Available online at www.sciencedirect.com
ScienceDirect
Discrete Mathematics 307 (2007) 1722–1736
DISCRETE MATHEMATICS
www.elsevier.com/locate/disc

On arithmetic and asymptotic properties of up–down numbers

Francis C.S. Brown[a], Thomas M.A. Fink[b], Karen Willbrand[c]

[a] *Ecole Normale Superieure, 45 Rue d'Ulm, Paris, France*
[b] *Institut Curie, 26 Rue d'Ulm, Paris, France*
[c] *Laboratoire de Physique Statistique, 24 Rue Lhomond, Paris, France*

Received 20 May 2006; received in revised form 29 July 2006; accepted 2 September 2006
Available online 14 November 2006

Abstract

Let $\sigma = (\sigma_1, \ldots, \sigma_N)$, where $\sigma_i = \pm 1$, and let $C(\sigma)$ denote the number of permutations π of $1, 2, \ldots, N + 1$, whose up–down signature $\operatorname{sign}(\pi(i + 1) - \pi(i)) = \sigma_i$, for $i = 1, \ldots, N$. We prove that the set of all up–down numbers $C(\sigma)$ can be expressed by a single universal polynomial Φ, whose coefficients are products of numbers from the Taylor series of the hyperbolic tangent function. We prove that Φ is a modified exponential, and deduce some remarkable congruence properties for the set of all numbers $C(\sigma)$, for fixed N. We prove a concise upper bound for $C(\sigma)$, which describes the asymptotic behaviour of the up–down function $C(\sigma)$ in the limit $C(\sigma) \ll (N + 1)!$.

Keywords: Up–down numbers; Euler–Bernoulli numbers; Random landscape; Tangent numbers

1. Introduction

Let $N \geqslant 1$, and let π be a permutation of $\{1, 2, \ldots, N + 1\}$. The up–down signature of π is defined to be the sequence $\sigma = (\sigma_1, \ldots, \sigma_N) \in \{1, -1\}^N$ of rises and falls of π. More precisely, the up–down signature σ is given by the formula:

$$\sigma_i = \operatorname{sign}(\pi(i + 1) - \pi(i)) \qquad \text{for } 1 \leqslant i \leqslant N.$$

Let $C(\sigma)$ denote the number of permutations π which have up–down signature σ. Some small values of the up–down numbers $C(\sigma)$ are listed in Table 1.

The enumeration of permutations with given up–down signatures is a long-standing combinatorial problem initiated by André [2], who computed the number of permutations with the alternating signature of length N: $A_N = C(+ - + - \ldots)$. The numbers A_N are called Euler–Bernoulli updown numbers and are given by the Taylor expansion of $\tan z + \sec z$. These numbers arose in the study of morsifications in singularity theory by Arnold [3], who also proved some surprising arithmetic properties for them. Many variants of these numbers have been studied extensively by Carlitz and Carlitz-Scoville (see e.g., [5,6]). The numbers $C(\sigma)$ for arbitrary σ can be regarded as a natural generalisation of the numbers A_N, but are altogether less well-understood. They have been studied in various combinatorial contexts

E-mail address: brown@clipper.ens.fr (F.C.S. Brown).

0012-365X/$ - see front matter
doi:10.1016/j.disc.2006.09.020

Table 1
The number $C(\sigma)$ of permutations on $N+1$ letters with given up–down signature σ of length N

$N=1$		$N=2$		$N=3$		$N=4$		$N=5$	
σ	$C(\sigma)$	σ	$C(\sigma)$	σ	$C(\sigma)$	σ	$C(\sigma)$	σ	$C(\sigma)$
$-$	1	$--$	1	$---$	1	$----$	1	$-----$	1
$+$	1	$-+$	2	$--+$	3	$---+$	4	$----+$	5
		$+-$	2	$-+-$	5	$--+-$	9	$---+-$	14
		$++$	1	$-++$	3	$--++$	6	$---++$	10
				$+--$	3	$-+--$	9	$--+--$	19
				$+-+$	5	$-+-+$	16	$--+-+$	35
				$++-$	3	$-++-$	11	$--++-$	26
				$+++$	1	$-+++$	4	$--+++$	10
						$+---$	4	$-+---$	14
						$+--+$	11	$-+--+$	40
						$+-+-$	16	$-+-+-$	61
						$+-++$	9	$-+-++$	35
						$++--$	6	$-++--$	26
						$++-+$	9	$-++-+$	40
						$+++-$	4	$-+++-$	19
						$++++$	1	$-++++$	5

We write $+$ for $+1$ and $-$ for -1. Since $C(\sigma)$ is symmetric on interchanging $+$ and $-$, only half of the values $C(\sigma)$ for $N=5$ are shown. The maximum values A_N in each column are asymptotically equal to $2^{N+3}\pi^{-(N+2)}(N+1)!$ (see [2]).

[4,7,11,14–16,18], and are related, for example, to the dimensions of irreducible representations of the symmetric group via the Littlewood–Richardson rule for the multiplication of Schur functors [9].

Now consider $N+1$ independent and identically distributed random variables $X_1, \ldots, X_{N+1}$, where the X_i are taken from a continuous distribution (i.e., if $i \neq j$, then $P(X_i = X_j) = 0$). Then the quantity

$$P(\sigma) = \frac{C(\sigma)}{(N+1)!} \tag{1.1}$$

is the probability that the random curve $X_1, \ldots, X_{N+1}$ has up–down signature σ. Thus another motivation for considering the numbers $C(\sigma)$ is because of the importance of one-dimensional random energy landscapes in statistical physics [19]. These arise in the study of spin glasses [8,13], protein folding [10] and drainage networks [10]. The numbers $P(\sigma)$ can also be used to define a test for randomness, which has been applied very effectively to the study of genetic microarray data in biology [1,21,22].

It is also known how to compute the probability that two random curves have the same up–down signature [12], and how to compute the expected values of a random permutation with any given up–down signature [16].

In this paper, we answer questions about the nature of the whole up–down sequence (or distribution) for a given length N, i.e., the entire set of up–down numbers $C(\sigma)$ (or $P(\sigma)$), for $\sigma \in \{1, -1\}^N$. This problem is far from simple because of the highly discontinuous nature of the up–down distribution (see Fig. 1). We approach the problem from two different angles. First of all, we show that there exists a universal polynomial Φ, whose coefficients are given by the Taylor expansion of the hyperbolic tangent function, which gives an explicit expression for the up–down function $C(\sigma)$ for signatures of arbitrary length (Theorem 2.4). This gives a concise description of the up–down distribution as the superposition of a small number of much simpler distributions, and gives an expression for each up–down number $C(\sigma)$ as an explicit linear combination of tangent (or Bernoulli) numbers. We also show that the polynomial Φ is in fact an exponential with respect to a certain modified product denoted $\star$ (Proposition 2.6). From this, we can deduce some remarkable congruence properties satisfied by the set of numbers $C(\sigma)$ (Corollary 2.7). This sheds light on the fine structure of the distribution $C(\sigma)$.

The second approach is to show how one can approximate the up–down distribution $P(\sigma)$ (and hence $C(\sigma)$) by considering it as a function of the lengths of its increasing or decreasing runs. We derive a simple upper bound for the quantities $P(\sigma)$ (Theorem 2.9), which gives the asymptotic behaviour of the up–down distribution in the tail $P(\sigma) \ll 1$. This sheds light on the coarse structure of the distribution $P(\sigma)$. In applications where the up–down numbers are used as a test for randomness, this is useful for establishing the non-randomness of a given data set.

