Discrete Applied Mathematics 161 (2013) 1448–1466

Contents lists available at SciVerse ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

The distribution of cycles in breakpoint graphs of signed permutations

Simona Grusea [a], Anthony Labarre [b,*]

[a] Institut de Mathématiques de Toulouse, INSA de Toulouse, Université de Toulouse, France

[b] Université Paris-Est, LIGM UMR CNRS 8049, Université Paris-Est Marne-la-Vallée, 5 boulevard Descartes, 77420 Champs-sur-Marne, France

ARTICLE INFO

Article history:
Received 21 April 2011
Received in revised form 22 January 2013
Accepted 12 February 2013
Available online 6 March 2013

Keywords:
Genome rearrangement
Hultman number
Permutation

ABSTRACT

Breakpoint graphs are ubiquitous structures in the field of genome rearrangements. Their cycle decomposition has proved useful in computing and bounding many measures of (dis)similarity between genomes, and studying the distribution of those cycles is therefore critical to gaining insight on the distributions of the genomic distances that rely on it. We extend here the work initiated by Doignon and Labarre [6], who enumerated unsigned permutations whose breakpoint graph contains k cycles, to *signed* permutations, and prove explicit formulae for computing the expected value and the variance of the corresponding distributions, both in the unsigned case and in the signed case. We also show how our results can be used to derive simpler proofs of other previously known results.

1. Introduction

The field of comparative genomics is concerned with quantifying similarity or divergence between organisms. Several measures have been proposed to that end, including pattern matching based approaches or edit distances relying on a given set of biologically relevant operations. A standard example of such a method, and a *de facto* standard in phylogenetics, is the approach based on *sequence alignment*, which is motivated by the observation that genomes evolve by point mutations and aims at explaining evolution by replacements, insertions or deletions of single nucleotides (see e.g. Li and Homer [18] for a recent account of sequence alignment techniques and their uses).

However, genomes also evolve by large-scale mutations that act on whole segments of the genome, as opposed to point mutations. Examples of such mutations include *reversals*, which reverse the order of elements along a segment, *transpositions*, which move segments to another location, and *translocations*, which exchange segments that belong to different chromosomes. Many models have been proposed for studying those *genome rearrangements*, which vary according to the kinds of mutations one wants to take into account, how these should be weighted, or which objects are best suited for representing genomes (see e.g. Fertin et al. [8] for an extensive survey). Nonetheless, a striking similarity between all these models is how heavily they rely on variants of a graph first introduced by Bafna and Pevzner [1], known as the *breakpoint graph*, and its decomposition into edge- or vertex-disjoint cycles, which has proved most useful in obtaining extremely tight bounds on many genome rearrangement distances, as well as formulae for computing the exact distance in several cases. The link between several genomic distances and the number of cycles in breakpoint graphs is discussed in more detail by Fertin et al. [8].

Many mathematical questions arise when studying genome rearrangement distances, particularly concerning their distributions, as well as related statistical parameters. Since quite a few such distances can be computed or approximated

* Corresponding author.
E-mail addresses: grusea@insa-toulouse.fr (S. Grusea), Anthony.Labarre@univ-mlv.fr (A. Labarre).

0166-218X/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2013.02.002

using the cycle decomposition of the breakpoint graph, investigating the distribution of such cycles appears as a natural, general and effective starting point to answering those questions. We will restrict our attention in this paper to the permutation model, which can be used when all genomes under comparison consist of exactly the same genes (but in a different order) without duplications. Breakpoint graphs can be associated to permutations, and the distribution of cycles in this case was first characterised by Doignon and Labarre [6], which later led Bóna and Flynn [3] to prove a very simple expression for the expected value of the *block-interchange distance* originally introduced by Christie [4].

However, it has often been argued that *signed permutations* provide a more realistic model of evolution, since signs can be used to represent on which strand a given DNA segment is located. Using this model, Székely and Yang [22] obtained bounds for the expectation and the variance of the number of cycles in the breakpoint graph of a random signed permutation. Using the finite Markov chain embedding technique, Grusea [11] obtained the distribution of the number of cycles in the breakpoint graph of a random signed permutation in the form of a product of transition probability matrices of a certain finite Markov chain. Her method allows to derive recurrence formulae and to compute this distribution numerically, but the computational complexity is quite high and limits the practical applications.

In this work, we obtain a new expression for computing the number of unsigned permutations whose breakpoint graph contains a given number of cycles, as well as what is to the best of our knowledge the first analytic expression for computing the number of *signed* permutations whose breakpoint graph contains a given number of cycles. The formula obtained in the signed case is complicated, but we obtain simpler formulae for a couple of restricted cases. We also use our results to derive elementary proofs of previously known results, including a binomial identity and the distribution of the number of cycles in the breakpoint graph of an unsigned permutation. We prove formulae for computing the expected value and the variance of the distribution of those cycles, both in the unsigned case and in the signed case.

2. Notations and definitions

We recall here a few notions that will be used throughout the paper. We assume the reader is familiar with graph theory (if not, see e.g. Diestel [5]), but nevertheless review a few useful definitions, if only to agree on notation. We will work with *non-simple* graphs, i.e. graphs that may contain *loops* (edges connecting a vertex to itself) as well as parallel edges. We will also work with both undirected and directed graphs, using $\{u, v\}$ to denote edges in the former case and (u, v) to denote arcs in the latter.

Definition 2.1. A *matching* M in a graph $G = (V, E)$ is a subset of pairwise vertex-disjoint edges of E. It is a *perfect matching* of $U \subseteq V$ if every vertex in U is incident to an edge in M.

Definition 2.2. A graph is *k-regular* if each of its vertices has degree k.

In particular, if G is a 2-regular graph, then it decomposes in a unique way into a collection of edge- and vertex-disjoint cycles, up to the ordering of cycles and to rotations of elements within each cycle (i.e., $(a, b, c, d) = (b, c, d, a)$), as well as directions in which cycles are traversed if G is undirected (i.e., $(a, b, c, d) = (d, c, b, a)$). This allows us to denote unambiguously $c(G)$ the number of cycles in G. The *length* of a cycle is the number of vertices it contains, and a *k-cycle* in G is a cycle of length k.

Definition 2.3. A graph is *Hamiltonian* if it contains a cycle visiting every vertex exactly once.

We now recall a few basic notions about permutations (for more details, see e.g. Björner and Brenti [2] and Wielandt [23]).

Definition 2.4. A *permutation* of $\{1, 2, \ldots, n\}$ is a bijective application of $\{1, 2, \ldots, n\}$ onto itself.

The *symmetric group* S_n is the set of all permutations of $\{1, 2, \ldots, n\}$, together with the usual function composition $\circ$, applied from right to left. We use lower case Greek letters to denote permutations, typically $\pi = \langle \pi_1\ \pi_2\ \cdots\ \pi_n \rangle$, with $\pi_i = \pi(i)$, and in particular write the *identity permutation* as $\iota = \langle 1\ 2\ \cdots\ n \rangle$.

Definition 2.5. The *graph* $\Gamma(\pi)$ of a permutation $\pi \in S_n$ has vertex set $\{1, 2, \ldots, n\}$, and contains an arc (i, j) whenever $\pi_i = j$.

Definition 2.4 implies that $\Gamma(\pi)$ is 2-regular and as such decomposes in a unique way into disjoint cycles (up to the ordering of cycles and to rotations of elements within each cycle), which we refer to as the *disjoint cycle decomposition* of π. It is also common to refer to a permutation as a k-cycle, if the only cycle of length greater than 1 that its graph contains has length k. Fig. 1 shows an example of such a decomposition. To lighten the presentation, we will shorten the notation $c(\Gamma(\pi))$ into $c(\pi)$, for a given permutation π.

Definition 2.6. The *conjugate* of a permutation π by a permutation σ, both in S_n, is the permutation $\sigma \circ \pi \circ \sigma^{-1}$, and can be obtained by replacing every element i in the disjoint cycle decomposition of π with σ_i.

Definition 2.7. A *signed permutation* is a permutation of $\{1, 2, \ldots, n\}$ where each element has an additional "+" or "−" sign.

The *hyperoctahedral group* $S_n^{\pm}$ is the set of all signed permutations of n elements, together with the usual function composition $\circ$, applied from right to left. It is not mandatory for a signed permutation to have negative elements, so $S_n \subset S_n^{\pm}$

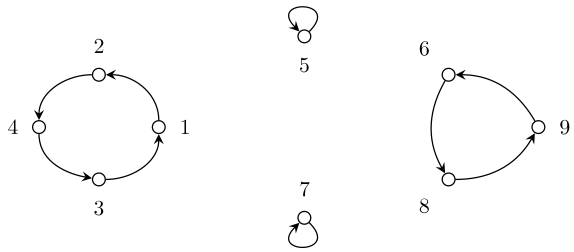

Fig. 1. The graph of the permutation $\pi = \langle 2\ 4\ 1\ 3\ 5\ 8\ 7\ 9\ 6 \rangle$.

since each permutation in S_n can be viewed as a signed permutation without negative elements. To lighten the presentation, we will conform to the tradition of omitting "+" signs for positive elements.

Finally, we recall the definition of the following graph introduced by Bafna and Pevzner [1], which turned out to be an extremely useful tool for studying and solving genome rearrangement problems and which will be central to our discussions.

Definition 2.8. Given a signed permutation π in $S_n^{\pm}$, transform it into an unsigned permutation π' in S_{2n} by mapping π_i onto the sequence $(2\pi_i - 1, 2\pi_i)$ if $\pi_i > 0$, or $(2|\pi_i|, 2|\pi_i| - 1)$ if $\pi_i < 0$, for $1 \le i \le n$. The *breakpoint graph* of π is the undirected bicoloured graph $\mathrm{BG}(\pi)$ with ordered vertex set $(\pi'_0 = 0, \pi'_1, \pi'_2, \ldots, \pi'_{2n}, \pi'_{2n+1} = 2n + 1)$ and whose edge set is the union of the following two perfect matchings of $V(\mathrm{BG}(\pi))$:

- black edges $\delta_B(\pi) = \{\{\pi'_{2i}, \pi'_{2i+1}\} \mid 0 \le i \le n\}$;
- grey edges $\delta_G = \{\{2i, 2i + 1\} \mid 0 \le i \le n\}$.

We will often use the notation $\mathrm{BG}(\pi) = \delta_B(\pi) \cup \delta_G$ to denote breakpoint graphs.

Genome rearrangement problems usually involve computing edit distances, i.e. the smallest number of moves needed to transform a genome into another one using only operations specified by a given set S. In the case of permutations, those distances are usually *left-invariant*, which intuitively means that genes can be relabelled so that either genome becomes ι without affecting the value of the distance to compute. Under this assumption, the pairwise genome rearrangement problem in $S_n^{\pm}$ can be viewed as a constrained sorting problem, and the intuition behind the breakpoint graph construction is that black edges are meant to represent the current situation (i.e. the ordering provided by π), while grey edges are meant to represent the target situation (i.e. the ordering provided by ι). Fig. 2 shows an example of a breakpoint graph. By definition, such a graph is a collection of even-length cycles that alternate black and grey edges. It can be easily seen that the example shown in Fig. 2 decomposes into two such cycles.

