THREE BIJECTIONS ON SET PARTITIONS

ROBERT PARVIAINEN

Abstract. We study three similar bijections on set partitions. The first gives a bijective proof of the equivalence of two statistics with a q-Stirling distribution, Milne's statistic and the intertwining number. The second proves the equivalence of a multivariate block size distribution to a covering statistic. The third demonstrates equivalence of the number of all set partitions up to a given size to set partitions of a larger size with second block a singleton.

1. Introduction

To partition a set of distinct elements into a given number of parts is of course a classic problem. The aim here is to give bijective demonstrations of three statements from this area of combinatorics.

In Section 3 we bijectively prove equivalence of two statistics counted by the q-Stirling numbers of the second kind. These are the intertwining number of Ehrenborg and Readdy, [2], and Milne's statistic, also known as the dual major index and the left smaller statistic [3, 6, 7].

A similar bijection is used in Section 4 to bijectively prove equivalence of a new statistic with a statistic on block sizes. The new statistic generalises the notion of doubles of Riordan, [5]. Riordan showed that the number of doubles, to be defined, in a set partition has the same distribution as the size of the block including 1, minus 1. We give a multivariate statistic that is equivalent to the distribution of $(b_1 - 1, b_2 - 1, \ldots)$, where b_i is the size of the ith block when blocks are ordered in the normal fashion.

The same ideas are used in Section 5 to show a simple but peculiar equivalence. The total number of partitions with k blocks of $\emptyset, [1], [2], \ldots, [n]$ is shown to be the same as the number of partitions of $[n+1]$ with $[k+1]$ blocks with second block of size at most 1.

2. Some definitions

A (set) partition of $[n] = \{1, 2, \ldots, n\}$ is a collection of non-empty, mutually disjoint subsets of $[n]$, or *blocks*, with union $[n]$. Elements in the blocks will be ordered in increasing order, and blocks ordered by their minimal elements. Write $P = B_1/B_2/\cdots/B_k$ for the partition of $[n]$ into the k blocks $B_1, B_2, \ldots, B_k$, where $\min B_1 < \min B_2 < \cdots < \min B_k$. Let Π_n^k denote the set of partitions of $[n]$ with k blocks, and Π_n all partitions of $[n]$.

It is common to represent a partition as a graph on the vertex set $[n]$, where the edge set are arcs that connect consecutive elements in blocks. The edge set $\{(i_1, j_1), \ldots, (i_m, j_m)\}$ is called the *standard representation*. The special embedding when the vertices are the points $1, 2, \ldots, n$ on the real line, and the edges are drawn as arcs above the line, will be called the *diagram* of the partition. Partition diagrams are shown in Figures 1 to 5 below. In what follows we will always consider the edge set in the standard representation to be ordered by the second co-ordinate, so that $j_1 < j_2 < \cdots < j_m$, and such that $i_k < j_k$.

We remark that if $n = 2m$ is even, *matchings* of $[2m]$ are the special case where all blocks have size 1 or 2, and if there are m blocks, the matching is *complete*.

The *restricted growth function* representation of a set partition will prove useful. It is a word of positive integers $w_1 \cdots w_n$ such that $w_1 = 1$ and $w_i \leq 1 + \max_{1 \leq j < i} w_j$. If $P = B_1/B_2/\cdots/B_k$

Date: November 8, 2007.

2000 *Mathematics Subject Classification.* Primary 05A18; Secondary 05A30.

Key words and phrases. Set partition, q-Stirling numbers, bijection.

is a partition we can define a restricted growth function $\omega(P) = w_1 \cdots w_n$ by letting $w_i = j$ if $i \in B_j$. This defines a bijection, as is easy to verify.

3. The intertwining number and Milne's statistic

Let b_i be the size of block i. Milne's, [3], statistic MIL on a partition P with k blocks is

$$\textsc{mil}(P) = 0 \cdot b_1 + 1 \cdot b_2 + \cdots (k-1) \cdot b_k.$$

With RG-words the definition is

$$\textsc{mil}(\omega(P)) = \sum_i (w_i - 1).$$

Milne's statistic was called the dual major index $\widehat{\mathrm{maj}}$ by Sagan, [6], and equals the ls (for left smaller) statistic of Wachs and White, [7].

