RESEARCH ANNOUNCEMENTS

The purpose of this department is to provide early announcement of significant new results, with some indications of proof. Although ordinarily a research announcement should be a brief summary of a paper to be published in full elsewhere, papers giving complete proofs of results of exceptional interest are also solicited. Manuscripts more than eight typewritten double spaced pages long will not be considered as acceptable. All papers to be communicated by a Council member should be sent directly to M. H. Protter, Department of Mathematics, University of California, Berkeley, California 94720.

BAXTER ALGEBRAS AND COMBINATORIAL IDENTITIES. I

BY GIAN-CARLO ROTA

Communicated August 15, 1968

1. **Introduction.** The spectacular results in the fluctuation theory of sums of independent random variables, obtained in the last 15 years by Andersen, Baxter, Bohnenblust, Foata, Kemperman, Spitzer, Takacs and others, have gradually led to the realization that the nature of the problem, as well as that of the methods of solution, is algebraic and combinatorial. After Baxter showed that the crux of the problem lay in simplifying a certain operator identity, several algebraic proofs (Atkinson, Kingman, Wendel) followed. It is the present purpose to carry this algebraization to the limit: the result we present amounts to a solution of the word problem for Baxter algebras. The solution is not presented as an algorithm, but by showing that every identity in a Baxter algebra is effectively equivalent to an identity of symmetric functions independent of the number of variables. Remarkably, the identities used so far in the combinatorics of fluctuation theory "translate" by the present method into classical symmetric function identities of easy verification. The present method is nevertheless also useful for guessing and proving new combinatorial identities: by way of example, it will be shown in the second part of this note how it leads to a generalization of the Bohnenblust-Spitzer formula for the action of arbitrary groups of permutations. A parallel theory of inequalities will be presented elsewhere.

2. **Definitions.** Let $\boldsymbol{A}$ be a commutative ring. A *Baxter operator* on $\boldsymbol{A}$ is a linear function P mapping all of $\boldsymbol{A}$ into itself and satisfying the identity

$$P(fPg) + P(gPf) = PfPg - P(fg). \tag{1}$$

The pair $(\boldsymbol{A}, P)$ will be called a *Baxter algebra.*

Let $\boldsymbol{O}$ be the category whose objects $(\boldsymbol{A}, T)$ are rings $\boldsymbol{A}$ together with operators $T: \boldsymbol{A} \rightarrow \boldsymbol{A}$ s.t. $T(f+g) = Tf + Tg$, and whose maps

$h: (\boldsymbol{A}, T) \to (\boldsymbol{A}', T')$ are ring endomorphisms $h: \boldsymbol{A} \to \boldsymbol{A}'$ such that $h(Tf) = T'h(f)$. (We sometimes use the same letter for operators on different rings.) The full subcategory of $\boldsymbol{O}$ whose objects are Baxter algebras will be denoted by $\boldsymbol{B}$. By standard results in universal algebra (cf. Cohn), both $\boldsymbol{O}$ and $\boldsymbol{B}$ contain free algebras on any nonempty set of generators. Recall that a (Baxter) algebra is *free* on a nonempty set S of generators if and only if any function $f: S \to \boldsymbol{A}$ extends uniquely to a morphism $\bar{f}: (\boldsymbol{A}(S), P) \to (\boldsymbol{A}, P)$ of the free (Baxter) algebra $(\boldsymbol{A}(S), P)$ on the set S. (Note: $\boldsymbol{A}(S)$ is not a polynomial ring, and it does not necessarily have an identity.)

3. **Main result.** Let F be a field of characteristic zero, having an infinite degree of transcendency over the rationals, and let $\boldsymbol{R}$ be the ring of all infinite sequences $u = (u_1, u_2, u_3, \cdots)$ with entries $u_n \in F$. Multiplication in $\boldsymbol{R}$ is defined componentwise, $uv = (u_1v_1, u_2v_2, u_3v_3, \cdots)$, addition as usual. For $k = 1, 2, \cdots, n$, choose an n-tuply infinite array $\{x_i^k: 1 \leqq k \leqq n,\ 1 \leqq i \leqq \infty\}$ of independent transcendentals in F, and set $x^k = (x_1^k, x_2^k, x_3^k, \cdots)$.