The *length* of a cycle in a breakpoint graph differs from the traditional graph-theoretical definition that we mentioned earlier: it is *half* the number of edges the cycle contains. Nevertheless, we will keep the terminology *k-cycle* to designate a cycle of length k, keeping in mind that its length is measured differently in the context of breakpoint graphs.

3. Cycle statistics

As is well-known (see e.g. Graham et al. [10]), the *unsigned Stirling number of the first kind* $\left[{n \atop k}\right]$ counts the number of permutations in S_n which decompose into k disjoint cycles:

$$\left[{n \atop k}\right] = |\{\pi \in S_n \mid c(\pi) = k\}|.$$

Recall also that those numbers arise as coefficients in the series expansion of the *rising factorial*

$$x^{\overline{n}} = x(x+1)\cdots(x+n-1) = \sum_{k=0}^{n} \left[{n \atop k}\right] x^k \tag{1}$$

and of the *falling factorial*

$$x^{\underline{n}} = x(x-1)\cdots(x-n+1) = \sum_{k=0}^{n} (-1)^{n-k} \left[{n \atop k}\right] x^k. \tag{2}$$

Signing the elements of a permutation does not change its disjoint cycle decomposition, so the number of *signed* permutations that decompose into k disjoint cycles is $2^n \left[{n \atop k}\right]$. We are interested in the following analogues of the Stirling number of the first kind, based on the cycle decomposition of the breakpoint graph.

Definition 3.1. The *Hultman number* $\mathcal{S}_H(n, k)$ counts the number of permutations in S_n whose breakpoint graph decomposes into k cycles:

$$\mathcal{S}_H(n, k) = |\{\pi \in S_n \mid c(\mathrm{BG}(\pi)) = k\}|.$$

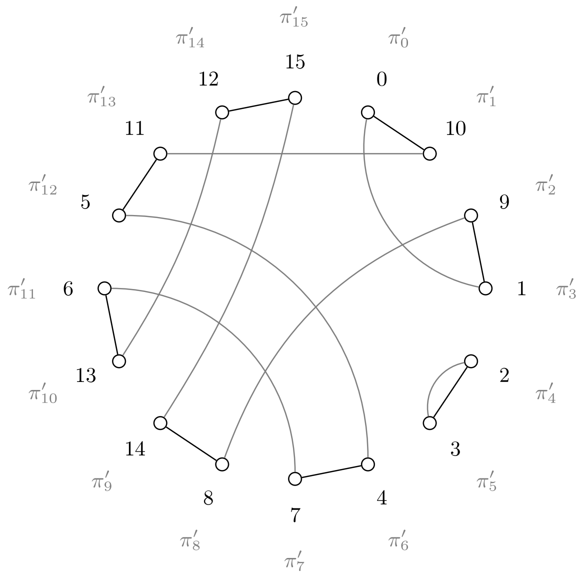

Fig. 2. The breakpoint graph of $\langle -5\ 1\ 2\ 4\ -7\ -3\ 6\rangle$.

The *signed Hultman number* $\mathcal{S}_H^{\pm}(n, k)$ counts the number of permutations in $S_n^{\pm}$ whose breakpoint graph decomposes into k cycles:

$$\mathcal{S}_H^{\pm}(n, k) = |\{\pi \in S_n^{\pm} \mid c(\mathrm{BG}(\pi)) = k\}|.$$

It is clear from Definition 2.8 that the number of cycles in any breakpoint graph is at least one and at most $n+1$. Hultman numbers were so named by Doignon and Labarre [6] after Axel Hultman, who first raised the question of computing those numbers [14]. The authors obtained an explicit but complicated formula for computing $\mathcal{S}_H(n, k)$, as well as formulae for enumerating permutations with a given "Hultman class" (the analogue of conjugacy classes of S_n based on the breakpoint graph). Bóna and Flynn [3] later observed that they can be computed using the following much simpler expression:

$$\mathcal{S}_H(n, k) = \begin{cases} \left[\begin{matrix} n+2 \\ k \end{matrix}\right] \Big/ \binom{n+2}{2} & \text{if } n-k \text{ is odd,} \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

based on a formula first obtained by Kwak and Lee [15].

In the next section, we present another way of obtaining an explicit formula for the unsigned Hultman numbers, which we will use in Section 7 to derive a new and simple proof of Eq. (3). In Section 5, we will prove the first explicit formula for computing the *signed* Hultman numbers.

4. A new formula for $\mathcal{S}_H(n, k)$

We will need the following results obtained by Hanlon et al. [12], whose notation we follow. For any fixed n in $\mathbb{N}_0$, let

$$Q_n^{\mathbb{C}}(h, \ell) = \mathbb{E}(\mathrm{Re}(\mathrm{tr}((VV^{\mathrm{t}})^n))),$$

where V is a random $h \times \ell$ matrix with independent standard complex normal entries, $\mathbb{E}$ denotes expectation, Re denotes real part, tr denotes trace and $^{\mathrm{t}}$ denotes matrix transposition. For the definition and the properties of the complex normal distribution, see for example Goodman [9].

Hanlon et al. [12] give two formulae for computing $Q_n^{\mathbb{C}}(h, \ell)$, both of which we will need. The first formula[1] is:

$$Q_n^{\mathbb{C}}(h, \ell) = \sum_{\omega \in S_n} h^{c(\omega)} \ell^{c(\omega \circ \omega_{(n)})}, \tag{4}$$

where $\omega_{(n)}$ is a fixed n-cycle in S_n. The second formula[2] is:

$$Q_n^{\mathbb{C}}(h, \ell) = \frac{1}{n} \sum_{i=1}^{n} (-1)^{i-1} \frac{(h+n-i)^{\underline{n}} (\ell+n-i)^{\underline{n}}}{(n-i)!(i-1)!}. \tag{5}$$

1 See Corollary 2.4 p. 158 of Hanlon et al. [12].
2 See Theorem 2.5 p. 158 of Hanlon et al. [12].

The link between the Hultman numbers and the previous results of Hanlon et al. [12] is obtained using the following result of Doignon and Labarre [6].

Corollary 4.1 (*[6]*). *$\mathcal{S}_H(n,k)$ counts the number of factorisations of a fixed $(n+1)$-cycle β into the product $\rho\circ\omega$, where ρ is an $(n+1)$-cycle and ω a permutation in S_{n+1} with $c(\omega)=k$.*

For a polynomial $P(x)$, let $[x^k]P(x)$ denote the coefficient of the monomial x^k in $P(x)$. We derive the following new expression for computing $\mathcal{S}_H(n,k)$.

Theorem 4.1. *For all n in $\mathbb{N}_0$, for all k in $\{1,2,\dots,n+1\}$:*

$$\mathcal{S}_H(n,k)=\frac{1}{n+1}\sum_{i=1}^{n+1}[h^k](h+n-i+1)^{\underline{n+1}}. \tag{6}$$

Proof. By Corollary 4.1, $\mathcal{S}_H(n,k)$ counts the number of factorisations of a fixed $(n+1)$-cycle β into the product $\rho\circ\omega$, with $c(\rho)=1$ and $c(\omega)=k$. This is clearly equivalent to enumerating factorisations of ρ^{-1} into the product $\omega\circ\beta^{-1}$ under the same conditions; therefore, setting $\omega_{(n+1)}$ to β^{-1} in Eq. (4), we observe that $\mathcal{S}_H(n,k)$ is the coefficient of the monomial $h^k\ell$ in the polynomial $Q^{\mathbb{C}}_{n+1}(h,\ell)$, hence by Eq. (5) equals:

$$\mathcal{S}_H(n,k)=\frac{1}{n+1}\sum_{i=1}^{n+1}(-1)^{i-1}\frac{[h^k](h+n-i+1)^{\underline{n+1}}\times[\ell](\ell+n-i+1)^{\underline{n+1}}}{(n-i+1)!(i-1)!}.$$

Since for every i in $\{1,2,\dots,k+1\}$ we have

$$\begin{aligned}[\ell](\ell+n-i+1)^{\underline{n+1}}&=[\ell](\ell+n-i+1)(\ell+n-i)\cdots(\ell+1)\ell(\ell-1)(\ell-2)\cdots(\ell-(i-1))\\&=(-1)^{i-1}(n-i+1)!(i-1)!,\end{aligned}$$

the above summation simplifies to the wanted expression, which completes the proof. □

Besides providing a new relation involving Hultman numbers, our new formula will prove useful in obtaining simple proofs of known results, as we will see in Sections 7 and 8. Moreover, we think that the interest of our formula also lies in the fact that the method used to prove it extends to the signed case.

5. An explicit formula for $\mathcal{S}^{\pm}_H(n,k)$

We now turn our attention to the problem of computing *signed* Hultman numbers, which we solve using ideas similar to those presented in the previous section. The result is obtained by characterising the 2-regular graphs that correspond to actual breakpoint graphs (Lemma 5.1), and then relating that characterisation to an enumeration result by Hanlon et al. [12].

5.1. Preliminaries

Following Hanlon et al. [12], for some fixed n in $\mathbb{N}_0$, let

$$Q^{\mathbb{R}}_n(h,\ell)=\mathbb{E}(\mathrm{tr}((VV^{t})^n)),$$

where V is again a random $h\times\ell$ matrix, but this time with independent standard *real* normal entries. Hanlon et al. [12] obtain two formulae for $Q^{\mathbb{R}}_n(h,\ell)$.

Let $\mathcal{F}_n$ denote the set of perfect matchings of $\{0,1,2,\dots,2n-1\}$. In particular, let $\varepsilon\in\mathcal{F}_n$ be the *identity perfect matching* $\{\{i,n+i\}\mid 0\le i\le n-1\}$. The first formula[3] for $Q^{\mathbb{R}}_n(h,\ell)$ is:

$$Q^{\mathbb{R}}_n(h,\ell)=\sum_{\delta\in\mathcal{F}_n}h^{c(\varepsilon\cup\delta)}\ell^{c(\delta\cup\delta_{(n)})}, \tag{7}$$

where $\delta_{(n)}$ is a fixed perfect matching such that $\varepsilon\cup\delta_{(n)}$ is Hamiltonian.

The second formula is based on partitions rather than on perfect matchings.

Definition 5.1 (*[19]*). A (integer) *partition* $\lambda=(\lambda_1,\lambda_2,\dots,\lambda_l)$ is a finite sequence of integers called *parts* such that $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_l\ge 0$. Its *length* is the number of non-zero parts it contains and if $\sum_{i=1}^{l}\lambda_i=n$, we call λ a *partition of* n, which we write as $\lambda\vdash n$.

[3] See Corollary 3.6 of Hanlon et al. [12].

We consider any two partitions to be equivalent if we obtain the same sequence when removing all parts that equal 0. The notation $\lambda = (1^{m_1} 2^{m_2} \dots r^{m_r})$ is also frequently used, and expresses the fact that exactly m_i parts of λ equal i. The reader must therefore bear in mind that when working with partitions, the notation a^b is more often to be understood in the previous meaning, and not as "a to the power b".

