A set partition identity via trees

Martin Klazar[1] and Vít Novák

Department of Applied Mathematics of Charles University
Malostranské náměstí 25
118 00 Praha 1
Czech Republic
klazar@kam.ms.mff.cuni.cz
novakvt@fzu.cz

Abstract

We consider two kinds of partitions having n blocks and an initial segment of positive integers as a ground set. Pretty partition has all blocks of size at most 2, does not induce the pattern $\dots a \dots b \dots b \dots a \dots$, and has no two consecutive numbers in the same block. Ugly partition differs only in that it does have some two consecutive numbers in the same block. Using rooted plane trees we construct, for any $n \geq 1$, a bijection matching pretty and ugly partitions.

1 Introduction

A *partition* u with n *blocks* is a set of n nonempty disjoint subsets of $X = \{1, 2, \dots, l\}$ whose union is X. We say that u is *abba-free* if there are no four distinct numbers $1 \leq i_1 < \dots < i_4 \leq l$ and no two distinct blocks A and B such that $i_1, i_4 \in A$ and $i_2, i_3 \in B$. Partitions having no two consecutive numbers in the same block are called *pretty*, otherwise they are *ugly*.

The purpose of this note is to prove bijectively the following identity.

Identity 1.1 *Among abba-free partitions with $n \geq 1$ blocks, each block of size 1 or 2, there is as many pretty partitions as ugly partitions.*

Any partition u can be written as a sequence $a_1 a_2 \dots a_l$ of labels given to the blocks: a_i is the label of the block B, $i \in B$. The *canonical form* of u is obtained when the blocks are ordered by their least elements as $B_1, B_2, \dots, B_n$ and B_i is labeled by i. We shall work with partitions in their sequential form.

[1]supported by grants GAČR 0194 and GAUK 194.

For instance, one way how to write $u = \{\{2,3,5\},\{1,6\},\{4\}\}$ as a sequence is $u = bccacb$ and the canonical form is $u = 122321$. For $n = 2$ the pretty and ugly partitions appearing in the identity are:

$$\{12, 121, 1212\} \text{ and } \{112, 122, 1122\}.$$

For $n = 3$ the two sets described in the identity have 11 elements.

The identity was discovered in [1] as a byproduct of formulae for generating functions enumerating *abba*-free partitions. In the next section we present a bijection proving the identity. Our main tool is an encoding of *abba*-free partitions by rooted plane trees.

2 The bijection

A *rooted plane tree* is a finite directed tree with all edges directed away from the distinguished vertex, called a *root*, and with a linear order on any set of children of a vertex. From now on we call them shortly *trees*.

We think of trees as plane pictures. We draw vertices as points, the root in the lowest position, and edges as straight segments directed up. The children of a vertex are drawn from left to right in accordance with the prescribed linear order. It is well known that there are $\binom{2n}{n}/(n+1)$ (Catalan number) different trees with n edges.

For $e = v_1v_2$ an edge in a tree T we refer to v_2 as to the *child* of v_1 and to v_1 as to the *parent* of v_2. A vertex with no child is called a *leaf*. A *layer* in T is the set of vertices with the same distance from the root. Suppose the vertices of T are ordered as $v_0, v_1, \ldots, v_n$ so that lower layers come first and in one layer left vertices come first. Hence v_0 is the root. Such an order is called *good ordering*. A vertex of T is called *solitary* (*young*) if it is the only vertex in its layer and its parent is the rightmost vertex in its layer (if it is a leaf whose parent is the root).

Let $\mathcal{S}(n)$ stand for the set of *abba*-free partitions with n blocks, each block of size 1 or 2. The subsets of pretty and ugly partitions are denoted by $\mathcal{P}(n)$ and $\mathcal{U}(n)$. The subset of partitions with two-element blocks only is $\mathcal{R}(n)$. The set of trees with n edges is denoted by $\mathcal{T}(n)$.