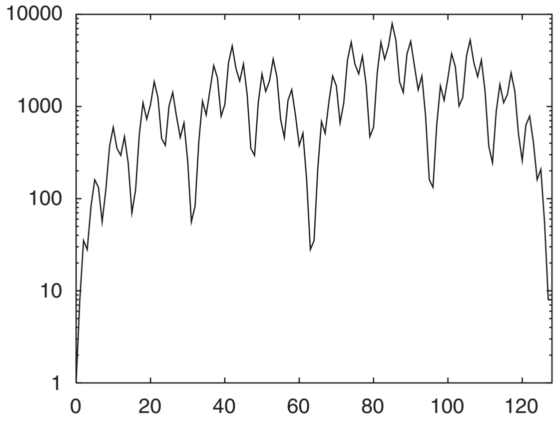

Fig. 1. The number of permutations $C(\sigma)$ as a function of the signature σ for $N=8$. A number on the horizontal axis represents a signature via its representation in binary (e.g., for $N=8$, $C(49)=C(00011001)=C(---++--+)=1016$). Only the first half of the up–down sequence is shown. The second half, corresponding to values between 128 and 256, is obtained by symmetry on interchanging $+$ and $-$.

The paper is organised as follows. In Section 2 we state our main results. In Section 3 we recall some well-known properties of the up–down numbers, and in Section 4 we give all the proofs of our results.

2. Statement of results

Let

$$\Sigma_N = \{(\sigma_1, \ldots, \sigma_N) : \sigma_i \in \{1, -1\}\}$$

denote the set of all up–down signatures of length N. Any function f on Σ_N can be expressed as a polynomial in N variables $s_1, \ldots, s_N$, where s_i takes values in $\{1, -1\}$. Since $s_i^2 = 1$ for all $1 \leqslant i \leqslant N$, it follows that f can be written as a sum of linear monomials. For example, any $\mathbb{Q}$-valued function on the set $\Sigma_2 = \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$ can be uniquely written in the form:

$$f(s_1, s_2) = a_0 + a_1 s_1 + a_2 s_2 + a_{1,2} s_1 s_2,$$

where $a_0, a_1, a_2, a_{1,2} \in \mathbb{Q}$. Let us define $c_N(s_1, \ldots, s_N)$ to be the polynomial function which interpolates the values of the up–down sequence $C(\sigma)$ for all σ of length N. By (1.1), the function interpolating $P(\sigma)$ is given by

$$p_N(s_1, \ldots, s_N) = \frac{1}{(N+1)!} c_N(s_1, \ldots, s_N).$$

The first few polynomials $c_1, \ldots, c_5$ are listed below, and can be used to reproduce all the entries in Table 1.

$$c_1 = 1, \quad c_2 = \tfrac{1}{2}(3 - s_1 s_2), \quad c_3 = 3 - s_1 s_2 - s_2 s_3,$$
$$c_4 = \tfrac{1}{2}(15 - 5(s_1 s_2 + s_2 s_3 + s_3 s_4) + 2 s_1 s_2 s_3 s_4),$$
$$c_5 = \tfrac{1}{2}(45 - 15(s_1 s_2 + s_2 s_3 + s_3 s_4 + s_4 s_5) + 6(s_1 s_2 s_3 s_4 + s_2 s_3 s_4 s_5) + 5 s_1 s_2 s_4 s_5).$$

We will show that the polynomials c_N (and hence p_N) can be obtained by truncating a certain universal polynomial Φ in an infinite number of variables $s_1, \ldots, s_N, \ldots$.

2.1. *The universal polynomial*

In order to consider all up–down sequences simultaneously, let Σ_∞ denote the set of all up–down sequences of arbitrary finite length followed by zeros:

$$\Sigma_\infty = \{(\sigma_1, \sigma_2, \ldots, \sigma_n, \ldots) : \quad \text{there exists } N \geqslant 1 \text{ such that } \sigma_i = 0 \text{ for all } i \geqslant N+1, \\ \text{and } \sigma_i \in \{1, -1\} \text{ for all } 1 \leqslant i \leqslant N\}. \tag{2.1}$$

Let

$$R_N = \mathbb{Q}[s_1, \ldots, s_N]/I_N,$$

where I_N is the ideal generated by the relations $s_1^2 = 1, \ldots, s_N^2 = 1$. Then R_N is naturally identified with the ring of $\mathbb{Q}$-valued functions on Σ_N. There are obvious inclusions $R_N \to R_{N+1}$ for all $N \geqslant 1$. The inductive limit

$$R = \lim_{N\to\infty} R_N \tag{2.2}$$

can naturally be identified with the ring of $\mathbb{Q}$-valued functions on Σ_∞. Any element $f \in R$ can be uniquely written as an infinite series of linear monomials

$$f(s_1, s_2, \ldots) = a_0 + \sum_{k \geqslant 1} \sum_{0<i_1<\cdots<i_k} a_{i_1,\ldots,i_k} s_{i_1} \ldots s_{i_k} \quad \text{where } a_{i_1,\ldots,i_k} \in \mathbb{Q}. \tag{2.3}$$

For any $N \geqslant 1$, we shall write

$$f_N(s_1, \ldots, s_N) = f(s_1, \ldots, s_N, 0, 0, \ldots),$$

for the series in R_N obtained by truncating f. The value of the function f on any signature $\sigma = (\sigma_1, \ldots, \sigma_N)$ of length N is then given by the finite sum $f_N(\sigma_1, \ldots, \sigma_N)$.

Definition 2.1. Let $A \subset \mathbb{N}$ denote any non-empty set of positive whole numbers, and let $n = |A|$. The set A can be uniquely partitioned into $k \geqslant 1$ maximal runs of $i_1, \ldots, i_k$ consecutive integers, where $i_1 + \cdots + i_k = n$. In other words,

$$A = A_1 \cup \cdots \cup A_k,$$

where

$$A_1 = \{a_1, a_1 + 1, \ldots, a_1 + i_1 - 1\}, \quad A_2 = \{a_2, a_2 + 1, \ldots, a_2 + i_2 - 1\}, \; \ldots,$$
$$A_k = \{a_k, a_k + 1, \ldots a_k + i_k - 1\},$$

such that $a_2 \geqslant a_1 + i_1 + 1, \ldots, a_k \geqslant a_{k-1} + i_{k-1} + 1$, so that there is a gap between the end of each consecutive sequence and the beginning of the next one. We call $(i_1, \ldots, i_k)$ the *run-type* of A. The run-type of $\{1, 2, 4, 5, 7\}$, for example, is $(2, 2, 1)$.

Definition 2.2. Let $k \geqslant 1$, and let $i_1, \ldots, i_k \in \mathbb{N}$. We define an infinite series $\gamma(i_1, \ldots, i_k) \in R$ by the formula:

$$\gamma(i_1, \ldots, i_k) = \sum_{\emptyset \neq A \subset \mathbb{N}} \prod_{a \in A} s_a, \tag{2.4}$$

where the sum is over all sets of positive integers A which have run-type $(i_1, \ldots, i_k)$. The series $\gamma(i_1, \ldots, i_k)$ is homogeneous of degree $i_1 + \cdots + i_k$.

