ON THE NETTO INVERSION NUMBER OF A SEQUENCE

DOMINIQUE FOATA

1. **Introduction.** Let $g=(x_1, x_2, \cdots, x_n)$ be an arbitrary sequence of real numbers and $\mathcal{C}$ the set of all sequences that can be formed from g by permutations. If $f=(x_{i_1}, x_{i_2}, \cdots, x_{i_n})$ is in $\mathcal{C}$, the *inversion number* $S(f)$ of f is defined as the number of couples (j, k) such that $1 \leqq j < k \leqq n$ and $x_{i_j} > x_{i_k}$ and the *index* $T(f)$ of f as the sum of all integers j such that $1 \leqq j \leqq n-1$ and $x_{i_j} > x_{i_{j+1}}$.

The function S seems to have been introduced by Netto [6] and rediscovered many times in statistics in the theory of *rank tests.* It also appears in the so-called two-sample problem under the name of *Wilcoxon-Mann-Whitney statistic* (see e.g. [1]).

MacMahon ([3], [4]) introduced the function T in the study of ordered partitions. Let q be a real or complex variable and $\boldsymbol{S}=\sum \{q^{S(f)}: f \in \mathcal{C}\}$ (resp. $\boldsymbol{T}=\sum \{q^{T(f)}: f \in \mathcal{C}\}$) be the generating function of S (resp. T). He then obtained [5] the surprising result that $\boldsymbol{S}$ and $\boldsymbol{T}$ have the same expression. Hence the fact that

(1.1) *for any nonnegative integer m there are in $\mathcal{C}$ as many sequences f such that $S(f)=m$ as sequences f' such that $T(f')=m$.*

It seems that no explicit one-to-one correspondence has been so far given between the set of sequences for which T is equal to m and the set of sequences for which S is equal to m. The purpose of the present paper is *to give the construction of such a correspondence.* This construction, without fully explaining the above result (1.1), allows us to introduce a new class of rearrangements of sequences and apply the same noncommutative algebraic methods as in [2].

In what follows, it will be more convenient to identify a sequence $f=(x_{i_1}, x_{i_2}, \cdots, x_{i_n})$ of $\mathcal{C}$ with the *associative monomial* or *word* $x_{i_1}x_{i_2} \cdots x_{i_n}$ of the free monoid X^* generated by $X=\boldsymbol{R}$, to extend the definition of S and T to all of X^* and to construct a *permutation* Φ of X^* satisfying

$$(1.2) \qquad S(\Phi(f)) = T(f) \quad \text{for all } f \in X^*$$

and such that if $f=x_{i_1}x_{i_2} \cdots x_{i_n}$, then $\Phi(f)=x_{v_1}x_{v_2} \cdots x_{v_n}$ where $(x_{v_1}, x_{v_2}, \cdots, x_{v_n})$ is a permutation of $(x_{i_1}, x_{i_2}, \cdots, x_{i_n})$.

The definitions and notations being given in §2, a set of permutations $(\gamma_x)_{x \in X}$ of X^* is introduced (§3) and the permutation Φ is defined in §4 by induction on the length of the words of X^*, i.e. for all $x \in X$ and $f \in X^*$, we set

Received by the editors April 26, 1966.

$$\Phi(x) = x \quad \text{and} \quad \Phi(fx) = \gamma_x(\Phi(f))x.$$

2. **Notations and definitions.** In what follows, X^* is the free monoid generated by a *totally ordered* set X. Each element f of X^* can be written as a *word* $f = x_1 x_2 \cdots x_n$ where $x_1, x_2, \cdots, x_n$ belong to X and are the n *letters* of the word and where n is a nonnegative integer, by definition equal to the *length* of f, denoted by λf. The word of length 0 is the *empty* word denoted by I. The words of length n ($n \geqq 0$) constitute a subset of X^* denoted by X_n and X_1 is identified with X. If f is the product of s ($s \geqq 2$) words $f_1, f_2, \cdots, f_s$ of X^*, we write $f = f_1 f_2 \cdots f_s$ and as in [7] we say that $f_1 f_2 \cdots f_s$ is a *factorization* of f. The word f is also a factorization of itself.

