$ non-shared arcs, corresponding to pairs $(x_1, y'_1), \ldots, (x_n, y'_n)$, where $x_1 < \cdots < x_n$, and $x_1, \ldots, x_n$ are rightmost vertices in their (unmarked) cells. Run the inverse of the FCA on the forest F, to obtain the subforest E by removing the non-shared arcs directed from $\mathsf{c}(x_\ell)$ to $\mathsf{c}(y'_\ell)$, $\ell = 1, \ldots, n$, and let τ be the corresponding FCP. Define $a' = y'_{\tau^{-1}(n)}$. If all arcs of F are shared, then let a' be the vertex in row 2 that is paired with the rightmost non-dependent vertex in row 1 (we call this the *non-FCA option*). In both cases, define $A = \mathsf{c}(a')$, and let $\rho_0 = A, \rho_1, \ldots, \rho_l$ be the vertices on the unique directed path in E from vertex A to the root vertex (ρ_l) of the component of E containing A. Thus $l \geqslant 0$, and the cell in row 1, column ρ_l is marked, and the cell in row 2, column ρ_ℓ is not marked, $\ell = 1, \ldots, l$. Also, the cell in row 1, column ρ_ℓ is not marked, $\ell = 0, \ldots, l-1$. Now define E' to be the subforest of F' containing the arcs $(\rho_\ell, \rho_{\ell+1})$, $\ell = 0, \ldots, l-1$. Suppose that F' has m arcs that are not in E', corresponding to pairs $(w'_1, z_1), \ldots, (w'_m, z_m)$, where $w_1 < \cdots < w_m$, and $w'_1, \ldots, w'_m$ are rightmost vertices in their (unmarked) cells. Run the inverse of the FCA on the forest F', to obtain the subforest E' by removing the arcs directed from $\mathsf{c}(w'_\ell)$ to $\mathsf{c}(z_\ell)$, $\ell = 1, \ldots, m$, and let κ be the corresponding FCP. Define $b = z_{\kappa^{-1}(m)}$. If $F' = E'$, let b be the vertex in row 1 that is paired with the rightmost non-dependent vertex in row 2 (again, we call this the *non-FCA option*). Let $\rho = (\rho_0, \rho_1, \ldots, \rho_l)$, which we call the *tail* of α. The tail length is l. We say that b is *in the tail* when the column containing vertex b is one of $\rho_0, \rho_1, \ldots, \rho_l$. The *type* of α is given by (l, ρ, a', b). If the cells in rows 1 and 2 of column ℓ in α have λ_ℓ vertices, $\ell = 1, \ldots, k$, then we say that α has *shape* $\lambda = (\lambda_1, \ldots, \lambda_k)$. Note that $\lambda_1 + \cdots + \lambda_k = s$, and that λ_ℓ is positive for all $\ell = 1, \ldots, k$, so λ is a *composition* of s with k parts.

For example, suppose that α is the full, vertical paired array given at the top of Fig. 8, with $s = 10$, $k = 7$, $i = 1$, $j = 0$, shape $(1, 2, 1, 1, 2, 2, 1)$, and rightmost forests F and F' given at the bottom of Fig. 8. The lines joining the pairs in α are of various types (and the same type of line is used in the rightmost forests when the pair corresponds to an edge in one or other of these forests): a thick solid line indicates a shared pair in the tail, a thick dashed line a shared pair not in the tail, a thick dashed and dotted line a pair contributing to F only, a thin solid line a pair contributing to F' only, and a thin dashed line a pair that contributes to neither of F, F'. When we run the inverse of the FCA to remove the arcs $(6, 2)$ and $(7, 6)$ from F, we obtain the ordered pair $(6, 2)$, and hence obtain $a' = 2'$ as indicated in Fig. 8, contained in column 2. Thus the tail is $\rho = (2, 1)$, of length $l = 1$. When we run the inverse of the FCA to remove the arcs $(2, 3)$, $(3, 4)$, $(4, 5)$, $(5, 6)$ and $(7, 5)$ from F', we obtain the 5-tuple $(3, 4, 5, 5, 6)$, and hence obtain $b = 8$ as indicated in Fig. 8, contained in column 6. Thus we conclude that α has type $(1, (2, 1), 2', 8)$.