Writing this out in full gives

$$\gamma(i_1, \ldots, i_k) = \sum \underbrace{(s_{a_1} s_{a_1+1} \ldots s_{b_1})}_{i_1} \underbrace{(s_{a_2} s_{a_2+1} \ldots s_{b_2})}_{i_2} \ldots \underbrace{(s_{a_k} s_{a_k+1} \ldots s_{b_k})}_{i_k}, \tag{2.5}$$

where $a_2 > b_1 + 1, \ldots, a_k > b_{k-1} + 1$. We have, for example,

$$\gamma(2, 2) = s_1 s_2 s_4 s_5 + s_1 s_2 s_5 s_6 + s_1 s_2 s_6 s_7 + s_1 s_2 s_7 s_8$$
$$+ \cdots + s_2 s_3 s_5 s_6 + s_2 s_3 s_6 s_7 + s_2 s_3 s_7 s_8 + \cdots + s_3 s_4 s_6 s_7 + \cdots + \cdots \; .$$

Now consider the Taylor expansion of the hyperbolic tangent function

$$\frac{\tanh z}{z} = 1 + \sum_{k \geqslant 1} T_k z^k, \tag{2.6}$$

where $T_{2k+1}=0$ for all k, and $T_2=-\frac{1}{3}$, $T_4=\frac{2}{15}$, $T_6=-\frac{17}{315}$, $T_8=\frac{62}{2835}$, and, in general,

$$T_{n-2}=\frac{2^n(2^n-1)B_n}{n!} \quad \text{for } n\geqslant 4, \tag{2.7}$$

where B_n is the nth Bernoulli number.

Definition 2.3. We define the *universal polynomial* $\Phi\in R$ to be the series

$$\Phi=1+\sum_{\emptyset\neq A\subset\mathbb{N}} T_{i_1}\dots T_{i_k}s_{A_1}\dots s_{A_k}, \tag{2.8}$$

where the sum is over all non-empty sets of positive integers A, whose run-type we denote $(i_1,\dots,i_k)$. As above, the corresponding partition is denoted by $A=A_1\cup\cdots\cup A_k$, and for any non-empty set $B\subset\mathbb{N}$, we write $s_B=\prod_{b\in B}s_b$.

Equivalently, we can write the universal polynomial in terms of γ-series:

$$\Phi=1+\sum_{k\geqslant 1}\sum_{i_1\geqslant 1,\dots,i_k\geqslant 1} T_{i_1}\dots T_{i_k}\gamma(i_1,\dots,i_k). \tag{2.9}$$

Theorem 2.4. *The universal polynomial describes the up–down sequences of length N for all* $N\geqslant 1$:

$$p_N(s_1,\dots,s_N)=2^{-N}\Phi_N(s_1,\dots,s_N), \tag{2.10}$$
$$c_N(s_1,\dots,s_N)=(N+1)!\,2^{-N}\Phi_N(s_1,\dots,s_N). \tag{2.11}$$

Therefore, if $\sigma=(\sigma_1,\dots,\sigma_N)$ is any signature of length N, with $\sigma_i\in\{1,-1\}$, then

$$C(\sigma)=(N+1)!\,2^{-N}\Phi_N(\sigma_1,\dots,\sigma_N).$$

Example 2.5. Consider the case $N=8$. The corresponding up–down sequence is a function on the 256 possible up–down signatures of length eight (see Fig. 1). One would expect the polynomial in $s_1,\dots,s_8$ which fits this complicated sequence to have a large number of terms. We have

$$\begin{aligned}\Phi_8=1&-\tfrac{1}{3}\gamma_8(2)+\tfrac{2}{15}\gamma_8(4)+\tfrac{1}{9}\gamma_8(2,2)-\tfrac{17}{315}\gamma_8(6)\\ &-\tfrac{2}{45}(\gamma_8(2,4)+\gamma_8(4,2))-\tfrac{1}{27}\gamma_8(2,2,2)+\tfrac{62}{2835}\gamma_8(8),\end{aligned} \tag{2.12}$$

where

$$\begin{aligned}\gamma_8(2)&=s_1s_2+s_2s_3+s_3s_4+s_4s_5+s_5s_6+s_6s_7+s_7s_8,\\ \gamma_8(4)&=s_1s_2s_3s_4+s_2s_3s_4s_5+s_3s_4s_5s_6+s_4s_5s_6s_7+s_5s_6s_7s_8,\\ \gamma_8(2,2)&=s_1s_2s_4s_5+s_1s_2s_5s_6+s_1s_2s_6s_7+s_1s_2s_7s_8+s_2s_3s_5s_6\\ &\quad+s_2s_3s_6s_7+s_2s_3s_7s_8+s_3s_4s_6s_7+s_3s_4s_7s_8+s_4s_5s_7s_8,\\ \gamma_8(6)&=s_1s_2s_3s_4s_5s_6+s_2s_3s_4s_5s_6s_7+s_3s_4s_5s_6s_7s_8,\\ \gamma_8(2,4)&=s_1s_2s_4s_5s_6s_7+s_1s_2s_5s_6s_7s_8+s_2s_3s_5s_6s_7s_8,\\ \gamma_8(4,2)&=s_1s_2s_3s_4s_6s_7+s_1s_2s_3s_4s_7s_8+s_2s_3s_4s_5s_7s_8,\\ \gamma_8(2,2,2)&=s_1s_2s_4s_5s_7s_8,\\ \gamma_8(8)&=s_1s_2s_3s_4s_5s_6s_7s_8.\end{aligned} \tag{2.13}$$

Recall that the subscript 8 means that the infinite series γ are truncated up to s_8. Quite remarkably, Theorem 2.4 predicts that

$$\begin{aligned}2\,c_8(s_1,\dots,s_8)=2835&-945\,\gamma_8(2)+378\,\gamma_8(4)+315\,\gamma_8(2,2)-153\,\gamma_8(6)\\ &-126\,(\gamma_8(2,4)+\gamma_8(4,2))-105\,\gamma_8(2,2,2)+62\,\gamma_8(8).\end{aligned} \tag{2.14}$$

The up–down distribution for $N=8$ is therefore completely described by the superposition of just 8 simpler distributions (2.13), which encode its symmetry in a subtle and concise way. By truncating further, we retrieve the polynomials $c_1,\dots,c_5$ listed earlier. Note that since $T_{2k+1}=0$, only γ's with even arguments can occur.

One can ask in general how many such γ's occur in c_N. We thank the referee for pointing out that this is given asymptotically by α^{N-1}, where $\alpha \approx 1.3247$ is the real root of $1 + x - x^3$ ([17], sequence A023434). This is still exponential, but grows considerably more slower than 2^N.

2.2. The universal polynomial as an exponential

The universal polynomial can be succinctly rewritten as follows. Let T denote the $\mathbb{Q}$-vector space which is generated by formal sums of the linear monomials

$$s_{i_1}s_{i_2}\dots s_{i_k} \quad \text{where } i_1 < i_2 < \dots < i_k.$$

As remarked earlier, T is isomorphic to the vector space underlying R (2.2). We now define a new product $\star : T \otimes T \to T$, which is defined on monomials by the formula

$$(s_{i_1}s_{i_2}\dots s_{i_k}) \star (s_{j_1}s_{j_2}\dots s_{j_\ell}) = \begin{cases} s_{i_1}s_{i_2}\dots s_{i_k}s_{j_1}s_{j_2}\dots s_{j_\ell} & \text{if } j_1 > i_k + 1, \\ s_{j_1}s_{j_2}\dots s_{j_\ell}s_{i_1}s_{i_2}\dots s_{i_k} & \text{if } i_1 > j_\ell + 1, \\ 0 & \text{otherwise,} \end{cases}$$

and extends in the obvious way to all series in T. The product $\star$ makes T into a commutative algebra with unit 1. We have, for example, $s_1s_2 \star s_2s_3 = 0$ but $s_1s_2 \star s_4s_5 = s_4s_5 \star s_1s_2 = s_1s_2s_4s_5$. We define the exponential map $\exp_\star : T \to T$ with respect to the product $\star$ by the formula:

$$\exp_\star(a) = 1 + a + \frac{1}{2!}(a \star a) + \frac{1}{3!}(a \star a \star a) + \cdots \quad \text{for all } a \in T.$$

