Z. Wahrscheinlichkeitstheorie verw. Gebiete
68, 205–242 (1984)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

A Variational Characterization of One-Dimensional Countable State Gibbs Random Fields

B.M. Gurevich

Laboratory of Statistical Methods, Laboratory Building "A",
Moscow State University, Lenin Hills, Moscow W 117234, UdSSR

We suggest a formulation of a variational principle for lattice Markov fields in one dimension so as to include the case where the specific entropy and the specific energy are both infinite.

1. Basic Definitions and Main Results

1.1. The general definition of a Gibbs random field can be found in [1]. We shall deal only with one-dimensional lattice Gibbs fields with a finite or countable number of states and a nearest neighbour interaction. In this case the above mentioned definition can be stated in the following equivalent form. Let V be a finite or countable set, let $X=V^{\mathbb{Z}}=\{x=(\ldots,x_{-1},x_0,x_1,\ldots)\colon x_i\in V,\ i\in\mathbb{Z}\}$, be the sequence space and $\mathscr{F}$ the σ-algebra generated by the cylinder subsets of X. Let U be a function defined on $V\times V$ with values in $\mathbb{R}\cup\{+\infty\}$. A probability measure μ defined on $\mathscr{F}$ is said to be a Gibbs random field with potential U if for every $k, l\in\mathbb{Z}$, $v_i\in V$, $0\leqq i\leqq l$, such that $\mu(x_k=v_0,\ x_{k+l}=v_l)>0$ the following equality holds

$$\mu(x_{k+1}=v_1,\ldots,x_{k+l-1}=v_{l-1}/x_k=v_0,\ x_{k+l}=v_l)$$
$$=[1/\Xi_l(v_0,v_l)]\exp\left[-\sum_{i=0}^{l-1}U(v_i,v_{i+1})\right], \tag{1.1}$$

$1/\Xi_l(v_0,v_l)$ being the normalizing factor (we set $\exp(-\infty)=0$).

According to an idea going back to Gibbs one can specify the homogeneous Gibbs fields with a given potential within the family of all homogeneous random fields with states in V by means of a variational principle. In case of a finite V it reads as follows. Let us denote by S the shift transformation acting on X and by $\mathscr{I}$ the set of S-invariant probability measures on $\mathscr{F}$. For $\mu\in\mathscr{I}$ let $h_\mu=h_\mu(S)$ be the specific entropy of μ, i.e. the entropy of S with respect to μ, and let $e_\mu(U)=\int U(x_0,x_1)\,d\mu$ (in our case this is the specific energy of μ, see [2]). Obviously, $0\leqq h_\mu<+\infty$, $-\infty<e_\mu(U)\leqq+\infty$ and so the quantity

$$\mathscr{P}(U, \mu) = h_\mu - e_\mu(U), \qquad \mu \in \mathscr{I}, \tag{1.2}$$

takes values in $\mathbb{R} \cup \{-\infty\}$ ($-\mathscr{P}(U, \mu)$ is called the specific free energy of μ, see [2]). The variational principle reads: a measure $\mu \in \mathscr{I}$ is a Gibbs random field with potential U if an only if

$$\mathscr{P}(U, \mu) = \sup_{\nu \in \mathscr{I}} \mathscr{P}(U, \nu).$$

In such a form the variational principle was proved by Spitzer [13] (a more general result dealing with multi-dimensional Gibbs fields can be found in [7]). Spitzer also proved the uniqueness of a homogeneous Gibbs field (for a finite V such a field always exists) and found its explicit form. More recently Kesten [5] stated some necessary and sufficient conditions of the existence and proved the uniqueness of a homogeneous Gibbs field when V is countable and $U < \infty$. The purpose of the present paper[1] is to extend both the variational principle and Kesten's results to the case of a countable V under the weakest possible conditions on U. Our method differs from that of the articles mentioned. It is based on certain considerations induced by the theory of countable state Markov chains combined with some entropy ideas in ergodic theory (see [11] for all the entropy notions and results used below).

1.2. *Definition.* A potential U will be called *indecomposable* if given $v', v'' \in V$ there exist $v_1, \ldots, v_n \in V$ such that $U(v_i, v_{i+1}) < \infty$, $1 \leqq i \leqq n-1$, where $v_1 = v'$, $v_n = v''$.

Indecomposability is the only condition on U assumed satisfied throughout the paper.

When V is finite the results of [13] can be easily extended to an arbitrary indecomposable potential. However, when passing to an infinite V one finds that even for $U < \infty$ there can exist measures $\mu \in \mathscr{I}$ with $h_\mu = e_\mu(U) = +\infty$. That is why (1.2) does not allow a direct generalization, i.e. some regularization is required. It can be apparently made in different ways. One of these ways was suggested by Ito and Mori [4], another by Walters [15]. In both cases some additional restrictions are imposed on the potential. A fruitful approach to the variational principle was suggested by Föllmer [2] and further developed by Preston [10] (see also Pirlot [9] and Künsh [6]). This approach works when there exists a homogeneous Gibbs field with U. We shall compare it with our formulation of the variational principle in the end of this section.

1.3. Our aim is to define a functional $\mathscr{P}(U, \cdot)$ with values in $\mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ such that: (a) when V is finite, $\mathscr{P}(U, \cdot)$ is defined by (1.2); (b) the above stated variational principle is true.