We are now ready to enumerate full, vertical paired arrays. Once again, in the proof we use the FCA and FCP from Section 2.

Theorem 5.1. *For $i, j \geqslant 0$, $k, s \geqslant 1$, we have*

$$f^{(s)}_{k,i,j} = s! \sum_{l=0}^{k-1} \binom{s-1-l}{k-1-l} \binom{k-1-l}{i} \binom{k-1-l}{j}.$$

Proof. Each paired array in $\mathcal{F}^{(s)}_{k,i,j}$ has a unique type and shape, and we can uniquely construct those of given type and shape as follows. For the given shape, we begin with a 2 by k array, with each cell containing an ordered set of vertices of prescribed size. Then we pair these vertices (all are mixed

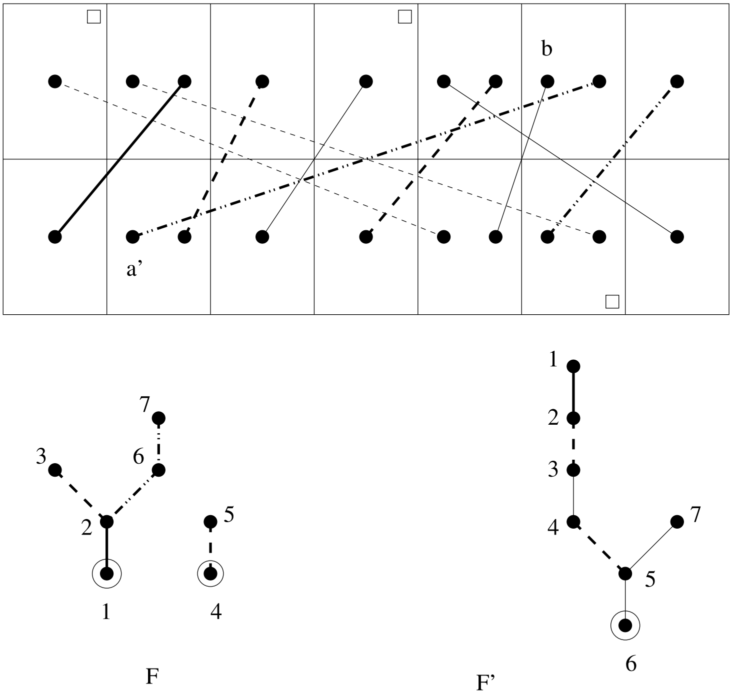

Fig. 8. A vertical paired array with rightmost forests.

pairs) in all possible ways for the given type in four stages. First, we pair the rightmost vertices as prescribed by the tail. Second, we use the FCA to pair the rightmost vertices in row 2 (note that b is safe, by construction). Third, we use the FCA to pair the unpaired rightmost vertices in row 1 (note that a' is safe, by construction). Fourth, we pair the remaining vertices arbitrarily. (In addition, along the way, we have to choose the marked cells in a consistent fashion.) In this way, for each composition λ of s with k parts and $l \geqslant 0$, we enumerate elements of $\mathcal{F}_{k,i,j}^{(s)}$ with shape λ and tail of length l. There are three cases:

Case 1 *(Vertex a' is not rightmost in its cell).* There are $s-k$ choices for a', which then fixes ρ_0. There are then $(k-1)_l$ choices for $\rho_1, \dots, \rho_l$, and then l of the pairs are determined. This leaves $s-l$ choices for b (any vertex not yet paired in row 1). Now mark the cell in row 1, column ρ_l. Also, if b is in the tail, mark the cell in row 2, column ρ_0, or if b is not in the tail, mark the cell in row 2 of the column that contains b (b is in row 1). Choose, from the $k-l-1$ cells not in the tail or already marked, i cells to mark in the top row, and j cells to mark in the bottom row. Now pair the unpaired rightmost vertices of unmarked cells in row 2 with vertices in row 1 to satisfy the forest condition, using vertex b as the safe position. There are $(s-l-1)_{k-j-l-2}$ possible choices for this, from the FCA. Suppose that there are n unmarked cells in row 1 whose rightmost vertices are not yet paired. Then pair these with vertices in row 2 to satisfy the forest condition, using vertex a' as the safe position. There are $(s-k+j)_{n-1}$ possible choices for this, from the FCA. Finally, there are $(s-k+j-n+1)!$ ways to pair the remaining vertices, arbitrarily. But $(s-k+j)_{n-1} \cdot (s-k+j-n+1)! = (s-k+j)!$, and we conclude that the number of elements in $\mathcal{F}_{k,i,j}^{(s)}$ with shape λ and tail of length l in this case is