Moreover if Y and Z are subsets of X^*, we designate by Y^* the submonoid generated by Y and by YZ the subset of the words $f = f'f''$ with $f' \in Y$ and $f'' \in Z$. Thus XX^* $(= X^*X)$ is the subset of words of positive length. Now since X is totally ordered, each $x \in X$ determines a partition of X in two subsets L_x and R_x. The set $L_x =]\leftarrow, x]$ (resp. $R_x =]x, \rightarrow[$) is formed with all $y \in X$ such that $y \leqq x$ (resp. $y > x$). Then for each $x \in X$ and $f = x_1 x_2 \cdots x_n \in X^*$, we denote by $l_x f$ (resp. $r_x f$) the number of subscripts j for which $1 \leqq j \leqq n$ and $x_j \leqq x$ (resp. $x < x_j$). Note that we always have $l_x f + r_x f = \lambda f$. If $l_x f = l_x f'$ for all $x \in X$ or if f' is a rearrangement of the letters of f, we set $\alpha(f) = \alpha(f')$.

Finally for $f = x_1 x_2 \cdots x_n \in X^*$, we set

$S(f)$ = number of couples (j, k) such that $1 \leqq j < k \leqq n$ and $x_j > x_k$.

$T(f)$ = sum of all integers j such that $1 \leqq j \leqq n-1$ and $x_j > x_{j+1}$.

$$\begin{aligned} f^\pi &= f \quad \text{if } n = 0 \text{ or } 1, \\ &= x_n\, x_1\, x_2 \cdots x_{n-1} \quad \text{if } n > 1. \end{aligned}$$

3. **The set of permutations $(\gamma_x)_{x \in X}$.** First, it is obvious that for every $x \in X$,

$$\{X^*L_x, X^*R_x\} \quad \text{and} \quad \{L_x X^*, R_x X^*\}$$

are two partitions of X^*X $(= XX^*)$. Moreover, let $f = x_1 x_2 \cdots x_n$ be a word of X^*L_x (resp. X^*R_x, $L_x X^*$, $R_x X^*$) and denote by $(r_1, r_2, \cdots, r_s)$ the increasing sequence of integers j ($1 \leqq j \leqq n$) such that $x_j \in L_x$ (resp. R_x, L_x, R_x). This sequence is not empty. Put $r_0 = 0$, $r_{s+1} = n+1$ and for $p = 1, 2, \cdots, s$

$$f_p = x_{r_{p-1}+1} x_{r_{p-1}+2} \cdots x_{r_p} \quad \text{if } f \in X^*L_x \text{ or } f \in X^*R_x$$

and

$$f_p = x_{r_p} x_{r_p+1} \cdots x_{r_{p+1}-1} \text{ if } f \in L_x X^* \text{ or } f \in R_x X^*.$$

Clearly, $f_1 f_2 \cdots f_s$ is the unique factorization of f where each $f_p \in R_x^* L_x$ (resp. $L_x^* R_x$, $L_x R_x^*$, $R_x L_x^*$).

This factorization will now be used for establishing a one-to-one correspondence between X^*L_x and L_xX^* on one hand, and X^*R_x and R_xX^* on the other hand and so defining a permutation γ_x of X^*. First we set $\gamma_x(I)=I$, then if $f_1 f_2 \cdots f_s$ is the factorization of a word $f \in X^*L_x$ (resp. X^*R_x) into words of $R_x^*L_x$ (resp. $L_x^*R_x$), we set

$$\gamma_x(f) = f_1^\pi f_2^\pi \cdots f_s^\pi. \tag{3.1}$$

We have $f_p^\pi \in L_x R_x^*$ (resp. $R_x L_x^*$) for $p=1, \cdots, s$; hence from above $f_1 f_2 \cdots f_p$ is the factorization of a unique word $\gamma_x(f) \in L_x X^*$ (resp. $R_x X^*$) into words of $L_x R_x^*$ (resp. $R_x L_x^*$). Finally, as $h \to h^\pi$ maps in a one-to-one manner $L_x^* R_x$ onto $R_x L_x^*$ and $R_x^* L_x$ onto $L_x R_x^*$, γ_x is a *permutation* of X^* and besides, for every $f \in X^*$, $\gamma_x(f)$ is a rearrangement of the letters of f, i.e.

$$\alpha(\gamma_x(f)) = \alpha(f). \tag{3.2}$$

Before introducing the permutation Φ, we give in the following lemma some properties of the functions S and T.