Proposition 2.6. *The universal polynomial Φ is an exponential*:

$$\begin{aligned} \Phi &= \exp_\star\left(\sum_{i \geqslant 1} T_i\,\gamma(i)\right) \\ &= \exp_\star(T_2\,\gamma(2)) \star \exp_\star(T_4\,\gamma(4)) \star \exp_\star(T_6\,\gamma(6)) \star \cdots . \end{aligned} \tag{2.15}$$

Proof. Let $\alpha_1, \dots, \alpha_i, \dots \in T$ such that $\alpha_i \star \alpha_i = 0$ for all $i \geqslant 1$. It is a simple exercise to show that

$$\exp_\star\left(\sum_{i \geqslant 1} \alpha_i\right) = 1 + \sum_i \alpha_i + \sum_{i<j} \alpha_i \star \alpha_j + \sum_{i<j<k} \alpha_i \star \alpha_j \star \alpha_k + \cdots$$

If we apply this argument to the sum of the infinite series of monomials $T_2\,s_1s_2$, $T_2\,s_2s_3, \dots, T_4\,s_1s_2s_3s_4$, $T_4\,s_2s_3s_4s_5$, and so on (recall that $T_{2k+1} = 0$), we deduce that

$$\exp_\star\left(\sum_{i \geqslant 1} T_i\,\gamma(i)\right) = 1 + \sum_{k \geqslant 1} \sum_{i_1 \geqslant 1, \dots, i_k \geqslant 1} T_{i_1} \dots T_{i_k}\gamma(i_1, \dots, i_k),$$

which proves identity (2.15), as required. $\square$

2.3. Congruence properties for all up–down sequences of fixed length

The universal polynomial can be used to deduce a number of surprising congruence properties which are satisfied by the entire up–down sequence of length N, for a fixed N. We give two of the most elegant such congruences.

Corollary 2.7. *Let p be any odd prime. For all signatures $\sigma = (\sigma_1, \dots, \sigma_N)$ of length $N = p - 1$,*

$$C(\sigma) \equiv \sigma_1 \dots \sigma_{p-1} \pmod{p}. \tag{2.16}$$

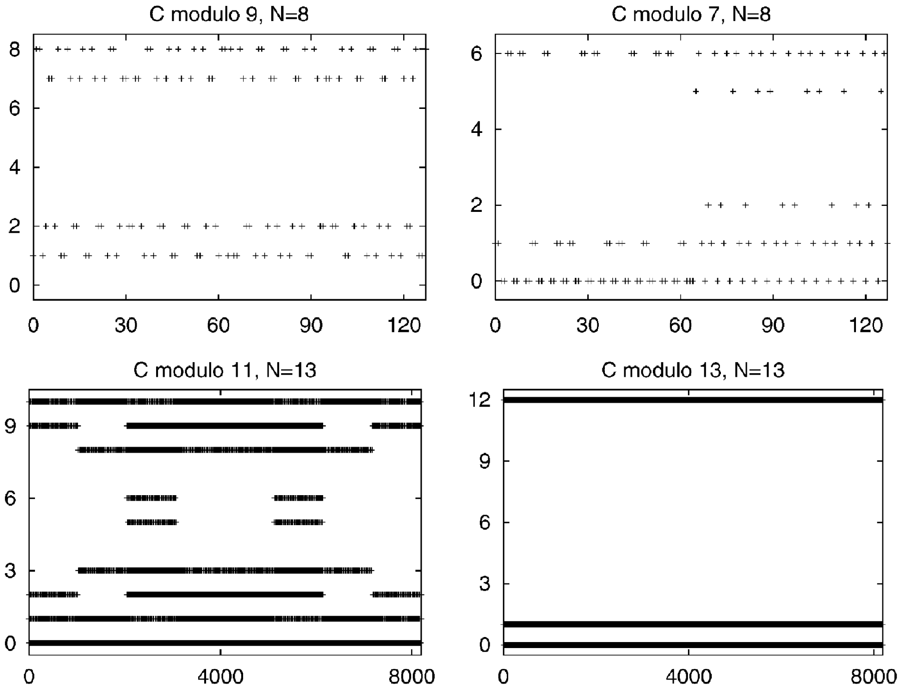

Fig. 2. Congruence properties of the up–down numbers can be deduced from the universal polynomial Φ. Top: The (first half of the) up–down numbers $C(\sigma)$ for signatures σ of length 8, plotted in binary order, modulo 9 (left), and modulo 7 (right). Bottom: A plot of $C(\sigma)$ modulo 11 (left) and modulo 13 (right) for all signatures of length 13. Eq. (2.17) predicts that $C(\sigma) \equiv 0, \pm 1 \pmod{13}$. The density of the points is such that they appear as a solid line.

In particular, $C(\sigma)$ *only takes the values* $\pm 1 \pmod p$. *Likewise, for all signatures* σ *of length* $N = p$,

$$2C(\sigma) \equiv (\sigma_1 + \sigma_p)\sigma_1 \ldots \sigma_p \pmod p, \tag{2.17}$$

and therefore $C(\sigma)$ *only takes the values* $0, \pm 1 \pmod p$.

The proof of these identities, given in Section 4.2, will follow from formula (2.15) using well-known congruence properties of Bernoulli numbers due to Kummer and Clausen–Von Staudt [20].

Example 2.8. Many more congruence properties can be derived from Φ as follows. For example, in the case $N = 8$, we can reduce (2.14) modulo 9 and 7 to give the simple relations:

$$2\,c_8(s_1, \ldots, s_8) \equiv -105\,\gamma_8(2,2,2) + 62\gamma_8(8) \pmod 9,$$
$$2\,c_8(s_1, \ldots, s_8) \equiv -153\,\gamma_8(6) + 62\gamma_8(8) \pmod 7.$$

This implies that

$$C(\sigma) \equiv (6\,\sigma_3\sigma_6 + 4)\sigma_1 \ldots \sigma_8 \pmod 9,$$
$$C(\sigma) \equiv (4(\sigma_1\sigma_2 + \sigma_1\sigma_8 + \sigma_7\sigma_8) + 3)\sigma_1 \ldots \sigma_8 \pmod 7,$$

for all signatures $\sigma = (\sigma_1, \ldots, \sigma_8)$ of length 8. It follows, for example, that $C(\sigma)$ can only be congruent to $\pm 1, \pm 2$ modulo 9 for all σ of length 8 (see Fig. 2).

2.4. *An upper bound for* $P(\sigma)$ *and* $C(\sigma)$

In some applications, it is necessary to approximate the distribution of C, or bound C from above. In order to do this, we need to rewrite a signature σ in terms of the lengths of its runs. Let $(i_1, i_2, \ldots, i_n)$ denote the signature with an island of i_1 pluses, followed by an island of i_2 minuses, and so on, where $i_1 + i_2 + \cdots + i_n = N$ is a composition of N. For example, (2, 3, 1) corresponds to $++---+$. One approach to finding approximations to $C(\sigma)$ is the near separability of the function P at maxima or minima. In other words, probabilistic considerations suggest the approximation:

$$P(i_1, \ldots, i_n) \simeq \frac{P(i_1, \ldots, i_\ell) P(i_{k+1}, \ldots, i_n)}{P(i_{k+1}, \ldots, i_\ell)}, \tag{2.18}$$

for $k < \ell < n$. Applying (2.18) repeatedly in the case where $\ell = k + 1$, we obtain

$$P(i_1, .., i_n) \simeq \frac{P(i_1, i_2) P(i_2, i_3) \ldots P(i_{n-1}, i_n)}{P(i_2) \ldots P(i_{n-1})}. \tag{2.19}$$

The right-hand side can be written explicitly in closed form by (3.3). It turns out that this approximation is an upper bound, which gives the following inequality.