In the rest of the note we shall construct a bijection F between the sets $\mathcal{P}(n)$ and $\mathcal{U}(n)$. First we restate the identity in terms of tree structures called *gap trees*. In the second step we construct the desired bijection, working with gap trees rather than with partitions.

From partitions to gap trees

We start with a bijection G between $\mathcal{R}(n)$ and $\mathcal{T}(n)$. Suppose $u = a_1a_2\dots a_{2n} \in \mathcal{R}(n)$ is in the canonical form. The tree $T = G(u)$ is constructed by processing u from left to right. In the beginning $i = 1$, $T_0 = p$, and $v = p$ where p is a single unlabeled vertex. In the general step T_{i-1} is a tree with unlabeled root and all other vertices labeled by positive integers and v is a vertex of T_{i-1}. If $a_i \neq a_j$ for all $1 \le j < i$ we derive T_i from T_{i-1} by adding a new child with the label a_i to the right of the children of v. Then we move to the next term of u, v remains the same. If a_i appears in u before we put T_i equal to T_{i-1}, v equal to the vertex labeled by a_i, and we move to the next term of u. The procedure terminates for $i = 2n$, we forget the labels and set $G(u) = T = T_{2n}$.

Lemma 2.1 *The mapping $G:\ \mathcal{R}(n) \to \mathcal{T}(n)$ is a bijection.*

Proof. The algorithm adds vertices in their good order and v traces T in the good order. Let us define the inverse of G. We take the vertices $(v_0, v_1, \dots, v_n)$ of $T \in \mathcal{T}(n)$ in their good order and write down for each v_i first the index i and then, left to right, the indices of its children. We set $G^{-1}(T)$ equal to the sequence obtained, the initial 0 deleted. Clearly, G and G^{-1} are inverses of one another. $\square$

The mapping G corresponds to the bredth-first search in T. We remark that *abab*-free partitions (the avoidance of *abab* is defined in a way analogous to that of *abba*) with n blocks, each of size 2, can be put in a bijective correspondence with $\mathcal{T}(n)$ as well. These partitions are proper bracketings with n brackets. The correspondence matching them with trees is based on the depth-first search and is well known.

A *gap* in a finite sequence $u = a_1a_2\dots a_l$ is the space between two consecutive terms or the space before a_1 or the space after a_l. The set of gaps $g(u)$ has $l+1$ elements. Suppose $u = a_1 \dots a_{2n} \in \mathcal{R}(n)$ and let $x = a_i = a_j$, $i < j$. The *first* (the *second*) *gap of* x is the gap following after a_i (after a_j). The *first gap of* u is the gap of u before a_1.

The *gaps of a vertex* v of a tree $T \in \mathcal{T}(n)$ are the wedge-shaped spaces into which the edges going up from v divide the neighborhood of v. A vertex with d children has $d+1$ gaps. In particular, any leaf has exactly one gap. The set $g(T)$ of all gaps has $2n+1$ elements. For $e = v_1v_2$ an edge of T we call the leftmost gap of v_2 the *top gap of* e and the gap of v_1 to the right of e the *bottom gap of* e. The *first gap of* T is root's leftmost gap.

The mapping G induces a bijection $G^* : \; g(u) \to g(G(u))$. Suppose $u = a_1 a_2 \dots a_{2n} \in \mathcal{R}(n)$ is in the canonical form. The first gap of u is sent to the first gap of $T = G(u)$. The first (the second) gap of an integer x is sent to the top (to the bottom) gap of the edge whose endvertex is the xth one in the good order, we remind that the root is the 0th vertex.

A *gap tree* is a pair (T, s) where T is a tree and $s : \; g(T) \to \mathbf{N}_0$ is an integer mapping. Its *size* is $|E(T)| + \sum s(g)$ where we sum over $g(T)$. The set of gap trees of size n is denoted by $\mathcal{GT}(n)$. A vertex is *solitary* (*young*) in (T, s) if it is solitary (young) in T and $s(g) = 0$ for its leftmost gap (for its only gap).