Notation: We introduce the notation $\mathcal{P}_n := \{\lambda \vdash n | \lambda_3 \leq 1\}$.

The second formula[4] for $Q_n^{\mathbb{R}}(h, \ell)$ is:

$$Q_n^{\mathbb{R}}(h, \ell) = \sum_{\lambda \in \mathcal{P}_n} c_\lambda(2) F_\lambda(h) F_\lambda(\ell), \tag{8}$$

where for any partition $\lambda \in \mathcal{P}_n$ of the form $(a, b, 1^{n-a-b})$, with either $a \geq b \geq 1$ or $a = n$ and $b = 0$:

- the function $F_\lambda : \mathbb{R} \to \mathbb{R}$ is defined as:

$$F_\lambda(x) = 2^{a-b}(x/2 + a - 1)^{\underline{a-b}}(x + 2b - 2)^{\underline{n-a+b}}, \tag{9}$$

- the coefficients $c_\lambda(2)$ are given as follows:

$$c_\lambda(2) = \frac{(-1)^{n+a-b+1} 2^{a-b+1} n(2a - 2b + 1)(a - 1)!}{(n + a - b + 1)^{\underline{2}}(n - a + b)^{\underline{2}}(n - a - b)!(2a - 1)!(b - 1)!}, \tag{10}$$

  if $a \geq b \geq 1$, and

$$c_\lambda(2) = \frac{2^n n!}{(2n)!}, \quad \text{if } \lambda = (n). \tag{11}$$

The numbers $c_\lambda(2)$ appear as coefficients in the expansion of the nth power-sum function in terms of zonal polynomials. For definitions and details, see for example Macdonald [19].

5.2. Characterising breakpoint graphs

Recall that a breakpoint graph is a 2-regular graph that is the union of two perfect matchings of $\{0, 1, \dots, 2n + 1\}$. We now make the connection between signed Hultman numbers and the previously mentioned results explicit.

Definition 5.2. A *configuration* is the union of two perfect matchings δ_B and δ_G of $\{0, 1, \dots, 2n+1\}$, where $\delta_G = \{\{2i, 2i+1\} \mid 0 \leq i \leq n\}$.

Note that the above definition only slightly generalises Definition 2.8, by allowing any choice of a perfect matching for δ_B, whereas there are implicit constraints on the choice of δ_B in the definition of the breakpoint graph. By definition, every breakpoint graph is a configuration, but not every configuration is a breakpoint graph, as we will see below shortly. The following notion will help us characterise configurations that are breakpoint graphs.

Definition 5.3. The *complement* of a configuration $C = \delta_B \cup \delta_G$, denoted by $\overline{C} = \delta_B \cup \overline{\delta_G}$, is obtained by replacing δ_G with $\overline{\delta_G} = \{\{2i - 1, 2i\} \mid 1 \leq i \leq n\} \cup \{\{0, 2n + 1\}\}$.

Before stating our characterisation of breakpoint graphs, we wish to stress that Elias and Hartman [7] previously used a similar but different notion of complementation (they replace δ_B with $\overline{\delta_B}$ – whose definition we will omit here – whereas we replace δ_G with $\overline{\delta_G}$) to characterise valid breakpoint graphs of *unsigned* permutations. This is not enough for our purpose, which is why we generalise their result below to encompass *signed* permutations as well.

Lemma 5.1. *A configuration $\delta_B \cup \delta_G$ is the breakpoint graph of some signed permutation π if and only if the complement configuration $\delta_B \cup \overline{\delta_G}$ is Hamiltonian.*

Proof. We can easily see that the complement $\overline{\mathrm{BG}(\pi)}$ of a breakpoint graph is Hamiltonian, since its edges are $\{\{\pi'_i, \pi'_{i+1}\} \mid 0 \leq i \leq 2n\} \cup \{\{0, 2n + 1\}\}$.

Conversely, if the complement $\delta_B \cup \overline{\delta_G}$ of a configuration is Hamiltonian, then we can recover the elements of an unsigned permutation $\pi' = \langle 0\ \pi'_1\ \pi'_2\ \cdots \pi'_{2n}\ 2n + 1\rangle$ by visiting the vertices along the Hamiltonian cycle as follows: take $0 = \pi'_0$ as starting point, and follow the edge in δ_B that is incident to 0, setting the value of π'_1 to the other endpoint of that edge. We then keep following the cycle $\delta_B \cup \overline{\delta_G}$, assigning the label of the ith encountered vertex to π'_i as we go, ending with $2n+1 = \pi'_{2n+1}$. Moreover, for every $0 \leq i \leq n - 1$, the edge $\{\pi'_{2i+1}, \pi'_{2i+2}\}$ belongs to $\overline{\delta_G}$, and hence the set $\{\pi'_{2i+1}, \pi'_{2i+2}\}$ is equal to $\{2j - 1, 2j\}$, for some $1 \leq j \leq n$. From the unsigned permutation π', we can therefore easily recover the corresponding signed permutation π in $S_n^{\pm}$, whose breakpoint graph is $\delta_B \cup \delta_G$. □

[4] See Theorem 5.4 of Hanlon et al. [12].

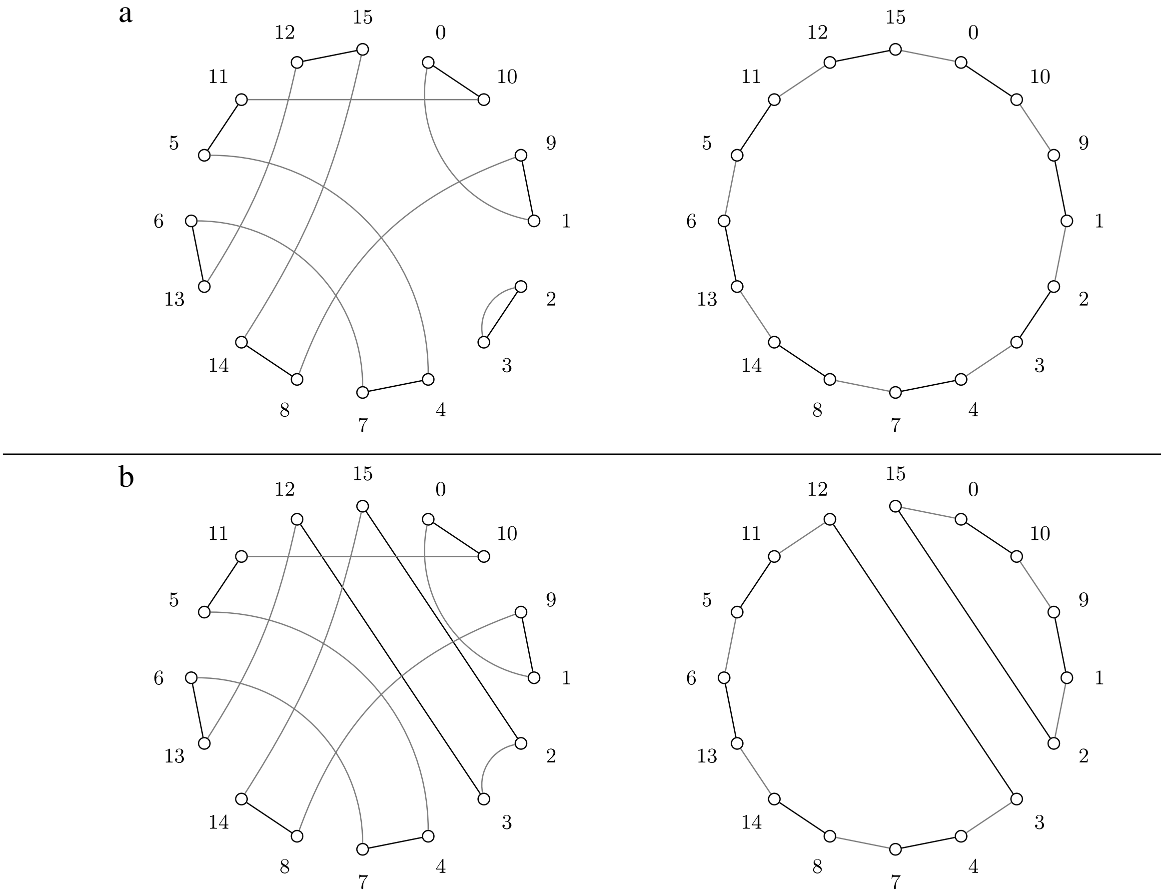

Fig. 3. (a) The complement of the breakpoint graph from Fig. 2 is Hamiltonian; (b) a configuration whose complement is not Hamiltonian.

Fig. 3(a) shows the complement of the breakpoint graph of Fig. 2, which is Hamiltonian. On the other hand, the complement of the configuration shown in Fig. 3(b) is not Hamiltonian.

5.3. Enumerating breakpoint graphs with k cycles

We now show that Eq. (7) remains valid when replacing the identity perfect matching ε with the perfect matching δ_G and choosing $\overline{\delta_G}$ as the fixed perfect matching $\delta_{(n+1)}$, which clearly satisfies the condition that $\delta_G \cup \overline{\delta_G}$ is Hamiltonian as required. The proof can be easily generalised to any choice of a perfect matching $\tau_{(n+1)}$ such that $\delta_G \cup \tau_{(n+1)}$ is Hamiltonian, but the following statement will be sufficient for our purposes.