Theorem 2.9. *For all* $i_1, \ldots, i_n \geqslant 1$,

$$P(i_1, \ldots, i_n) \leqslant \frac{(i_2 + 1) \ldots (i_{n-1} + 1)}{(i_1 + i_2 + 1) \ldots (i_{n-1} + i_n + 1)} \frac{1}{i_1! \ldots i_n!}. \tag{2.20}$$

By (1.1), we can multiply through by $(i_1 + \cdots + i_n + 1)!$ to obtain a similar upper bound for $C(i_1, \ldots, i_n)$.

Remark 2.10. Eq. (2.20) gives the limiting behaviour of P in the tail $P \ll 1$. If the number of islands n is very small, or if there is a very large island i_k, then certainly the right-hand side of (2.20), and therefore $P(i_1, \ldots, i_n)$ itself, will be small. This is relevant when using P as a test for randomness. However, the converse is far from true, and the question of when $P(i_1, \ldots, i_n)$ is small is considerably more subtle. Note that the denominators $(i_k + i_{k+1} + 1)$ in Eq. (2.20) take into account not just the island sizes i_k but also first-order dependencies between adjacent islands. One can speculate that (2.20) is something like the dominant term of an asymptotic formula expressing $P(i_1, \ldots, i_n)$ in terms of the island sizes i_k.

Remark 2.11. Eq. (2.20) is most accurate when $i_2, \ldots, i_n \gg 1$. One can obtain a complementary upper bound for any signatures ρ and τ:

$$P(\rho, 1, \tau) \leqslant P(\rho) P(\tau). \tag{2.21}$$

This inequality follows immediately from Eq. (3.4). In [1], the up–down probabilities $P(\sigma)$ were used as a test for randomness and applied to genetic microarray data. By combining inequalities (2.21) and (2.20), one could easily show by hand that many such gene expression curves were non-random.

3. Recurrence relations for the up–down numbers

We recall two well-known recurrence relations satisfied by the up–down numbers. The first is linear, the second is quadratic. We will write, for instance, $(\alpha, j) = (i_1, \ldots, i_n, j)$, where a Roman letter denotes an island of $+$'s or $-$'s, and a Greek letter denotes any signature $(i_1, \ldots, i_n)$ which consists of several islands (see Section 2.4).

3.1. A linear recursion for $C(\sigma)$

The numbers $C(\sigma)$ satisfy the following linear recursion relation, which is the same recursion as that for multinomial coefficients:

$$C(i_1, \ldots, i_n) = C(i_1 - 1, \ldots, i_n) + C(i_1, i_2 - 1, \ldots, i_n) + \cdots + C(i_1, \ldots, i_n - 1) \tag{3.1}$$

subject to the boundary conditions $C(0, \alpha) = C(\alpha)$, $C(\alpha, 0) = C(\alpha)$, and

$$C(\alpha, i, 0, j, \beta) = C(\alpha, i + j, \beta). \tag{3.2}$$

Eq. (3.1) can be derived in the following way (see also [5]). In a permutation of $1, 2, \ldots, N + 1$ with signature $(i_1, \ldots, i_n)$, the largest element, $N + 1$, must occur at the end of a sequence of pluses. If we remove it, we obtain a permutation of length N with signature $(i_1, \ldots, i_{2k-1} - 1, i_{2k}, \ldots,)$ or $(i_1, \ldots, i_{2k-1}, i_{2k} - 1, \ldots)$. It follows that there is a one to one correspondence between the set of all permutations with signature $(i_1, \ldots, i_n)$ and the union of all permutations with signatures $(i_1 - 1, \ldots, i_n)$, ..., $(i_1, \ldots, i_n - 1)$, which proves (3.1).

Although there is no simple formula for $C(i_1, \ldots, i_n)$ when $n \geqslant 3$, one can show (using the previous recurrence relation, for example) that

$$C(i) = 1 \quad \text{and} \quad C(i, j) = \binom{i+j}{i}.$$

Using the fact that $P(\sigma) = C(\sigma)/(N+1)!$, for all signatures σ of length N, we deduce that

$$P(i) = \frac{1}{(i+1)!} \quad \text{and} \quad P(i, j) = \frac{1}{(i+j+1)} \frac{1}{i!} \frac{1}{j!}. \tag{3.3}$$

3.2. A quadratic relation for $P(\sigma)$.

The second recurrence relation we will require is most simply written in terms of P. Let σ and μ be arbitrary signatures. Then there is the quadratic relation

$$P(\sigma)P(\mu) = P(\sigma + \mu) + P(\sigma - \mu), \tag{3.4}$$

where $\sigma + \mu$ denotes the concatenation of the signatures σ, $+$ and μ, and $\sigma - \mu$ is the concatenation of the signatures σ, $-$ and μ. In order to obtain (3.4), we interpret $P(\sigma)$ as being the probability that a random curve has signature σ. The equation holds because a random curve $X_1, \ldots, X_{N+1}$ decouples into two independent sections $X_1, \ldots, X_m$ and $X_{m+1}, \ldots, X_{N+1}$ if one makes no assumption about the relative values of the points X_m and X_{m+1} where the curves join.

Remark 3.1. By rewriting (3.4) in terms of $\Phi_N = 2^N p_N$, and considering the special case when $\mu = \emptyset$, we obtain the identity

$$\Phi_N(\sigma) = \tfrac{1}{2}\,(\Phi_{N+1}(\sigma+) + \Phi_{N+1}(\sigma-)),$$

for any signature σ of length N. This identity implies a self-similarity for the scaled up–down curves Φ_N: the values of the up–down sequence of level N are given by the average of adjacent values of the up–down sequence of level $N+1$.

Lemma 3.2. *The quantity $P(i_1, \ldots, i_n)$ is given by the exact formula*

$$\sum_{r_n=0}^{i_n} \sum_{r_{n-1}=0}^{i_{n-1}+r_n} \cdots \sum_{r_2=0}^{i_2+r_3} \frac{(-1)^{r_2+r_3+\cdots+r_n}}{(i_n - r_n)!\,(i_{n-1} + r_n - r_{n-1})! \ldots (i_2 + r_3 - r_2)!(i_1 + r_2 + 1)!}.$$

Proof. Let α denote any signature, and let $j, k \geqslant 1$. Applying Eq. (3.4) with $\sigma = (\alpha, j)$, $\mu = (k-1)$ implies that $P(\alpha, j, k) + P(\alpha, j+1, k-1) = P(\alpha, j)P(k-1)$. Applying this formula inductively and writing $P(k-1) = 1/k!$, we obtain

$$P(\alpha, j, k) = \sum_{r=0}^{k} (-1)^r \frac{P(\alpha, j+r)}{(k-r)!}. \tag{3.5}$$

This expresses the P-value of an arbitrary signature in terms of P-values of signatures which have a strictly smaller number of islands. Applying this formula inductively to the signature $(i_1, \ldots, i_n)$, one obtains the formula in the lemma. $\square$

Remark 3.3. Using (1.1), the lemma gives an exact formula for $C(\sigma)$ in terms of multinomial coefficients, but which has the disadvantage of being inefficient to compute. A similar formula is given in [16, Eq. (6)]. There are other known methods for computing $C(\sigma)$. For example, one can express $C(\sigma)$ as the determinant of a matrix consisting of binomial coefficients (see [11,14] and the refinement in [9]). There is also a simple iterative algorithm for computing $C(\sigma)$ as a sum of numbers which are all positive [7,18], but, unlike the formula given in the lemma, this does not give a formula in closed form. The universal polynomial Φ gives a completely different way to compute the up–down numbers $C(\sigma)$.