(3.3) LEMMA. *For each $f \in X^*$ and $x \in X$,*

$$S(fx) = S(f) + r_x f, \tag{3.4}$$

$$S(\gamma_x(f)) = S(f) - r_x f \quad \textit{if } f \in X^*L_x, \tag{3.5}$$

$$S(\gamma_x(f)) = S(f) + l_x f \quad \textit{if } f \in X^*R_x, \tag{3.6}$$

$$T(fx) = T(f) \quad \textit{if } f \in X^*L_x, \tag{3.7}$$

$$T(fx) = T(f) + \lambda f \quad \textit{if } f \in X^*R_x. \tag{3.8}$$

PROOF. Let $f = x_1 x_2 \cdots x_n \in X^*$.

First, (3.4) holds for the inversion number of fx is equal to the inversion number of f, plus the number of subscripts j $(1 \leqq j \leqq n)$ such that $x_j > x$, i.e. $r_x f$.

Now if $f \in R_x^* L_x$, we can write $f = f' x_n$ $(f' \in R_x^*,\ x_n \leqq x)$; thence

$$r_x f = r_{x_n} f = r_{x_n} f' = \lambda f'. \tag{3.9}$$

But $\gamma_x(f) = f^\pi = x_n f'$. Thus $S(\gamma_x(f))$ is equal to the inversion number of f' plus if $n>1$, the number of subscripts j $(1 \leqq j \leqq n-1)$ such that $x_n > x_j$, which is 0 since $f' \in R_x^*$, i.e.

$$\begin{aligned} S(\gamma_x(f)) &= S(x_n f') = S(f') \\ &= S(f' x_n) - r_{x_n} f' \quad \text{from (3.4)} \\ &= S(f) - r_x f \quad \text{from (3.9)}; \end{aligned}$$

(3.5) is then true for the words $f \in R_x^* L_x$. Finally, if $f \in X^* L_x$, let $f_1 f_2 \cdots f_s$ be its factorization into words of $R_x^* L_x$. By applying γ_x to f, we obtain $\gamma_x(f) = f_1^\pi f_2^\pi \cdots f_s^\pi$ and clearly the inversion number of f is decreased by $r_x f_1 + r_x f_2 + \cdots + r_x f_s$, i.e. $r_x f$.

(3.6) has an analogous proof. We simply notice that applying γ_x to a word f of $L_x^* R_x$, increases the inversion number by $\lambda f - 1$, or $l_x f$.

When $f \in X^* L_x$, the last letter x_n of f is less than or equal to x and the indices of f and fx are the same. Hence (3.7) holds.

On the contrary, if $f \in X^* R_x$, then $x_n > x$ and we get

$$T(fx) = T(f) + \lambda f.$$

That is (3.8), which completes the proof of the lemma.

4. **The combinatorial theorem.** By induction on the length of words $f \in X^*$, we then define Φ in the following way:

$$\Phi(f) = f \quad \text{if } \lambda f \leqq 1 \tag{4.1}$$

and

$$\Phi(fx) = \gamma_x(\Phi(f))x \quad \text{for all } x \in X. \tag{4.2}$$

We then have

(4.3) THEOREM. *The map* $\Phi: X^* \to X^*$

(4.4) *is a permutation; and we have*

$$\alpha(\Phi(f)) = \alpha(f) \tag{4.5}$$

$$S(\Phi(f)) = T(f) \textit{ identically.} \tag{4.6}$$

PROOF. It is sufficient to verify that for all $n \geqq 0$ the restriction Φ_n of Φ to X_n is a permutation of X_n satisfying (4.5) and (4.6). This is obvious for $n \leqq 1$ since by definition Φ_0 (or Φ_1) is the identity map. On the other hand from the definition of Φ we have for $n > 0$,

$$\Phi_{n+1}(fx) = \gamma_x(\Phi_n(f))x, \tag{4.7}$$

valid for all $f \in X_n$ and $x \in X$.

