COMBINATORIAL RESULTS IN FLUCTUATION THEORY

By Charles Hobby[1] and Ronald Pyke[2]

University of Washington

1. Introduction and summary. The main purpose of the paper is to prove a combinatorial extension of a recent theorem of Baxter [6] concerning partial sums of random variables. The method of proof is closely related to that used by Sparre Andersen [1] in his basic 1949 paper which initiated the combinatorial approach to problems in fluctuation theory. The method does not seem to have been used since.

The central idea behind Sparre Andersen's method, as it is used below, is conceptually very simple. It consists of verifying the validity of two operations on the finite sequences of real numbers under consideration. The first operation is referred to as "shrinking", and the second as "counting". In order to prove an invariant combinatorial result for such finite sequences of numbers, one shows first that if the result is true for a given sequence, it remains true as one decreases (or shrinks) the smallest number in the sequence (the shrinking Lemma 2.1) and, second, that if one inserts a sufficiently small number into a sequence for which the theorem holds, then the result also holds for the new sequence (the counting Lemma 2.2).

Section 2 contains the fundamental combinatorial theorem concerning the joint behavior of the number of partial sums greater than zero and the number of them less than the last partial sum. Section 3 presents a probabilistic framework for the results of Section 2, as well as some further results.

2. A combinatorial theorem. Let $c = (c_1, c_2, \cdots, c_n)$ be a fixed sequence of n real numbers. For convenience, assume $0 < c_1 < c_2 < \cdots < c_n$. Let $\mathcal{S}_c$ denote the set of all $2^n n!$ sequences that may be formed from c by using all possible permutations, and all possible attachments of a $+$ or $-$ sign to each coordinate. For any sequence $x = (x_1, x_2, \cdots, x_n)$ of real numbers set $s_0(x) = 0$, $s_j(x) = x_1 + x_2 + \cdots + x_j$ for $j = 1, 2, \cdots, n$. We will refer to both the sequence x and the sequence of partial sums, $(s_0, s_1, \cdots, s_n)$ as a *path*. With each sequence $x \,\varepsilon\, \mathcal{S}_c$, associate the ordered pair of integers (m, k) where m is the number of positive partial sums $s_i(x)$, and k is the number of positive reversed partial sums $s_n(x) - s_i(x)$. The sequence x is then said to be of *type* (m, k). We set $v_n(m, k; c)$ equal to the number of sequences $x \,\varepsilon\, \mathcal{S}_c$ which are of type (m, k). Clearly $0 \leqq v_n(m, k; c) \leqq 2^n n!$. For any sequence $y = (y_1, y_2, \cdots, y_n)$ of

Received November 29, 1962; revised May 1, 1963.

[1] This author's research was supported in part by a grant from the National Science Foundation, NSF-G19915.

[2] This author's research was supported in part by the Office of Naval Research, Contract Nonr 477(11)-042038. Reproduction in whole or in part is permitted for any purpose of the United States Government.

real numbers, let $N_n(y)$ and $M_n(y)$ denote respectively the number of positive partial sums $s_i(y)$ and the number of positive reversed partial sums $s_n(y) - s_i(y)$, $(i = 0, 1, \cdots, n)$. Set $L_n(y)$ equal to the first subscript j for which $s_j(y) = \max\{s_i(y); 0 \leqq i \leqq n\}$. It is desirable to exclude the possibility of some of the partial sums being zero, or equivalently, to postulate that all of the partial sums are distinct. To this end we make the

DEFINITION. The sequence c is said to have *property D* if for every sequence $x \in \mathcal{S}_c$, $s_i(x) \neq 0$ for $i > 0$.

It is straightforwardly checked that if c possesses property D, then $v_n(m, n - m; c) = 0$ for $0 \leqq m \leqq n$, and $\sum_{m,k=0}^{n} v_n(m, k; c) = 2^n n!$, for $0 \leqq m, k \leqq n$.

The main result to be proved in this section is that $v_n(m, k; c)$ is a constant for all choices of c which have property D. The invariance of the counting function $v_n(m, k; c)$ is obtained by an argument which first proves a partial invariance and then uses this to proceed inductively to compute the explicit form of $v_n(m, k; c)$ and to complete the invariance argument.