4. Proofs

4.1. Proof of Theorem 2.4

Let E_{N+1} denote the number of permutations on $N+1$ letters which have an even number of rises. By symmetry, this is also the number of permutations with an even number of falls. This in turn is equal to the number of permutations whose up–down signature $(\sigma_1, \ldots, \sigma_N)$ satisfies $\sigma_1 \ldots \sigma_N = 1$.

Lemma 4.1. *We have*

$$E_{N+1} = \frac{(N+1)!}{2}(1+T_N).$$

Proof. Let $A_{n,r}$ denote the number of permutations on n letters with r rises, where $1 \leqslant r \leqslant n$. It is well-known [6] that the quantities $A_{n,r}$ are the Eulerian numbers, whose generating series is given by

$$F(x, y) = \frac{e^{xy} - e^x}{ye^x - e^{xy}} = \sum_{n=1}^{\infty} \frac{x^n}{n!} \sum_{r=0}^{n} A_{n,r} y^r.$$

The generating series for permutations with an even number of rises is therefore given by

$$\sum_{n=1}^{\infty} E_n \frac{x^n}{n!} = \frac{1}{2}(F(x,1) + F(x,-1)) = \frac{1}{2}\left(\frac{x}{1-x} + \tanh(x)\right),$$

where $F(x, 1)$ is to be interpreted as $\lim_{y \to 1} F(x, y)$. Comparing the coefficients of $x^{N+1}/(N+1)!$ yields

$$E_{N+1} = \frac{(N+1)!}{2}(1+T_N). \qquad \square$$

Lemma 4.2. *Let $N \geqslant 2$. For all $1 \leqslant n \leqslant N$,*

$$\Phi_N(s_1, \ldots, s_{n-1}, 0, s_{n+1}, \ldots, s_N) = \Phi_{n-1}(s_1, \ldots, s_{n-1})\Phi_{N-n}(s_{n+1}, \ldots, s_N),$$

where $\Phi_0 = 1$.

Proof. If we work in the algebra T, the exponential formula (Proposition 2.6) gives

$$\Phi(s_1, \ldots, s_{n-1}, 0, s_{n+1}, \ldots) = \exp_\star\left(\sum_{i \geqslant 1} T_i\, \gamma(i)(s_1, \ldots, s_{n-1}, 0, s_{n+1}, \ldots)\right).$$

By definition of the sums $\gamma(i)(s_1, s_2, \ldots) = \sum_{k \geqslant 1} s_k s_{k+1} \ldots s_{k+i-1}$, this is

$$\exp_\star\left(\sum_{i \geqslant 1} T_i\, \gamma(i)(s_1, \ldots, s_{n-1}, 0, 0, \ldots) + \sum_{i \geqslant 1} T_i\, \gamma(i)(s_{n+1}, s_{n+2}, \ldots)\right).$$

By the multiplicativity of the exponential, this is a product:

$$\exp_\star\left(\sum_{i \geqslant 1} T_i\, \gamma(i)(s_1, \ldots, s_{n-1}, 0, 0, \ldots)\right) \star \exp_\star\left(\sum_{i \geqslant 1} T_i\, \gamma(i)(s_{n+1}, s_{n+2}, \ldots)\right)$$
$$= \Phi(s_1, \ldots, s_{n-1}, 0, 0, \ldots) \star \Phi(s_{n+1}, s_{n+2}, \ldots).$$

We have proved that

$$\Phi(s_1,\ldots,s_{n-1},0,s_{n+1},\ldots)=\Phi(s_1,\ldots,s_{n-1},0,0,\ldots)\star\Phi(s_{n+1},s_{n+2},\ldots),$$

in the algebra T. But the definition of the product $\star$ coincides with the ordinary product for monomials which are sufficiently far apart:

$$s_{i_1}\ldots s_{i_r}\star s_{j_1}\ldots s_{j_k}=s_{i_1}\ldots s_{i_r}s_{j_1}\ldots s_{j_k}$$

if $i_1<\cdots<i_r\leqslant n-1$ and $n+1\leqslant j_1<\cdots<j_k$. It follows that the identity

$$\Phi(s_1,\ldots,s_{n-1},0,s_{n+1},\ldots)=\Phi(s_1,\ldots,s_{n-1},0,0,\ldots)\Phi(s_{n+1},s_{n+2},\ldots)$$

holds in the algebra R. The result follows on truncating. The lemma can also be proved by direct computation using the definition of the universal polynomial (Eq. (2.9)). $\square$

Proof of Theorem 2.4. For all $N\geqslant 1$, there exists a polynomial $p_N(s_1,\ldots,s_N)\in R_N$ such that $P(\sigma)=p_N(\sigma_1,\ldots,\sigma_N)$ for all signatures $\sigma=(\sigma_1,\ldots,\sigma_N)$. We can write p_N uniquely as a linear monomial in $s_1,\ldots,s_N$ with coefficients in $\mathbb{Q}$. First of all, the quadratic relation (3.4) implies that

$$\begin{aligned}p_{n-1}(s_1,\ldots,s_{n-1})\,p_{N-n}(s_{n+1},\ldots,s_N)&=p_N(s_1,\ldots,s_{n-1},1,s_{n+1},\ldots,s_N)\\&\quad+p_N(s_1,\ldots,s_{n-1},-1,s_{n+1},\ldots,s_N),\end{aligned}$$

for all $1\leqslant n\leqslant N$. This can be rewritten as

$$2\,p_N(s_1,\ldots,s_{n-1},0,s_{n+1},\ldots,s_N)=p_{n-1}(s_1,\ldots,s_{n-1})\,p_{N-n}(s_{n+1},\ldots,s_N). \tag{4.1}$$

Suppose by induction that $p_n=2^{-n}\,\Phi_n$ for all $1\leqslant n<N$. Then Lemma 4.2 implies that the polynomial $2^{-N}\Phi_N$ satisfies identity (4.1) also. It follows from the induction hypothesis that p_N and $2^{-N}\,\Phi_N$ coincide whenever at least one of the s_i's is 0. Since only linear monomials are involved, this implies that $p_N-2^{-N}\Phi_N$ is a multiple of $s_1\ldots s_N$. In order to compute the coefficient of the term $s_1\ldots s_N$, let

$$S=\{(s_1,\ldots,s_N):\ s_i\in\{\pm 1\}\text{ such that }s_1\ldots s_N=1\}.$$

For any $1\leqslant i_1<\cdots<i_k\leqslant N$, where k is strictly smaller than N, we have

$$\sum_{(s_1,\ldots,s_N)\in S} s_{i_1}\ldots s_{i_k}=0.$$

It follows that taking the sum over all signatures in S picks out the constant term 1 and the leading term $s_1\ldots s_N$ only. It therefore suffices to show that

$$\sum_{(s_1,\ldots,s_N)\in S} p_N(s_1,\ldots,s_N)=\sum_{(s_1,\ldots,s_N)\in S} 2^{-N}\Phi_N(s_1,\ldots,s_N). \tag{4.2}$$

The left-hand side is the probability that the signature $\sigma=(\sigma_1,\ldots,\sigma_N)$ of a random curve satisfies $\sigma_1\ldots\sigma_N=1$. This is just $E_{N+1}/(N+1)!$, where E_{N+1} is the number of permutations on $N+1$ letters which have an even number of rises. The right-hand side is

$$|S|\,2^{-N}(1+\text{coeff. of } s_1\ldots s_N \text{ in } \Phi_N)=2^{-1}(1+T_N).$$

By Lemma 4.1, both sides of (4.2) agree, which completes the induction step. We conclude that $p_N=2^{-N}\Phi_N$, as required. $\square$

