From Hertzsprung's problem to pattern-rewriting systems

Anders Claesson

7 April 2021

Abstract

Drawing on a problem posed by Hertzsprung in 1887, we say that a given permutation $\pi \in \mathcal{S}_n$ contains the Hertzsprung pattern $\sigma \in \mathcal{S}_k$ if there is factor $\pi(d+1)\pi(d+2)\cdots\pi(d+k)$ of π such that $\pi(d+1)-\sigma(1) = \cdots = \pi(d+k)-\sigma(k)$. Using a combination of the Goulden-Jackson cluster method and the transfer-matrix method we determine the joint distribution of occurrences of any set of (incomparable) Hertzsprung patterns, thus substantially generalizing earlier results by Jackson et al. on the distribution of ascending and descending runs in permutations. We apply our results to the problem of counting permutations up to pattern-replacement equivalences, and using pattern-rewriting systems—a new formalism similar to the much studied string-rewriting systems—we solve a couple of open problems raised by Linton et al. in 2012.

Keywords: Hertzsprung's problem, cluster method, pattern, permutation, rewriting system

1 Introduction

Severin Carl Ludvig Hertzsprung (1839–1893) was a Danish senior civil servant with a graduate degree in astronomy from the University of Copenhagen. In 1887 a letter authored by him titled "En kombinationsopgave" appeared in Tidsskrift for mathematik [16]. The first paragraph reads

> *At bestemme Antallet af Maader, hvorpaa Tallene* 1, 2, 3, 4,. . . , n *kunne opstilles i Række saaledes, at Differensen mellem to ved Siden af hinanden staaende Tal overalt skal være numerisk forskjellig fra 1.*

Or, translated, the problem is to determine the number of ways in which the numbers 1, 2, 3, 4, . . . , n can be arranged such that the difference between two adjacent numbers is different from 1. For $n = 4$ there are only two such arrangements, namely 2413 and 3142. An alternative interpretation is that we are asked for the number of ways n kings can be placed on an n by n chessboard, one on each row and column, so that no two attack each other. Indeed, Kaplansky [21]—who rediscovered Hertzsprung's problem in 1944—called it the n-kings problem.

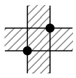
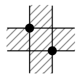
In Hertzsprung's solution—which in his own words is only slightly elegant—the problem is first generalized by defining $u_{n,k}$ as the number of arrangements such that the numbers 1, 2, ..., k satisfy the requirement that two numbers whose difference is 1 cannot be placed next to each other. Hertzsprung derives a recurrence relation for the numbers $u_{n,k}$ to which he guesses a solution and subsequently verifies that it indeed satisfies the recurrence. The solution to the original problem is then given by

$$u_{n,n} = n! + \sum_{k=1}^{n}(-1)^k \sum_{i=1}^{k} \binom{k-1}{i-1}\binom{n-k}{i} 2^i (n-k)!$$

The fist few numbers of this sequence are 1, 1, 0, 0, 2, 14, 90, 646, 5242, 47622, which is A002464 in the OEIS [17]. A more elegant, but less explicit, solution (see e.g. p. 737 of Flajolet and Sedgewick [12]) is obtained using generating functions: $u_{n,n}$ is the coefficient of x^n in the expansion of

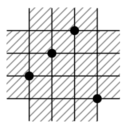
$$\sum_{m\geq 0} m!\left(\frac{x-x^2}{1+x}\right)^m. \tag{1}$$

A reader familiar with mesh patterns [7] may have noticed that Hertzsprung's problem is to count permutations in $\mathcal{S}_n$ that avoid the two patterns

 and 

We will use the following terminology. Let π be a word (e.g. a permutation written in one line notation). A word β is said to be a *factor* of π if there are words α and γ such that $\pi = \alpha\beta\gamma$. If α is empty we also say that β is a *prefix* of π. If, in addition, γ is nonempty then β is said to be a *proper prefix* of π. Similarly, if γ is empty we say that β is a *suffix* of π, and that β is a *proper suffix* if, in addition, α is nonempty.

Let $\tau \in \mathcal{S}_k$ and $\pi \in \mathcal{S}_n$. We call τ a *Hertzsprung factor* of π if there is an integer c and a factor $\beta = b_1b_2\dots b_k$ of π such that $\tau(i) - b_i = c$ for each $i \in [k]$. In this context we also say that β is an *occurrence* of τ and sometimes we write $\beta \simeq \tau$. As an example, 213 is a Hertzsprung factor of $\pi = 1546372$; indeed, $546 \simeq 213$ is an occurrence of 213 in π. We similarly define *Hertzsprung prefix* and *Hertzsprung suffix*. In terms of mesh patterns, τ is a Hertzsprung factor of π if π contains the pattern (τ, H_k), where H_k is the mesh $\{0, 1, \dots, k\}^2 \setminus \{0, k\}^2$. For instance, 2341 is a Hertzsprung factor of π if and only if π contains the mesh pattern

$(2341, H_4) =$ 

Mesh patterns of the form (τ, H_k) will be referred to as *Hertzsprung patterns*. Since all mesh patterns in this paper will be of this form we will, by slight abuse of notation, identity the pattern (τ, H_k) with τ. While the Hertzsprung terminology is original with this paper, there are a number of results on these patterns under different names in the literature. Monotone Hertzsprung patterns have the longest history and are usually called (increasing or decreasing) runs in the literature [1, 18, 19, 20, 31]. Myers [26] seem to have been the first to study non-monotone Hertzsprung patterns. She calls them rigid patterns. Bóna [5] talks of permutations very tightly avoiding or containing a pattern. Let us now detail some of these results.

The identity permutation (as well as the corresponding Hertzsprung pattern) will be denoted by $\mathrm{id}_k = 12\ldots k$. Its reverse will be denoted $\overline{\mathrm{id}}_k = k\ldots 21$. Jackson and Read [19] showed that the generating functions for respectively id_k-avoiding and $\{\mathrm{id}_k, \overline{\mathrm{id}}_k\}$-avoiding permutations are

$$\sum_{m\geq 0} m!\left(\frac{x-x^k}{1-x^k}\right)^m \quad \text{and} \quad \sum_{m\geq 0} m!\left(\frac{x-2x^k+x^{k+1}}{1-x^k}\right)^m.$$

The special case $k=2$ in the latter formula is the generating function (1).

Some Hertzsprung patterns $\tau \in \mathcal{S}_k$ can overlap with themselves in the sense that there is a permutation σ of length less than $2k$ such that τ is both a proper Hertzsprung prefix and a proper Hertzsprung suffix of σ. Myers [26] calls such patterns extendible, we will call them *self-overlapping*. The monotone patterns, id_k and $\overline{\mathrm{id}}_k$, are both self-overlapping. The pattern 2143 is also self-overlapping, while 132 is not. Myers derives a formula for counting the number of permutations of $[n]$ with a prescribed number m of occurrences of any single non-self-overlapping pattern $\tau \in \mathcal{S}_k$:

$$\sum_i (-1)^{m-i}\binom{i}{m}\binom{n-(k-1)i}{i}\big(n-(k-1)i\big)! \tag{2}$$

Let $\mathcal{S}_n(\tau)$ denote the set of permutations in $\mathcal{S}_n$ that avoid the Hertzsprung pattern τ. Bóna [5] showed that $|\mathcal{S}_n(\tau)| \leq |\mathcal{S}_n(\mathrm{id}_k)|$ for each non-self-overlapping pattern τ. He also explains why most patterns are non-self-overlapping. Both Myers and Bóna ask if it is possible to determine $|\mathcal{S}_n(\tau)|$ for any self-overlapping pattern τ other than id_k and $\overline{\mathrm{id}}_k$.

Let T be an *antichain* of Hertzsprung patterns. That is, every pair of distinct patterns in T are incomparable in the sense that one is not a Hertzsprung factor of the other. Let $\mathcal{S}_n(T) = \cap_{\tau\in T}\mathcal{S}(\tau)$ be the set of T*-avoiding permutations* in $\mathcal{S}_n$. In Section 2 we show that there is a rational function $R(x) \in \mathbb{Q}(x)$ such that

$$\sum_{n\geq 0} |\mathcal{S}_n(T)|x^n = \sum_{m\geq 0} m!R(x)^m.$$

Moreover, there is an efficient way of determining $R(x)$ from the set T. We in fact show something stronger.

It is known [18, 20] that the number of permutations of $[n]$ with exactly ℓ occurrences of id_k is the coefficient of $u^\ell x^n$ in

$$\sum_{m\geq 0} m!x^m\left(\frac{1-ux-(1-u)x^{k-1}}{1-ux-(1-u)x^k}\right)^m. \tag{3}$$

Suppose that $T = \{\tau_1, \tau_2, \ldots, \tau_k\}$ is an antichain of Hertzsprung patterns. We show that there is a rational function $R(u_1,\ldots,u_k;x)$ in $\mathbb{Q}(u_1,\ldots,u_k,x)$ such that

$$\sum_{\pi\in\mathcal{S}} u_1^{\tau_1(\pi)}u_2^{\tau_2(\pi)}\cdots u_k^{\tau_k(\pi)}x^{|\pi|} = \sum_{m\geq 0} m!R(u_1,u_1,\ldots,u_k;x)^m,$$

where $\mathcal{S} = \cup_n\mathcal{S}_n$ and $\tau_i(\pi)$ denotes the number of occurrences of τ_i in π.