Lemma 5.2. *For any n in* $\mathbb{N}_0$*:*

$$Q^{\mathbb{R}}_{n+1}(h, \ell) = \sum_{\tau \in \mathcal{F}_{n+1}} h^{c(\delta_G \cup \tau)} \ell^{c(\tau \cup \overline{\delta_G})}. \tag{12}$$

Proof. First, let us note that every perfect matching ϕ in $\mathcal{F}_{n+1}$ can be seen as a fixed-point-free involution, i.e. a permutation of $\{0, 1, 2, \ldots, 2n+1\}$ that decomposes into a collection of 2-cycles only, by viewing each edge of ϕ as a 2-cycle. Therefore, conjugating ϕ by any permutation of the same number of elements is a well-defined operation that simply renames the endpoints of the given edges. Let μ be the permutation defined by

$$\mu : \{0, 1, \ldots, 2n+1\} \to \{0, 1, \ldots, 2n+1\} : i \mapsto \mu(i) = \begin{cases} i/2 & \text{if } i \text{ is even,} \\ \dfrac{i + 2n + 1}{2} & \text{otherwise.} \end{cases}$$

As the example in Fig. 4 shows, δ_G can be mapped onto $\varepsilon = \mu \circ \delta_G \circ \mu^{-1}$, and we fix $\delta_{(n+1)} = \mu \circ \overline{\delta_G} \circ \mu^{-1}$. Finally, observe that given any two perfect matchings ϕ_1 and ϕ_2 in $\mathcal{F}_{n+1}$, the graphs $\mu \circ \phi_1 \circ \mu^{-1} \cup \mu \circ \phi_2 \circ \mu^{-1}$ and $\phi_1 \cup \phi_2$ are isomorphic,

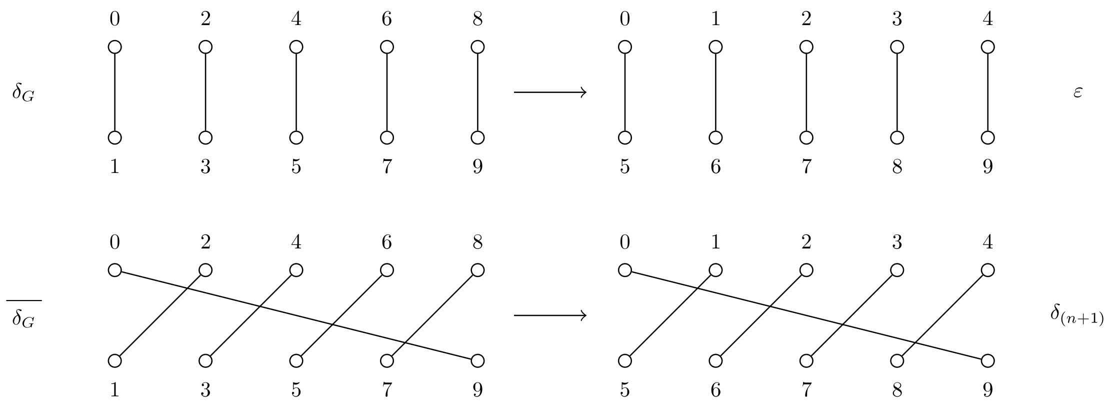

Fig. 4. Mapping δ_G (resp. $\overline{\delta_G}$) onto ε (resp. $\delta_{(n+1)}$) by conjugating them by $\mu = \langle 0\,5\,1\,6\,2\,7\,3\,8\,4\,9\rangle$.

and hence $c(\mu\circ\phi_1\circ\mu^{-1}\cup\mu\circ\phi_2\circ\mu^{-1}) = c(\phi_1\cup\phi_2)$. Taking $\delta = \mu\circ\tau\circ\mu^{-1}$, the following relations hold:

- $c(\varepsilon\cup\delta) = c(\mu\circ\delta_G\circ\mu^{-1}\cup\mu\circ\tau\circ\mu^{-1}) = c(\delta_G\cup\tau)$,
- $c(\delta\cup\delta_{(n+1)}) = c(\mu\circ\tau\circ\mu^{-1}\cup\mu\circ\overline{\delta_G}\circ\mu^{-1}) = c(\tau\cup\overline{\delta_G})$,
- $c(\varepsilon\cup\delta_{(n+1)}) = c(\mu\circ\delta_G\circ\mu^{-1}\cup\mu\circ\overline{\delta_G}\circ\mu^{-1}) = c(\delta_G\cup\overline{\delta_G}) = 1$,

and the formula in the statement follows from the above relations, the bijectivity of conjugation, and Eq. (7). □

Lemma 5.1 implies that enumerating signed permutations of n elements whose breakpoint graph decomposes into k alternating cycles is equivalent to enumerating perfect matchings τ in $\mathcal{F}_{n+1}$ verifying $c(\delta_G\cup\tau) = k$ and $c(\tau\cup\overline{\delta_G}) = 1$, where δ_G is defined in Definition 2.8 and $\overline{\delta_G}$ is defined in Definition 5.3. Using Lemma 5.2, we thus obtain the following.

Remark 5.1. For every k in $\{1, 2, \ldots, n+1\}$, $\mathcal{S}_H^{\pm}(n,k)$ is the coefficient of the monomial $h^k\ell$ in $Q^{\mathbb{R}}_{n+1}(h,\ell)$.

The second expression for $Q^{\mathbb{R}}_{n+1}(h,\ell)$ given in Eq. (8) allows us to obtain the following explicit formula for $\mathcal{S}_H^{\pm}(n,k)$.

Theorem 5.1. *For all n in $\mathbb{N}_0$, for all k in $\{1, 2, \ldots, n+1\}$:*

$$\mathcal{S}_H^{\pm}(n,k) = \sum_{\lambda\in\mathcal{P}_{n+1}} c_\lambda(2)\times[h^k]F_\lambda(h)\frac{(-1)^{n-a-b}2^{a-b-1}(2b)!(a-1)!(n-a-b+2)!}{(2b-1)b!}, \tag{13}$$

where $\lambda \in \mathcal{P}_{n+1}$, and where the function $F_\lambda(\cdot)$ as well as the coefficients $c_\lambda(2)$ follow the definitions previously given in Section 5.1.[5]

Proof. Remark 5.1 and Eq. (8) yield

$$\mathcal{S}_H^{\pm}(n,k) = \sum_{\lambda\in\mathcal{P}_{n+1}} c_\lambda(2)\times[h^k]F_\lambda(h)\times[\ell]F_\lambda(\ell). \tag{14}$$

For a partition $\lambda\in\mathcal{P}_{n+1}$ of the form $(a, b, 1^{n-a-b+1})$, with $a\geq b\geq 1$ or $a = n+1$, $b = 0$, it is easy to see that

$$[\ell]F_\lambda(\ell) = \frac{(-1)^{n-a-b}2^{a-b-1}(2b)!(a-1)!(n-a-b+2)!}{(2b-1)b!}. \tag{15}$$

Indeed:

1. if $\lambda = (a, b, 1^{n-a-b+1})$, with $a\geq b\geq 1$, we have

$$\begin{aligned}F_\lambda(\ell) &= 2^{a-b}(\ell/2+a-1)(\ell/2+a-2)\cdots(\ell/2+b)(\ell+2b-2)(\ell+2b-3)\cdots(\ell+1)\\ &\quad\times\ell(\ell-1)\cdots(\ell-(n-a-b+2)).\end{aligned}$$

The coefficient of ℓ in the above expression equals

$$\begin{aligned}[\ell]F_\lambda(\ell) &= 2^{a-b}\frac{(a-1)!}{(b-1)!}\times(2b-2)!(-1)^{n-a-b+2}(n-a-b+2)!\\ &= \frac{(-1)^{n-a-b}2^{a-b-1}(2b)!(a-1)!(n-a-b+2)!}{(2b-1)b!}.\end{aligned}$$

[5] With the slight modification that n needs to be replaced with $n+1$, since $\lambda\in\mathcal{P}_{n+1}$.

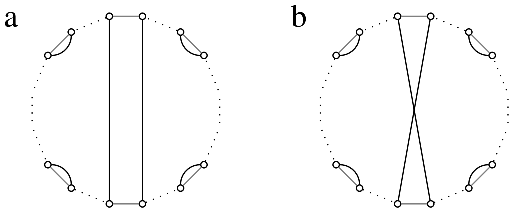

Fig. 5. The two forms of 2-cycles that may arise in a configuration. Only four 1-cycles are shown in each graph, but there can be any number of them.

2. if $\lambda = (n+1)$, i.e. $a = n+1$ and $b = 0$, we have
$$F_{(n+1)}(\ell) = 2^{n+1}(\ell/2+n)^{\underline{n+1}}(\ell-2)^{\underline{0}}$$
$$= 2^{n+1}(\ell/2+n)(\ell/2+n-1)\cdots(\ell/2+1)\ell/2,$$
so $[\ell]F_{(n+1)}(\ell) = 2^n n!$, which verifies Eq. (15).

The proof then follows from Eqs. (14) and (15). □

We conclude this section with Table 1, which shows a few experimental values of the signed Hultman numbers. These values were previously obtained by the first author using the method described in a previous paper of hers [11].

Note that for $k = 1$, the sequence defined by $\mathcal{S}_H^{\pm}(n, 1)$ for $n = 1, 2, \ldots$ corresponds to sequence A001171 in the On-Line Encyclopedia of Integer Sequences [20]. As we will see in the next section, other known sequences also appear in that table.

6. Special cases

The expression obtained in Theorem 5.1 allows us to compute $\mathcal{S}_H^{\pm}(n, k)$ for all valid values of n and k, but we must acknowledge that even though the formula is suited for practical use, it is unfortunately quite complicated and difficult to manipulate. Simpler expressions do however exist for some particular cases, as we will show below. We will rely a lot on Lemma 5.1 in this section, and decide to use a slightly different layout for the breakpoint graph: labels are omitted for clarity, and grey edges rather than black edges are now laid out on a circle, so that computing the complement of a given configuration simply amounts to shifting grey edges sideways by one position. In order to make verifications easier for the reader, we also draw edges in the complement as dotted edges. The following particular cases are easy to verify:

1. $\mathcal{S}_H^{\pm}(n, k) = 0$ for all $k < 1$ and all $k > n+1$ (trivial);
2. $\mathcal{S}_H^{\pm}(n, n+1) = 1$, since the only permutation whose breakpoint graph decomposes into $n+1$ cycles is ι;
3. $\mathcal{S}_H^{\pm}(n, n) = \binom{n+1}{2}$, since enumerating such permutations comes down to counting breakpoint graphs whose cycles all have length 1, except for one that has length 2. This in turn is equivalent to enumerating the ways in which one can connect any two of the $n+1$ grey edges by black edges so as to obtain a valid configuration (with respect to Lemma 5.1); as can be verified on Fig. 5, only one of the two possible choices of black edges (namely, configuration (b)) is valid, and the equality follows from the fact that there are $\binom{n+1}{2}$ possible ways to select two grey edges out of $n+1$.

We now show how one can obtain a simple and explicit formula for $\mathcal{S}_H^{\pm}(n, n-1)$. Although the formula is quite simple, we hope that the proof will convince the reader of the shortcomings of a case analysis in this setting.