4.2. *Proof of Corollary 2.7*

First recall the theorem due to Clausen–Von Staudt [20, Theorem 5.10], which states that for all $k\geqslant 2$,

$$B_{2k}-\sum_{(p-1)|2k}\frac{1}{p}\in\mathbb{Z}, \tag{4.3}$$

where the sum ranges over all primes p such that $p-1$ divides $2k$. Now, the coefficients which occur in the polynomial $c_N(s_1, \dots, s_N)$ are

$$\frac{(N+1)!}{2^N} T_{2k-2} = \left(\frac{(N+1)!}{2^N}\right)\left(\frac{2^{2k}(2^{2k}-1)B_{2k}}{(2k)!}\right) \quad \text{for } 4 \leqslant 2k \leqslant N+2. \tag{4.4}$$

Now let p be an odd prime, and suppose that $N = p-1$. If $4 \leqslant 2k \leqslant N-2$, the prime p does not occur in the denominator of B_{2k} by (4.3), and therefore

$$\frac{(N+1)!}{2^N} T_{2k-2} \equiv 0 \pmod{p}.$$

It remains to compute the coefficients (4.4) for $2k = N$ and $2k = N+2$. In the first case, we have

$$\frac{(N+1)!}{2^N} T_{N-2} = \left(\frac{(N+1)!}{2^N}\right)\left(\frac{2^N(2^N-1)B_N}{N!}\right) = p\, B_{p-1}(2^{p-1}-1).$$

But $2^{p-1} - 1 \equiv 0 \pmod{p}$, and $p\, B_{p-1} \equiv 1 \pmod{p}$ by (4.3). It follows that this coefficient vanishes modulo p also. Therefore all terms in the polynomial c_N vanish modulo p except the leading term, and we are left with

$$c_N(s_1, \dots, s_N) \equiv (N+1)!\, 2^{-N}\, T_N\, \gamma_N(N) \pmod{p},$$

where $\gamma_N(N)$ consists of the single term $s_1 \dots s_N$. By Eq. (4.4), we have

$$\frac{(N+1)!}{2^N} T_N = \left(\frac{p!}{2^{p-1}}\right)\left(\frac{2^{p+1}(2^{p+1}-1)B_{p+1}}{(p+1)!}\right) \equiv 12\, B_{p+1} \pmod{p}.$$

The congruences for Bernoulli numbers discovered by Kummer [20, Corollary 5.14] imply, in particular, that $2\, B_{p+1} \equiv (p+1)B_2 \pmod{p}$, and so $12\, B_{p+1} \equiv 1 \pmod{p}$. We conclude that

$$c_N(s_1, \dots, s_N) \equiv \gamma_N(N) = s_1 \dots s_N \pmod{p},$$

as required.

The result when $N = p$ holds for similar reasons, since all the terms of c_N vanish modulo p except the leading term. The coefficient of this term is

$$\frac{(N+1)!}{2^N} T_{N-1} = \left(\frac{(p+1)!}{2^p}\right)\left(\frac{2^{p+1}(2^{p+1}-1)B_{p+1}}{(p+1)!}\right) \equiv 6\, B_{p+1} \equiv 2^{-1} \pmod{p}.$$

This proves that

$$2\, c_N(s_1, \dots, s_N) \equiv \gamma_N(N-1) = s_1 \dots s_{N-1} + s_2 \dots s_N \pmod{p},$$

as required, and completes the proof of Corollary 2.7.

4.3. Proof of Theorem 2.9

We first prove some general inequalities relating up–down numbers for different signatures of equal length. A similar-looking inequality was proved by Niven [14] to prove that the value of $C(\sigma)$ is greatest on the alternating signature $\sigma = + - + - \dots$.

Lemma 4.3. *Let α denote any signature, and let $a, b, c \in \mathbb{N}$ such that $a \geqslant c$. Then*

$$C(\alpha, a-c+1, b, c) \geqslant C(\alpha, a+1, b).$$

Proof. This inequality is easily proved by induction with respect to the total length $\ell = |\alpha| + a + b + 1$, where $|\alpha|$ is the length of the signature α. The details are left to the reader. The induction step is given by rewriting the left-hand side using relation (3.1):

$$C(\alpha, a-c, b, c) + C(\alpha, a-c+1, b-1, c) + C(\alpha, a-c+1, b, c-1)$$

plus terms of the form $C(\alpha', a-c+1, b, c)$, where α' is a signature of shorter length than α. Likewise, the right-hand side can be written as

$$C(\alpha, a, b) + C(\alpha, a+1, b-1),$$

plus terms of the form $C(\alpha', a+1, b)$. If we assume that the inequality holds for α with all smaller values of a, b, c, then $C(\alpha, a-c+1, b-1, c) \geqslant C(\alpha, a+1, b-1)$ (this is the case $(\alpha, a, b-1, c)$), and $C(\alpha, a-c+1, b, c-1) \geqslant C(\alpha, a, b)$ (this is the case $(\alpha, a-1, b, c-1)$). If we assume that the inequality holds for all signatures α' of shorter length than α, and a, b, c, then $C(\alpha', a-c+1, b, c) \geqslant C(\alpha', a+1, b)$. This is enough to complete the induction step, and hence the proof. The initial cases $b=0$, $c=0$ are both trivial by (3.2). The case $a=c$ is proved using an inductive argument similar to the one given above. □

Proposition 4.4. *Let α denote any signature, and let $a, b, c \in \mathbb{N}$ such that $a \geqslant c \geqslant 1$. Then for all $0 \leqslant n \leqslant c-1$,*

$$C(\alpha, a-n, b, c) \geqslant C(\alpha, a+1, b, c-n-1). \tag{4.5}$$

Proof. The proof is by induction on the total length

$$\ell(\alpha, a, b, c, n) = |\alpha| + a + b + c - n,$$

where $|\alpha|$ denotes the length of the signature α. Let $a', b', c', n' \in \mathbb{N}$ such that $a' \geqslant c' \geqslant 1$ and $0 \leqslant n' \leqslant c'-1$. Suppose that (4.5) is true for all:

$$a \leqslant a', \quad b \leqslant b', \quad c \leqslant c', \quad n \geqslant n',$$

and all α satisfying $|\alpha| \leqslant |\alpha'|$ such that

$$a \geqslant c \geqslant 1, \quad c-1 \geqslant n \quad \text{and} \quad \ell(\alpha, a, b, c, n) < \ell(\alpha', a', b', c', n').$$

Then we will prove (4.5) for a', b', c' and n'. First of all, let us assume that $b' > 0$ and $c' - n' - 1 > 0$. This implies that $a' - n' \geqslant 1$. By (3.1),

$$C(\alpha, a'-n', b', c') = C(\alpha, a'-n'-1, b', c') + C(\alpha, a'-n', b'-1, c') + C(\alpha, a'-n', b', c'-1)$$

plus terms of the form $C(\alpha', a'-n', b', c')$, where α' is strictly shorter than α. Each term in the right-hand side can be bounded below by the induction hypothesis. The middle term is bounded below as follows:

$$C(\alpha, a'-n', b'-1, c') \geqslant C(\alpha, a'+1, b'-1, c'-n'-1)\ . \tag{4.6}$$

Similarly, on setting $a = a'-1$, $c = c'-1$, $b = b'$, $n = n'-1$, we obtain

$$C(\alpha, (a'-1)-(n'-1), b', c'-1) \geqslant C(\alpha, (a'-1)+1, b', (c'-1)-(n'-1)-1)$$

i.e.,

$$C(\alpha, a'-n', b', c'-1) \geqslant C(\alpha, a', b', c'-n'-1). \tag{4.7}$$

Finally, we have

$$C(\alpha, a'-(n'+1), b', c') \geqslant C(\alpha, a'+1, b', c'-(n'+1)-1),$$

which is just

$$C(\alpha, a'-n'-1, b', c') \geqslant C(\alpha, a'+1, b', c'-n'-2). \tag{4.8}$$

Adding the three inequalities (4.6), (4.7) and (4.8) together, we obtain

$$\begin{aligned} &C(\alpha, a'-n'-1, b', c') + C(\alpha, a'-n', b'-1, c') + C(\alpha, a'-n', b', c'-1) \\ &\quad \geqslant C(\alpha, a', b', c'-n'-1) + C(\alpha, a'+1, b'-1, c'-n'-1) + C(\alpha, a'+1, b', c'-n'-2). \end{aligned}$$

After adding inequalities of the form $C(\alpha', a'-n', b', c') \geqslant C(\alpha', a'+1, b', c'-n'-1)$, and rewriting the left- and right-hand sides using (3.1), we obtain

$$C(\alpha, a'-n', b', c') \geqslant C(\alpha, a'+1, b', c'-n'-1)$$

which proves (4.5) for a', b', c' and n'.