In Section 3 we apply the machinery developed in Section 2 to the problem of counting permutations up to certain pattern-replacement equivalences introduced by Linton, Propp, Roby and West [25]. We provide a new formalism that we call a pattern-rewriting systems, which are similar to the much studied string-rewriting systems. Using these we solve a couple of open problems posed by Linton et al.

2 Enumeration

Let T be an antichain of Hertzsprung patterns. We will apply the Goulden-Jackson cluster method [14] to count permutations with respect to the number of occurrences of patterns in T. The context of the original formulation of this method is the free monoid over a finite set. The method has, however, been successfully adopted and applied to permutations by Goulden and Jackson themselves and more recently by Dotsenko and Khoroshkin [10] and Elizalde and Noy [11].

A *marked permutation* is a pair (π, M) where π is a permutation and M is a subset of all occurrences in π of patterns from T. The members of M are called *marked occurrences* and π is called the *underlying permutation* of (π, M). As an example, let $T = \{123\}$, $\pi = 1234567 \in \mathcal{S}_7$ and $M = \{123, 234, 567\}$. The marked permutation (π, M) is then depicted below.

$$1\ 2\ 3\ 4\ 5\ 6\ 7$$

Let (π, M) be a marked permutation such that $|\pi| \geq 2$ and M cover π in the sense that each letter of π is in some occurrence. We say that $\alpha, \beta \in M$ *overlap* if they have at least one letter in common. Clearly, being overlapping is a symmetric relation and we can view M as an undirected graph with edge set consisting of all (α, β) such that α and β overlap. If this graph is connected, then we say that (π, M) is a T-*cluster*, or simply a *cluster* when T is know from context. The marked permutation in the example above is not a cluster, but 1234 is.

Let π be a permutation of $[k]$ and let $\sigma_1, \sigma_2, \ldots, \sigma_k$ be nonempty permutations. The *inflation* [2] of π by $\sigma_1, \sigma_2, \ldots, \sigma_k$ is $\pi[\sigma_1, \sigma_2, \ldots, \sigma_k] = \sigma_1'\sigma_2' \ldots \sigma_k'$, where $\sigma'_{\pi^{-1}(i)}$ is obtained from $\sigma_{\pi^{-1}(i)}$ by adding the constant $|\sigma_{\pi^{-1}(1)}| + |\sigma_{\pi^{-1}(2)}| + \cdots + |\sigma_{\pi^{-1}(i-1)}|$ to each of its letters. An example should make this clear,

$$231[1, 213, 21] = 3\ 546\ 21$$

Note that σ_i' is an occurrence of σ_i in $\pi[\sigma_1, \sigma_2, \ldots, \sigma_k]$ by construction. The inflation operation naturally generalizes to marked permutations. Here is a permutation with some marked occurrences of 123 and 132:

$$\begin{aligned} 213456\big[&1\ 2\ 3\ 4\ 5,\ 1\ 2\ 3\ 5\ 4,\ 1,\ 1,\ 1\ 3\ 2,\ 1\big] \\ &= 6\ 7\ 8\ 9\ 10\ 1\ 2\ 3\ 5\ 4\ 11\ 12\ 13\ 15\ 14\ 16 \end{aligned} \tag{4}$$

It is clear that a marked permutation is a cluster if and only if its underlying permutation has length at least 2 and it cannot be expressed as the inflation of two or more nonempty marked permutation. Moreover, any marked permutation can be uniquely written $\pi\big[(\sigma_1, M_1), (\sigma_2, M_2), \ldots, (\sigma_k, M_k)\big]$ where π is a permutation and each (σ_i, M_i) is either $(1, \emptyset)$ or a cluster.

For each nonempty permutation π let there be an associate indeterminate u_π and let $\boldsymbol{u} = (u_1, u_{12}, u_{21}, u_{123}, u_{132}, \ldots)$ be the sequence of all such indeterminates. Let T be an antichain of Hertzsprung patterns. For any marked permutation (π, M), define the monomial

$$\boldsymbol{u}_M = \prod_{\alpha \in M} u_{\mathrm{st}(\alpha)},$$

where $\mathrm{st}(\alpha)$ denotes the standardization of α. That is, $\mathrm{st}(\alpha)$ is the permutation of $\{1, 2 \ldots, |\alpha|\}$ obtained from α by replacing its smallest letter by 1, its next smallest letter by 2, etc. If M is the set of marked occurrences in (4) then $\boldsymbol{u}_M = u_{123}^4 u_{132}^2$. The *cluster generating function* associated with T is defined by

$$C(\boldsymbol{u}; x) = \sum_{(\pi,M)} \boldsymbol{u}_M x^{|\pi|},$$

where the sum is over all T-clusters. Similarly, let $F(\boldsymbol{u}; x) = \sum_{(\pi,M)} \boldsymbol{u}_M x^{|\pi|}$, where the sum is over all marked permutations. Further, let

$$\boldsymbol{u}^{T(\pi)} = \prod_{\tau \in T} u_\tau^{\tau(\pi)},$$

where $\tau(\pi)$ denotes the number of occurrences of τ in π. Then

$$\sum_{\pi \in \mathcal{S}} x^{|\pi|} (\mathbf{1} + \boldsymbol{u})^{T(\pi)} = \sum_{m \geq 0} m! \, (x + C(\boldsymbol{u}; x))^m .$$

Indeed, the left-hand side is the generating function $F(\boldsymbol{u}; x)$ of marked permutations: for each $\tau \in T$ and for each of the $\tau(\pi)$ occurrences of τ in π there is a choice to be made, mark it with a u_τ or leave it unmarked. That the right-hand side equals $F(\boldsymbol{u}; x)$ follows from the unique representation of marked permutations as the inflation of a permutation with clusters and $(1, \emptyset)$. On replacing $\boldsymbol{u}$ by $\boldsymbol{u} - \mathbf{1}$ we have the following result.

Theorem 2.1. *Let T be an antichain of Hertzsprung patterns. Then*

$$\sum_{\pi \in \mathcal{S}} \boldsymbol{u}^{T(\pi)} x^{|\pi|} = \sum_{m \geq 0} m! \, (x + C(\boldsymbol{u} - \mathbf{1}; x))^m .$$

In particular,

$$\sum_{n \geq 0} |\mathcal{S}_n(T)| x^n = \sum_{m \geq 0} m! \, (x + C(-\mathbf{1}; x))^m .$$

If $T = \{\tau\}$ and τ is a non-self-overlapping pattern, then the only T-cluster is $(\tau, \{\tau\})$ and $C(u, x) = ux^{|\tau|}$, where $u = u_\tau$. Hence we arrive at the following generating function version of Myers's formula (2).

Corollary 2.2. *For any non-self-overlapping pattern τ,*

$$\sum_{\pi \in \mathcal{S}} u^{\tau(\pi)} x^{|\pi|} \; = \; \sum_{m \geq 0} m! \left(x + (u-1) x^{|\tau|} \right)^m .$$

Example 2.3. The distribution of the number of occurrences of 132 is given by

$$\sum_{\pi \in \mathcal{S}} u^{132(\pi)} x^{|\pi|} = \sum_{m \geq 0} m! \left(x + (u-1) x^3 \right)^m .$$

Letting $u = 0$ we get a generating function for 132-avoiding permutations:

$$\sum_{n \geq 0} |\mathcal{S}_n(132)| x^n = \sum_{m \geq 0} m! \left(x - x^3 \right)^m .$$

It follows that

$$|\mathcal{S}_n(132)| = \sum_{i=0}^{\lfloor n/3 \rfloor} (-1)^i (n-2i)! \binom{n-2i}{i}, \tag{5}$$

which is the case $\tau = 132$, $m = 0$ and $k = 3$ of Myers's formula (2).

Example 2.4. For $T = \{123\}$ the underlying permutation of any cluster is id_k for some $k \geq 3$. A cluster of length k is built from clusters of length $k-2$ or length $k-1$ by marking the suffix $(k-2, k-1, k)$ of id_k. Formally, if $\mathcal{C}_k$ denotes the set of clusters whose underlying permutation is id_k then $\mathcal{C}_2 = \emptyset$, $\mathcal{C}_3 = \{(123, \{123\})\}$, and $\mathcal{C}_k$ consists of all clusters

$$\big(\mathrm{id}_k,\ M \cup \{(k-2, k-1, k)\}\big),$$

where $(\mathrm{id}_{k-1}, M) \in \mathcal{C}_{k-1}$ or $(\mathrm{id}_{k-2}, M) \in \mathcal{C}_{k-2}$. It follows that

$$C(u, x) = ux^3 + uxC(u, x) + ux^2 C(u, x),$$

in which $u = u_{123}$ tracks marked factors and x tracks the length of the cluster. Thus the number of clusters of fixed length is a Fibonacci number, the generating function for clusters is $C(u, x) = ux^3/(1 - u(x + x^2))$, and we have rediscovered a special cases of (3):

$$\begin{aligned} \sum_{\pi \in \mathcal{S}} u^{132(\pi)} x^{|\pi|} &= \sum_{m \geq 0} m!\, (x + C(u-1, x))^m \\ &= \sum_{m \geq 0} m! \left(x + \frac{(u-1)x^3}{1 - (u-1)(x + x^2)} \right)^m . \end{aligned}$$

Having seen a couple of examples we now return to the general case. Let T be any antichain of Hertzsprung patterns. We will show that the cluster generating function associated with T is counting walks in a digraph and thus it is rational. This can be seen as an application of the transfer-matrix method [32].