The distribution of MIL is given by the q-Stirling numbers of the second kind, $S_q(n,k)$. Another statistic with the same distribution is the intertwining number INT defined by Ehrenborg and Readdy, [2]. For two integers i and j let $I(i,j)$ be the interval of integers from the smallest to the largest of i and j, inclusively. Now the definition of INT goes like this. For a partition $P = B_1/B_2/\cdots/B_k$, let

$$\textsc{int}(P) = \sum_{i<j} |\{(b_r, b_s) \in B_i \times B_j : I(r,s) \cap (B_i \cup B_j)\}|$$

An interesting property of INT is that the ordering of blocks are irrelevant.

A more graphic way to define INT is to count crossings in a modified diagram. To all minimal elements add parallel infinite lines from the vertices leaving in say a 45 degree angle upwards and to the left. To all maximal elements parallel infinite lines from the vertices leaving in a 45 degree angle upwards and to the right. The intertwining number is the number of crossings in the new diagram. An example is given Figure 1.

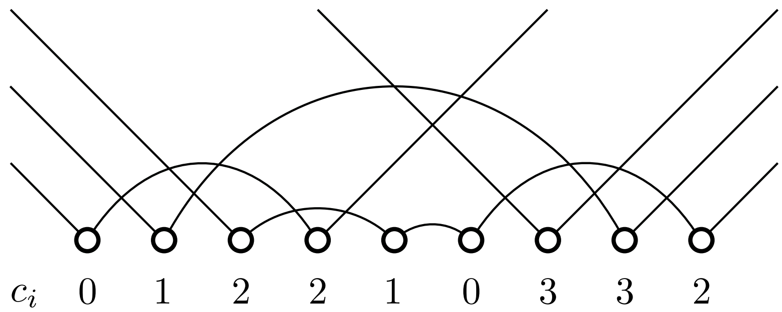

Figure 1. The modified diagram of the set partition 14/28/3569/7 and its crossing weight sequence c_i. The intertwining number of the partition is 14.

Now we want to give each vertex in the diagram a weight (because the sequence of weights will, bar a triviality, be a restricted growth function). Note that each vertex has an arc *or* a line leaving in the leftward direction. Starting with vertex 1 and moving to the right, let c_i be the number of crossings on that arc or line, that *have not already been counted*. Figure 1 provides illustration. Now

$$\textsc{int}(P) = \sum_{i=1}^{n} c_i.$$

Define a mapping ψ by letting $\psi(P)$ be the word $(c_1+1)(c_2+1)\cdots(c_n+1)$.

Lemma 1. *The mapping ψ is a bijection from set partitions to restricted growth functions.*

Proof. First we show that $\psi(P)$ is a restricted growth function, that is, that $c_1 = 0$ (obvious) and that $c_i \le 1 + \max_{1\le j<i} c_j$. Induction is appropriate here.

Suppose c_1 to c_{i-1} fulfils the restricted growth condition and consider vertex i. If vertex $i-1$ is minimal, i.e. a line is leaving the vertex to the left, then the line or arc leaving vertex i to the left can produce at most $c_{i-1}+1$ new crossings (with the same lines/arcs as the previous line, plus the one leaving vertex $i-1$ to the right), so $c_i \le 1 + c_{i-1}$. If vertex $i-1$ is not minimal, but vertex $i-2$ is, the line or arc leaving vertex i to the left can produce at most $c_{i-2}+1$ new crossings (with the same lines/arcs as the previous line, plus the ones leaving vertex $i-2$ and vertex $i-1$ to the right, minus the one connecting to vertex $i-1$ from the left). The general case where vertices $i-1$ to $i-j$ are not minimal, but $i-j-1$ is follows similarly. In each case, $c_i \le 1 + \max_{1\le j<i} c_j$.