$$(s-k)(k-1)_l(s-l)(s-l-1)_{k-j-l-2}(s-k+j)!\binom{k-l-1}{i}\binom{k-l-1}{j}. \tag{3}$$

(It is straightforward to check that these cardinalities are correct when we use the non-FCA options also.)

Case 2 *(Vertex a' is rightmost in its cell, and b is in the tail).* The number of choices for b and ρ is $s\binom{k-1}{l}(l+1)! - l(k)_{l+1}$. Then a', and l of the pairs, are uniquely determined. The cell in row 1, column ρ_l is marked. The cell in row 2, column ρ_0 is marked for two reasons – because b is in the tail, and because the cell contains a', rightmost, so that a' will be paired with a non-dependent vertex in row 1. The rest of the enumeration proceeds as in Case 1, and we conclude that the number of elements in $\mathcal{F}^{(s)}_{k,i,j}$ with shape λ and tail of length l in this case is

$$\left(s\binom{k-1}{l}(l+1)! - l(k)_{l+1}\right)(s-l-1)_{k-j-l-2}(s-k+j)!\binom{k-l-1}{i}\binom{k-l-1}{j}. \tag{4}$$

Case 3 *(Vertex a' is rightmost in its cell, and b is not in the tail).* The number of choices for b and ρ is $s(k-1)_{l+1}$, and then a' together with l of the pairs are uniquely determined. In this case three different cells must now be marked: in row 1, column ρ_l; in row 2, column ρ_0; in row 2, the column that contains b. There are then $\binom{k-l-1}{i}\binom{k-l-2}{j-1}$ ways to choose which other cells are marked, and the rest of the enumeration proceeds as in Cases 1 and 2. We conclude that the number of elements in $\mathcal{F}^{(s)}_{k,i,j}$ with shape λ and tail of length l in this case is

$$s(k-1)_{l+1}(s-l-1)_{k-j-l-2}(s-k+j)!\binom{k-l-1}{i}\binom{k-l-2}{j-1}. \tag{5}$$

Adding (3), (4), and (5), and simplifying, we obtain that the total number of elements in $\mathcal{F}^{(s)}_{k,i,j}$ with shape λ and tail of length l is

$$s(k-1)_l(s-l-1)!\binom{k-l-1}{i}\binom{k-l-1}{j},$$

and the result follows, by summing over $l \geqslant 0$ and multiplying by $\binom{s-1}{k-1}$, the number of choices for λ. □

In the next result, we give an explicit enumeration for vertical paired arrays, by applying Theorem 5.1. The proof is quite technical, involving generating functions and a hypergeometric summation.

Theorem 5.2. *For $i, j \geqslant 0$, $k, s \geqslant 1$, we have*

$$v^{(s)}_{k,i,j} = \frac{(s+i)!(s+j)!}{(s+i+j)!}\binom{s+i+j}{k-1}\left[\binom{k-1}{i}\binom{k-1}{j} - \binom{k-1}{s+i}\binom{k-1}{s+j}\right].$$

Proof. If a column in a vertical paired array has no vertices, then at least one of the cells in rows 1 and 2 of that column must be marked, from the nonempty condition. Thus, suppose that a vertical paired array with s mixed pairs and k columns, has $k-m$ columns with no vertices and m with a positive number of vertices (the same number in both rows of such columns). Of the $k-m$ columns with no vertices, suppose that a are marked in row 1 only, b are marked in row 2 only, and that $c-m$ are marked in both row 1 and 2 (we use this parameterization for convenience in determining the summations below). Then we have

$$v^{(s)}_{k,i,j} = \sum_{\substack{a,b,c\geqslant 0\\ a+b+c=k}} \frac{k!}{a!\,b!\,c!} S_{a,b,c}, \tag{6}$$