THEOREM 2.1. *For all non-negative integers n, m, k satisfying $n \geqq 1$, $m + k < n$, and for all sequences $c = (c_1, c_2, \cdots, c_n)$, $0 < c_1 < c_2 < \cdots < c_n$, which possess property D,*

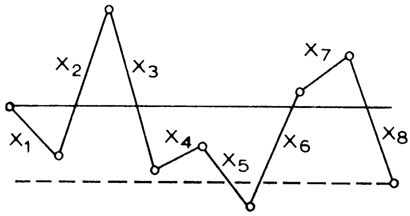
$$v_n(m, k; c) = \binom{2m}{m}\binom{2k}{k} 2^{n-1-2m-2k}(n-1)! \tag{2.1}$$

Notice that the condition $m + k < n$ is equivalent to specifying that $s_n < 0$. Further, if $x = (x_1, \cdots, x_n)$ is of type (m, k), then $(-x_1, -x_2, \cdots, -x_n)$ is of type $(n - m, n - k)$ and $(x_n, x_{n-1}, \cdots, x_1)$ is of type (k, m). It follows that $v_n(m, k; c) = v_n(n - m, n - k; c) = v_n(n - k, n - m; c) = v_n(k, m; c)$. Thus (2.1) leads to a formula for $v_n(m, k; c)$ for all integers m, k satisfying $0 \leqq m, k \leqq n, m + k \neq n$.

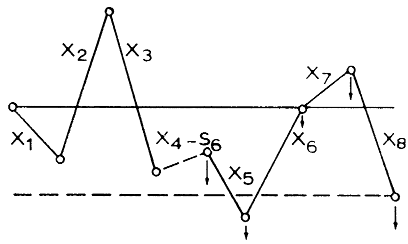
The proof of Theorem 2.1 uses two lemmas. For any n-tuple $c = (c_1, c_2, \cdots, c_n)$ for which $0 < c_1 < \cdots < c_n$, and any $\delta \in (0, c_1)$, define the shrinking transformation, T_δ, by $T_\delta(x) = (c_1 - \delta, c_2, \cdots, c_n)$. Write $\mathcal{S}_{c,\delta}$ for $\mathcal{S}_{T_\delta(c)}$. If $x \in \mathcal{S}_c$, denote by x^δ the element in $\mathcal{S}_{c,\delta}$ which is formed from $T_\delta(c)$ in the same way as x is formed from c. It will suffice to prove that $v_n(m, k; c)$ does not change when c is replaced by $T_\delta(c)$, provided that c and $T_\delta(c)$ both have property D. However, the argument used below to prove this result can be modified to show that $v_n(m, k; c)$ is unchanged if c is replaced by *any* other sequence of n numbers possessing property D.

LEMMA 2.1. *If a sequence $c = (c_1, c_2, \cdots, c_n)$, $0 < c_1 < \cdots < c_n$, possesses property D, then $v_n(m, k; c) = v_n(m, k; T_\delta(c))$ for all integers $0 \leqq m, k < n$, $m + k < n$ and all $\delta \in (0, c_1)$ for which $T_\delta(c)$ possesses property D.*

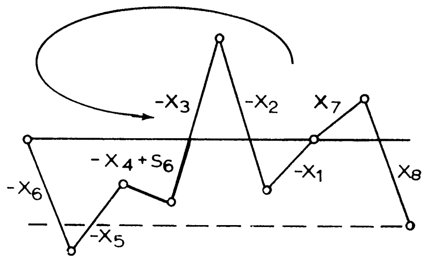
PROOF. We must obtain a 1-1 correspondence, ϕ_δ, between $\mathcal{S}_c$ and $\mathcal{S}_{c,\delta}$, such that for all $x \in \mathcal{S}_c$, x and $\phi_\delta(x)$ are of the same type. The correspondence we shall define is illustrated in Figure 1. Observe first that there are only finitely many values of δ, say $\delta_1 < \delta_2 < \cdots$, for which $T_\delta(c)$ does not possess property D.



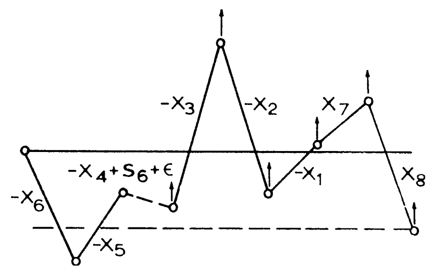
This path $(x_1, \cdots, x_7)$ is of type (2,1). The smallest increment is x_4.



Switch the first 6 steps of the path as shown.



Shrink x_4 until type changes, namely, until $s_6 = 0$.



Continue shrinking, and the path returns to type (2,1).

FIG. 1

These are simply the values of δ for which at least one partial sum of a member of $\mathcal{S}_{(c_2, \cdots, c_n)}$ is equal to $c_1 - \delta$. It follows from the definition of δ_1 that x and x^δ are of the same type for every $x \varepsilon \mathcal{S}_c$ if $0 \leqq \delta < \delta_1$. It will therefore only be necessary to consider $d \varepsilon (\delta_1, \delta_2)$ and to exhibit the mapping ϕ_d for this choice of d. For all other values of d, ϕ_d will be a composition of a finite number of mappings of this form.