We next describe the inverse. Suppose we are given a restricted growth function $(c_1 + 1)\cdots(c_n+1)$. First draw two lines from each vertex, one up to the left, one up to the right. Go through the vertices in order. For the first vertex do nothing. For the second vertex, if $c_2 = 1$ do nothing. If $c_2 = 0$ draw an arc between vertices 1 and 2, and erase the line from vertex 1 to the right and the line from vertex 2 to the left. For the ith vertex, if $c_i = 1 + \max\{c_1, c_2, \ldots c_{i-1}\}$ do nothing. Otherwise connect vertex i to the unique vertex $1 \le j < i$ such that the arc crosses c_i lines, and erase the line from vertex j to the right and the line from vertex i to the left.

This is always possible, because while at vertex i, there are exactly $\max\{c_1, c_2, \ldots c_i\}$ lines passing over the vertex, from left to right. This holds by induction. From the description given above, it holds for $i = 1$ and $i = 2$. Next, assume that there are $\max\{c_1, c_2, \ldots c_{i-1}\}$ lines passing over vertex $i-1$. If we do nothing at vertex i, that is if $c_i = 1+\max\{c_1, c_2, \ldots c_{i-1}\}$, an extra line (from vertex $i-1$) will pass over vertex i, bringing the total to $\max\{c_1, c_2, \ldots c_{i-1}\}+1$ which equals $\max\{c_1, c_2, \ldots c_i\}$ as required. In the other case, the number of lines passing over vertex i is unchanged.

By construction the resulting diagram is the diagram of a partition with crossing weight sequence $(c_1, \cdots, c_n)$. □

The two bijections between set partitions and restricted growth functions combines to form a bijection from and to set partitions. This bijection preserves the number of blocks.

Lemma 2. *The set partitions P and $\omega^{-1}(\psi(P))$ have the same set of minimal elements. In particular, they have the same number of blocks.*

Proof. The number of blocks of P equals the number of distinct elements in the restricted growth function $\omega(P)$. The minimal elements are the positions in $\omega(P)$ where the k distinct elements occur for the first time. That is the set $\{i : w_i = 1 + \max_{1\le j<i} w_j\}$.

In the description of ψ^{-1} in the proof of lemma 1, this is exactly the set of vertices which have a line, as opposed to an arc, leaving to the left. But these are just the minimal elements. □

Combining previous results gives the main result of this section.

Theorem 3. *The mapping $\phi_1(\cdot) = \omega^{-1}(\psi(\cdot))$ is bijection that preserves the number of blocks. Further, $\mathrm{INT}(P) = \mathrm{MIL}(\phi_1(P))$.*

3.1. **Equivalence by the ECO-method.** We note that the equivalence can also be proved using the ECO-method, with the notions of weighted succession rules, weighted generating trees and weighted production matrices introduced in [4]. In fact, that method even suggests the bijection, in the sense that it gives a suggestion as to which partition to map to which. We do not give details here, but note that both structures, partitions weighted by the intertwining number and number of blocks, and partitions weighted by Milne's statistic and number of blocks, have the same weighted generating tree. The weighted production matrix (for both cases) is, with $[n]_q = 1 + q + \cdots + q^{n-1}$,

$$\mathbf{P_1} = \begin{pmatrix} q[1]_q & t & 0 & 0 & \cdots \\ 0 & q[2]_q & t & 0 & \cdots \\ 0 & 0 & q[3]_q & t & \cdots \\ 0 & 0 & 0 & q[4]_q & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The generating functions can be expressed using $\mathbf{P}_1$, see [1] and [4].

Theorem 4. *The ordinary generating function $F_1(q,t;z)$ for set partitions with respect to any statistic distributed as $S_q(n,k)$ where size is marked by z and number of blocks by t, is*

$$F_1(q,t;z) = (1,0,0,0,\ldots)(\mathbf{I} - z\mathbf{P}_1)^{-1}(1,1,1,1,\ldots)^\top,$$

where $\mathbf{P}_1$ is defined above and $\mathbf{I}$ is the identity matrix.