For $\sigma, \tau \in T$ define the set of permutations $\mathcal{O}(\sigma, \tau)$ by stipulating that $\pi \in \mathcal{O}(\sigma, \tau)$ if and only if

- σ is a proper Hertzsprung prefix of π,
- τ is a proper Hertzsprung suffix of π and
- $|\pi| < |\sigma| + |\tau|$.

If $\mathcal{O}(\sigma, \tau)$ is nonempty, then we say that σ and τ *overlap*. In particular, $\mathcal{O}(\tau, \tau)$ is nonempty if and only if τ is self-overlapping. Let

$$\Omega(\sigma, \tau) = \sum_{\pi \in \mathcal{O}(\sigma, \tau)} x^{|\pi| - |\sigma|}.$$

Construct an edge weighted digraph D_T with vertices $V = \{\epsilon\} \cup T$, edge set $V \times V$ and weight function $w : V \times V \to \mathbb{Q}[\boldsymbol{u}, x]$ defined by

$$w(\sigma, \tau) = \begin{cases} 0 & \text{if } \tau = \epsilon, \\ u_\tau x^{|\tau|} & \text{if } \sigma = \epsilon, \\ u_\tau \Omega(\sigma, \tau) & \text{otherwise.} \end{cases}$$

Example 2.5. Let $T = \{123, 132\}$. We have $\mathcal{O}(132, 123) = \mathcal{O}(132, 132) = \emptyset$,

$$\mathcal{O}(123, 123) = \{1234, 12345\} \quad \text{and} \quad \mathcal{O}(123, 132) = \{12354\}.$$

For ease of notation, let $u = u_{123}$ and $v = u_{132}$. The digraph D_T together with its adjacency matrix are depicted below.

$$\begin{tikzcd} 123 \arrow[loop left, "ux+ux^2"] \arrow[rr, "vx^2"] & & 132 \\ & \epsilon \arrow[ul, "ux^3"] \arrow[ur, "vx^3"'] & \end{tikzcd} \qquad \begin{bmatrix} 0 & ux^3 & vx^3 \\ 0 & ux+ux^2 & vx^2 \\ 0 & 0 & 0 \end{bmatrix}$$

Here, and in what follows, edges that are known to have weight 0 are omitted.

Theorem 2.6. *Let T be an antichain of Hertzsprung patterns. Let A be the adjacency matrix of the digraph D_T defined above. Then*

$$C(\boldsymbol{u}; x) \;=\; \frac{1}{\det(1-A)} \sum_{i=1}^{|T|} (-1)^i \det\bigl(1 - A : i+1, 1\bigr),$$

where $(B : i, j)$ denotes the matrix obtained by removing the ith row and jth column of B. In particular, $C(\boldsymbol{u}; x)$ belongs to the field of rational functions $\mathbb{Q}(x)(u_\tau : \tau \in T)$.

Proof. The generating function for (weighted) directed walks in D_T starting at ϵ and ending in any other vertex equals $C(\boldsymbol{u}; x)$ by construction of D_T. Thus $C(\boldsymbol{u}; x)$ is given by summing all but the first entry of the first row of

$$(1-A)^{-1} = 1 + A + A^2 + A^3 + \cdots$$

If B is any invertible matrix, then Cramer's rule gives

$$(B^{-1})_{ij} = (-1)^{i+j} \det(B; j, i) / \det(B),$$

from which the result follows. $\square$

By applying Theorem 2.6 to $T = \{132\}$ and $T = \{123\}$ we find that the cluster generating functions are respectively ux^3 and $ux^3/(1 - u(x + x^2))$, which agree with examples 2.3 and 2.4. More generally, we have the following corollary.

Corollary 2.7. *For any Hertzsprung pattern τ,*

$$\sum_{\pi \in \mathcal{S}} u^{\tau(\pi)} x^{|\pi|} = \sum_{m \geq 0} m! \left(x + \frac{(u-1)x^{|\tau|}}{1 - (u-1)\Omega(\tau, \tau)} \right)^m.$$

Example 2.8. Hertzsprung's problem generalized to counting permutations with respect to the joint distribution of the Hertzsprung patterns 12 and 21 has the solution

$$\sum_{\pi \in \mathcal{S}} u^{12(\pi)} v^{21(\pi)} x^{|\pi|} \;=\; \sum_{m \geq 0} m! \left(\frac{x - (u-1)(v-1)x^3}{(1 + x - ux)(1 + x - vx)} \right)^m.$$

More generally, the joint distribution of id_k and $\overline{\mathrm{id}}_\ell$ is given by

$$\sum_{m\geq 0} m!\left(x+\frac{(u-1)x^k}{1-(u-1)([k]_x-1)}+\frac{(v-1)x^\ell}{1-(v-1)([\ell]_x-1)}\right)^m,$$

where $[k]_x = (1-x^k)/(1-x) = 1+x+\cdots+x^{k-1}$. Generalizing Hertzsprung's problem in another direction we find that the generating function for permutations avoiding all patterns in $\mathcal{S}_3$ is

$$\begin{aligned}&\sum_{m\geq 0} m!\left(\frac{-2x^6+3x^5-3x^4-5x^3+x^2+x}{x^4+x^3-x^2-x-1}\right)^m\\ &= 1+x+2x^2+4x^4+34x^5+298x^6+2434x^7+21374x^8\\ &\quad+205300x^9+2161442x^{10}+24804386x^{11}+\cdots\end{aligned}$$

Theorem 2.1 together with Theorem 2.6 solves the problem of determining the distribution of a set of Hertzsprung patterns. Or, at least they solve the problem as long as the patterns are short. The weight function is defined in terms of the sets $\mathcal{O}(\sigma,\tau)$ and those sets are a priori expensive to compute. We shall now detail why, in fact, the weight function can be efficiently computed.

The polynomials $\Omega(\sigma,\tau)$ are akin to the correlation polynomials of Guibas and Odlyzko [15]. Consider $\pi\in\mathcal{O}(\sigma,\tau)$. Let δ be the factor of π in which its prefix corresponding to σ intersects with its suffix corresponding to τ. The permutation 978563421 is, for instance, a member of $\mathcal{O}(53412, 563421)$ and $\delta = 5634$. Let $i=|\delta|$. Note that the smallest element in δ is either $|\sigma|-i+1$ or $|\tau|-i+1$. Indeed, either σ is a (literal) prefix of π or τ is a (literal) suffix of π. In the first case, $\min\delta = |\sigma|-i+1$. In the second case, $\min\delta=|\tau|-i+1$. Thus we arrive at the formula

$$\Omega(\sigma,\tau)=\sum_{i\geq 1}\chi_i(\sigma,\tau)x^{|\tau|-i},$$

where $\chi_i(\sigma,\tau)\in\{0,1\}$ is computed as follows. Consider the last i letters $a_1a_2\dots a_i$ of σ and the first i letters $b_1b_2\dots b_i$ of τ. If the differences a_1-b_1, a_2-b_2, $\dots$, a_i-b_i are all equal and their common difference is $|\sigma|-i$ or $-(|\tau|-i)$, then $\chi_i(\sigma,\tau)=1$; otherwise, $\chi_i(\sigma,\tau)=0$. We can organize this computation in a table as in the following example.

$$\begin{array}{lll} \sigma: & 53412 & \\ \tau: & \;\;563421 & 1 \\ & \;\;\;\;563421 & 0 \\ & \;\;\;\;\;\;563421 & 1 \\ & \;\;\;\;\;\;\;\;563421 & 0 \end{array}$$

Writing χ_j for $\chi_j(\sigma,\tau)$ we have $(\chi_1,\chi_2,\chi_3,\chi_4)=(0,1,0,1)$ and $\Omega(\sigma,\tau)=x^2+x^4$.