Let $\mathfrak{D}$ consist of all $x \varepsilon \mathcal{S}_c$ such that x and x^d are of different types. For each $x \varepsilon \mathfrak{D}$ there is a subscript $r = r(x)$ at which either $s_r(x)$ and $s_r(x^d)$ or $s_n(x) - s_r(x)$ and $s_n(x^d) - s_r(x^d)$ have opposite signs. One needs to show that this subscript $r = r(x)$ is unique, and that exactly one of the above cases obtains. Assuming for the moment that these facts are true, define $\phi_d(x) = (\phi(x))^d$, where ϕ is the 1-1 mapping of $\mathcal{S}_c$ onto itself given by

$$
\begin{aligned}
\phi(x) &= x \text{ if } x \not\varepsilon \mathfrak{D} \\
&= (-x_r, \cdots, -x_1, x_{r+1}, \cdots, x_n) \text{ if } s_r(x)s_r(x^d) < 0, \\
&= (x_1, \cdots, x_r, -x_n, \cdots, -x_{r+1}) \text{ if } [s_n(x) - s_r(x)] \\
&\qquad \cdot [s_n(x^d) - s_r(x^d)] < 0.
\end{aligned}
\tag{2.2}
$$

It is clear that ϕ_d is 1-1.

It remains to show that $r = r(x)$ is unique, and that x and $\phi_d(x)$ are always of the same type. These two results are intuitively fairly obvious from Figure 1.

However, a formal proof of the second result requires a careful treatment of several cases, so it is included here. The proof parallels that in [1].

To show that x and $\phi_d(x)$ are of the same type, observe first that if $\text{sgn}[s_n(x) - s_r(x)] \neq \text{sgn}[s_n(x^d) - s_r(x^d)]$, then for $y = (-x_n, -x_{n-1}, \cdots, -x_1)$, one obtains $r(y) = n - r(x)$ and $\text{sgn}[s_{r(y)}(y)] \neq \text{sgn}[s_{r(y)}(y^d)]$. Furthermore, x and $\phi_d(x)$ are of the same type, say (m, k), if and only if both y and $\phi_d(y)$ are of type $(n - k, n - m)$. Therefore, without loss of generality, we assume that $\text{sgn}[s_r(x)] \neq \text{sgn}[s_r(x^d)]$. Then by (2.2) $\phi(x) = (-x_r, \cdots, -x_1, x_{r+1}, \cdots, x_n)$. Consequently

$$
\begin{aligned}
s_i(\phi(x)) &= s_{r-i}(x) - s_r(x) \quad \text{if } 0 \leqq i \leqq r \\
&= s_i(x) - 2s_r(x) \quad \text{if } r \leqq i \leqq n.
\end{aligned} \tag{2.3}
$$

Observe that $r(x) = r(\phi(x))$. We shall require the following equalities, where $i \neq 0, i \neq r$:

$$
\begin{aligned}
&\text{sgn}[s_i(x)] = \text{sgn}[s_i(x^d)], \qquad \text{sgn}[s_i(x) - s_r(x)] = \text{sgn}[s_i(x)], \\
&\qquad \text{sgn}[s_n(x) - s_i(x) - s_r(x)] = \text{sgn}[s_n(x) - s_i(x)] \quad (1 \leqq i < r).
\end{aligned} \tag{2.4}
$$

The second equality follows from the first and the fact that $\text{sgn}[s_r(x)] \neq \text{sgn}[s_r(x^d)]$, since at least one of the equalities $\text{sgn}[s_i(x) - s_r(x)] = \text{sgn}[s_i(x)]$ or $\text{sgn}[s_i(x^d) - s_r(x^d)] = \text{sgn}[s_i(x^d)]$ must hold. A similar argument will yield the third equality if one observes that $s_n(\phi(x)) - s_{r-i}(\phi(x)) = s_n(x) - s_i(x) - s_r(x)$. Hence $\text{sgn}[s_n(x) - s_i(x) - s_r(x)] = \text{sgn}[s_n(x^d) - s_i(x^d) - s_r(x^d)]$.

If $1 \leqq i < r$, it follows from (2.3) and (2.4) that $\text{sgn}[s_i(\phi_d(x))] = \text{sgn}[s_i(\phi(x))] = \text{sgn}[s_{r-i}(x) - s_r(x)] = \text{sgn}[s_{r-i}(x)]$. Similarly, if $r < i \leqq n$, then $\text{sgn}[s_i(\phi_d(x))] = \text{sgn}[s_i(\phi(x))] = \text{sgn}[s_i(\phi(x)) - s_r(\phi(x))] = \text{sgn}[s_i(x) - s_r(x)] = \text{sgn}[s_i(x)]$. Finally, $\text{sgn}[s_r(\phi_d(x))] = -\text{sgn}[s_r(\phi(x))] = \text{sgn}[s_r(x)]$. Therefore x and $\phi_d(x)$ have the same number of positive partial sums.

