THE SUPPORT OF DISCRETE EXTREMAL MEASURES WITH GIVEN MARGINALS

J. L. Denny

Let X and Y be countable sets. Let P be a probability on $X \times Y$ with marginals (projections) P_X on X and P_Y on Y. The set $M(P)$ of all probability measures on $X \times Y$ with marginals P_X and P_Y is convex. The purpose of this note is to characterize the supports of the extremal measures of $M(P)$, that is, the extreme points of $M(P)$. We show that each support of an extremal measure is the union of the graphs of two functions f and g, f mapping a subset of X into Y and g mapping a subset of Y into X, so that for each integer n the composition $(g \circ f)^n$ has no fixed points. Conversely, each probability P supported on the union of such graphs is an extremal measure of $M(P)$.

Now assume only that μ is a signed measure on $X \times Y$ and define μ_X, μ_Y, and $M(\mu)$ analogously. It is also true that each extremal $\nu \in M(\mu)$ assigns all its mass to the union of two such graphs. In this situation there may not be any extremal measures. However, even when extremal measures do exist the above converse is not true, for nonfinite positive μ.

As an application we characterize the extreme points of the set of Markov matrices with a given stationary invariant probability measure. When the Markov matrix is finite ($N \times N$) and doubly stochastic the extreme points are the permutation matrices. This result is known as the Birkhoff-von Neumann theorem (see, for example, p. 189 of [9]) and is a corollary to results in this paper. As pointed out by Letac [5], when P is a probability measure $M(P)$ is compact in the weak* topology and so the Choquet representation is applicable (see [2] or [8]). For other applications we refer to the survey paper [7] and for the case of non-discrete measures we mention [1] and [10]. This paper uses a method of Letac [5] which he employed to obtain another characterization of the supports of discrete measures.

The following is the usual definition of composition of functions. Some notation is needed. Let $A \subset X$ and $B \subset Y$ with at least one of A or B nonempty. Let $f: A \to Y$ and $g: B \to X$. Let $D_1 = A \cap f^{-1}(B)$. If D_1 is nonempty set $(g \circ f)^1(x) = g(f(x))$, $x \in D_1$. Assuming D_{n-1} and $(g \circ f)^{n-1}$ are defined set

$$D_n = D_{n-1} \cap ((g \circ f)^{n-1})^{-1}[D_1]$$

and $(g \circ f)^n(x) = (g \circ f)^1((g \circ f)^{n-1}(x))$, $x \in D_n$.

Definition. The pair of functions (f,g) is aperiodic if for each n, $x \in D_n$ implies $(g \circ f)^n(x) \neq x$.

THEOREM 1. *Let P be a probability on $X \times Y$ and let $Q \in M(P)$. The following two assertions are equivalent:*

Received May 1, 1978.

Michigan Math. J. 27 (1980).