Proposition 6.1. *For all $n \geq 1$, we have $\mathcal{S}_H^{\pm}(n, n-1) = 5\binom{n+1}{4} + 4\binom{n+1}{3}$.*

Proof. Note that $\mathcal{S}_H^{\pm}(n, n-1)$ is the number of permutations whose breakpoint graph contains either one 3-cycle or two 2-cycles, all other cycles having length 1 in both cases:

1. the number of permutations satisfying the first condition is the number of ways to connect three grey edges in the breakpoint graph in such a way that the complement configuration is Hamiltonian (see Lemma 5.1). As Fig. 6 shows, there are eight possible ways to create such a configuration, only four of which are valid (namely, configurations (a)–(d)). The reader can easily verify that the other configurations are invalid by checking that there is no Hamiltonian cycle in the graph whose edges are the union of the black edges δ_B and of the dotted edges $\overline{\delta_G}$.
We obtain the rightmost term in the wanted expression by noting that only four of the eight possible 3-cycles are valid, and there are $\binom{n+1}{3}$ ways to select three grey edges out of $n+1$.
2. the number of permutations satisfying the second condition can be constructed by choosing four grey edges, then connecting them by pairs while ensuring that the resulting configuration is valid. Fig. 7 shows all possible configurations with two cycles of length two.
The reader can again easily verify the validity of all configurations by checking whether there is a Hamiltonian cycle in the graph whose edges are the union of the black edges δ_B and of the dotted edges $\overline{\delta_G}$. Only five possible configurations

Table 1
A few values of $\mathcal{S}^{\pm}_H(n, k)$.

n	k											
	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1										
2	4	3	1									
3	20	21	6	1								
4	148	160	65	10	1							
5	1 348	1 620	701	155	15	1						
6	15 104	19 068	9 324	2 247	315	21	1					
7	198 144	264 420	138 016	38 029	5 908	574	28	1				
8	2 998 656	4 166 880	2 325 740	692 088	124 029	13 524	966	36	1			
9	51 290 496	74 011 488	43 448 940	13 945 700	2 723 469	344 961	27 930	1 530	45	1		
10	979 732 224	1 459 381 440	897 020 784	305 142 068	64 711 856	8 996 295	850 905	53 262	2 310	55	1	
11	20 661 458 688	31 674 232 128	20 241 273 264	7 255 047 116	1 640 552 028	249 029 717	26 004 330	1 910 403	95 304	3 355	66	1

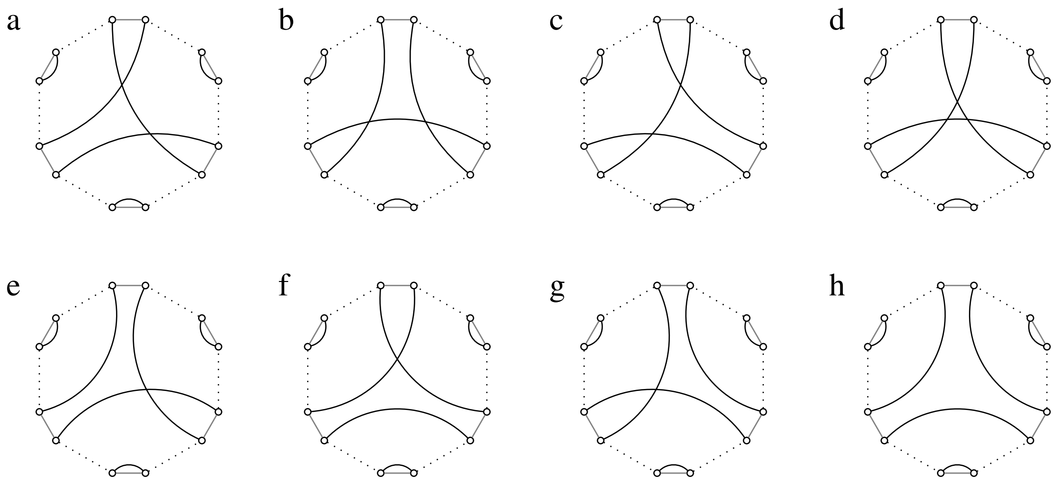

Fig. 6. All possible forms of 3-cycles that may arise in a configuration. Only three 1-cycles are shown in each graph, but there can be any number of them.

with two 2-cycles are valid (namely, configurations (b),(f), (i), (k) and (l)) out of the twelve shown in Fig. 7, and there are $\binom{n+1}{4}$ ways to select two pairs of grey edges out of $n+1$, which yields the leftmost term in the wanted expression and completes the proof. □

7. Simpler proofs of previous results

Theorem 4.1 allows us to obtain a new proof of Bóna and Flynn's formula (Eq. (3)).

Corollary 7.1 ([3]). *For all n in* $\mathbb{N}_0$*:*

$$\mathcal{S}_H(n,k) = \begin{cases} \left[{n+2 \atop k}\right] \Big/ \binom{n+2}{2} & \text{if } n-k \text{ is odd,} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The key idea of the proof is the fact that, for every $i = 1, 2, \ldots, n+1$, we have

$$(h+n-i+1)^{\underline{n+1}} = \frac{1}{n+2}\left((h-i+1)^{\overline{n+2}} - (h-i)^{\overline{n+2}}\right), \tag{16}$$

since

$$\begin{aligned}\frac{1}{n+2}\left((h-i+1)^{\overline{n+2}} - (h-i)^{\overline{n+2}}\right) &= \frac{1}{n+2}\left((h-i+1)\cdots(h+n-i+2) - (h-i)\cdots(h+n-i+1)\right) \\ &= \frac{1}{n+2}(h-i+1)\cdots(h+n-i+1)\left((h+n-i+2)-(h-i)\right) \\ &= (h+n-i+1)^{\underline{n+1}}.\end{aligned}$$

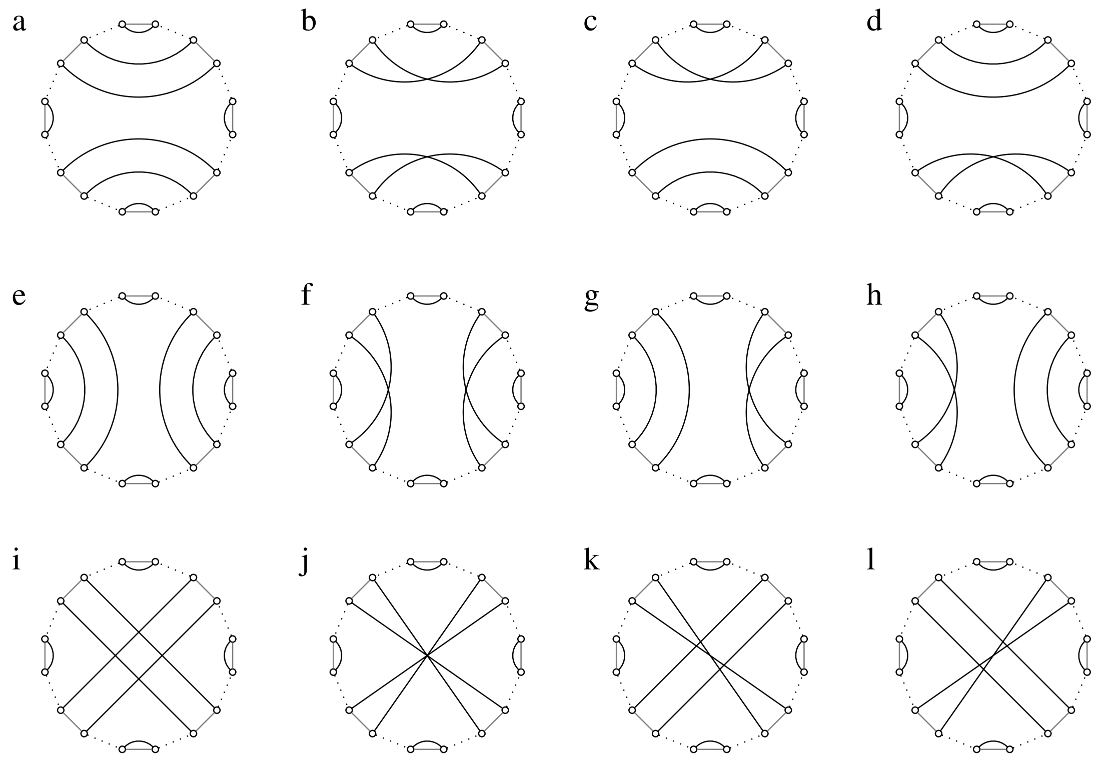

Fig. 7. All possible pairs of 2-cycles that may arise in a configuration. Only four 1-cycles are shown in each graph, but there can be any number of them.

Summing over i in Eq. (16), we obtain:

$$\frac{1}{n+1}\sum_{i=1}^{n+1}(h+n-i+1)^{\underline{n+1}} = \frac{1}{(n+1)(n+2)}\sum_{i=1}^{n+1}\left((h-i+1)^{\overline{n+2}}-(h-i)^{\overline{n+2}}\right)$$
$$= \frac{1}{(n+1)(n+2)}\left(h^{\overline{n+2}}-(h-n-1)^{\overline{n+2}}\right)$$
$$= \frac{1}{(n+1)(n+2)}\left(h^{\overline{n+2}}-h^{\underline{n+2}}\right).$$

By Eqs. (1) and (2), the coefficient of h^k in $h^{\overline{n+2}}$ is $\left[{n+2 \atop k}\right]$ and the coefficient of h^k in $h^{\underline{n+2}}$ is $(-1)^{n-k}\left[{n+2 \atop k}\right]$. Using Eq. (6), we conclude that

$$\mathcal{S}_H(n,k) = \begin{cases} \dfrac{2}{(n+1)(n+2)}\begin{bmatrix} n+2 \\ k \end{bmatrix} & \text{if } n-k \text{ is odd,} \\ 0 & \text{otherwise,} \end{cases}$$

which completes the proof. □

Theorem 4.1 also allows us to obtain a simple proof of a binomial identity previously obtained by Sury et al. [21].

Corollary 7.2 (*[21]*)**.** *For all n in $\mathbb{N}_0$:*

$$\sum_{i=0}^{n}\frac{(-1)^i}{\binom{n}{i}} = (1+(-1)^n)\frac{n+1}{n+2}.$$

Proof. Setting k to 1 in Eq. (6) yields

$$\mathcal{S}_H(n,1) = \frac{1}{n+1}\sum_{i=1}^{n+1}(-1)^{i-1}(n-i+1)!(i-1)! = \frac{n!}{n+1}\sum_{i=0}^{n}\frac{(-1)^i}{\binom{n}{i}}.$$

On the other hand, as previously observed[6] by Doignon and Labarre [6], we have:

$$\mathcal{S}_H(n,1) = \begin{cases} \dfrac{2n!}{n+2} & \text{if } n \text{ is even,} \\ 0 & \text{otherwise,} \end{cases}$$

which completes the proof. □

8. Expected value and variance of the Hultman numbers

In order to gain more insight into the distribution of the Hultman numbers, we will now investigate the question of computing the expected value and variance of the number of cycles in breakpoint graphs, both for unsigned and for signed permutations.

It will also be interesting to see how these values compare to the expected value and variance of the number of cycles in the usual disjoint cycle decomposition of a uniform random unsigned permutation π in S_n. We recall here (see e.g. Wilf [24]) the exact values of these quantities:

$$\mathbb{E}(c(\pi)) = H_n,$$
$$\mathrm{Var}(c(\pi)) = H_n - \sum_{k=1}^{n} \frac{1}{k^2},$$

as well as their asymptotic behaviour when $n \to \infty$:

$$\mathbb{E}(c(\pi)) = \log(n) + \gamma + o(1), \tag{17}$$

$$\mathrm{Var}(c(\pi)) = \log(n) + \gamma - \frac{\pi^2}{6} + o(1), \tag{18}$$

where H_n denotes the nth harmonic number $H_n = \sum_{i=1}^{n} \frac{1}{i}$ and γ denotes the Euler–Mascheroni constant. As usual, $o(1)$ denotes a quantity that converges to 0 as $n \to \infty$.

8.1. The unsigned case

Bóna and Flynn [3] already proved a formula for computing the expected number of cycles in the breakpoint graph of a uniform random unsigned permutation. In this section we provide a new proof of their result and also give an explicit formula for the variance of this distribution. We start by computing the generating function of the Hultman numbers.