The exponential generating function $G_1(q,t;z)$ is

$$G_1(q,t;z) = (1,0,0,0,\ldots)\exp(z\mathbf{P}_1)(1,1,1,1,\ldots)^\top.$$

4. Block sizes and covers

In [5] Riordan briefly considers the number of doubles (adjacent letters which agree) in restricted growth functions, and shows that it is equivalent to the number of 1s minus 1. Here we generalise this to include all block sizes, not only the first.

The generalised statistic is most conveniently defined by the diagram. Let cov_i count the number of arcs with covers vertices from i different blocks. That is the number of pairs (i_k, j_k) in the edge set $\{(i_1,j_1),\ldots,(i_m,j_m)\}$ such that the interval $i_k+1,\ldots,j_k-1$ contains elements from i different blocks. Say that such an arc is an i-cover.

In the restricted growth function representation, cov_i are defined like this. For word w let $\text{cov}_i(w)$ count the number of occurrences of av_ia, where a is any letter and v_i is a word with exactly i distinct letters, all different from a. So $\text{cov}_0(w)$ counts the number of equalities of adjacent letters, and $\text{cov}_1(w)$ counts the number of occurrences of two same letters separated by one or more occurrences of another letter. For example, if $w = 122321241$ then $\text{cov}_0(w) = 1$, $\text{cov}_1(w) = 2$ and $\text{cov}_2(w) = 2$. See also the right drawing in Figure 2 for the diagram version.

The main result is this, which will be proved by defining an appropriate bijection.

Theorem 5. *The number of partitions $P = B_1/B_2/\cdots/B_k$ of $[n]$ with k blocks of sizes $b_1, b_2, \ldots, b_k$ equals the number of set partitions of $[n]$ with k blocks and $\text{cov}_0 = b_1 - 1, \text{cov}_1 = b_2 - 1, \ldots, \text{cov}_{k-1} = b_k - 1$.*

In the spirit of the previous bijection, we define two weight sequences on the vertices of the partition diagram. These uniquely define the partition, and the recipes for this describe the bijection. Examples of the two sequences are shown in Figure 2.

First let the weight of vertex i be 1 if it is the smallest element of its block. Otherwise give it weight q_j if it is in block j. Denote this sequence $\mathbf{x} = (x_1, x_2, \ldots x_n)$, and say that it is the partition's $\mathbf{x}$ sequence. We have $x_1 \cdots x_n = q_1^{b_1-1} \cdots q_k^{b_k-1}$. A procedure to recover the partition from $\mathbf{x}$ goes like this. Read $\mathbf{x}$ from left to right. If $x_i = 1$ do nothing. If $x_i = q_j$ draw an arc to the next vertex on the left with weight q_j, if such a vertex exists. If not, draw an arc to the jth vertex on the left with weight 1. Call this procedure recipe X.

Also define an $\mathbf{y}$ sequence. Let the weight y_i of vertex i be q_{j+1} if it is the endpoint of an arc covering vertices from j different blocks, and 1 if it is not the endpoint of an arc. We have $y_1 \cdots y_n = q_1^{\text{COV}_0} \cdots q_k^{\text{COV}_{k-1}}$. The partition is recovered by reading $\mathbf{y}$ from left to right. If $y_i = 1$ skip. If $y_i = q_{j+1}$ go left until we have passed vertices from j blocks different from i, stopping at the first vertex which is not already the starting point of an arc. Draw an arc from this vertex to vertex i. Call this procedure recipe Y.

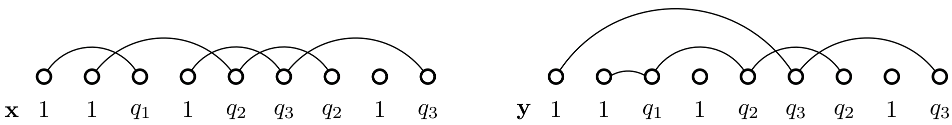

FIGURE 2. To the left, the partition $P_1 = 13/257/369/8$ and its $\mathbf{x}$ sequence. To the right the partition $P_2 = 169/2357/4/8$ and its $\mathbf{y}$ sequence — it is the image of P_1 under bijection ϕ_2.