Any sequence $u \in \mathcal{S}(n)$ can be encoded by a gap tree $H(u) = (T, s)$ of size n as follows. We decompose u into (u^*, t) where $u^* \in \mathcal{R}(m)$ is the subsequence of 2-element blocks and $t : \; g(u^*) \to \mathbf{N}_0$ counts the numbers of 1-element blocks in the gaps of u^*. We set $T = G(u^*)$ and $s(G^*(g)) = t(g)$ for any $g \in g(u^*)$. For an example illustrating H see Fig. 2.

Lemma 2.2 *The above mapping $H : \; \mathcal{S}(n) \to \mathcal{GT}(n)$ is a bijection. Moreover, it maps pretty partitions to those and only those gap trees which have no solitary vertex.*

Proof. Check the definitions. $\square$

Thus the desired bijection $F : \; \mathcal{P}(n) \to \mathcal{U}(n)$ is constructed as soon as we exhibit a bijection matching gap trees of size n without solitary vertices with those having at least one solitary vertex. We denote the former set as $\mathcal{GT}_0(n)$ and the latter set as $\mathcal{GT}_1(n)$.

Bijections for gap trees

It is was not too obvious to us how to match the elements of $\mathcal{GT}_0(n)$ and $\mathcal{GT}_1(n)$. However, we could easily see the bijection between the sets $\mathcal{GT}^0(n)$ and $\mathcal{GT}^1(n)$. The former set consists of gap trees of size n with no young vertex and the latter set of gap trees of size n with at least one young vertex.

Lemma 2.3 *There is a bijection $I : \; \mathcal{GT}^1(n) \to \mathcal{GT}^0(n)$.*

Proof. Suppose (T, s) is a gap tree with young vertices, let v be the leftmost one. We transform (T, s) into a gap tree $I((T, s)) = (U, t)$ of the same size and with no young vertex. Let (T_0, s) be the gap subtree of (T, s) rooted in the root r which is lying to the right of v. There is to distinguish two cases.

1. If there is nothing to the right of v — (T_0, s) consists of r only and $s(g) = 0$ for g the rightmost gap of r in T — we delete v and put $t(h) = s(g') + 1$ where h is now the rightmost gap of r in U and g' was the second rightmost gap of r in T (h arises by merging g and g'). The values of t on other gaps equal to those of s.

2. If there is anything to the right of v — (T_0, s) has more than one vertex or $s(g) > 0$ — (T_0, s) is cut off from r (r gets duplicated for a while) and is glued to v. We set $t(h) = 0$, on other gaps t retains the values of s.

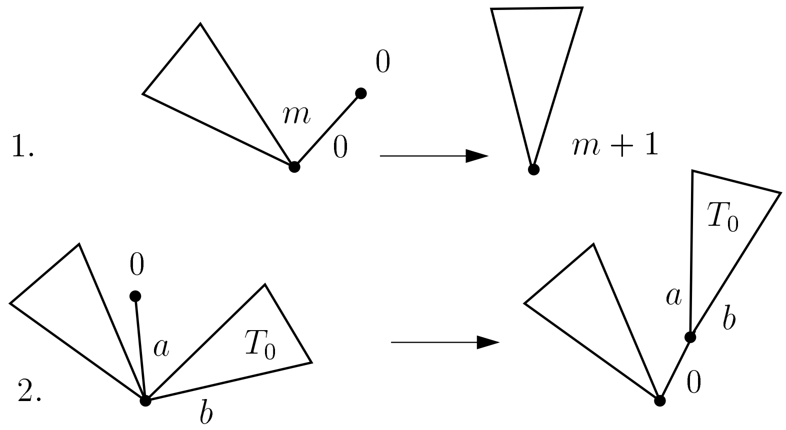

Figure 1: The bijection I.