Example 2.9. Penney's game [29, 15] is a well-known coin tossing game played by two players. Player I picks a sequence of heads and tails, say $\alpha\in\{H,T\}^k$. Player I shows α to Player II, who subsequently picks a word $\beta\in\{H,T\}^k$. A fair coin is tossed until either α or β appears as a consecutive subsequence of outcomes. The player whose sequence appears first wins. Thus a winning sequence of outcomes for Player I is a word $\omega\in\{H,T\}^n$ such that α is a suffix of ω and ω otherwise avoid both α and

β. There is a simple formula [15][Formula (1.5)], due to Conway [13], expressing the odds that Player II will win in terms of correlation polynomials. Further, this game has the curious property of being nontransitive: Assuming $k \geq 3$, no matter what α Player I picks, Player II can pick β so that his probability of winning exceeds 1/2. Consider the following related question for permutations. Let T be an antichain of Hertzsprung patterns and fix a pattern $\alpha \in T$. How many permutations in $\mathcal{S}_n$ avoid T except for a single occurrence of α at the end of the permutation? Let us call this number $f^\alpha(n)$. Let $F^\alpha(\boldsymbol{u}; x)$ be the generating function for marked permutations that end in an unmarked occurrence of α. Then

$$F^\alpha(\boldsymbol{u}; x) = C^\alpha(\boldsymbol{u}; x) \sum_{m \geq 1} m! \, (x + C(\boldsymbol{u}; x))^{m-1},$$

where $C(\boldsymbol{u}; x)$ is the usual cluster generating function and $C^\alpha(\boldsymbol{u}; x)$ is the generating function for marked permutations (π, M) such that the last k letters $\gamma = \pi(n-k+1)\dots\pi(n-1)\pi(n)$ of π form an unmarked occurrence of α and $(\pi, M \cup \{\gamma\})$ is a T-cluster. By construction,

$$F^\alpha(-\mathbf{1}; x) = \sum_{n \geq 0} f^\alpha(n) x^n.$$

This formula is, however, only useful if we can efficiently compute $C^\alpha(\boldsymbol{u}; x)$. Define the digraph D_T^α with vertices $V = \{\epsilon, \hat{\alpha}\} \cup T$, in which $\hat{\alpha}$ is a distinct copy of α, and weight function $\hat{w} : V \times V \to \mathbb{Q}[\boldsymbol{u}, x]$ given by

$$\hat{w}(\sigma, \tau) = \begin{cases} w(\sigma, \tau) & \text{if } \sigma \neq \hat{\alpha} \text{ and } \tau \neq \hat{\alpha}, \\ 0 & \text{if } \sigma = \hat{\alpha}, \\ x^{|\alpha|} & \text{if } \sigma = \epsilon \text{ and } \tau = \hat{\alpha}, \\ \Omega(\sigma, \alpha) & \text{if } \sigma \notin \{\epsilon, \hat{\alpha}\} \text{ and } \tau = \hat{\alpha}. \end{cases}$$

Here, w—in the first clause—is the weight function of the digraph D_T. Let A denote the adjacency matrix of D_T^α. In particular, if $T = \{\alpha, \beta\}$ then $V = \{\epsilon, \alpha, \beta, \hat{\alpha}\}$ and the adjacency matrix is

$$A = \begin{bmatrix} 0 & ux^{|\alpha|} & vx^{|\beta|} & x^{|\alpha|} \\ 0 & u\Omega(\alpha, \alpha) & v\Omega(\alpha, \beta) & \Omega(\alpha, \alpha) \\ 0 & u\Omega(\beta, \alpha) & v\Omega(\beta, \beta) & \Omega(\beta, \alpha) \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Now, $C^\alpha(\boldsymbol{u}; x)$ is the generating function for walks in D_T^α that start at ϵ and end in $\hat{\alpha}$. By the transfer-matrix method and Cramer's rule,

$$C^\alpha(\boldsymbol{u}; x) = (-1)^{|T|+1} \det\big(1 - A : |T| + 2, 1\big) / \det(1 - A).$$

3 Pattern-rewriting systems

The Robinson-Schensted-Knuth (RSK) algorithm associates a permutation $\pi \in \mathcal{S}_n$ with a pair $\big(P(\pi), Q(\pi)\big)$ of standard Young tableaux having the same shape. For $\pi, \sigma \in \mathcal{S}_n$, let us write $\pi \equiv \sigma$ if $P(\pi) = P(\sigma)$. In particular,

$$132 \equiv 312 \quad \text{and} \quad 213 \equiv 231.$$

Knuth [22] showed that, when appropriately extended, these two elementary transformations suffice to determine whether $\pi \equiv \sigma$. Indeed, $\pi \equiv \sigma$ if and only if one can be obtained from the other through a sequence of transformations of the form

$$\dots acb \dots \equiv \dots cab \dots \quad \text{and} \quad \dots bac \dots \equiv \dots bca \dots$$

where $a < b < c$. For instance, $\mathit{423}516 \equiv \mathit{243}516$, $24\mathit{3516} \equiv 24\mathit{3156}$ and, by transitivity, $423516 \equiv 243156$. A relation that is similar to Knuth's called the *forgotten equivalence* has also been studied [28].

Linton, Propp, Roby and West [25] initiated the systematic study of equivalence relations induced by pattern replacement. They considered three types patterns: (i) unrestricted (classical) patterns; (ii) consecutive patterns (adjacent positions); and (iii) patterns in which both positions and values are required to be adjacent. Knuth's equivalence as well as the forgotten equivalence belong to the second type. Patterns of the third type are Hertzsprung patterns. A number of authors have continued the work of Linton et al. For instance, Pierrot, Rossin and West [30] and Kuszmaul [23] have studied equivalences of type (ii). Kuszmaul and Zhou [24] concentrated on equivalences of type (iii). It is case (iii) that we shall also focus on.

We will approach the equivalence problem by considering the identities as unidirectional rewrite rules. We do not assume familiarity with the literature on rewriting systems and will give a brief introduction to the topic. Excellent sources for further reading are the books *Term rewriting and all that* by Baader and Nipkow [3] and *String-rewriting systems* by Book and Otto [6].

An *abstract rewriting system* is simply a set together with a binary relation. We are interested in a special kind of rewriting systems that we call pattern-rewriting systems. To stress their similarity with the well known string-rewriting systems we first introduce those systems.

Let Σ be a finite alphabet and let Σ^* denote the set of (finite) strings with letters from Σ. A *string-rewriting system*—also known as a Semi-Thue system—R on Σ is a subset of $\Sigma^* \times \Sigma^*$. Each pair (u, v) of R is a *rewrite rule* and is often written $u \to v$. The *rewrite relation* $\to_R$ on Σ^* induced by R is defined as follows. For any $w, w' \in \Sigma^*$ we have $w \to_R w'$ if and only if there is $x \to y$ in R such that $w = uxv$ and $w' = uyv$ for some $u, v \in \Sigma^*$.

A *pattern-rewriting system* is a subset $R \subseteq \cup_n \mathcal{S}_n \times \mathcal{S}_n$. The *rewrite relation* on $\mathcal{S}$ induced by R is defined as follows. For any $\pi, \pi' \in \mathcal{S}$, we have $\pi \to_R \pi'$ if and only if there is $\alpha \to \beta$ in R such that we can write $\pi = \sigma\hat{\alpha}\tau$ and $\pi' = \sigma\hat{\beta}\tau$, where $\hat{\alpha} \simeq \alpha$ and $\hat{\beta} \simeq \beta$. As an example, if $R = \{123 \to 132\}$, then $1234 \to_R 1324$ and $1234 \to_R 1243$. The domain of R is denoted by $\mathrm{dom}(R) = \big\{\alpha : (\alpha, \beta) \in R\big\}$. The reflexive transitive closure of $\to_R$ is denoted by $\xrightarrow{*}_R$. The equivalence closure (i.e. the reflexive symmetric transitive closure) of $\to_R$ is denoted by $\equiv_R$. We omit the subscript R from $\to_R$, $\xrightarrow{*}_R$ and $\equiv_R$ when the context prevents ambiguity from being introduced. We use $\leftarrow$ to denote the inverse of $\to$ (the relation $\{(\beta, \alpha) : \alpha \to \beta\}$). Similarity, $\xleftarrow{*}$ denotes the inverse of $\xrightarrow{*}$.

We shall now introduce some definitions and basic results that apply to abstract rewriting system in general and pattern-rewriting systems in particular. An element x is said to be *in normal form* if there is no y such that $x \to y$. We say that y is *a* *normal form of* x if $x \xrightarrow{*} y$ and y is in normal form. The relation $\to$ is said to be *terminating* if there are no infinite chains $x_0 \to x_1 \to x_2 \to \cdots$. If $\to$ is terminating then every element has a normal form. A relation $\to$ is *confluent* if, for all x,

$$y_1 \xleftarrow{*} x \xrightarrow{*} y_2 \implies \exists z.\ y_1 \xrightarrow{*} z \xleftarrow{*} y_2.$$

The relation $\to_R$ with $R = \{123 \to 132\}$ from the previous example is not confluent. Indeed, $1324 \leftarrow 1234 \to 1243$ in which both 1324 and 1243 are in normal form. The relation induced by $\{132 \to 123\}$ is, however, confluent (see Example 3.6 below).

It can be shown that being confluent is equivalent to satisfying

$$x \equiv y \iff \exists z.\ x \xrightarrow{*} z \xleftarrow{*} y.$$

This is known as the *Church-Rosser* property and it provides a test for equivalence. If we assume that $\to$ is confluent (or, equivalently, satisfies the Church-Rosser property) and that $\to$ is terminating, then every element x has a unique normal form which we will write $x{\downarrow}$. Under these assumptions we arrive at a practical test for equivalence.