The question is now, if a set partition has a given $\mathbf{x}$ (or $\mathbf{y}$) sequence, is this valid as an $\mathbf{y}$ (or $\mathbf{x}$) sequence? The answer is yes in both cases, as we now argue.

Assume we are given an $\mathbf{y}$ sequence, and use recipe X. The only way X fails is if a vertex i has weight q_j but there are less than j vertices with weight 1 to the left of it. But this cannot happen, because if i has $\mathbf{y}$-weight $y_i = q_j$ there are vertices from $j-1$ blocks between vertex i and the starting point of the arc ending at i. Each of these blocks plus the block vertex i belongs to has its smallest element to the left of i, and these all have weight 1.

The proof that we can use recipe Y on an $\mathbf{x}$ sequence is similar.

Define a mapping ϕ_2 on partitions like this. Let $\phi_2(P)$ be the partition produced by recipe Y on Ps $\mathbf{x}$-sequence. It is clear from the arguments above that ϕ_2 is a bijection with the desired properties.

Theorem 6. *The mapping ϕ_2 is a bijection such that if a set partition $P = B_1/B_2/\cdots/B_k$ has k blocks with sizes $b_1, b_2, \ldots, b_k$ then $\phi_2(P)$ also has k blocks, and* $\mathrm{COV}_0 = b_1 - 1, \mathrm{COV}_1 = b_2 - 1, \ldots, \mathrm{COV}_{k-1} = b_k - 1$.

4.1. **A simple consequence.** Define a generating function $F(\mathbf{q}; t, z) = F(q_1, q_2, \ldots; t, z)$ for set partitions where size of block i, or COV_{i+1}, is marked by q_i and size is marked by z and number of blocks by t. By considering partitions where the first m blocks are of size 1, the next result follows immediately.

Theorem 7. *The number of partitions of $[n+1]$ with no doubles into $k+1$ parts equals the number of partitions of $[n]$ into k parts. Further, the number of partitions of $[n+2]$ with no doubles and no 1-covers into $k+2$ parts equals the number of partitions of $[n]$ into k parts. In general,*

$$F(\underbrace{0, \ldots, 0}_{m \text{ zeroes}}, q_{m+1}, q_{m+2}, \ldots; t, z) = \sum_{i=0}^{m-1} (tz)^i + (tz)^m F(q_{m+1}, q_{m+2}, \ldots; t, z).$$

4.2. **Equivalence by the ECO-method.** Here too we can find a simple weighted succession rule common for both structures. It is described by the following weighted production matrix:

$$\mathbf{P}_2 = \begin{pmatrix} q_1 & t & 0 & 0 & \cdots \\ 0 & q_1+q_2 & t & 0 & \cdots \\ 0 & 0 & q_1+q_2+q_3 & t & \cdots \\ 0 & 0 & 0 & q_1+q_2+q_3+q_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The generating functions can be expressed just as in Theorem 4.

Theorem 8. *The ordinary generating function $F_2(\mathbf{q}, t; z)$ for set partitions where size of block i, or COV_{i+1}, is marked by q_i and size is marked by z and number of blocks by t, is*

$$F_2(\mathbf{q}, t; z) = (1, 0, 0, 0, \ldots)(\mathbf{I} - z\mathbf{P}_2)^{-1}(1, 1, 1, 1, \ldots)^\top,$$

where $\mathbf{P}_2$ is defined above and $\mathbf{I}$ is the identity matrix.