If $r \leqq i \leqq n$, then $\text{sgn}[s_n(\phi_d(x)) - s_i(\phi_d(x))] = \text{sgn}[s_n(\phi(x)) - s_i(\phi(x))] = \text{sgn}[s_n(x) - s_i(x)]$. If $1 \leqq i \leqq r$, then $\text{sgn}[s_n(\phi_d(x)) - s_i(\phi_d(x))] = \text{sgn}[s_n(\phi(x)) - s_i(\phi(x))] = \text{sgn}[s_n(x) - s_{r-i}(x) - s_r(x)] = \text{sgn}[s_n(x) - s_{r-i}(x)]$. Thus x and $\phi_d(x)$ have the same number of positive reversed partial sums. This completes the proof of Lemma 2.1.

In the above lemma it is proved that $v_n(m, k; c)$ remains constant as one shrinks the smallest coordinate of c. The second step in the proof of Theorem 2.1 is to show that one can recursively evaluate $v_n(m, k; c)$ when the smallest coordinate of c is sufficiently small.

LEMMA 2.2. *Let $c = (c_1, c_2, \cdots, c_n)$, $0 < c_1 < \cdots < c_n$, be a sequence possessing property D. Let $c_0 > 0$ be less than the absolute value of every non-zero partial sum of every sequence in $\mathcal{S}_c$. Set $c' = (c_0, c_1, \cdots, c_n)$. Then c' possesses property D and*

$$
\begin{aligned}
v_{n+1}(m, k; c') &= (2m - 1)v_n(m - 1, k; c) \\
&\quad + (2k - 1)v_n(m, k - 1; c) + 2(n - m - k)v_n(m, k; c)
\end{aligned} \tag{2.5}
$$

for all integers $0 \leqq m, k \leqq n, m + k < n$. *It is understood that* $v_n(i, j; c) = 0$ *for negative values of* $i, j, n - i$, *or* $n - j$.

PROOF. Clearly c' possesses property D. To check that (2.5) is valid, observe simply that there are always $2(n + 1)$ ways that c_0 can be inserted into a sequence in $\mathcal{S}_c$ to form a sequence in $\mathcal{S}_{c'}$. Namely, c_0 can be placed in any of $n + 1$ locations and with either of two signs. If c_0 is inserted with either sign following the ith coordinate of a sequence in $\mathcal{S}_c$ for which $s_i > 0$, then $s_i \pm c_0 > 0$ and so the resulting sequence in $\mathcal{S}_{c'}$ has one additional positive partial sum. The same will happen if c_0 is inserted with a positive sign at the beginning of the sequence. In both of these cases the number of positive reversed partial sums remains unchanged. Thus one constructs $2m - 1$ sequences of type (m, k) in $\mathcal{S}_{c'}$ from each sequence of type $(m - 1, k)$ in $\mathcal{S}_c$. This justifies the first term on the right hand side of (2.5). The other terms are explained in a similar manner.

PROOF OF THEOREM 2.1. We proceed by induction. The theorem is obvious for $n = 1$. Assume the theorem is true for some value $n \geqq 1$, and consider any sequence $c = (c_1, c_2, \cdots, c_{n+1})$ of length $n + 1$ which satisfies the assumptions of the theorem. By Lemma 2.1 it is possible to shrink the smallest element c_1, and leave the problem invariant. Therefore, decrease c_1 until it is less than the absolute value of every non-zero partial sum of every sequence in $\mathcal{S}_c$. By Lemma 2.2, this implies that (2.5) is valid. Applying the induction hypothesis to (2.5) yields

$$v_{n+1}(m, k; c) = 2^{n-1-2m-2k}(n - 1)!\left\{(2m - 1)2^2\binom{2m - 2}{m - 1}\binom{2k}{k} + (2k - 1)2^2\binom{2m}{m}\binom{2k - 2}{k - 1} + 2(n - m - k)\binom{2m}{m}\binom{2k}{k}\right\}$$

from which the desired evaluation of $v_{n+1}(m, k; c)$ may be derived.

As mentioned earlier, the shrinking method could be extended to show the constancy of $v_n(m, k; c)$ over c without using the counting lemma. Then to evaluate the constant, one could attempt to obtain an exact count for some conveniently chosen c. (Perhaps a sequence c for which $s_i(c) < c_{i+1}$ could be used, as it was by Sparre Andersen [1]). However, the above method seems more straightforward.