Lemma 3.1 ([3]). *If $\to$ is terminating and confluent, then $x \equiv y \Leftrightarrow x{\downarrow} = y{\downarrow}$.*

Corollary 3.2. *If R is a terminating and confluent pattern-rewriting system, then the set of* $\operatorname{dom}(R)$*-avoiding permutations in $\mathcal{S}_n$ is a complete set of representatives for $\mathcal{S}_n/{\equiv}$, the set of equivalence classes of $\equiv$.*

For a particular relation, our job is thus reduced to proving that it terminates and that it is confluent. Let us start by discussing strategies for proving confluence. A relation $\to$ is *locally confluent* if

$$y_1 \leftarrow x \to y_2 \implies \exists z.\ y_1 \xrightarrow{*} z \xleftarrow{*} y_2.$$

Newman's lemma [3, 27] establishes that any locally confluent and terminating relation is confluent. This is an improvement, but to test a pattern-rewriting system for local confluence we must a priori still consider all permutations $\pi \in \mathcal{S}$ and all pairs $\pi \to \sigma$ and $\pi \to \tau$. Note, however, that if $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$ are rules in R and $\pi = \pi_1\hat{\alpha}_1\pi_2\hat{\alpha}_2\pi_3$ where $\hat{\alpha}_1 \simeq \alpha_1$ and $\hat{\alpha}_2 \simeq \alpha_2$, then

$$\begin{tikzcd}
& \pi_1\hat{\alpha}_1\pi_2\hat{\alpha}_2\pi_3 \arrow[dl] \arrow[dr] & \\
\pi_1\hat{\beta}_1\pi_2\hat{\alpha}_2\pi_3 \arrow[dr] & & \pi_1\hat{\alpha}_1\pi_2\hat{\beta}_2\pi_3 \arrow[dl] \\
& \pi_1\hat{\beta}_1\pi_2\hat{\beta}_2\pi_3 &
\end{tikzcd}$$

in which $\hat{\beta}_1 \simeq \beta_1$ and $\hat{\beta}_2 \simeq \beta_2$. Thus, to prove local confluence, we can restrict attention to permutations of the form $\pi = \pi_1\gamma\pi_2$ and pairs of, not necessarily distinct, rules $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$ in R such that α_1 is a proper Hertzsprung prefix of γ, α_2 is a proper Hertzsprung suffix of γ and $|\gamma| < |\alpha_1| + |\alpha_2|$. The situation is summarized in the following lemma.

Lemma 3.3. *Let R be a terminating pattern-rewriting system. Let*

$$\mathcal{O}(R) = \bigcup_{(\alpha_1,\alpha_2)} \mathcal{O}(\alpha_1, \alpha_2),$$

where the union is over all pairs $(\alpha_1, \alpha_2) \in \mathrm{dom}(R) \times \mathrm{dom}(R)$. If, for all $\pi \in \mathcal{O}(R)$,

$$\rho_1 \leftarrow \pi \rightarrow \rho_2 \implies \exists\sigma.\ \rho_1 \xrightarrow{*} \sigma \xleftarrow{*} \rho_2,$$

then $\rightarrow$ is confluent.

As a corollary to Lemma 3.3, confluence of a finite and terminating pattern-rewriting system is decidable.

A relation $\rightarrow$ is called *globally finite* if for each element x there are finitely many elements y such that $x \xrightarrow{+} y$, where $\xrightarrow{+}$ denotes the transitive closure of $\rightarrow$. A relation $\rightarrow$ is called *acyclic* if there is no element x such that $x \xrightarrow{+} x$. It is well known that any acyclic and globally finite relation is terminating. Let R be a pattern-rewriting system. Since each rule in R preserves the length of a permutation, it is clear that $\rightarrow_R$ is globally finite. Thus to prove that $\rightarrow_R$ terminates it suffices to prove that that it is acyclic. Let us call $f : \mathcal{S} \rightarrow \mathbb{N}$ an *increasing statistic* with respect to R if $\pi \rightarrow_R \sigma$ implies $f(\pi) < f(\sigma)$. Clearly, if there exists such a function f then $\rightarrow_R$ is acyclic.

Lemma 3.4. *Let R be a pattern-rewriting system. If there exists an increasing statistic with respect to R, then $\rightarrow_R$ is terminating.*

Having established the general framework we will now turn to applications. A summary of the cases we shall consider can be found in Table 1; EQ2 through EQ6 were introduced by Linton et al. [25] and in the case of Hertzsprung patterns they left the enumeration of equivalence classes as open problems. Kuszmaul and Zhou [24] have characterized the equivalence classes induced by the three separate relations $123 \equiv 132$, $123 \equiv 321$ and $123 \equiv 132 \equiv 321$. These are referred to as EQ2, EQ3 and EQ4 in Table 1. Kuszmaul and Zhou provided explicit formulas for the number of equivalence classes modulo EQ2 and EQ3 and noted that the formula is the same in both cases. Indeed, it is formula (5), which is a special case of Myers's formula (2). For each of EQ1 through EQ7 we compute generating functions for the number of equivalence classes in $\mathcal{S}_n$. In particular we solve the two remaining open problems, EQ5 and EQ6, posed by Linton et al. The generating function for EQ7 is also new. Our approach is quite general and could be applied other pattern-replacement equivalences of the Hertzsprung type. The reason for choosing EQ1–EQ6 is that those equivalences have a history, and that we need to limit the scope of our investigation. We choose EQ7 for two reasons. It is a natural generalization of Stanley's equivalence (EQ1) and it illustrates that large systems are within reach of our methods.

For reference, numerical data on the number of equivalence classes modulo each of EQ2 through EQ7 is given in Table 2 of the appendix.

Example 3.5 (EQ1)**.** Stanley [33] considered the equivalence relation on $\mathcal{S}_n$ generated by the interchange of any two adjacent elements $\pi(i)$ and $\pi(i+1)$ such that $|\pi(i) - \pi(i+1)| = 1$. In our terminology it is the equivalence $12 \equiv 21$. Let us express

	Equivalences	Rules	OEIS	Reference
EQ1	$12 \equiv 21$	$21 \to 12$ $231 \to 312$	A013999	Stanley [33]
EQ2	$123 \equiv 132$	$132 \to 123$	A212580	Kuszmaul & Zhou [24]
EQ3	$123 \equiv 321$	$321 \to 123$ $2341 \to 4123$	A212580	Kuszmaul & Zhou [24]
EQ4	$123 \equiv 132$ $\equiv 213$	$132 \to 123$ $213 \to 123$	A212581	Kuszmaul & Zhou [24]
EQ5	$123 \equiv 132$ $\equiv 321$	$132 \to 123$ $321 \to 123$ $2341 \to 4123$	A212432	Example 3.9
EQ6	$123 \equiv 132$ $\equiv 213$ $\equiv 321$	$132 \to 123$ $213 \to 123$ $321 \to 123$ $2341 \to 4123$	A212433	Example 3.10
EQ7	$123 \equiv 132$ $\equiv 213$ $\equiv 231$ $\equiv 312$ $\equiv 321$	$132 \to 123$ $213 \to 123$ $231 \to 123$ $312 \to 123$ $321 \to 123$ $2341 \to 4123$ $34512 \to 45123$ $54123 \to 45123$ $6745123 \to 7456123$		Example 3.11

Table 1: Equivalences and corresponding pattern-rewriting systems

Stanley's equivalence relation using a rewrite system. Consider $\{21 \to 12\}$. Note that $231 \leftarrow 321 \to 312$ in which both 231 and 312 are in normal form. Hence this system is not confluent. We can however make it into a confluent system by adding a rule. Let $R = \{21 \to 12, 231 \to 312\}$. To see that the rewrite rule induced by R terminates consider the following permutation statistic. Let $\Sigma_{12}(\pi)$ be the sum of positions of occurrences of the Hertzsprung pattern 12. In other words, $\Sigma_{12}(\pi)$ is the sum of all $i \in [n-1]$ such that $\pi(i+1) = \pi(i)+1$. Clearly, Σ_{12} is increasing with respect to R and hence $\to_R$ terminates. To show confluence we will use Lemma 3.3. We have $\mathcal{O}(21,21) = \{321\}$, $\mathcal{O}(21,231) = \mathcal{O}(231,231) = \emptyset$, $\mathcal{O}(231,21) = \{3421\}$ and hence $\mathcal{O}(R) = \{321, 3421\}$. Confluence thus follows from the two diagrams below.

$$\begin{tikzcd}
& 3421 \arrow[dl] \arrow[ddd] \\
4231 \arrow[d] & \\
4312 \arrow[dr] & \\
& 3412
\end{tikzcd}
\qquad
\begin{tikzcd}
321 \arrow[d] \arrow[dr] & \\
231 \arrow[r] & 312
\end{tikzcd}$$

By Corollary 3.2, each equivalence class under $12 \equiv 21$ contains exactly one permutation avoiding $\{21, 231\}$. Hence the generating function for the number of equivalence classes can be computed using Theorem 2.1 together with Theorem 2.6.

