(in Russian), 2nd ed., Moscow-Leningrad, 1950, p. 289). A similar remark applies to Theorem 7 (b), if we appeal to a theorem of Frostman and Nevanlinna, R. (cf. Nevanlinna, *Eindeutige analytische Funktionen*, Berlin, 1936, p. 198) instead of the Riesz uniqueness theorem.

[21] Plessner, *loc. cit.*, p. 224.

[22] Relative to $S(e^{i\theta})$.

ON A GENERALIZATION OF CLASSICAL PROBABILITY THEORY, I. MARKOFF CHAINS

BY RICHARD BELLMAN

THE RAND CORPORATION, SANTA MONICA, CALIF.

Communicated by H. S. Vandiver, July 27, 1953.

1. *Introduction.*—Classical probability, in its discrete form, deals in part with the region defined by the inequalities

$$x_i \geq 0, \quad \sum_{i=1}^{N} x_i = 1 \tag{1}$$

and with the transformations which preserve this region, the Markoff transformations

$$x_i' = \sum_{j=1}^{N} a_{ij} x_j, \qquad i = 1, 2, \ldots, N, \tag{2}$$

whose coefficients are subjected to the conditions

$$a_{ij} \geq 0, \quad \sum_{i=1}^{N} a_{ij} = 1, \qquad j = 1, 2, \ldots, N. \tag{3}$$

In recent years, a non-classical theory has proved of value in quantum mechanics. This is a theory based upon the transformations that preserve the quadratic form, $\sum_{i=1}^{N} x_i \bar{x}_i$, which is to say the orthogonal transformations.

The purpose of this paper is to present a further extension of the concept of a probability. Apart from the intrinsic mathematical interest of an investigation of this type, we have been principally motivated by the hope that allied concepts may play some future role in mathematical physics, by the expectation that use of the Chapman-Kolmogoroff equations, in a fashion similar to the classical theory, will uncover new classes of partial differential equations with interesting properties, properties which will be reflections of simple probabilistic relations, and finally by the belief that other classes of hypercomplex numbers may similarly be utilized to provide a closer link, of mutual advantage, between analysis and algebra.

Proofs of the results announced here and extensions in the directions indicated will be presented subsequently.

Let us observe that in the generalization contained herein we have retained time as a scalar variable. There are further generalizations in which the time variable itself is replaced by a hypercomplex number.

2. *Generalized Probabilities.*—Our basic quantities will be real symmetric matrices, which we shall denote by X_i. A finite set of matrices $X = \{X_i\}$, all of which are semi-definite and satisfy the further restriction

$$\sum_{i=1}^{N} \operatorname{tr} X_i = 1 \tag{4}$$

will be called a probability distribution.

A set of N^2M real matrices, $A = \{A_{ijk}\}$, $i, j = 1, 2, \ldots, N$, $k = 1, 2, \ldots, M$, will be called a transition matrix. A new set of matrices $Y = \{Y_i\}$ may be derived from the previous set by means of the linear transformation

$$Y_i = \sum_{k=1}^{M} \sum_{j=1}^{N} A_{ijk} X_j A_{ijk}^T, \qquad i = 1, 2, \ldots, N, \tag{5}$$

(A^T denotes the transpose of A).

As is well known, the matrix AXA^T is semidefinite whenever X is. Consequently this transformation preserves the semidefinite character of the set. The analog of the Markoff condition in (3) is

LEMMA. *The necessary and sufficient condition that* $\sum_{i=1}^{N} \operatorname{tr} Y_i = 1$ *for all X such that* $\sum_{i=1}^{N} \operatorname{tr} X_i = 1$ *is that*

$$\sum_{k=1}^{M} \sum_{i=1}^{N} A_{ijk} A_{ijk}^T = I, \qquad = 1, 2, \ldots, N \tag{6}$$

where I is the identity matrix.

It is easy to see that the product of two transformations of the type is again a transformation of this type.

3. *Generalized Markoff Chains.*—If we consider the state of a system at time n to be described by the set $X(n) = \{X_i(n)\}$, the generalized Markoff chain has the form

$$X_i(n+1) = \sum_{k=1}^{M} \sum_{j=1}^{N} A_{ijk} X_j A_{ijk}^T, \qquad i = 1, 2, \ldots, N, \tag{7}$$

where

$$\sum_{k=1}^{M} \sum_{i=1}^{N} A_{ijk} A_{ijk}^T = I.$$

As in the scalar theory, the question arises as to whether or not lim $X_i(n)$ exists as $n \to \infty$ and if so, the dependence of this quantity upon the initial state $X(0)$.

The following result holds.

THEOREM. *If A_{ijk} is positive definite for all i, j, and k, in general we will have* $\lim_{n \to \infty} X_i(n) = X_i$, *where X_i is independent of $X(0)$.*

By the phrase "in general" we mean apart from situations in which certain special relations exist among the A_{ijk} which will permit $\sum_{j, k} A_{ijk}X_jA^T_{ijk}$ to be semidefinite for semidefinite X_i.

There are several different proofs of this result, extensions of the proofs used in the scalar case. Each is a slight variation since the one-dimensional concept of positivity has to be modified in an appropriate fashion. An essential tool is the concept of the adjoint transformation represented by

$$Z_i = \sum_{k=1}^{M} \sum_{i=1}^{N} A_{ijk}X_iA^T_{ijk}.$$

SOME FUNCTIONAL EQUATIONS IN THE THEORY OF DYNAMIC PROGRAMMING

BY RICHARD BELLMAN

THE RAND CORPORATION, SANTA MONICA, CALIF.

Communicated by M. Morse, July 29, 1953

1. *Introduction.*—The purpose of this paper is to present some results concerning representative types of functional equations occurring in the theory of dynamic programming. Since considerations of space forbid our entering into any discussion of the origin and interpretation of all of the equations appearing below, we shall discuss one, the "gold-mining" equation, in some detail and refer to previous notes[2-4] and a forthcoming monograph, "An Introduction to the Theory of Dynamic Programming," soon to be published by The RAND Corporation, for further details.

2. *Functional Equations.*—In this section we state some results whose proof is contained in the monograph cited above.

THEOREM 1. (*The Gold-Mining Equation.*) *Consider the equation*

$$f(x, y, z) = \text{Max}\begin{bmatrix} \text{A:} & p_1f(0, y, x + z) + p_2f(c_2x, y, z + c_1x) + p_3\phi(z) \\ \text{B:} & q_1f(x, 0, y + z) + q_2f(x, d_2y, z + d_1y) + q_3\phi(z) \end{bmatrix}, \quad (1.1)$$

$x, y, z \geq 0$, with $f(0, 0, z) = \phi(z)$, and