So assume that Φ_n is a permutation of X_n satisfying $\alpha(\Phi_n(f)) = \alpha(f)$ and $S(\Phi_n(f)) = T(f)$ identically. Then $\gamma_x \circ \Phi_n$ is also a permutation of X_n and satisfies $\alpha(\gamma_x(\Phi_n(f))) = \alpha(f)$ identically according to (4.7).

Hence, $fx \to \Phi_{n+1}(fx) = \gamma_x(\Phi_n(f))x$ is a permutation of the subset of the words of X_n ending by x and also $\alpha(\Phi_{n+1}(fx)) = \alpha(fx)$.

Since $X_{n+1} = \bigcup_{x \in X} X_n \{x\}$, it then follows that Φ_{n+1} is a permutation of X_{n+1} satisfying

$$\alpha(\Phi_{n+1}(f)) = \alpha(f) \text{ identically.}$$

Property (4.6) is then a consequence of the lemma. For

$$\begin{aligned} S(\Phi_{n+1}(fx)) &= S(\gamma_x(\Phi_n(f))x) \\ &= S(\gamma_x(\Phi_n(f))) + r_x\gamma_x(\Phi_n(f)) \quad \text{(according to (3.4))} \\ &= S(\gamma_x(\Phi_n(f))) + r_x f \end{aligned}$$

since $\gamma_x(\Phi_n(f))$ is only a rearrangement of the letters of f.

Two cases are to be considered.

(i) $f \in X^*L_x$. Then

$$\begin{aligned} S(\gamma_x(\Phi_n(f))) &= S(\Phi_n(f)) - r_x\Phi_n(f) \quad \text{(according to (3.5))} \\ &= S(\Phi_n(f)) - r_x f. \end{aligned}$$

Hence,

$$\begin{aligned} S(\Phi_{n+1}(fx)) &= S(\Phi_n(f)) \\ &= T(f) \qquad \text{(by induction)} \\ &= T(fx) \quad \text{(according to (3.7)).} \end{aligned}$$

(ii) $f \in X^*R_x$. Then

$$\begin{aligned} S(\gamma({}_x\Phi_n(f))) &= S(\Phi_n(f)) + l_x\Phi_n(f) \quad \text{(according to (3.6))} \\ &= S(\Phi_n(f)) + l_x f. \end{aligned}$$

Hence,

$$\begin{aligned} S(\Phi_{n+1}(fx)) &= S(\Phi_n(f)) + l_x f + r_x f \\ &= T(f) + \lambda f \quad \text{(by induction)} \\ &= T(fx) \quad \text{(according to (3.8)).} \end{aligned}$$

This establishes the theorem.

References

1. F. N. David and D. E. Barton, *Combinatorial chance*, Griffin, London, 1962.

2. D. Foata, *Étude algébrique de certains problèmes d'analyse combinatoire et du calcul des probabilités*, Publ. Inst. Statist. Univ. Paris **14** (1965), 81–241.

3. P. A. MacMahon, *Combinatory analysis*, Vol. 1, Cambridge Univ. Press, Cambridge, 1915.

4. ———, *The indices of permutations and the derivation therefrom of functions of a single variable associated with the permutations of any assemblage of objects*, Amer. J. Math **35** (1913), 281–322.

5. ———, *Two applications of general theorems in combinatory analysis*, Proc. London Math. Soc. **15** (1916), 314–321.

6. E. Netto, *Lehrbuch der Combinatorik*, Chelsea, New York, 1901.

7. M. P. Schützenberger, *On a factorisation of free monoids*, Proc. Amer. Math. Soc. **16** (1965), 21–24.

Institut de Recherche Mathématique Avancée, Université de Strasbourg