The Annals of Mathematical Statistics
1972, Vol. 43, No. 1, 303-307

A VARIATIONAL CHARACTERIZATION OF FINITE MARKOV CHAINS

By Frank Spitzer
Cornell University

Summary. We prove a one-dimensional version of a theorem of O. E. Lanford and D. Ruelle [4] in equilibrium statistical mechanics which characterizes Gibbs states as those states which minimize the free energy. It is shown that a stationary Markov process is the most random (has the large stentropy) among all strictly stationary processes with the same energy. This energy is based on a potential U which determines the transition matrix M of the Markov chain with minimal free energy.

Let $\mathscr{T}$ be the class of *translation invariant states* (strictly stationary processes) defined as follows. The statement $\mu \in \mathscr{T}$ will mean that $(\Omega, \mathscr{F}, \mu)$ is a probability space, with $\Omega = \{1, 2, \cdots, n\}^{\mathbb{Z}}$, $\mathscr{F}$ is the σ-field generated by the cylinder sets of Ω, and μ is a countably additive probability measure on $(\Omega, \mathscr{F})$ which is invariant under the shift T which is defined by $(T\omega)_k = \omega_{k+1}$, $k \in \mathbb{Z}$. If I_N is the cylinder $\{1, 2, \cdots, n\}^{[1,2,\cdots,N]}$, then the cylinder set probabilities will be denoted

$$\mu_N(i) = \mu\{\omega : \omega_1 = i_1, \cdots, \omega_N = i_N\},\ i = (i_1, i_2, \cdots, i_N) \in I_N,\ N \geqq 1 .$$

For each $\mu \in \mathscr{T}$ and each $N \geqq 1$ we define the entropy (randomness in $[1, N]$) by

$$S_N(\mu) = -\sum_{i \in I_N} \mu_N(i) \log \mu_N(i) ,$$

with convention $0 \log 0 = 0$. Clearly $0 \leqq S_N(\mu) \leqq N \log N$, and it is known (the proof goes just as in 7.2.3. of [5] in the case $n = 2$) that the *specific entropy*

$$s(\mu) = \lim_{N\to\infty} N^{-1} S_N(\mu) , \qquad \mu \in \mathscr{T} \tag{1}$$

exists and is an *affine*, *upper semi-continuous* function on $\mathscr{T}$ (in the vague topology of $\mathscr{T}$ as a subset of the continuous positive linear functionals on $C(\Omega)$).

Now we introduce a nearest neighbor pair *potential* $U(i, j)$, $1 \leqq i, j \leqq n$, to be thought of as an energy of interaction between ω_k and ω_{k+1} when $\omega_k = i$, and $\omega_{k+1} = j$. Then the average energy in an interval $[1, N]$ is

$$E_N(\mu) = \sum_{i \in I_N} \mu_N(i) \sum_{k=1}^{N-1} U(i_k, i_{k+1}) .$$

Clearly the limit

$$e_U(\mu) = \lim_{N\to\infty} N^{-1} E_N(\mu) = \sum_{r=1}^{n} \sum_{s=1}^{n} \mu_2(r, s) U(r, s) \tag{2}$$

Received June 3, 1971.