Let $\mu_{n,r}(m_1, k_1, m_2, k_2; c)$ denote the number of sequences $x \in \mathcal{S}_c$ for which $(x_1, x_2, \cdots, x_r)$ is of type (m_1, k_1) and $(x_{r+1}, x_{r+2}, \cdots, x_n)$ is of type (m_2, k_2). Because of Theorem 2.1, we write $v_n(m, k)$ in place of $v_n(m, k; c)$.

COROLLARY 2.1. *For all integers* m_1, m_2, k_1, k_2, r *and* n, *satisfying* $0 \leqq r \leqq n$, *and any sequence* c *possessing property* D,

$$\mu_{n,r}(m_1, k_1, m_2, k_2; c) = \binom{n}{r} v_r(m_1, k_1)v_{n-r}(m_2, k_2). \tag{2.6}$$

PROOF. This result is a consequence of the fact that all sequences $x \in \mathcal{S}_c$ for which $(x_1, \cdots, x_r)$ is of type (m_1, k_1) and $(x_{r+1}, \cdots, x_n)$ is of type (m_2, k_2)

can be constructed by first selecting r coordinates of c in one of the $\binom{n}{r}$ possible ways and then forming the required type of sequence from the two resulting subsequences.

One application of this corollary is to obtain, by performing the appropriate summation in (2.6), the combinatorial form of the result of Sparre Andersen given as Theorem 5 in [2]. A second consequence of Theorem 2.1 is

COROLLARY 2.2. *Let $v_n^{(i)}(m, k; c)$ be the number of $x \in \mathcal{S}_c$ which have m partial sums $s_j(x)$, for $j \geqq i$, greater than $s_i(x)$ and k reversed partial sums $s_n(x) - s_j(x)$, for $j \geqq i$, less than $s_n(x)$. Then if c possesses property D and $i < n$,*

$$v_n^{(i)}(m, k; c) = 2^i i! \binom{n}{i} v_{n-i}(m, k) = n(n - i)^{-1} v_n(m, k).$$

Define $T_n(x) = N_n(x) + M_n(x)$ and $J_n(x) = \max\{n - T_n(x), T_n(x) - n\} - 1$. These quantities have a uniform distribution as is stated in

COROLLARY 2.3. *If c possesses property D, then*

(a) $\operatorname{card}\{x \in \mathcal{S}_c : T_n(x) = k\} = 2^{n-1}(n - 1)!,\ 0 \leqq k \leqq 2n,\ k \neq n;$

(b) $\operatorname{card}\{x \in \mathcal{S}_c : J_n(x) = k\} = 2^n(n - 1)!,\ 0 \leqq k < n.$

PROOF. These results are stated as a corollary of Theorem 2.1 since indeed they may be obtained from (2.1) by the appropriate summations. However, it is possible to derive them directly by simply combinatorial arguments. To prove (a) let $c = (c_1, c_2, \cdots, c_n)$ be a sequence possessing property D. For each $x \in \mathcal{S}_c$, form a new sequence of length $2n$, namely $x^* = (x_1, x_2, \cdots, x_n, -x_1, -x_2, \cdots, -x_n)$. Observe that for $0 \leqq j \leqq n$, $s_j(x^*) > 0$ if and only if $s_j(x) > 0$, while $s_{n+j}(x^*) > 0$ if and only if $s_n(x) - s_j(x) > 0$. Hence $T_n(x)$ is equal to $N_{2n}(x^*)$ if $N_{2n}(x^*) < n$ and is equal to $N_{2n}(x^*) + 1$ if $N_{2n}(x^*) \geqq n$. Moreover, $s_{2n}(x^*) = 0$. This latter condition, together with property D, implies by a known result of Andersen [3] and Spitzer [8] that $N_{2n}(x^*)$ ranges over each of the $2n$ integers $0, 1, 2, \cdots, 2n - 1$ exactly once as each of the $2n$ possible cyclic permutations are applied to x^*. This completes the proof of (a). The proof of (b) is then immediate. (Actually, these proofs show that the results are valid when only a special type of cyclic permutation with sign change is used, namely the inverted cyclic permutations introduced in [7].)

The geometrical representation of a path given in Figure 1 indicates that Theorem 2.1 is valid for sequences c of vectors instead of real numbers. In fact the following generalization is immediate.