Stanley defines a permutation $\pi = a_1a_2\cdots a_n \in \mathcal{S}_n$ as *salient* if we never have $a_i = a_{i+1}+1$ $(1 \le i \le n-1)$ or $a_i = a_{i+1}+2 = a_{i+2}+1$ $(1 \le i \le n-2)$. In terms of Hertzsprung patterns, π is salient if and only if it is $\{21, 312\}$-avoiding. Stanley shows that each equivalence class under $12 \equiv 21$ contains exactly one salient permutation. On account of the preceding paragraph, this implies that the two sets $\{21, 231\}$ and $\{21, 312\}$ are Wilf-equivalent (their sets of avoiders are equipotent).

Stanley counts the salient permutations by a direct inclusion-exclusion argument. The resulting generating function is

$$\sum_{m\geq 0} m!\,(x(1-x))^m\,.$$

We, of course, obtain the same answer. In fact, theorems 2.1 and 2.6 show that the joint distribution of 21 and 231 is the same as the joint distribution of 21 and 312. In both cases the cluster generating function is $x^2(vx+u)/(1-ux)$, where $u = u_{21}$ and $v = u_{231}$ or $v = u_{312}$.

Example 3.6 (EQ2)**.** Consider the pattern-rewriting system $R = \big\{132 \to 123\big\}$. As usual, let $\to$ also denote the rewrite relation induced by R. The equivalence closure of $\to$ is equal to the equivalence relation induced by $123 \equiv 132$. Note that $\to$ is terminating. Indeed, the number of occurrences of 123 is an increasing statistic (Σ_{12} is also an increasing statistic). Moreover, since $\mathcal{O}(132,132)$ is empty, $\to$ is trivially confluent. By Corollary 3.2, the number of equivalence classes is $|\mathcal{S}_n(132)|$ for which we derived a formula in Example 2.3.

Example 3.7 (EQ3)**.** The relation induced by $123 \equiv 321$ is the equivalence closure of the rewrite relation $\to$ induced by $R = \{321 \to 123, 2341 \to 4123\}$. This relation is terminating since Σ_{123}, the sum of positions of occurrences of 123, is an increasing statistic. We shall show that it is confluent as well. The relation induced by $\{321 \to 123\}$ is not confluent, that is why we include the rule $2341 \to 4123$. We have $\mathcal{O}(321, 2341) = \mathcal{O}(2341, 2341) = \emptyset$,

$$\mathcal{O}(321, 321) = \{4321, 54321\} \quad \text{and} \quad \mathcal{O}(2341, 321) = \{456321\}.$$

Confluence thus follows from Lemma 3.3 and the three diagrams below.

$$\begin{tikzcd}
 & 456321 \arrow[dl] \arrow[dddd] \\
634521 \arrow[d] & \\
652341 \arrow[d] & \\
654123 \arrow[dr] & \\
 & 456123
\end{tikzcd}
\qquad
\begin{tikzcd}
4321 \arrow[d] \arrow[dr] & \\
2341 \arrow[r] & 4123
\end{tikzcd}
\qquad
\begin{tikzcd}
54321 \arrow[d] \arrow[dr] \arrow[drr] & & \\
34521 \arrow[r] & 52341 \arrow[r] & 54123
\end{tikzcd}$$

Linton et al. [25] noted that the number of equivalence classes under $123 \equiv 132$ and $123 \equiv 321$ appears to be the same. This was later proved by Kuszmaul and Zhou [24]. The $\{321, 2341\}$-cluster generating function is

$$\frac{u_{321}u_{2341}x^5 - u_{2341}x^4 - u_{321}x^3}{u_{321}x^2 + u_{321}x - 1}.$$

Recall that the generating function for 132-clusters is $u_{132}x^3$. While these generating functions are quite different, on specializing u_{2341}, u_{321} and u_{132} to -1 we get $-x^3$ in both instances, resulting in an alternative proof of the observation made by Linton et al. In particular, the two sets of patterns $\{132\}$ and $\{321, 2341\}$ are Wilf-equivalent. Finding a combinatorial proof of this fact remains an open problem.

Example 3.8 (EQ4)**.** Let $R = \{132 \to 123, 213 \to 123\}$. Its equivalence closure coincides with the equivalence relation induced by $123 \equiv 132 \equiv 213$. That $\to$ terminates follows from Σ_{123} being an increasing statistic. Note that

$$\mathcal{O}(R) = \mathcal{O}(132, 213) \cup \mathcal{O}(213, 132) = \{1324, 21354\}.$$

Using either $132 \to 123$ or $213 \to 123$ we have $1324 \to 1234$ in which 1234 is in normal form. The diagram for 21354 is rather simple too:

$$\begin{tikzcd}
 & 21354 \arrow[dl] \arrow[dr] & \\
12354 \arrow[dr] & & 21345 \arrow[dl] \\
 & 12345 &
\end{tikzcd}$$

Thus, by Corollary 3.2, the number of equivalence classes under $123 \equiv 132 \equiv 213$ is the same as the number of permutations of $[n]$ that avoid 132 and 213, and the corresponding generating function is

$$\sum_{n \geq 0} |\mathcal{S}_n(132, 213)| x^n = \sum_{m \geq 0} m! \left(\frac{(1+x)^2(1-x)x}{x^2 + x + 1} \right)^m.$$

The coefficient of x^n is given by Kuszmaul and Zhou's formula [24, Theorem 2.32]

$$\sum_{i=0}^{\lfloor n/3 \rfloor} (-2)^i \binom{n-2i}{i} (n-2i)!$$

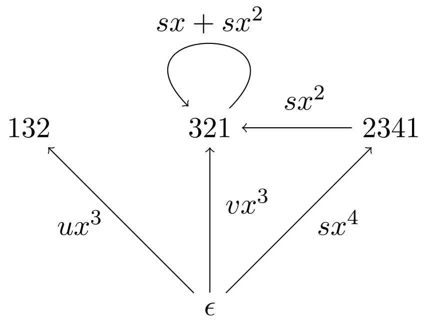
Example 3.9 (EQ5)**.** Consider the pattern-rewriting system

$$R = \{132 \to 123, 321 \to 123, 2341 \to 4123\}.$$

The equivalence closure of $\to$ is induced by $123 \equiv 132 \equiv 321$. There is no overlap of the pattern 132 with 123 or 2341. Hence

$$\mathcal{O}(R) = \mathcal{O}(321, 321) \cup \mathcal{O}(2341, 321) = \{4321, 54321, 456321\}$$

and the proofs of confluence and termination are the same as in Example 3.7 (EQ3). Thus the number of equivalence classes under $\equiv$ equals the number of permutations avoiding the patterns in $\mathrm{dom}(R) = \{132, 321, 2341\}$. To count such permutations we construct the corresponding digraph. Letting $u = u_{132}$, $v = u_{321}$ and $s = u_{2341}$ it is depicted below.



The cluster generating function is

$$C(u, v, s; x) = \frac{(uv + vs)x^5 + (uv - s)x^4 - (u + v)x^3}{vx^2 + vx - 1}$$

and on applying Theorem 2.1 we find that

$$\sum_{n \geq 0} |\mathcal{S}_n(132, 321, 2341)| x^n = \sum_{m \geq 0} m! \left(x(1 - 2x^2) \right)^m .$$

Example 3.10 (EQ6)**.** Consider the pattern-rewriting system

$$R = \{132 \to 123, 213 \to 123, 321 \to 123, 2341 \to 4123\}.$$

Its equivalence closure coincides with the equivalence relation induced by $123 \equiv 132 \equiv 213 \equiv 321$. That $\to$ terminates follows from Σ_{123} being an increasing statistic with respect to R. Note that $\mathcal{O}(R) = \{1324, 4321, 21354, 54321, 456321\}$. In Example 3.7 (EQ3) we have seen diagrams for 4321, 54321 and 456321. This leaves 1324 and 21354, but we have seen diagrams for those permutations in Example 3.8 (EQ4). As before, we use Theorem 2.6 to compute the cluster generating function. In the end we find that the generating function for the number of equivalence classes under $123 \equiv 132 \equiv 213 \equiv 321$ is

$$\sum_{n \geq 0} |\mathcal{S}_n(132, 213, 321, 2341)| x^n = \sum_{m \geq 0} m! \left(\frac{-x^5 - 2x^4 - 2x^3 + x^2 + x}{x^2 + x + 1} \right)^m .$$

Example 3.11 (EQ7)**.** As detailed in Example 3.5 (EQ1), Stanley characterized and counted the equivalence classes induced by $12 \equiv 21$. In other words, the equivalences induced by considering all permutations in $\mathcal{S}_2$ as equivalent. We shall consider the corresponding problem for $\mathcal{S}_3$. That is, the equivalence induced by

$$123 \equiv 132 \equiv 213 \equiv 231 \equiv 312 \equiv 321.$$

Let R consist of the rules $\alpha \to 123$, for $\alpha \in \mathcal{S}_3 \setminus \{123\}$, together with the following four rules that are needed to make the relation confluent:

$$2341 \to 4123,\ \ 34512 \to 45123,\ \ 54123 \to 45123,\ \ 6745123 \to 7456123.$$