The transformation is depicted schematicly on the above figure (0, a, b, and m stand for the values of s and t on the corresponding gaps). All young vertices are destroyed. To reconstruct (T, s) from (U, t) we check first whether $t(h) > 0$ for h the rightmost gap of the root of U. If yes we proceed backwards via (1) otherwise via (2). Hence I is a bijection. $\square$

It remains to work out the bijections between $\mathcal{GT}_0(n)$ and $\mathcal{GT}^0(n)$, and $\mathcal{GT}_1(n)$ and $\mathcal{GT}^1(n)$. We prove more. We present a bijection $J : \mathcal{GT}(n) \to \mathcal{GT}(n)$ that maps a gap tree with k solitary vertices to a gap tree with k young vertices.

We need few definitions. For T a tree the rightmost branch $(x_1, x_2, \ldots, x_k)$, $x_1 = r$, the final leaf x_{k+1} is omitted, is called the *right side of* T. The *top side of* T $(y_1, y_2, \ldots, y_l)$ consists of the vertices $y_1, \ldots, y_m$ of the highest layer, ordered from right to left, and of the vertices $y_{m+1}, \ldots, y_l$ of the highest but one layer which lie to the right of y_1's parent, again taken from right to left. An *encoding sequence* is a sequence $((a_1, b_1), \ldots, (a_m, b_m))$ of pairs of positive integers satisfying

$$b_1 = 1 \text{ and } b_i \leq a_{i-1} + b_{i-1} - 1 \text{ for } i = 2, \ldots, m.$$

Its *size* is $a_1 + a_2 + \dots a_m$. The one term sequence $((0,1))$ is defined to be an encoding sequence too.

Lemma 2.4 *There is a bijection $J: \mathcal{GT}(n) \to \mathcal{GT}(n)$ that maps a gap tree with k solitary vertices to a gap tree with k young vertices.*

Proof. Suppose $z = ((a_1, b_1), \dots, (a_m, b_m))$ is an encoding sequence. We show two ways to decode it and to obtain a tree T with $a_1 + a_2 + \dots + a_m$ edges.

The sequence $z = ((0,1))$ is decoded in both ways as the one vertex tree. We start the first decoding with drawing, from bottom to top, a path of a_1 edges. We denote this initial tree as T_1. In the general step, to derive T_{i+1} from T_i, we draw from bottom to top and to the right of T_i a path P of a_{i+1} edges starting in the b_{i+1}th vertex of the right side of T_i. P is clearly the final segment of the right side of T_{i+1}. On the end we set $T = T_m$. We denote this decoding as J_1. The order in which the edges of T are drawn is called the J_1-*order*.

The second decoding is a similar one, the difference being that T_1 is the broom of a_1 edges (the root has a_1 children, all of them are leaves) and that in the general step we join to the b_{i+1}th vertex of the top side of T_i a broom of a_{i+1} edges. Their endpoints become the initial segment of the top side of T_{i+1}. Each broom is drawn from right to left. This decoding is denoted as J_2, the J_2-*order* is defined analogously.

Both decodings are bijections from the set of encoding sequences of size n to $\mathcal{T}(n)$. Hence $J_3 = J_1 \circ J_2^{-1}$ is a bijection on $\mathcal{T}(n)$. Since solitary (young) vertices correspond in J_2 (in J_1) exactly to the terms $(1,1)$ of the encoding sequence, we conclude that J_3 has the property stated in the lemma. It remains to extend it to $\mathcal{GT}(n)$.

We define the bijection $J_3^*: g(T) \to g(J_3(T))$ as follows. Suppose g is the top (the bottom) gap of the mth edge, in the J_2-order, of T. We set $J_3^*(g)$ equal to the top (to the bottom) gap of the mth edge, in the J_1-order, of $J_3(T)$. The first gap of T is sent, of course, to the first gap of $J_3(T)$.