Call x^k a *free sequence.*

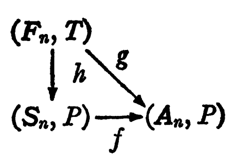
Define the operator P in $\boldsymbol{R}$ by

$$(2) \qquad P: (u_1, u_2, u_3, \cdots) \to (0, u_1, u_1 + u_2, u_1 + u_2 + u_3, \cdots).$$

It was remarked by Baxter—and it is easily verified—that P, thus defined, is a Baxter operator. Let $\boldsymbol{S}_n$ be the Baxter subalgebra of $(\boldsymbol{R}, P)$ generated by the free sequences $x^1, x^2, \cdots, x^n$, that is, the smallest subring of $\boldsymbol{R}$ containing the x^k and closed under the action of P; the object $(\boldsymbol{S}_n, P)$ will be called the *standard Baxter algebra* on n generators. We can now state our main result:

THEOREM. *The standard Baxter algebra* $(\boldsymbol{S}_n, P)$ *on n generators* $(n > 0)$ *is free.*

MAIN STEPS ON PROOF. (1) Let $f_1, f_2, \cdots, f_n$ be the generators of a free Baxter algebra $(\boldsymbol{A}_n, P)$, let $t_1, t_2, \cdots, t_n$ be the generators of a free algebra $(\boldsymbol{F}_n, T)$ in the category $\boldsymbol{O}$. Then the map $g: t_i \to f_i$, $1 \leqq i \leqq n$, extends uniquely to a morphism, again denoted by g, of $(\boldsymbol{F}_n, T)$ to $(\boldsymbol{A}_n, P)$, in the category $\boldsymbol{O}$. It must be shown that the morphism g factors through $(\boldsymbol{S}_n, P)$, as in the diagram below.



where $h: t_i \to x^i$ and $f: x^i \to f_i$, and where h and f are morphisms in $\boldsymbol{O}$.

It then follows that f is a morphism in $\boldsymbol{B}$. The morphism h is uniquely determined by the requirement $h: t_i \to x^i$; by one of the Noether theorems (cf. Cohn) all one has to show is that whenever $h(p)=0$ for some $p \in \boldsymbol{F}_n$, then $g(p)=0$.

(2) The ring $\boldsymbol{F}_n$ is a polynomial ring (without identity) in infinitely many generators; if u is a generator and $u \neq t_i$, then u can be uniquely written in the form $u=Tv$, where v is a product of generators, or *monomial*. From this, it follows that we can associate to every monomial m in $\boldsymbol{F}_n$ an integer $\operatorname{occ}(m)$, counting the number of occurrences of T in the (unique) expression of m in terms of T and the t_i. An element $p \in \boldsymbol{F}_n$ is a unique sum of monomials; set $\operatorname{occ}(p)$ to be the maximum of $\operatorname{occ}(m)$, as m ranges over all monomial summands of p. The proof is by induction over the integer $\operatorname{occ}(p)$.

If m and m' are monomials, $\operatorname{occ}(mm')=\operatorname{occ}(m)+\operatorname{occ}(m')$ and $\operatorname{occ}(Tm)=\operatorname{occ}(m)+1$.

(3) Every monomial $m \in \boldsymbol{F}_n$ differs by an element of the kernel of g from a sum of terms, each of which is of one of the forms: aTb, a, Tb, where a and b are monomials and $\operatorname{occ}(a)=0$, and $\operatorname{occ}(b)<\operatorname{occ}(m)$. Indeed, if m is not of any of these forms, then the expression of m as a product of generators contains at least two terms Tc and Td, where c and d are monomials. By identity (1), $TcTd$ differs from $T(cTd+dTc+cd)$ by an element of the kernel of g. Clearly $\operatorname{occ}(cTd+dTc+cd)<\operatorname{occ}(TcTd)$. Induction completes the proof.

(4) Hence, every $p \in \boldsymbol{F}_n$ is of the form $p=q+r$, where $g(r)=0$ and q is a sum of monomials, each one of which is of one of the forms listed in (3). To complete the proof, it must be shown that if $h(q)=0$, then $q=0$. Now, if $\operatorname{occ}(z)>0$ for any $z \in \boldsymbol{F}_n$, then the first entry of the sequence $h(z)$ is 0. Hence, if w is the sum of all monomials of q not containing any occurrence of T, then $h(w)=0$. But w is a polynomial in the t_i, hence $h(w)$ is a polynomial in the x^i; hence $w=0$. Thus q is a sum of terms of the form aTb and Tb.

(5) Let y be the sum of all summands of q of the form aTb, where $\operatorname{occ}(a)=0$, and $a\neq 0$. The second entry of the sequence $h(a)$ is a polynomial in x_2^i, $1 \leqq i \leqq n$; the second entry of $h(Tb)=Ph(b)$ is a polynomial in the x_1^i, $1 \leqq i \leqq n$. It follows that $h(y)=0$ only if $y=\sum_j aTb_j$, and $h(\sum_j Tb_j)=0$.

(6) It is easily verified by induction that if $Pv=0$ for $v \in \boldsymbol{S}_n$, then $v=0$.

(7) It follows from (4) and (5) that if $h(q)=0$, and $q\neq 0$, then there exist two terms n and s in $\boldsymbol{F}_n$, such that $h(Tn)=h(Ts)=0$, and $\operatorname{occ}(n)<\operatorname{occ}(q)$, $\operatorname{occ}(s)<\operatorname{occ}(q)$. But $h(Tn)=Ph(n)=0$. By (6), we infer $h(n)=0$. Similarly, $h(s)=0$. Since $\operatorname{occ}(n)<\operatorname{occ}(q)$, the induction hypothesis gives $n=0$, and the proof is complete.