The exponential generating function $G_2(\mathbf{q}, t; z)$ is

$$G_2(\mathbf{q}, t; z) = (1, 0, 0, 0, \ldots) \exp(z\mathbf{P}_2)(1, 1, 1, 1, \ldots)^\top.$$

5. All partitions and partitions with small second block

A peculiar equivalence can be proven with a bijection similar to the previous two. Adopt the convention that there is one way to partition the empty set, and that that partition has zero blocks. Now, the equivalence goes like this.

Theorem 9. *The combined number of partitions of $\emptyset, [1], [2], \ldots, [n]$ into k blocks equals the number of partitions of $[n+1]$ into $[k+1]$ blocks, with second block, if it exists, of size one. Equivalently, by Theorem 5 it equals the number of partitions of $[n+1]$ into $[k+1]$ blocks, with no 1-covers.*

To describe the bijection that will verify this equivalence, we will consider a slight variation of restricted growth functions. Namely, when considering all partitions of $\emptyset, [1], [2], \ldots, [n]$, prefix enough zeroes to make the length equal to $n+1$. Denote the set of all these restricted grow functions $\mathcal{W}_{n+1}$. Denote the set of partitions with second block of sizes 0 and 1 by $\mathcal{V}_n$.

A bijection is now easily described. For each word $w = w_1 w_2 \ldots w_n$ in $\mathcal{W}_n$ define a word $v = \phi_3(w)$ as follows. Go through the word letter by letter. First there is a sequence of zeroes. If $w_i = 0$ let $v_i = 1$. The next letter, at index k say, must be 1. Let $v_k = 2$. For the remainder, let $v_i = w_i + 1$ if $w_i > 1$ and $v_i = 1$ if $w_i = 1$.

It is evident that $v \in \mathcal{V}_n$, and that the number of distinct letters in w equals the number of distinct letters in v. The inverse is just as easily described. Brief version: map 1 to 0 until the first occurrence of a 2, then to 1, and map $k > 1$ to $k - 1$.

Corollary 10. *The number of partitions of $[n]$ with $[k]$ blocks whose second block is of size 0 or 1 (equivalently, with no 1-covers) is*

$$\sum_{i=k-1}^{n-1} S(i, k-1),$$

where $S(n,k)$ are the Stirling numbers of the second kind.

In this case the three classes of objects are generated by the weighted production matrix

$$\mathbf{P}_3 = \mathbf{P}_2|_{q_1=1, q_2=0, q_3=1, q_4=1, \ldots} = \begin{pmatrix} 1 & t \cdot (1, 0, 0, \cdots) \\ (0, 0, \cdots)^\top & \mathbf{P_2} \end{pmatrix},$$

a fact from which (once demonstrated of course) the equivalence also follows.

Acknowledgement

The Author acknowledge the support received from the ARC Centre of Excellence for Mathematics and Statistics of Complex Systems (MASCOS).

References

[1] E. Deutsch, L. Ferrari, and S. Rinaldi. Production matrices. *Advances in Applied Mathematics*, 34:101–122, 2005.

[2] R. Ehrenborg and M. Readdy. Juggling and applications to q-analogues. *Discrete Mathematics*, 157(1–3):107–125, 1996.

[3] S. C. Milne. Restricted growth functions, rank row matchings of partition lattices, and q-Stirling numbers. *Advances in Mathematics*, 43:173–196, 1982.

[4] R. Parviainen. Parametric production matrices and weighted succession rules: a Dyck path example. *Journal of Integer Sequences*, 10(6):Article 07.6.1, 2007.

[5] J. Riordan. A budget of rhyme scheme counts. *Annals of the New York Academy of Sciences*, 319(1):455–465, May 1979.

[6] B. E. Sagan. A maj statistic for set partitions. *European Journal of Combinatorics*, 12:69–79, 1991.

[7] M. Wachs and D. White. p, q-Stirling numbers and set partition statistics. *Journal of Combinatorial Theory, Series A*, 56:27–46, 1991.

ARC Centre of Excellence for Mathematics and Statistics of Complex Systems, 139 Barry Street, The University of Melbourne, Victoria, 3010

E-mail address: robertp@ms.unimelb.edu.au