where

$$S_{a,b,c} = \sum_{m=0}^{s} \binom{c}{m} f^{(s)}_{m,i-a-c+m,j-b-c+m}.$$

But, from Theorem 5.1, we have

$$\begin{aligned} S_{a,b,c} &= s \sum_{s-1 \geqslant m-1 \geqslant l \geqslant 0} \binom{c}{m}\binom{s-1-l}{m-1-l}\binom{m-1-l}{a+c-i-1-l}\binom{m-1-l}{b+c-j-1-l} \\ &= s!\left[y^{a+c-i-1} z^{b+c-j-1}\right] \sum_{s-1 \geqslant m-1 \geqslant l \geqslant 0} \binom{c}{m}\binom{s-1-l}{m-1-l}(yz)^l \big((1+y)(1+z)\big)^{m-1-l}, \end{aligned}$$

where we use the notation $[A]B$ to denote the coefficient of A in the expansion of B. Now

$$\binom{c}{m} = \binom{-m-1}{c-m}(-1)^{c-m} = \left[x^c\right]x^m(1-x)^{-m-1},$$

which gives

$$S_{a,b,c} = s!\left[x^c y^{a+c} z^{b+c}\right] G(x,y,z), \tag{7}$$

where, summing over m by the binomial theorem, we have

$$\begin{aligned} G(x,y,z) &= \frac{xy^{i+1}z^{j+1}}{(1-x)^2} \sum_{l=0}^{s-1} \left(\frac{xyz}{1-x}\right)^l \left(1 + \frac{x(1+y)(1+z)}{1-x}\right)^{s-l-1} \\ &= \frac{xy^{i+1}z^{j+1}}{(1-x)^2} \frac{(1+\frac{x(1+y)(1+z)}{1-x})^s - (\frac{xyz}{1-x})^s}{1+\frac{x(1+y)(1+z)}{1-x} - \frac{xyz}{1-x}} \\ &= \frac{xy^{i+1}z^{j+1}}{(1-x)^{s+1}} \frac{(1+x(y+z+yz))^s - (xyz)^s}{1+x(y+z)}. \end{aligned}$$

But, changing variables in (7), we have

$$S_{a,b,c} = s!\left[x^0 y^{a+c} z^{b+c}\right] x^k G\left(x, \frac{y}{x}, \frac{z}{x}\right),$$

so from (6) we obtain

$$\begin{aligned} v^{(s)}_{k,i,j} &= s!\left[x^0 y^k z^k\right] x^k (1+y+z)^k G\left(x, \frac{y}{x}, \frac{z}{x}\right) \\ &= s!\left[x^0 y^k z^k\right] x^{k-i-j-s-1} y^{i+1} z^{j+1} \frac{(1+y+z)^{k-1}}{(1-x)^{s+1}} \left(\left(x(1+y+z)+yz\right)^s - (yz)^s\right) \\ &= R_1 - R_2, \end{aligned}$$

where

$$\begin{aligned} R_2 &= s!\left[x^{i+j+s-k+1} y^{k-i-1} z^{k-j-1}\right] \frac{(1+y+z)^{k-1}}{(1-x)^{s+1}} (yz)^s \\ &= s!\binom{2s+i+j-k+1}{s} \frac{(k-1)!}{(k-s-i-1)!(k-s-j-1)!(2s+i+j-k+1)!} \\ &= \frac{(s+i)!(s+j)!}{(s+i+j)!} \binom{s+i+j}{k-1}\binom{k-1}{s+i}\binom{k-1}{s+j}, \end{aligned}$$

$$
\begin{aligned}
R_1 &= s!\big[x^{i+j+s-k+1}y^{k-i-1}z^{k-j-1}\big]\frac{(1+y+z)^{k-1}}{(1-x)^{s+1}}\big(x(1+y+z)+yz\big)^s \\
&= s!\big[x^{i+j+s-k+1}y^{k-i-1}z^{k-j-1}\big]\sum_{m\geqslant 0}\binom{s}{m}x^{s-m}(1+y+z)^{k+s-m-1}(yz)^m(1-x)^{-s-1} \\
&= s!\sum_{m\geqslant 0}\binom{s}{m}\binom{s+i+j+m-k+1}{s} \\
&\quad \times \frac{(s-m+k-1)!}{(k-m-i-1)!(k-m-j-1)!(s+i+j+m-k+1)!}.
\end{aligned}
$$