The preceding Theorem states that every identity in an arbitrary Baxter algebra is true, whenever it has been verified in the standard Baxter algebra, where it reduces to an identity between symmetric functions.

4. An example. Consider the algebra (S_1, P) generated by $x=(x_1, x_2, x_3, \cdots)$. Easy computations (cf. Baxter) show that the $(k+1)$th entry of the vector $s_n=P(x^n)$ (x^n is the nth power of x) is the symmetric function $x_1^n+x_2^n+\cdots+x_k^n$. Similarly, the kth entry of the vector $a_n=P(xP(\cdots(Px)\cdots))$ with n occurrences of P, is the nth elementary symmetric function a_n in the variables $x_1, \cdots, x_k$. Waring's formula states

$$\sum_{n\geq 0} a_n(-t)^n = \exp\left(-\sum_{i\geq 1}\frac{t^i}{i}s_i\right), \tag{3}$$

and is valid in the algebra of formal power series in the variable t with coefficients in S_1. By the Main Theorem, the same formula is valid in an arbitrary Baxter algebra, giving the Baxter-Spitzer formula:

$$\sum_{n\geq 0}(-t)^n P(xP(xP\cdots(Px)\cdots)) = \exp(-P\log(1-tx)^{-1}). \tag{4}$$

Further identities, as well as applications to the Baxter algebras arising in probability and combinatorial theory, will be given in the second and third parts.

Bibliography

E. Sparre Andersen

On the fluctuations of sums of random variables, Math. Scand 1 (1953), 263–285.

On the fluctuations of sums of random variables. II, Math. Scand. 2 (1954), 195–223.

F. V. Atkinson

Some aspects of Baxter's functional equation, J. Math. Anal. Appl. 7 (1963), 1–30.

G. Baxter

An analytic problem whose solution follows from a simple algebraic identity, Pacific J. Math. 10 (1960), 731–742.

An operator identity, Pacific J. Math. 8 (1958), 649–663.

Combinatorial methods in fluctuation theory, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete. 1 (1963), 263–270.

D. Blackwell

Extension of a renewal theorem, Pacific J. Math. 3 (1960), 315–320.

P. M. Cohn

Universal algebra, Harper and Row, New York, 1965.

Ch. Hobby and R. Pyke

Combinatorial results in fluctuation theory, Ann. Math. Statist. 34 (1963), 1232–1242.

J. P. IMHOF

On ladder indices and random walk, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete. **9** (1967), 10–15.

J. KIEFER AND J. WOLFOWITZ

On the characteristics of the general queueing process with applications to random walk, Ann. Math. Statist. **27** (1956), 147–161.

J. F. C. KINGMAN

Spitzer's identity and its use in probability theory, J. London Math. Soc. **37** (1962), 309–316.

On the algebra of queues, Methuen's Review Series in Applied Probability, vol. 6, 1966, p. 1–44.

F. POLLACZEK

Théorie analytique des problèmes stochastiques relatifs à un groupe de lignes téléphoniques avec dispostif d'attente, Memor. Sci. Math. **150** (1961).

SIDNEY C. PORT

An elementary probability approach to fluctuation theory, J. Math. Anal. Appl. **6** (1963), 109–151.

M. P. SCHÜTZENBERGER

Certain infinite formal products and their combinatorial applications, Colloq. Combinatorial Methods in Probability Theory, (Aarhus), 1962, pp. 58–63.

S. SHERMAN

Fluctuation and periodicity, J. Math. Anal. Appl. **9** (1964), 231–235.

FRANK SPITZER

A combinatorial lemma and its application to probability theory, Trans. Amer. Math. Soc. **82** (1965), 323–339.

Principles of random walk, Van Nostrand, Princeton, N. J., 1964.

J. G. WENDEL

Brief proof of a theorem of Baxter, Math. Scand. **11** (1962), 107–108.

Order statistics of partial sums, Ann. Math. Statist. **31** (1960), 1031–1044.

Spitzer's formula; a short proof, Proc. Amer. Math. Soc. **9** (1958), 905–908.

D. FOATA

Thèse, Faculté de Sciences, Paris, 1964.

J. H. B. KEMPERMAN

Changes of sign in cummulative sums, Proceedings of the Netherlands Academy of Sciences **64** (1961), 291–313.

L. TAKACS

Combinatorial methods in the theory of stochastic processes, Wiley, New York, 1967.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139
THE ROCKEFELLER UNIVERSITY, NEW YORK, NEW YORK 10021
UNIVERSITY OF COLORADO, BOULDER, COLORADO 80302