Let $\mathcal{H}$ be an arbitrary Hilbert space over the reals. For $x \in \mathcal{H}$, let $K(x)$ denote the open sphere with center $x/2$ and radius $\|x/2\|$, that is, $K(x) = \{y \in \mathcal{H}: \|y - x/2\| < \|x/2\|\}$. Let $c = (c_1, c_2, \cdots, c_n)$ be a sequence of elements of $\mathcal{H}$ for which $s_n(c) \neq 0$. (Here the partial sum notation is the obvious extension of that introduced earlier.). For each $i = 1, 2, \cdots, n$, write $c_i = c_i^{\perp} + c_i^{p}$ where $c_i^{\perp}$ and c_i^{p} are, respectively, the perpendicular and projection of c_i with respect to the subspace spanned by $s_n(c)$. Define $\mathcal{S}_c^*$ to be the collection of all $2^n n!$ sequences of length n which can be formed from c by permutations and by forming the "conjugate" $c_i^\perp - c_i^p$. A path $x \varepsilon \mathcal{S}_c^*$ is said to be of *type* (m, k) if m is the number of subscripts i for which $s_n(x) \varepsilon K(s_n(x) - s_i(x))$, and k is the number of subscripts i for which $s_n(x) \varepsilon K(s_i(x))$.

THEOREM 2.2. *For all non-negative integers n, m, k satisfying $n \geqq 1$, $m + k < n$, and for all sequences $c = (c_1, \cdots, c_n)$ with $c_i \varepsilon \mathcal{H}$ and with $(c_1^p, c_2^p, \cdots, c_n^p)$ possessing property D,*

$$v_n(m, k; c) = \binom{2m}{m}\binom{2k}{k} 2^{n-2m-2k}(n - 1)!. \tag{2.7}$$

PROOF. This theorem is an immediate consequence of Theorem 2.1, as may be seen as follows. In order that $s_n(x) \varepsilon K(s_n(x) - s_i(x))$, it is necessary and sufficient that $\|s_n(x) - [s_n(x) - s_i(x)]/2\|^2 < \|[s_n(x) - s_i(x)]/2\|^2$ or that

$$\|s_i^p(x) + s_n^p(x)\| < \|s_n^p(x) - s_i^p(x)\|. \tag{2.8}$$

A similar relationship may be obtained for $s_n(x) \varepsilon K(s_i(x)$. Since the subspace of $\mathcal{H}$ spanned by $s_n(c)$ is 1-dimensional, it is isomorphic to the real line. The proof is then completed by observing that for the case of real c_i, (2.8) is equivalent to stating that either $s_i(x) < 0 < s_n(x)$ or $s_n(x) < 0 < s_i(x)$. Furthermore, as x varies over $\mathcal{S}_c^*$, the partial sums $s_i^p(x)$ vary over the partial sums of sequences in $\mathcal{S}_{c^p}$ where $c^p = (c_1^p, \cdots, c_n^p)$.

It is worth pointing out that the definition of type (m, k) could have been described geometrically as follows. Construct hyperplanes through $s_0(x)$ and $s_n(x)$, perpendicular to the line joining $s_0(x)$ and $s_n(x)$. Then a path is of type (m, k) if $s_n(x)$ and m of the partial sums $s_i(x)$, $1 \leqq i < n$, are on opposite sides of the hyperplane through $s_0(x)$, while $s_0(x)$ and k of the $s_i(x)$, $1 \leqq i < n$, are on opposite sides of the hyperplane through $s_n(x)$.

Let $\mathcal{L}$ be any subspace of $\mathcal{H}$. For $x \varepsilon \mathcal{H}$, write $x = x^\perp + x^p$ for the unique decomposition of x determined by $\mathcal{L}$, and set $x^* = x^\perp - x^p$. For any sequence $y = (y_1, y_2, \cdots, y_n)$ of elements of $\mathcal{H}$, let $t(y)$ denote the conjugated-cyclic permutation of y given by $t(y) = (y_2, y_3, \cdots, y_n, y_1^*)$. Let $c = (c_1, c_2, \cdots, c_n)$ be a sequence of elements of $\mathcal{H}$, and let

$$\mathcal{C}_c^* = \{c, t(c), t^2(c), \cdots, t^{n-1}(c)\}.$$

For $x \varepsilon \mathcal{C}_c^*$, define $J_n(x)$ to be the number of subscripts i for which $s_i(x)^p \varepsilon K(s_n(x)^p)$.

THEOREM 2.3. *Assume that for each $x \varepsilon \mathcal{C}_c^*$ and all $i = 1, \cdots, n - 1$, it is not true that $s_i(x)^p \perp [s_n(x)^p - s_i(x)^p]$. Then, for each $k = 0, 1, \cdots, n - 1$, there exists exactly one element $x \varepsilon \mathcal{C}_c^*$ for which $J_n(x) = k$.*

The proof of this result is exactly similar to that of Theorem 1 in [7], even though the latter theorem was stated for finite-dimensional spaces only.