Lemma 8.1. *For all $n \in \mathbb{N}_0$, we have:*

$$F(x) = \sum_{k=0}^{n+1} \mathcal{S}_H(n,k)x^k = \frac{x^{\overline{n+2}} - x^{\underline{n+2}}}{2\binom{n+2}{2}}.$$

Proof. The derivation is straightforward:

$$\begin{aligned}
\sum_{k=0}^{n+1} \mathcal{S}_H(n,k)x^k &= \frac{1}{\binom{n+2}{2}} \sum_{k=0}^{n+1} \frac{\left[{n+2 \atop k}\right] - (-1)^{n+2-k}\left[{n+2 \atop k}\right]}{2} x^k \quad \text{(by Eq. (3))} \\
&= \frac{1}{2\binom{n+2}{2}} \left(\sum_{k=0}^{n+2} \left[{n+2 \atop k}\right] x^k - \sum_{k=0}^{n+2} (-1)^{n+2-k} \left[{n+2 \atop k}\right] x^k \right) \\
&= \frac{x^{\overline{n+2}} - x^{\underline{n+2}}}{2\binom{n+2}{2}} \quad \text{(by Eqs. (1) and (2)).} \quad \square
\end{aligned}$$

Knowing the generating function allows us to easily derive the expected value and the variance of the number of cycles in the breakpoint graph of a random unsigned permutation. For this purpose, we first need to compute some derivatives of the generating function.

[6] The result can also be easily derived from Eq. (3).

Lemma 8.2. *For all $n \in \mathbb{N}_0$, we have:*

$$F(1) = n!,$$
$$F'(1) = \frac{1}{2\binom{n+2}{2}} \left\{(n+2)!H_{n+2} + (-1)^{n-1}n!\right\},$$
$$F''(1) = \frac{1}{2\binom{n+2}{2}} \left\{(n+2)!\left(H_{n+2}^2 - \sum_{k=1}^{n+2}\frac{1}{k^2}\right) + 2(-1)^n n!(H_n - 1)\right\}.$$

Proof. We obtain the three expressions separately.

1. For the first expression, note that, by definition, $F(1) = \sum_{k=1}^{n+1} \mathcal{S}_H(n,k)$, which is simply the total number of permutations of n elements and therefore equals $n!$
2. We simplify the computation of $F'(x)$ by writing $x^{\underline{n+2}} = (x-1)g(x)$, with

$$g(x) = x\prod_{i=2}^{n+1}(x-i).$$

With this notation we have

$$F(x) = \frac{x^{\overline{n+2}} - (x-1)g(x)}{2\binom{n+2}{2}}.$$

We thus obtain

$$F'(x) = \frac{1}{2\binom{n+2}{2}}\left(x^{\overline{n+2}}\sum_{i=0}^{n+1}\frac{1}{x+i} - g(x) - (x-1)g'(x)\right).$$

At $x = 1$ we have $1^{\overline{n+2}} = (n+2)!$ and $g(1) = (-1)^n n!$, and hence the stated formula for $F'(1)$ follows.
3. Finally, the second derivative of F is given by

$$F''(x) = \frac{1}{2\binom{n+2}{2}}\left(x^{\overline{n+2}}\sum_{0\le i\ne j\le n+1}\frac{1}{(x+i)(x+j)} - 2g'(x) - (x-1)g''(x)\right).$$

The above sum evaluated at $x = 1$ equals

$$\begin{aligned}\sum_{0\le i\ne j\le n+1}\frac{1}{(1+i)(1+j)} &= \sum_{i,j=0}^{n+1}\frac{1}{(1+i)(1+j)} - \sum_{i=0}^{n+1}\frac{1}{(1+i)^2}\\ &= \left(\sum_{i=0}^{n+1}\frac{1}{1+i}\right)^2 - \sum_{i=0}^{n+1}\frac{1}{(1+i)^2}\\ &= H_{n+2}^2 - \sum_{k=1}^{n+2}\frac{1}{k^2}.\end{aligned}$$

We also have

$$g'(x) = g(x)\left(\frac{1}{x} + \sum_{i=2}^{n+1}\frac{1}{x-i}\right),$$

and thus

$$g'(1) = g(1)\left(1 - \sum_{i=2}^{n+1}\frac{1}{i-1}\right) = (-1)^n n!(1-H_n).$$

Using these expressions in the formula for $F''(x)$ above, evaluated at $x = 1$, gives the formula in the statement. □

The recovery of the expected number of cycles in the breakpoint graph of a random unsigned permutation, previously obtained by Bóna and Flynn [3], is now an easy task.

Theorem 8.1 (*[3]*). *For all $n \in \mathbb{N}_0$, the expected number of cycles in the breakpoint graph of a uniform random unsigned permutation π of n elements is*

$$\mathbb{E}(c(\mathrm{BG}(\pi))) = H_n + \frac{1}{\lfloor (n+2)/2 \rfloor}.$$

Proof. As is well-known (see e.g. Wilf [24]), the expected value can be obtained from the generating function $F(x)$ by the formula $F'(1)/F(1)$. Using the formulae for $F(1)$ and $F'(1)$ obtained in Lemma 8.2, we obtain that the expected value $\mathbb{E}(c(\mathrm{BG}(\pi)))$ equals

$$\frac{F'(1)}{F(1)} = H_{n+2} + \frac{(-1)^{n-1}}{(n+1)(n+2)},$$

which is easily seen to be equivalent to the expression in the statement. □

Furthermore, knowing the generating function also allows us to compute the variance of the number of cycles in the breakpoint graph. We prove the following result.

Theorem 8.2. *For all $n \in \mathbb{N}_0$, the variance of the number of cycles in the breakpoint graph of a uniform random unsigned permutation π of n elements is*

$$\mathrm{Var}(c(\mathrm{BG}(\pi))) = H_{n+2} - \sum_{k=1}^{n+2} \frac{1}{k^2} + \frac{(-1)^n(2H_{n+2} + 2H_n - 3)}{(n+1)(n+2)} - \frac{1}{((n+1)(n+2))^2}.$$

Proof. The variance can be obtained from the generating function $F(x)$ by the following formula (see e.g. Wilf [24]):

$$(\log F)'(1) + (\log F)''(1) = \frac{F'(1)}{F(1)} + \frac{F''(1)}{F(1)} - \left(\frac{F'(1)}{F(1)}\right)^2.$$

Using the formulae for $F(1)$, $F'(1)$ and $F''(1)$ obtained in Lemma 8.2, we obtain that $\mathrm{Var}(c(\mathrm{BG}(\pi)))$ equals

$$\begin{aligned}\frac{F'(1)}{F(1)} + \frac{F''(1)}{F(1)} - \left(\frac{F'(1)}{F(1)}\right)^2 &= H_{n+2} + \frac{(-1)^{n-1}}{(n+1)(n+2)} + H_{n+2}^2 - \sum_{k=1}^{n+2} \frac{1}{k^2} + \frac{2(-1)^n(H_n - 1)}{(n+1)(n+2)} \\ &\quad - \left(H_{n+2} + \frac{(-1)^{n-1}}{(n+1)(n+2)}\right)^2 \\ &= H_{n+2} - \sum_{k=1}^{n+2} \frac{1}{k^2} + \frac{(-1)^n(2H_{n+2} + 2H_n - 3)}{(n+1)(n+2)} - \frac{1}{((n+1)(n+2))^2}. \quad \square\end{aligned}$$

It is interesting to see how the mean and variance behave for large n.

Remark 8.1. The expected value and variance of the number of cycles in the breakpoint graph of a uniform random unsigned permutation π in S_n have the following asymptotical behaviour when $n \to \infty$:

$$\begin{aligned}\mathbb{E}(c(\mathrm{BG}(\pi))) &= \log(n) + \gamma + o(1), \\ \mathrm{Var}(c(\mathrm{BG}(\pi))) &= \log(n) + \gamma - \frac{\pi^2}{6} + o(1).\end{aligned}$$

Proof. For the expected value, the result simply follows from the fact that $\mathbb{E}(c(\mathrm{BG}(\pi))) = H_n + o(1)$ and $H_n = \log(n) + \gamma + o(1)$.

For the variance, first note that $\mathrm{Var}(c(\mathrm{BG}(\pi))) = H_{n+2} - \sum_{k=1}^{n+2} \frac{1}{k^2} + o(1)$.

By further using the fact that $\log(n+2) = \log(n) + o(1)$ and the well-known result $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$, the stated asymptotic formula follows. □

Interestingly, we recover exactly the same asymptotical behaviour as for the number of cycles in the usual disjoint cycle decomposition (recall Eqs. (17) and (18)).

8.2. The signed case

We now turn to the problem of computing the expected value and the variance of the number of cycles in the breakpoint graph of a uniform random signed permutation. As in the unsigned case, we start with the computation of the generating function for the signed Hultman numbers.

Lemma 8.3. *We have*

$$G(x) = \sum_{k=1}^{n+1} S_H^{\pm}(n,k)x^k = \sum_{\lambda \in \mathcal{P}_{n+1}} c_\lambda(2) F_\lambda(x) F'_\lambda(0),$$

where $c_\lambda(2)$ and F_λ are defined as in Eq. (9).

Proof. Recall (Remark 5.1) that $S_H^{\pm}(n,k)$ is the coefficient of the monomial $h^k\ell$ in the polynomial $Q_{n+1}^{\mathbb{R}}(h,\ell)$. If we take now $h = x$ and consider $Q_{n+1}^{\mathbb{R}}(x,\ell)$ as a polynomial only in the variable ℓ, we note that the coefficient of the monomial ℓ is obtained by summing up all the terms $S_H^{\pm}(n,k)x^k$, for $k = 1, \ldots, n+1$. Therefore, $G(x)$ equals the coefficient of ℓ in $Q_{n+1}^{\mathbb{R}}(x,\ell)$, and hence

$$G(x) = \frac{\partial}{\partial \ell} Q_{n+1}^{\mathbb{R}}(x,\ell)\bigg|_{\ell=0}.$$

The formula in the statement easily follows from Eq. (8). □

In order to compute the expected value and the variance of the number of cycles in the breakpoint graph of a random signed permutation, we will need the following preliminary lemma.

Lemma 8.4. *Let $n \geq 1$ and $\lambda \in \mathcal{P}_{n+1}$ a partition of $n+1$ of the form $(a, b, 1^{n-a-b+1})$.*

1. *In the case where $a \geq b \geq 1$, we have:*

$$F'_\lambda(0) = \frac{(-1)^{n-a-b} 2^{a-b} (a-1)!(2b-2)!(n-a-b+2)!}{(b-1)!},$$
$$F'_\lambda(1) = \frac{(-1)^{n-a-b+1}(2a-1)!(b-1)!(n-a-b+1)!}{2^{a-b}(a-1)!},$$
$$F''_\lambda(1) = F'_\lambda(1)\left\{2H_{2a-1} - 2H_{n-a-b+1} - H_{a-1} + H_{b-1}\right\}.$$

2. *In the case where $\lambda = (n+1)$, we have:*

$$F'_{(n+1)}(0) = 2^n n!,$$
$$F'_{(n+1)}(1) = \frac{(2n+1)!}{2^n n!}(H_{2n+1} - H_n/2),$$
$$F''_{(n+1)}(1) = \frac{(2n+1)!}{2^n n!}\left\{\left(H_{2n+1} - \frac{H_n}{2}\right)^2 - \sum_{k=0}^{n} \frac{1}{(2k+1)^2}\right\}.$$

Proof. We handle both cases separately.