We need to check the initial cases when $b=0$, $c=n+1$, or $|\alpha|=0$. If $b=0$, then (4.5) is trivial, since, by (3.2), $C(\alpha, a-n, 0, c) = C(\alpha, a+c-n) = C(\alpha, a+1, 0, c-n-1)$. If $n=c-1$, then (4.5) reduces to the inequality of Lemma 4.3. The case when $|\alpha|=0$ clearly holds from the induction argument given above. Likewise, the case $a=c$ is also covered by the argument above. □

Corollary 4.5. *For any signature α, and $a, b, c \in \mathbb{N}$ such that $a \geqslant c$, we have*

$$C(\alpha, a, b, c) \geqslant C(\alpha, a+1, b, c-1).$$

Equivalently, $P(\alpha, a, b, c) \geqslant P(\alpha, a+1, b, c-1)$.

Remark 4.6. The corollary implies that $C(\alpha, i, j, k)$ is maximised (for values of $i \geqslant k$ such that $i+k$ is fixed) when i and k are most nearly equal.

Proof of Theorem 2.9. Let γ denote any up–down signature. We write $\gamma = (\beta, r+1)$, where $r \geqslant 0$. It is clear that

$$P(\beta, r)P(j+1, k, j) = P(\beta, r)P(j, k, j+1).$$

Using relation (3.4), this implies that

$$P(\beta, r+1, j+1, k, j) + P(\beta, r, j+2, k, j) = P(\beta, r+1, j, k, j+1) + P(\beta, r, j+1, k, j+1).$$

Corollary 4.5 implies that $P(\beta, r, j+2, k, j) \leqslant P(\beta, r, j+1, k, j+1)$, on setting $\alpha = (\beta, r)$, $a = j+1$, $b=k$, and $c = j+1$. Substituting into the previous equality implies that

$$P(\beta, r+1, j+1, k, j) \geqslant P(\beta, r+1, j, k, j+1).$$

Recalling that $\gamma = (\beta, r+1)$, this is just

$$P(\gamma, j+1, k, j) \geqslant P(\gamma, j, k, j+1), \tag{4.9}$$

which, by adding $P(\gamma, j, k+1, j)$ to both sides, implies that

$$P(\gamma, j, k+1, j) + P(\gamma, j, k, j+1) \leqslant P(\gamma, j+1, k, j) + P(\gamma, j, k+1, j).$$

By (3.4), this is equivalent to the inequality:

$$P(\gamma, j, k)P(j) \leqslant P(\gamma, j)P(k, j).$$

It follows from (3.3) that

$$P(\gamma, j, k) \leqslant \frac{P(\gamma, j)P(j, k)}{P(j)} = P(\gamma, j)\frac{j+1}{j+k+1}\frac{1}{k!}.$$

Applying this inequality inductively to the up–down sequence $(i_1, \ldots, i_n)$, we obtain

$$\begin{aligned} P(i_1, \ldots, i_n) &\leqslant P(i_1, \ldots, i_{n-1})\frac{i_{n-1}+1}{i_{n-1}+i_n+1}\frac{1}{i_n!} \\ &\leqslant \cdots \leqslant \frac{(i_2+1)\ldots(i_{n-1}+1)}{(i_1+i_2+1)\ldots(i_{n-1}+i_n+1)}\frac{1}{i_1!\ldots i_n!}, \end{aligned}$$

which is precisely inequality (2.20). □

Acknowledgements

The second author wishes to thank B. Derrida for some interesting discussions and suggestions. K. Willbrand was supported by a Fondation des Treilles.

References

[1] S.E. Ahnert, K. Willbrand, F.C.S. Brown, T.M.A. Fink, Unbiased pattern detection in microarray data series, Bioinformatics 22 (12) (2006) 1471–1476.
[2] D. André, Sur les Permutations Alternées, J. Math. Pures Appl. 7 (1881) 167–184.
[3] V. Arnold, Bernoulli–Euler updown numbers associated with function singularities, their combinatorics and arithmetic, Duke Math. J. 63 (2) (1990) 537–555.
[4] A. Björner, M. Wachs, Generalized quotients in Coxeter groups, Trans. Amer. Math. Soc. 308 (1988) 1–37.
[5] L. Carlitz, Permutations with prescribed pattern, II, Math. Nachr. 83 (1978) 101–126.
[6] L. Carlitz, R. Scoville, Generalized Eulerian numbers: combinatorial applications, J. Reine Angew. Math. 265 (1974) 110–137.
[7] N.G. de Bruijn, Permutations with given ups and downs, Nieuw Arch. 3 (XVIII) (1970) 61–65.
[8] B. Derrida, E. Gardner, Metastable states of a spin glass chain at zero temperature, J. Phys. 47 (1986) 959–965.
[9] H.O. Foulkes, Enumeration of permutations with prescribed up–down and inversion sequences, Discrete Math. 15 (1976) 235–252.
[10] H. Frauenfelder, et al., (Eds.), Landscape Paradigms in Physics and Biology, North Holland, Amsterdam, 1997.
[11] P.A. MacMahon, Combinatorial Analysis, vol. I, Cambridge University Press, Cambridge, 1915.
[12] C.L. Mallows, L.A. Shepp, Enumerating pairs of permutations with the same up–down form, Discrete Math. 54 (1985) 301–311.
[13] M. Mézard, G. Parisi, M. Virasoro, Spin Glass Theory and Beyond, North Holland, Amsterdam, 1987.
[14] I. Niven, A combinatorial problem of finite sequences, Nieuw Arch. Wisk. 3 (XVI) (1968) 116–123.
[15] B. Shapiro, M. Shapiro, A. Vainshtein, Periodic de Bruijn triangles: exact and asymptotic results, Discrete Math. 298 (1–3) (2005) 321–333.
[16] G. Szpiro, The number of permutations with a given signature and the expectations of their elements, Discrete Math. 226 (2001) 423–430.
[17] The On-Line Encyclopedia of Integer Sequences. ⟨http://www.research.att.com:/~njas/sequences/⟩.
[18] G. Viennot, Permutations ayant une forme donnée, Discrete Math. 26 (1979) 279–284.
[19] D.I. Warren, E. Seneta, Peaks and Eulerian numbers in a random sequence, J. Appl. Probab. 33 (1996) 101–114.
[20] L.C. Washington, Introduction to cyclotomic fields, second ed., Graduate Texts in Mathematics, vol. 83, Springer, Berlin, 1997.
[21] K. Willbrand, 'Approches inspirées de la physique statistique et de la théorie de l'information pour l'analyse et la modélisation de données issues des puces à ADN', Thèse de Doctorat, Université Paris 7, 2003, 96pp.
[22] K. Willbrand, F. Radvanyi, J.-P. Nadal, J.-P. Thiery, T.M.A. Fink, Identifying genes from up–down properties of microarray expression series, Bioinformatics 21 (2005) 3859–3864.