One can check that Σ_{12} is an increasing statistic and hence $\to$ terminates. Confluence follows from verifying local confluence for each permutation in $\mathcal{O}(R)$, which is the union of the sets

$$\begin{gathered}
\mathcal{O}(123, 213) = \{1324\},\ \ \mathcal{O}(213, 132) = \{21354\},\ \ \mathcal{O}(231, 312) = \{45312\},\\
\mathcal{O}(231, 321) = \{45321\},\ \ \mathcal{O}(231, 54123) = \{6754123\},\ \ \mathcal{O}(312, 231) = \{4231\},\\
\mathcal{O}(312, 6745123) = \{86745123\},\ \ \mathcal{O}(321, 312) = \{54312\},\\
\mathcal{O}(321, 321) = \{4321, 54321\},\ \ \mathcal{O}(321, 54123) = \{654123, 7654123\}\\
\mathcal{O}(2341, 312) = \{456312\},\ \ \mathcal{O}(2341, 321) = \{456321\},\\
\mathcal{O}(2341, 54123) = \{67854123\},\ \ \mathcal{O}(34512, 231) = \{456231\},\\
\mathcal{O}(34512, 6745123) = \{8\,9\,10\,6\,7\,4\,5\,1\,2\,3\},\ \ \mathcal{O}(54123, 2341) = \{652341\},\\
\mathcal{O}(54123, 34512) = \{7634512\},\ \ \mathcal{O}(6745123, 2341) = \{78562341\},\\
\mathcal{O}(6745123, 34512) = \{896734512\}.
\end{gathered}$$

We omit the details. Thus, the two cardinalities $|\mathcal{S}_n/{\equiv}|$ and $|\mathcal{S}_n(\mathrm{dom}(R))|$ are equal. The adjacency matrix of D_T (with $T = \mathrm{dom}(R)$) is

$$\begin{bmatrix}
0 & u_0x^3 & u_1x^3 & u_2x^3 & u_3x^3 & u_4x^3 & u_5x^4 & u_6x^5 & u_7x^5 & u_8x^7\\
0 & 0 & u_1x & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & u_0x^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & u_3x^2 & u_4x^2 & 0 & 0 & u_7x^4 & 0\\
0 & 0 & 0 & u_2x & 0 & 0 & 0 & 0 & 0 & u_8x^5\\
0 & 0 & 0 & 0 & u_3x^2 & u_4x^2 + u_4x & 0 & 0 & u_7x^4 + u_7x^3 & 0\\
0 & 0 & 0 & 0 & u_3x^2 & u_4x^2 & 0 & 0 & u_7x^4 & 0\\
0 & 0 & 0 & u_2x & 0 & 0 & 0 & 0 & 0 & u_8x^5\\
0 & 0 & 0 & 0 & 0 & 0 & u_5x & u_6x^2 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & u_5x & u_6x^2 & 0 & 0
\end{bmatrix}$$

and the corresponding generating function is

$$\sum_{m\geq 0} m!\left(\frac{-x^5 - 3x^4 - 4x^3 + x^2 + x}{x^2 + x + 1}\right)^m.$$

Questions, conjectures and remarks

Consider two Hertzsprung patterns $\sigma, \tau \in \mathcal{S}_k$ as equivalent if $|\mathcal{S}_n(\sigma)| = |\mathcal{S}_n(\tau)|$ for all $n \geq 0$. How many equivalence classes are there? In other words, what is the number of Wilf-classes? By Corollary 2.7 it is the same as the number of distinct "autocorrelation polynomials" $\Omega(\sigma,\sigma)$ with $\sigma \in \mathcal{S}_k$. For $k = 1, 2, \ldots, 7$ we find

$$\begin{gathered}
\{\, 0 \,\} \\
\{\, x \,\} \\
\{\, 0,\, x^2 + x \,\} \\
\{\, 0,\, x^2,\, x^3,\, x^3 + x^2 + x \,\} \\
\{\, 0,\, x^3,\, x^4,\, x^4 + x^3 + x^2 + x \,\} \\
\{\, 0,\, x^3,\, x^4,\, x^4 + x^2,\, x^5,\, x^5 + x^4,\, x^5 + x^4 + x^3 + x^2 + x \,\} \\
\{\, 0,\, x^4,\, x^5,\, x^6,\, x^6 + x^3,\, x^6 + x^5,\, x^6 + x^5 + x^4 + x^3 + x^2 + x \,\}
\end{gathered}$$

and the sequence for the number of Wilf-classes starts

$$1, 1, 2, 4, 4, 7, 7, 11, 12, 18, 17, 25, 27, 38, 38 \quad (k = 1, 2, \ldots, 15)$$

For $k \geq 3$ these numbers coincide with those of sequence A304178 in the OEIS, which leads us to the conjecture below. For a palindrome $w = c_1 \ldots c_n \in \{0,1\}^n$ let $P(w) = \{i \in [n] : c_1 \ldots c_i \text{ is a palindrome}\}$ be the set of palindrome prefix lengths of w; e.g. $P(0100010) = \{1, 3, 7\}$.

Conjecture 3.12. *Let $a_k = |\{\Omega(\sigma,\sigma) : \sigma \in \mathcal{S}_k\}|$ be the number of Wilf-classes of Hertzsprung patterns of length k. Let b_k be the number of distinct sets $P(w)$ for palindromes $w \in \{0,1\}^k$. Then $a_k = b_{k+1}$ for $k \geq 3$.*

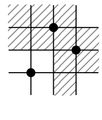
Recall that Bóna [5] showed that $|\mathcal{S}_n(\tau)| \leq |\mathcal{S}_n(\mathrm{id}_k)|$ for each non-self-overlapping pattern $\tau \in \mathcal{S}_k$. Data collected using Corollary 2.7 suggest that the inequality holds regardless of whether τ is self-overlapping or not.

Conjecture 3.13. *We have $|\mathcal{S}_n(\tau)| \leq |\mathcal{S}_n(\mathrm{id}_k)|$ for all Hertzsprung patterns $\tau \in \mathcal{S}_k$ and all natural numbers n.*

In view of Theorem 2.1 it is natural to ask if there are nontrivial mesh-patterns p other than the Hertzsprung patterns for which there is a rational function $R(x)$ such that $\sum_{n\geq 0} |\mathcal{S}(p)|x^n = \sum_{m\geq 0} m!R(x)^m$? It appears that the answer is yes.

Conjecture 3.14. *We have*

$$\sum_{n\geq 0} |\mathcal{S}_n(p)| x^n = \sum_{m\geq 0} m! \left(\frac{x}{1+x^2}\right)^m, \quad \textit{where } p =$$



This conjecture is based on computing the numbers $|\mathcal{S}_n(p)|$ for $n \leq 14$. They are 1, 1, 2, 5, 20, 103, 630, 4475, 36232, 329341, 3320890, 36787889, 444125628, 5803850515, and 81625106990. At the time of writing this sequence is not in the OEIS [17]. If the conjecture is true then there is, however, a close connection with a sequence in the OEIS, namely A177249. It is defined by letting a_n be number of permutations of $[n]$ whose disjoint cycle decompositions have no adjacent transpositions, that is, no cycles of the form $(i, i+1)$. Let $F(x) = \sum_{n\geq 0} |\mathcal{S}_n(p)|x^n$ be the sought series and let $A(x) = \sum_{n\geq 0} a_n x^n$. Brualdi and Deutsch [8] have shown that $(1 + x^2)A(x) = F(x)$.

Our conjecture is thus equivalent to $|\mathcal{S}_n(p)| = a_n + a_{n-2}$ for $n \geq 2$. Alternatively, one may note that the compositional inverse of $x/(1+x^2)$ is $xC(x^2)$, where $C(x)$ is the generating function for the Catalan numbers. Thus, our conjecture is also equivalent to $[x^n]F(xC(x^2)) = n!$, which could lead to a novel decomposition of permutations.

As defined, pattern-rewriting systems are limited to transforming Hertzsprung patterns. This is an artificial limitation though. Let $\mathcal{M}_k = \mathcal{P}\big([0,k] \times [0,k]\big)$ be the set of $(k+1) \times (k+1)$ meshes. Define a *pattern-rewriting system* as a set of rules $R \subseteq \bigcup_{k\geq 0}(\mathcal{S}_k \times \mathcal{M}_k \times \mathcal{S}_k)$. Each rule (α, M, β) is extended to $\mathcal{S} = \cup_n \mathcal{S}_n$ by allowing that any occurrence of the mesh pattern (α, M) is rewritten as an occurrence of β. In particular, this encompasses the three types of equivalences studied by Linton et al., $M = \emptyset$, $M = [1, k-1] \times [0, k]$ and $M = H_k$.

Using Lemma 3.3 we are able to automate the proof of confluence for any finite and terminating pattern-rewriting system. Our current strategy to prove that a pattern-rewriting system terminates requires us to manually come up with an increasing statistic. This was easy for the small systems that we have considered but may not be easy in general. Can there be a mechanical test? Or is this property undecidable? It is known that termination of string-rewriting systems is undecidable, even assuming that the system is length-preserving [9].