1. Let us first examine the case where $\lambda = (a, b, 1^{n+1-a-b})$ and $a \geq b \geq 1$. In order to simplify the proof, we write $F_\lambda(x) = x(x-1)h_\lambda(x)$, where $h_\lambda(x)$ is obtained and defined as follows:

$$\begin{aligned} F_\lambda(x) &= 2^{a-b}(x/2+a-1)^{\underline{a-b}}(x+2b-2)^{\underline{n+1-a+b}} \quad \text{(see Eq. (9)[7])} \\ &= 2^{a-b}(x/2+a-1)^{\underline{a-b}}(x+2b-2)(x+2b-1)\cdots(x+1)x(x-1) \\ &\quad \times (x-2)(x-3)\cdots(x-2+b-n+a) \\ &= x(x-1)\underbrace{2^{a-b}(x/2+a-1)^{\underline{a-b}}(x+2b-2)^{\underline{2b-2}}(x-2)^{\underline{n-a-b+1}}}_{=h_\lambda(x)}. \end{aligned}$$

   (a) Using the above notation, we have
   $$F'_\lambda(0) = -h_\lambda(0) = (-1)2^{a-b}(a-1)^{\underline{a-b}}(2b-2)!(-2)^{\underline{n-a-b+1}},$$
   from which we easily obtain the wanted expression.

   (b) We also have
   $$\begin{aligned} F'_\lambda(1) &= h_\lambda(1) = 2^{a-b}\,(a-1/2)^{\underline{a-b}}\,(2b-1)^{\underline{2b-2}}(-1)^{\underline{n-a-b+1}} \\ &= 2^{a-b}\,(a-1/2)^{\underline{a-b}}\,(2b)!(-1)^{\underline{n-a-b+1}}, \end{aligned}$$

[7] Recall, as explained in the statement of Theorem 5.1, that we must replace n with $n+1$.

and obtaining the formula for $F'_\lambda(1)$ given in the statement is a simple matter, using the fact that

$$(a-1/2)^{\underline{a-b}} = \frac{(2a-1)(2a-3)\cdots(2b+1)}{2^{a-b}} = \frac{1}{2^{a-b}}\frac{(2a-1)!}{(a-1)!2^{a-1}}\frac{b!2^b}{(2b)!} = \frac{(2a-1)!b!}{2^{a-b}2^{a-b-1}(a-1)!(2b)!}.$$

(c) In order to simplify the computation of the second derivative, we will write $F_\lambda(x) = (x-1)g_\lambda(x)$, where

$$g_\lambda(x) = \underbrace{2^{a-b}\,(x/2+a-1)^{\underline{a-b}}}_{=\alpha_\lambda(x)}\,\underbrace{(x+2b-2)^{\underline{2b-1}}}_{=\beta_\lambda(x)}\,\underbrace{(x-2)^{\underline{n-a-b+1}}}_{=\gamma_\lambda(x)}.$$

With this notation, it is easy to see that $F''_\lambda(1) = 2g'_\lambda(1)$, with

$$g'_\lambda(1) = \alpha'_\lambda(1)\beta_\lambda(1)\gamma_\lambda(1) + \alpha_\lambda(1)\beta'_\lambda(1)\gamma_\lambda(1) + \alpha_\lambda(1)\beta_\lambda(1)\gamma'_\lambda(1).$$

Note that

$$\alpha'_\lambda(1) = \alpha_\lambda(1)\left(\frac{1}{2a-1}+\frac{1}{2a-3}+\cdots+\frac{1}{2b+1}\right) = \alpha_\lambda(1)\{H_{2a-1}-H_{2b}-(H_{a-1}-H_b)/2\},$$

$$\beta'_\lambda(1) = \beta_\lambda(1)\sum_{k=1}^{2b-1}\frac{1}{k} = \beta_\lambda(1)H_{2b-1},$$

$$\gamma'_\lambda(1) = -\gamma_\lambda(1)\sum_{k=1}^{n-a-b+1}\frac{1}{k} = -\gamma_\lambda(1)H_{n-a-b+1},$$

and

$$\alpha_\lambda(1) = \frac{(2a-1)!b!}{(2b)!2^{a-b-1}(a-1)!},$$
$$\beta_\lambda(1) = (2b-1)!,$$
$$\gamma_\lambda(1) = (-1)^{n-a-b+1}(n-a-b+1)!.$$

Combining all of the above, we obtain:

$$g'_\lambda(1) = \alpha_\lambda(1)\beta_\lambda(1)\gamma_\lambda(1)\{H_{2a-1}-H_{2b}-(H_{a-1}-H_b)/2+H_{2b-1}-H_{n-a-b+1}\} = \frac{(-1)^{n-a-b+1}(2a-1)!(b-1)!(n-a-b+1)!}{2^{a-b}(a-1)!}\{H_{2a-1}-H_{n-a-b+1}-(H_{a-1}-H_{b-1})/2\}$$

and we finally deduce the formula in the statement.

2. We now turn to the case where $\lambda = (n+1)$, i.e. $a = n+1$ and $b = 0$.

(a) Following the definition[8] of $F_\lambda(x)$ given earlier, we have

$$F_{(n+1)}(x) = 2^{n+1}\,(x/2+n)^{\underline{n+1}} = x\prod_{k=1}^{n}(x+2k).$$

We thus obtain

$$F'_{(n+1)}(x) = \prod_{k=1}^{n}(x+2k) + F_{(n+1)}(x)\sum_{k=1}^{n}\frac{1}{x+2k},$$

which easily gives the wanted expressions when evaluated at $x = 0$ and $x = 1$.

(b) For the second derivative, we obtain

$$F''_{(n+1)}(x) = F_{(n+1)}(x)\sum_{0\le i\ne j\le n}\frac{1}{(x+2i)(x+2j)},$$

hence

$$F''_{(n+1)}(1) = \frac{(2n+1)!}{2^n n!}\left\{\left(\sum_{k=0}^{n}\frac{1}{2k+1}\right)^2 - \sum_{k=0}^{n}\frac{1}{(2k+1)^2}\right\},$$

and the formula in the statement follows. □

Knowing the generating function G from Lemma 8.3, we can easily obtain the expected value of the number of cycles in the breakpoint graph of a random signed permutation of n elements.

8 Again, we replace n with $n+1$ in the definition.

Theorem 8.3. *The expected value of the number of cycles in the breakpoint graph of a uniform random signed permutation* $\pi^{\pm}$ *of* n *elements is*

$$\mathbb{E}(c(\mathrm{BG}(\pi^{\pm}))) = H_{2n+1} - \frac{H_n}{2} - \sum_{(a,b)\in\mathcal{A}_n} r_n(a,b),$$

where $\mathcal{A}_n = \{(a,b)\in\mathbb{N}^2 : a \geq b \geq 1, a+b \leq n+1\}$ *and*

$$r_n(a,b) = \frac{(-1)^{n+a-b}(n+1)(2a-2b+1)(a-1)!(2b-2)!(n-a-b+2)!}{2^{n-a+b-1}n!(b-1)!(n+a-b+2)^{\underline{2}}(n-a+b+1)^{\underline{2}}}.$$

Proof. As recalled in the proof of Theorem 8.1, we have $\mathbb{E}(c(\mathrm{BG}(\pi^{\pm}))) = G'(1)/G(1)$. Note that, by definition, $G(1) = \sum_{k=1}^{n+1} S_H^{\pm}(n,k)$, which equals the number of signed permutations of n elements, i.e. $2^n n!$. By Lemma 8.3, the expected number of cycles in the breakpoint graph of a random signed permutation is

$$\mathbb{E}(c(\mathrm{BG}(\pi^{\pm}))) = \frac{1}{2^n n!}\sum_{\lambda\in\mathcal{P}_{n+1}} c_\lambda(2)F'_\lambda(1)F'_\lambda(0).$$

Using the formulae for $F'_\lambda(1)$ and $F'_\lambda(0)$ derived in Lemma 8.4 and the expression for the coefficients[9] $c_\lambda(2)$ given in Eqs. (10) and (11), the formula in the statement follows. □

The generating function G allows us also to compute the variance of the number of cycles in the breakpoint graph of a random signed permutation.

Theorem 8.4. *The variance of the number of cycles in the breakpoint graph of a uniform random signed permutation* $\pi^{\pm}$ *of* n *elements is*

$$\begin{aligned}\mathrm{Var}(c(\mathrm{BG}(\pi^{\pm}))) = {} & H_{2n+1} - \frac{H_n}{2} - \sum_{k=0}^{n}\frac{1}{(2k+1)^2} - \left(\sum_{(a,b)\in\mathcal{A}_n} r_n(a,b)\right)^2 \\ & + \sum_{(a,b)\in\mathcal{A}_n} r_n(a,b)\{2H_{2n+1} - H_n - 2H_{2a-1} + 2H_{n-a-b+1} + H_{a-1} - H_{b-1} - 1\},\end{aligned}$$

where $\mathcal{A}_n$ *and the coefficients* $r_n(a,b)$ *are defined as in Theorem 8.3.*

Proof. As recalled in the proof of Theorem 8.2, the variance can be obtained from the generating function G by evaluating the function $(\log G)'(x) + (\log G)''(x)$ at $x = 1$. Therefore, the variance of the number of cycles in the breakpoint graph of a random signed permutation equals

$$\begin{aligned}\frac{G'(1)}{G(1)} + \frac{G''(1)}{G(1)} - \left(\frac{G'(1)}{G(1)}\right)^2 &= \frac{G'(1)+G''(1)}{G(1)} - (\mathbb{E}(c(\mathrm{BG}(\pi^{\pm}))))^2 \\ &= \frac{1}{2^n n!}\sum_{\lambda\in\mathcal{P}_{n+1}} c_\lambda(2)(F'_\lambda(1)+F''_\lambda(1))F'_\lambda(0) - (\mathbb{E}(c(\mathrm{BG}(\pi^{\pm}))))^2. \quad \text{(using Lemma 8.3)}\end{aligned}$$

Using the formulae for $F'_\lambda(1)$, $F''_\lambda(1)$ and $F'_\lambda(0)$ given in Lemma 8.4, we obtain that the variance equals

$$\begin{aligned}& H_{2n+1} - \frac{H_n}{2} - \sum_{k=0}^{n}\frac{1}{(2k+1)^2} + \left(H_{2n+1} - \frac{H_n}{2}\right)^2 - (\mathbb{E}(c(\mathrm{BG}(\pi^{\pm}))))^2 \\ & \quad - \sum_{(a,b)\in\mathcal{A}_n} r_n(a,b)\left\{2H_{2a-1} - 2H_{n-a-b+1} - H_{a-1} + H_{b-1} + 1\right\},\end{aligned}$$

which equals the wanted expression once $\mathbb{E}(c(\mathrm{BG}(\pi^{\pm})))$ is replaced with the value derived in Theorem 8.3. □

As in the unsigned case, we will study the behaviour of the mean and variance for large values of n. To that end, we will first prove the following lemma.