3. Probabilistic interpretations and applications. Let $X_1, X_2, \cdots, X_n$ be random variables defined on a probability space $(\Omega, \mathcal{A}, P)$. The random vector $X = (X_1, X_2, \cdots, X_n)$ is said to be (i) *exchangeable* if its distribution function (d.f.) is invariant under all permutations of the coordinate variables and (ii) *symmetric* if its d.f. is invariant under all sign attachments to the coordinate variables. Set $S_0 = 0$, $S_i = X_1 + \cdots + X_i$, for $i = 1, 2, \cdots, n$. The random vector $X = (X_1, X_2, \cdots, X_n)$ is said to possess *property D* if $P[S_i = 0] = 0$ for $i = 1, \cdots, n$. Let N_n and M_n denote, respectively, the number of subscripts i for which S_i is positive and the number for which $S_n - S_i$ is positive. Set $p_n(m, k) = P[N_n = m, M_n = k]$. The main results of Section 2 are rephrased in probabilistic language in the following theorem.

THEOREM 3.1. *If the random vector X is exchangeable and symmetric, and possesses property D. then*

$$p_n(m, k) = (2n)^{-1} \binom{2m}{m}\binom{2k}{k} 4^{-m-k} \qquad (m + k < n) \tag{3.1}$$

and

$$p_n(m, n - m) = 0, \qquad p_n(m, k) = p_n(k, m) = p_n(n - m, n - k)$$

for all integers m, k, n with $0 \leqq m, k \leqq n$. *Furthermore*

$$P[N_n = m] = \binom{2m}{m}\binom{2n - 2m}{n - m} 4^{-n}. \tag{3.2}$$

PROOF. Let $Z = (Z_1, Z_2, \cdots, Z_n)$, $0 < Z_1 < Z_2 < \cdots < Z_n$, be an ordering of $|X_1|, |X_2|, \cdots, |X_n|$. Theorem 3.1 is then an immediate consequence of Theorem 2.1 applied to the relationship:

$$P[N_n = m, M_n = k \mid Z = z] = v_n(m, k; z)/2^n n! \tag{3.3}$$

It follows from the form of the expressions in (3.1) and (3.2) that $p_n(m, k) = p_n(0, 0)P[N_{m+k} = m]$ if $m + k < n$. This is related to the type of result contained in Theorem 2 of [2]. Also one sees that $np_n(m, k) = (n + 1)p_{n+1}(m, k)$. If one specifies in Theorem 3.1 that $X_1, X_2, \cdots, X_n$ are independent and identically distributed r.v.'s, whose common d.f. is continuous and symmetric about zero, then one obtains the statement of the theorem which Baxter derived in [6] using analytic methods.

All of the above results have direct application to stochastic processes. For example, let $\{Z_t : t \geqq 0\}$ be a separable and measurable stochastic process, for which every vector of increments $(Z_{t_1} - Z_{t_1+h}, \cdots, Z_{t_k} - Z_{t_k+h})$ for $0 \leqq t_1 < t_1 + h \leqq t_2 < t_2 + h \leqq \cdots < t_k + h$ and $k = 1, 2, \cdots$, satisfies the assumptions of Theorem 3.1. For fixed $T > 0$, define U_T and W_T to be, respectively, the Lebesgue measure of the sets $\{t \in (0, T]: Z_t > 0\}$ and $\{t \in (0, T]: Z_t < Z_T\}$. If it is assumed further, for example, that the Z-process is continuous in probability at each $t \in (0, T]$, then the usual limiting approximation argument yields $P[U_T \leqq u, W_T \leqq w] = (2/\pi T)(uw)^{\frac{1}{2}}$ for all $u, w \geqq 0$, $u + w \leqq T$. These generalized arcsine laws were derived by Baxter [6] for infinitely divisible processes.

4. Remarks. A natural generalization of the concept of a path being of type (m, k), which was introduced in Section 2, is to say that the path $x = (x_1, x_2, \cdots, x_n)$ is of type $(m, k; i, j)$ if there are m partial sums above $s_i(x)$ and k partial sums below $s_j(x)$. Clearly types (m, k) and $(m, k; 0, n)$ are equivalent. In [2], Sparre Andersen showed that the distribution of the number of partial sums above s_i is invariant, as was the number of (positive) partial sums above s_0. Unfortunately, the natural conjecture, that the number of sequences $x \varepsilon \mathcal{S}_c$ of a specified type $(m, k; i, j)$ does not depend upon the sequence c, is false. In fact, if one counts only the number of sequences of type $(0, 0; i, j)$, namely those having their maximum and minimum at s_i and s_j respectively, one finds that the number depends upon the sequence c. For example, consider $n = 3$ and sequences $c = (1, 3, 5)$ and $c' = (2, 3, 4)$. There are three elements of $\mathcal{S}_c$ of type $(0, 0; 0, 1)$, namely $(-5, 3, 1)$, $(-5, 1, 3)$ and $(-5, 3, -1)$, while only one element of $\mathcal{S}_{c'}$ is of this type, namely $(-4, 3, -2)$. Thus even though the individual distributions of the location of the maximum and minimum partial sums are invariant, the joint distribution of these locations depends on the sequence c. This, in turn, implies that the set of all of the ranks of the partial sums cannot have an invariant distribution.