The starting point of this article was Hertzsprung's problem and we have implicitly considered it the simplest nontrivial instance of a larger class of problems. There is arguably a simpler, yet nontrivial, instance though: permutations avoiding the single Hertzsprung pattern 12. The study of the corresponding counting sequence appears to have started with Euler. We have

$$\sum_{\pi\in\mathcal{S}} u^{12(\pi)} x^{|\pi|} = \sum_{m\geq 0} m! \left(\frac{x}{1-(u-1)x}\right)^m.$$

One may also note that any permutation π can be uniquely written as an inflation

$$\pi = \sigma[\mathrm{id}_{k_1}, \mathrm{id}_{k_2}, \ldots, \mathrm{id}_{k_m}],$$

where σ avoids the Hertzsprung pattern 12. Allowing ourselves to use the terminology of L-species [4] this amounts to the isomorphism $S' = F' \cdot E$, where E, S and F are the L-species of sets, permutations and 12-avoiding permutations, respectively. Indeed, suppose $n \geq 1$ and write $\pi \in S_n$ as $\pi = \sigma[\mathrm{id}_{k_1}, \mathrm{id}_{k_2}, \ldots, \mathrm{id}_{k_m}] = \pi_1\pi_2\cdots\pi_m$. Let $\hat{\sigma}$ be the permutation obtained by taking the first element of each π_i and let the set A consist of all the remaining elements. Then $\pi \mapsto (\hat{\sigma}, A)$ is the sought isomorphism. Clearly $\hat{\sigma} \simeq \sigma$ and $12(\pi) = |A|$. For instance,

$$3142[12, 1, 1234, 12] = 451678923 \;\mapsto\; \big(4162,\ \{3,5,7,8,9\}\big).$$

Since $S(x) = (1-x)^{-1}$ and $E(x) = e^x$ it follows from $S'(x) = F'(x)E(x)$ that $F'(x) = e^{-x}/(1-x)^2$. Replacing the species E with the $\mathbb{Z}[u]$-weighted species E_u in which each set A has weight $u^{|A|}$ and, similarly, replacing the species S with S_w where the weight function is $w(\pi) = u^{12(\pi)}$, we have $S'_w = F' \cdot E_u$ and hence

$$\sum_{n\geq 0}\Bigg(\sum_{\pi\in\mathcal{S}_n} u^{12(\pi)}\Bigg)\frac{x^n}{n!} \;=\; 1 + \int_0^x \frac{e^{(u-1)t}}{(1-t)^2}\,dt.$$

Even simple Hertzsprung patterns have interesting properties!

Appendix

Table 2 gives the number of equivalence classes modulo EQ2–EQ7 for $1 \leq n \leq 20$.

References

[1] Morton Abramson and W. O. J. Moser. Permutations without rising or falling w-sequences. *Ann. Math. Statist.*, 38:1245–1254, 1967.

[2] M. H. Albert and M. D. Atkinson. Simple permutations and pattern restricted permutations. *Discrete Mathematics*, 300(1):1 – 15, 2005.

[3] Franz Baader and Tobias Nipkow. *Term rewriting and all that.* Cambridge University Press, Cambridge, 1998.

[4] F. Bergeron, G. Labelle, and P. Leroux. *Combinatorial species and tree-like structures*, volume 67 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1998. Translated from the 1994 French original by Margaret Readdy, With a foreword by Gian-Carlo Rota.

[5] Miklós Bóna. Where the monotone pattern (mostly) rules. *Discrete mathematics*, 308(23):5782–5788, 2008.

[6] Ronald V. Book and Friedrich Otto. *String-rewriting systems.* Texts and Monographs in Computer Science. Springer-Verlag, New York, 1993.

[7] Petter Brändén and Anders Claesson. Mesh patterns and the expansion of permutation statistics as sums of permutation patterns. *Electronic Journal of Combinatorics*, 18(2):P5, 2011.

[8] Richard A. Brualdi and Emeric Deutsch. Adjacent q-cycles in permutations. *Ann. Comb.*, 16(2):203–213, 2012.

[9] A-C Caron. Linear bounded automata and rewrite systems : Influence of initial configurations on decision properties. In S. Abramsky and T. S. E. Maibaum, editors, *TAPSOFT '91*, pages 74–89, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.

[10] Vladimir Dotsenko and Anton Khoroshkin. Shuffle algebras, homology, and consecutive pattern avoidance. *Algebra & Number Theory*, 7(3):673–700, 2013.

[11] Sergi Elizalde and Marc Noy. Clusters, generating functions and asymptotics for consecutive patterns in permutations. *Advances in Applied Mathematics*, 49(3-5):351–374, 2012.

[12] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics.* cambridge University press, 2009.

[13] Martin Gardner. On the paradoxical situations that arise from nontransitive relations. *Scientific American*, 231(4):120–125, 1974.

	EQ2 & EQ3 (A212580)	EQ4 (A212581)	EQ5 (A212432)
1	1	1	1
2	2	2	2
3	5	4	4
4	20	17	16
5	102	89	84
6	626	556	536
7	4458	4011	3912
8	36144	32843	32256
9	328794	301210	297072
10	3316944	3059625	3026112
11	36755520	34104275	33798720
12	443828184	413919214	410826624
13	5800823880	5434093341	5399704320
14	81591320880	76734218273	76317546240
15	1228888215960	1159776006262	1154312486400
16	19733475278880	18681894258591	18604815528960
17	336551479543440	319512224705645	318348065548800
18	6075437671458000	5782488507020050	5763746405053440
19	115733952138747600	110407313135273127	110086912964367360
20	2320138519554562560	2218005876646727423	2212209395234979840

	EQ6 (A212433)	EQ7	
1	1	1	
2	2	2	
3	3	1	
4	13	6	
5	71	40	
6	470	330	
7	3497	2664	
8	29203	23258	
9	271500	222154	
10	2786711	2326410	
11	31322803	26568950	
12	382794114	328995136	
13	5054810585	4392819522	
14	71735226535	62935547966	
15	1088920362030	963253101304	
16	17607174571553	15688298164890	
17	302143065676513	270944692450742	
18	5484510055766118	4946387077324072	
19	104999034898520903	95184319122508074	
20	2114467256458136473	1925732716758497918	

Table 2: Enumeration of equivalence classes EQ2–EQ7

[14] Ian P. Goulden and D. M. Jackson. An inversion theorem for cluster decompositions of sequences with distinguished subsequences. *Journal of the London Mathematical Society*, 2(3):567–576, 1979.

[15] Leonidas J. Guibas and Andrew M. Odlyzko. String overlaps, pattern matching, and nontransitive games. *Journal of Combinatorial Theory, Series A*, 30(2):183–208, 1981.

[16] Severin Hertzsprung. En kombinationsopgave. *Tidsskrift for mathematik*, 5:13–17, 1887.

[17] OEIS Foundation Inc. The Online Encyclopedia of Integer Sequences. `https://oeis.org`, 2020.

[18] D. M. Jackson and Romas Aleliunas. Decomposition based generating functions for sequences. *Canadian Journal of Mathematics*, 29(5):971–1009, 1977.

[19] D. M. Jackson and R. C. Read. A note on permutations without runs of given length. *Aequationes mathematicae*, 17(1):336–343, 1978.

[20] D. M. Jackson and J. W. Reilly. Permutations with a prescribed number of p-runs. *Ars Combinatoria*, 1(1):297–305, 1976.

[21] I. Kaplansky. Symbolic solution of certain problems in permutations. *Bulletin of the American Mathematical Society*, 50(12):906–914, 1944.

[22] Donald E. Knuth. Permutations, matrices, and generalized Young tableaux. *Pacific J. Math.*, 34:709–727, 1970.

[23] William Kuszmaul. Counting permutations modulo pattern-replacement equivalences for three-letter patterns. *Electron. J. Combin.*, 20(4):Paper 10, 58, 2013.

[24] William Kuszmaul and Ziling Zhou. New results on families of pattern-replacement equivalences. *Discrete Mathematics*, 343(7):111878, 2020.

[25] Steven Linton, James Propp, Tom Roby, and Julian West. Equivalence classes of permutations under various relations generated by constrained transpositions. *Journal of Integer Sequences*, 15(2):3, 2012.

[26] Amy N. Myers. Counting permutations by their rigid patterns. *Journal of Combinatorial Theory Series A*, 99(2):345–357, 2002.

[27] M. H. A. Newman. On theories with a combinatorial definition of "equivalence". *Annals of Mathematics*, 43(2):223–243, 1942.

[28] Jean-Christophe Novelli and Anne Schilling. The forgotten monoid. *RIMS Kokyuroku Bessatsu B8 (arXiv:0706.2996)*, 2007.

[29] Walter Penney. 95. Penney-ante. *Journal of Recreational Mathematics*, 2:241, 1969.

[30] Adeline Pierrot, Dominique Rossin, and Julian West. Adjacent transformations in permutations. *Discrete Mathematics & Theoretical Computer Science*, DMTCS Proceedings vol. AO, 23rd International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2011), 2011.

[31] John Riordan. Permutations without 3-sequences. *Bulletin of the American Mathematical Society*, 51(10):745–748, 1945.

[32] Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.

[33] Richard P. Stanley. An equivalence relation on the symmetric group and multiplicity-free flag h-vectors. *J. Comb.*, 3(3):277–298, 2012.