Now, for this latter sum over $m \geqslant 0$, we observe that the ratio of the $(m+1)$st term to the mth term is a rational function of m, which implies that it is a hypergeometric sum. In particular, using the standard notation for hypergeometric series, we have

$$
\begin{aligned}
R_1 &= \frac{(s+k-1)!}{(k-i-1)!(k-j-1)!(i+j-k+1)!}\,{}_3F_2\left(\begin{matrix} i+1-k,\ j+1-k,\ -s \\ i+j-k+2,\ 1-s-k \end{matrix};1\right) \\
&= \frac{(s+k-1)!}{(k-i-1)!(k-j-1)!(i+j-k+1)!}\,\frac{\binom{s+i}{s}\binom{s+j}{s}}{\binom{s+k-1}{s}\binom{s+i+j-k+1}{s}} \\
&= \frac{(s+i)!(s+j)!}{(s+i+j)!}\binom{s+i+j}{k-1}\binom{k-1}{i}\binom{k-1}{j},
\end{aligned}
$$

where the second last equality follows from the Pfaff–Saalschütz Theorem for ${}_3F_2$ hypergeometric summations (see, *e.g.*, Theorem 2.2.6 on page 69 of [1]). The result follows immediately. □

The proof of Theorem 1.2 follows immediately from Corollary 4.3 and Theorem 5.2.

Acknowledgments

This work was supported by a Discovery Grant from NSERC (IG) and by a USRA Award from NSERC (WS). We are grateful to Andu Nica and an anonymous referee for helpful discussions and suggestions.

References

[1] G.E. Andrews, R. Askey, R. Roy, Special Functions, Encyclopedia Math. Appl., vol. 71, Cambridge University Press, 1999.
[2] U. Haagerup, S. Thorbjornsen, Random matrices with complex Gaussian entries, Expo. Math. 21 (2003) 293–337.
[3] I.P. Goulden, A. Nica, A direct bijection for the Harer–Zagier formula, J. Combin. Theory Ser. A 111 (2005) 224–238.
[4] J. Harer, D. Zagier, The Euler characteristic of the moduli space of curves, Invent. Math. 85 (1986) 457–486.
[5] C. Itzykson, J.-B. Zuber, Matrix integration and combinatorics of modular groups, Comm. Math. Phys. 134 (1990) 197–207.
[6] D.M. Jackson, On an integral representation for the genus series for 2-cell embeddings, Trans. Amer. Math. Soc. 344 (1994) 755–772.
[7] S. Kerov, Rook placements on Ferrers boards, and matrix integrals (English translation), J. Math. Sci. (N. Y.) 96 (1999) 3531–3536.
[8] M. Kontsevich, Intersection theory on the moduli space of curves and the matrix Airy function, Comm. Math. Phys. 147 (1992) 1–23.
[9] S.K. Lando, A.K. Zvonkin, Graphs on Surfaces and Their Applications, Springer-Verlag, Berlin, Heidelberg, 2004.
[10] B. Lass, Démonstration combinatoire de la formule de Harer–Zagier, C. R. Acad. Sci. Paris, Sér. I 333 (2001) 155–160.
[11] J.A. Mingo, A. Nica, Annular noncrossing permutations and partitions, and second-order asymptotics for random matrices, Int. Math. Res. Not. 28 (2004) 1413–1460.
[12] R.C. Penner, Perturbative series and the moduli space of Riemann surfaces, J. Differential Geom. 27 (1988) 35–53.
[13] R.P. Stanley, Enumerative Combinatorics, vol. 2, Cambridge University Press, 1999.
[14] W.T. Tutte, Graph Theory, Addison–Wesley, London, 1984.
[15] D. Zagier, On the distribution of the number of cycles of elements in symmetric groups, Nieuw Arch. Wiskd. 13 (1995) 489–495.
[16] A. Zvonkin, Matrix integrals and map enumeration: An accessible introduction, in: Combinatorics and Physics, Marseilles, 1995, Math. Comput. Modelling 26 (1997) 281–304.