Next, one might consider, for fixed m and k, $(m + k \leqq n)$, the number of sequences $x \varepsilon \mathcal{S}_c$, for which $s_m(x)$ is the maximum partial sum and $s_{m+k}(x)$ is the minimum among the partial sums $s_{m+1}(x), s_{m+2}(x), \cdots, s_n(x)$. However, this number depends on the choice of c, as is shown by the example constructed above (for $m = 0, k = 1$).

If, instead of looking for the minimum, we ask that $s_{m+k}(x)$ be the maximum of $s_{m+1}(x), s_{m+2}(x), \cdots, s_n(x)$, then a positive result is obtained. Formally, let us say that a sequence $x = (x_1, x_2, \cdots, x_n)$ is of *location type* (m, k) if $L_n(x) = m$ and $L_{n-m}(x_{m+1}, \cdots, x_n) = k$. Then one may prove

THEOREM 4.1. *If $c = (c_1, \cdots, c_n)$ satisfies property D, then the number of sequences $x \varepsilon \mathcal{S}_c$ for which x is of location type (m, k), $m + k \leqq n$, is given by*

$$\binom{n}{m}\binom{n-m}{k} v_m(m) v_k(0, k-1) v_j(0) = \binom{2m}{m}\binom{2k}{k}\binom{2j}{j} n!\, 2^{-n}(2k-1)^{-1} \tag{4.1}$$

where $j = n - m - k$.

PROOF. This result follows immediately from Theorem 2.1 if one observes that a sequence $x \varepsilon \mathcal{S}_c$ is of location type (m, k) if and only if $(x_1, \cdots, x_m)$ is of type (i, m) for some $1 \leqq i \leqq m$, $(x_{m+1}, \cdots, x_{m+k})$ is of type $(0, k - 1)$, and $(x_{m+k+1}, \cdots, x_n)$ is of type $(0, i)$ for some $0 \leqq i \leqq n - m - k$. These subsequences can be chosen from c in $\binom{n}{m}\binom{n-m}{k}$ ways, while the other terms in (4.1) can be seen to represent the number of ways of obtaining the desired types if one recalls that $v_n(0) = v_n(n) = \sum_{i=1}^{n} v_n(i, n)$ for all n.

The technique of shrinking and counting which was introduced by Sparre Andersen in 1949 and is used to prove Theorem 2.1 above may be used to obtain combinatorial results which are somewhat different in nature from Theorem 2.1. One example of this is the result of Sparre Andersen (Theorem 5 of [4]), which gives the number of paths generated by the $n!$ permutations of a given sequence, which have exactly k, $(k = 0, 1, \cdots, n)$, coincidences with their respective convex minorants. (See [4] for a precise statement of the result.) The proof could proceed as follows. Consider a path, such as in Figure 1. Shrink one of the path's segments until a coincidence is about to be either lost or gained (i.e. until 3 partial sums lie on a straight line.) Then interchange the two adjacent subpaths between these 3 coincidence points, and then continue the shrinking process. It is geometrically clear that the number of coincidences remains invariant. One could then proceed to make the actual count. This gives a very straightforward proof, (although not the shortest (e.g. see [5])), which has the advantage that it leads immediately to the generalization of the problem to sequences of two-dimensional vectors with positive first coordinates (as studied in [5] and in other papers referred to therein).

REFERENCES

[1] Andersen, E. Sparre (1949). On the number of positive sums of random variables. *Skand. Aktuarietidskr.* **32** 27–36.

[2] Andersen, E. Sparre (1953). On sums of symmetrically dependent random variables. *Skand. Aktuarietidskr.* **36** 123–138.

[3] Andersen, E. Sparre (1953). On the fluctuation of sums of random variables. *Math. Scand.* **1** 263–285.

[4] Andersen, E. Sparre (1954). On the fluctuation of sums of random variables II. *Math. Scand.* **2** 195–223.

[5] Barton, D. E. and Mallows, C. L. (1961). The randomisation basis of the problem of the amalgamation of weighted means. *J. Roy. Stat. Soc. Ser. B* **23** 423–433.

[6] Baxter, Glen (1962). On a generalization of the finite arcsine law. *Ann. Math. Statist.* **33** 909–915.

[7] Hobby, Charles and Pyke, Ronald (1963). Combinatorial results in multi-dimensional fluctuation theory. *Ann. Math. Statist.* **34** 402–404.

[8] Spitzer, Frank (1956). A combinatorial lemma and its application to probability theory. *Trans. Amer. Math. Soc.* **82** 323–339.