Lemma 8.5. *As* $n \to \infty$, *we have*

$$\sum_{(a,b)\in\mathcal{A}_n} |r_n(a,b)| = \frac{1}{\log(n)} \times o(1).$$

---

[9] Again, we replace n with $n+1$ in the definitions.

Proof. If we denote $k = a - b$, we have $0 \leq k \leq n - 1$. Furthermore, as $a + b = 2b + k$, the condition $a + b \leq n + 1$ translates into $2b \leq n - k + 1$ and the above sum becomes

$$\sum_{k=0}^{n-1} \frac{2^{k-n+1}(n+1)(2k+1)}{n!(n+k+2)^{\underline{2}}(n-k+1)^{\underline{2}}} \sum_{b=1}^{\lfloor (n-k+1)/2 \rfloor} \frac{(k+b-1)!(2b-2)!(n-k-2b+2)!}{(b-1)!}$$
$$= \sum_{k=0}^{n-1} \frac{2^{k-n+1}(n+1)(2k+1)}{(n+k+2)^{\underline{2}}(n-k+1)(k+1)\binom{n}{k+1}} \sum_{b=1}^{\lfloor (n-k+1)/2 \rfloor} \frac{\binom{k+b-1}{k}}{\binom{n-k}{2b-2}}$$
$$\leq \sum_{k=0}^{n-1} \frac{2^{k-n+1}}{(n+k+2)\binom{n}{k+1}} \sum_{b=1}^{\lfloor (n-k+1)/2 \rfloor} \binom{k+b-1}{k}$$
$$= \sum_{k=0}^{n-1} \frac{2^{k-n+1}\binom{k+\lfloor (n-k+1)/2 \rfloor}{k+1}}{(n+k+2)\binom{n}{k+1}} \quad \left(\text{using } \sum_{j=k}^{n} \binom{j}{k} = \binom{n+1}{k+1}\right).$$

We further observe that

$$\sum_{(a,b)\in\mathcal{A}_n} |r_n(a,b)| \leq \sum_{k=0}^{n-1} \frac{2^{k-n+1}}{n+k+2} \leq \frac{1}{n+2} \sum_{k=0}^{n-1} \frac{1}{2^{n-1-k}} = 2\left(1 - \frac{1}{2^n}\right)\frac{1}{n+2},$$

and the result in the statement easily follows. □

Based on this lemma, we can now obtain the following.

Remark 8.2. When $n \to \infty$, the expected value and variance of the number of cycles in the breakpoint graph of a uniform random signed permutation $\pi^{\pm}$ of n elements have the following asymptotical behaviour:

$$\mathbb{E}(c(\mathrm{BG}(\pi^{\pm}))) = \frac{\log(n)}{2} + \frac{\gamma}{2} + \log(2) + o(1),$$
$$\mathrm{Var}(c(\mathrm{BG}(\pi^{\pm}))) = \frac{\log(n)}{2} + \frac{\gamma}{2} + \log(2) - \frac{\pi^2}{8} + o(1).$$

Note that, in the limit when $n \to \infty$, the mean and variance in the signed case are of the same order ($\log(n)$) as in the unsigned case, but they differ by a factor of 1/2.

9. Conclusions

In this paper, we proved the first explicit formula for enumerating signed permutations whose breakpoint graph contains a given number of cycles, and proved simpler expressions for particular cases. We also obtained a new expression for enumerating unsigned permutations whose breakpoint graph contains a given number of cycles, and used both formulae to derive simpler proofs of some other previously known results. Getting more insight into breakpoint graphs and their cycle decomposition is particularly relevant to edit distances used in the field of genome rearrangements, and we hope that our results can help shed light on their distributions, expected values and variances. There are several interesting directions in which our work could be extended, which we outline and motivate below.

Just like one can define conjugacy classes in the symmetric and hyperoctahedral groups, we could investigate conjugacy classes with respect to the breakpoint graph. This was already initiated by Doignon and Labarre [6], who referred to them as "Hultman classes" and provided explicit formulae for enumerating those classes in the case of unsigned permutations. More work remains to be done in the unsigned case: indeed, the work done by Bóna and Flynn [3] provides us with a very nice formula for computing the distribution of cycles, but no simpler expression than the complicated ones obtained by Doignon and Labarre [6] is yet known for enumerating Hultman classes or their cardinalities. Moreover, no work so far has been done in order to enumerate Hultman classes in the signed setting, and obtaining an expression for enumerating the so-called "simple permutations", which are defined in this context as permutations whose breakpoint graph contains no cycle of length greater than 2, seems especially interesting (for more information about the importance of those permutations in genome rearrangements, see Hannenhalli and Pevzner [13] and Labarre and Cibulka [17]).

The expression we obtained for the signed Hultman numbers is quite useful in practice, since it allows us to obtain the distribution of those numbers for large values of n. Unfortunately, it does not seem easy to use in order to gain insights and have an intuitive interpretation of the shape of the distribution, which would be useful in order to know how this distribution can be approximated or how it grows as n increases. Finding simpler generating functions, recurrence relations or nicer formulae would be useful in that regard and in order to obtain more information on the properties of this distribution.

The connection between the cycle structure of breakpoint graphs and factorisations of even permutations (Corollary 4.1) proved useful not only in characterising the distribution of those cycles and of the related cycle types, but also provided the foundations of a simple and generic method for obtaining lower bounds on *any* "revertible" edit distance between unsigned permutations (see Labarre [16] for more details). Is there any way to use the results and connections obtained in Section 5 in order to obtain similar results for signed permutations?

Finally, recall that permutations are just one way of modelling genomes. One natural direction would be to investigate the distribution of cycles in the breakpoint graph of other structures, like set systems or "fragmented" permutations (see again Fertin et al. [8] for an overview of existing models).

Acknowledgements

The first author was partially supported by the ANR MAEV under contract ANR-06-BLAN-0113. Both authors also wish to thank the group "Evolution Biologique et Modélisation", LATP, Université de Provence, where part of this research was performed, as well as Mathilde Bouvel for bringing Ref. [15] to their attention.

References

[1] V. Bafna, P.A. Pevzner, Genome rearrangements and sorting by reversals, SIAM Journal on Computing (ISSN: 0097-5397) 25 (2) (1996) 272–289.
[2] A. Björner, F. Brenti, Combinatorial descriptions, in: Combinatorics of Coxeter Groups, in: Graduate Texts in Mathematics, vol. 231, Springer-Verlag, 2005 (Chapter 8).
[3] M. Bóna, R. Flynn, The average number of block interchanges needed to sort a permutation and a recent result of Stanley, Information Processing Letters 109 (16) (2009) 927–931.
[4] D.A. Christie, Sorting permutations by block-interchanges, Information Processing Letters (ISSN: 0020-0190) 60 (4) (1996) 165–169.
[5] R. Diestel, Graph Theory, third ed., in: Graduate Texts in Mathematics, vol. 173, Springer-Verlag, Berlin, ISBN: 978-3-540-26182-7, 2005, ISBN 3-540-26182-6.
[6] J.-P. Doignon, A. Labarre, On Hultman numbers, Journal of Integer Sequences 10 (6) (2007) 13. Article 07.6.2.
[7] I. Elias, T. Hartman, A 1.375-approximation algorithm for sorting by transpositions, IEEE/ACM Transactions on Computational Biology and Bioinformatics (ISSN: 1545-5963) 3 (4) (2006) 369–379.
[8] G. Fertin, A. Labarre, I. Rusu, E. Tannier, S. Vialette, Combinatorics of Genome Rearrangements, in: Computational Molecular Biology, The MIT Press, 2009.
[9] N.R. Goodman, Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction), The Annals of Mathematical Statistics 34 (1) (1963) 152–177.
[10] R.L. Graham, D.E. Knuth, O. Patashnik, Concrete Mathematics: A Foundation for Computer Science, second ed., Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, ISBN: 0201558025, 1994.
[11] S. Grusea, On the distribution of the number of cycles in the breakpoint graph of a random signed permutation, IEEE/ACM Transactions on Computational Biology and Bioinformatics (ISSN: 1545-5963) 8 (5) (2011) 1411–1416.
[12] P.J. Hanlon, R.P. Stanley, J.R. Stembridge, Some combinatorial aspects of the spectra of normally distributed random matrices, Contemporary Mathematics 138 (1992) 151–174.
[13] S. Hannenhalli, P.A. Pevzner, Transforming cabbage into turnip: polynomial algorithm for sorting signed permutations by reversals, Journal of the ACM 46 (1) (1999) 1–27.
[14] A. Hultman, Toric permutations, Master's Thesis, Department of Mathematics, KTH, Stockholm, Sweden, 1999.
[15] J.H. Kwak, J. Lee, Genus polynomials of dipoles, Kyungpook Mathematical Journal 33 (1) (1993) 115–125.
[16] A. Labarre, Edit distances and factorisations of even permutations, in: D. Halperin, K. Mehlhorn (Eds.), Proceedings of the Sixteenth Annual European Symposium on Algorithms (ESA), in: Lecture Notes in Computer Science, vol. 5193, Springer-Verlag, ISBN: 978-3-540-87743-1, 2008, pp. 635–646.
[17] A. Labarre, J. Cibulka, Polynomial-time sortable stacks of burnt pancakes, Theoretical Computer Science (ISSN: 0304-3975) 412 (8–10) (2011) 695–702.
[18] H. Li, N. Homer, A survey of sequence alignment algorithms for next-generation sequencing, Briefings in Bioinformatics 11 (5) (2010) 473–483. URL: http://bib.oxfordjournals.org/content/11/5/473.abstract.
[19] I.G. Macdonald, Symmetric Functions and Hall Polynomials, second ed., in: Oxford Mathematical Monographs, Oxford University Press, 1998.
[20] N.J.A. Sloane, The on-line encyclopedia of integer sequences, 2012. Published electronically at http://oeis.org/.
[21] B. Sury, T. Wang, F.-Z. Zhao, Identities involving reciprocals of binomial coefficients, Journal of Integer Sequences 7 (12) (2004). Article 04.2.8.
[22] L. Székely, Y. Yang, On the expectation and variance of the reversal distance, Acta Universitatis Sapientiae, Mathematica 1 (1) (2009) 5–20.
[23] H. Wielandt, Finite Permutation Groups, Academic Press, New York, 1964. Translated from German by R. Bercov.
[24] H.S. Wilf, Generatingfunctionology, third ed., A. K. Peters, Ltd., Natick, MA, USA, ISBN: 1568812795, 2006.