Finally, let $(T, s) \in \mathcal{GT}(n)$. We define $J((T,s)) = (J_3(T), t)$ where $t(J_3^*(g)) = s(g)$ for any $g \in g(T)$. Clearly g is the leftmost gap of a solitary vertex in T iff $J_3^*(g)$ is the gap of a young vertex in $J_3(T)$. Thus J has the property stated. $\square$

Our construction of the bijection $F: \mathcal{P}(n) \to \mathcal{U}(n)$ is complete: $F = H^{-1} \circ J^{-1} \circ I^{-1} \circ J \circ H$. We illustrate it for a specific partition on Fig. 2. In the top row the encoding sequence is $((2,1),(2,2),(1,2))$ and in the bottom row $((1,1),(1,1),(2,1),(1,1))$. In I^{-1} we proceed backwards via (2).

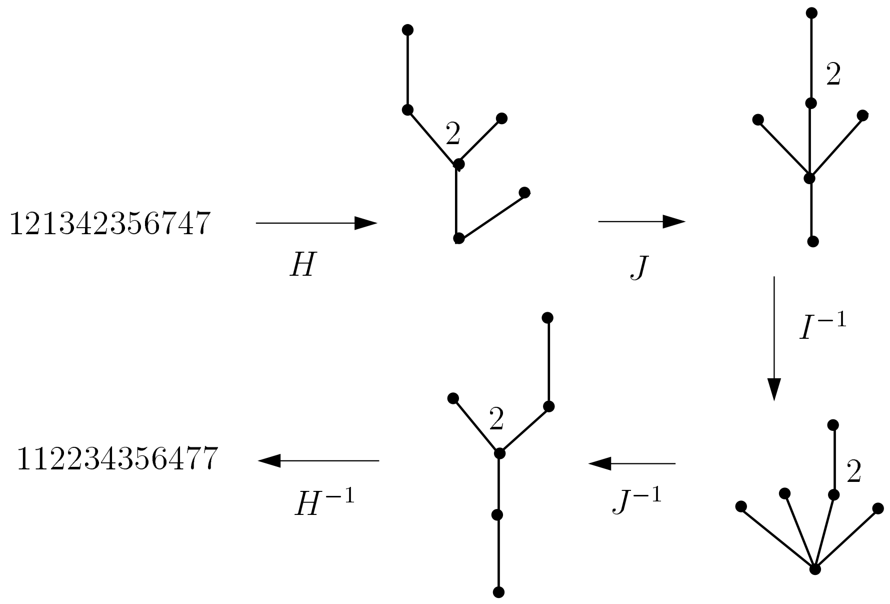

Figure 2: The bijection F.

3 Concluding remarks

The reader may wonder about the numbers $a_n = |\mathcal{P}(n)| = |\mathcal{U}(n)|$,

$$\{a_n\}_{n\geq 1} = \{1, 3, 11, 45, 197, 903, 4279, 20793, 103049, \ldots\}.$$

These are Schröder numbers [3], A1003 in [4], one of their explicit forms [2] is

$$a_n = \sum_{l=0}^{n-1} \frac{2^l}{n-l}\binom{n}{l+1}\binom{n-1}{l}.$$

The interested reader will find more references and expressions for Schröder numbers in [2] or in [4].

Our construction could be translated back to partitions but we prefer tree structures because they enable visual insight in the whole matter. We plan to prove along similar lines two other identities of [1] concerning *abba*-free and *abab*-free partitions.

References

[1] M. Klazar, On *abab*-free and *abba*-free set partitions, *Europ. J. of Combinatorics* **17** (1996), 53–68.

[2] M. Klazar, On numbers of Davenport-Schinzel sequences, submitted.

[3] E. Schröder, Vier combinatorische Probleme, *Zeitschrift für Mathematik und Physik* **15** (1870), 361–376.

[4] N. J. A. Sloane and collaborators, The On-Line Encyclopedia of Integer Sequences, `sequences@research.att.com`