To simplify some arguments below we shall define $\mathscr{P}(U, \mu)$ only on the set, call it $\mathscr{E}$, of those $\mu \in \mathscr{I}$ under which S is ergodic (if $\mu \in \mathscr{E}$ it is called an ergodic measure). This restriction is justified by the fact that when μ is non-ergodic, the

[1] Some of the results of this paper are published without proof in [3].

right side of (1.2) admits the integral representation corresponding to the ergodic decomposition of μ (see also Theorem D(ii) below).

In order to regularize (1.2) let us try to truncate both terms involved, then take their difference and finally remove the truncation.

For each $v\in V$ we set

$$\mathscr{E}(v)=\{\mu\in\mathscr{E}: \mu(x_0=v)>0\}.$$

Obviously, $\mathscr{E}=\bigcup_{v\in V}\mathscr{E}(v)$.

Denote by $\Gamma(v)$ the set of sequences $\gamma=(v_0, v_1, \ldots, v_l)$, $l=l(\gamma)$, such that $v_0=v_l=v$ and $v\neq v_i\in V$ for $1\leqq i\leqq l-1$. For $\gamma\in\Gamma(v)$ and for each $k\in\mathbb{Z}$ such that $0\leqq k\leqq l(\gamma)-1$ we set

$$A(v,\gamma,k)=\{x\in X: (x_{-k}, x_{-k+1}, \ldots, x_{-k+l(\gamma)})=\gamma\}. \tag{1.3}$$

It is clear that $A(v,\gamma,k)$ and $A(v,\gamma',k')$ intersect only when $\gamma=\gamma'$, $k=k'$. Moreover, the sets $A(v,\gamma,k)$, $\gamma\in\Gamma(v)$, $1\leqq k\leqq l(\gamma)-1$, form a countable partition, $\alpha(v)$, of the set

$$A(v)=\bigcup_{\gamma\in\Gamma(v)}\bigcup_{k=1}^{l(\gamma)-1} A(v,\gamma,k) \tag{1.4}$$

into measurable subsets. Introduce an arbitrary numbering on $\Gamma(v)$ and denote by $\alpha_n(v)$ the partition of $A(v)$ into the subsets $A(v,\gamma_i,k)$, $1\leqq i\leqq n$, $0\leqq k\leqq l(\gamma_i)-1$, and $A(v)\setminus A(v,n)$, where

$$A(v,n)=\bigcup_{i=1}^{n}\bigcup_{k=0}^{l(\gamma_i)-1} A(v,\gamma_i,k).$$

When $\mu\in\mathscr{E}(v)$ we have $\mu(A(v))=1$ and so $\alpha(v)$, $\alpha_n(v)$ can be viewed as partitions of the whole space $(X,\mathscr{F},\mu)$.

For every $v\in V$ and every $\mu\in\mathscr{E}(v)$ we set

$$\mathscr{P}(U,\mu,v)=\limsup_{n\to\infty}\,[h_\mu(S,\alpha_n(v))-\int_{A(v,n)} U(x_0,x_1)\,d\mu] \tag{1.5}$$

and we now define our functional $\mathscr{P}(U,\cdot)$ by

$$P(U,\mu)=\inf_{v:\,\mu\in\mathscr{E}(v)}\mathscr{P}(U,\mu,v). \tag{1.6}$$

Remark. One can in fact prove that $\mathscr{P}(U,\mu,v)$ does not depend on v on the set of those $v\in V$ for which $\mu\in\mathscr{E}(v)$. But we will not dwell upon this in the present paper.

1.4. **Proposition.** *If V is finite and $\mu\in\mathscr{E}$, then $P(U,\mu)$ defined as above satisfies* (1.2).

Proof. From the definition of $\alpha(v)$ it follows that if $\mu\in\mathscr{E}(v)$, then $\alpha(v)$ is a generator for S, so that ([11], Theorem 7.3)

$$h_\mu(S)=\lim_{n\to\infty} h_\mu(S,\alpha_n(v)).$$

Moreover, $h_\mu(S) \leqq \log \operatorname{card} V < \infty$. The function $U(x_0, x_1)$ is bounded below so that

$$\lim_{n\to\infty} \int_{A(v,n)} U(x_0, x_1)\,d\mu = \int_X U(x_0, x_1)\,d\mu = e_\mu(U),$$

where the limit equals $+\infty$ when $\mu(U(x_0, x_1) = +\infty) > 0$. The above implies that if V is finite then for every $v \in V$ such that $\mu \in \mathscr{E}(v)$ the content of the square brackets in (1.5) tends to $\mathscr{P}(U, \mu)$. This finishes the proof.

1.5. For precise formulation of the main results of this paper some notions and results related to non-negative matrices are required.

Definitions. Let Q be a non-negative function on $V \times V$. It can be clearly regarded as a matrix. We call it *indecomposable* if given $v, v' \in V$ there exist $v_i \in V$, $1 \leqq i \leqq n$, such that $v_1 = v$, $v_n = v'$, and $Q(v_i, v_{i+1}) > 0$ for $i = 1, \ldots, n-1$. We set

$$\lambda(Q) = \sup_{Q'} \lambda(Q'), \tag{1.7}$$

where the supremum is taken over all $Q' = Q|_{V' \times V'}$, the restrictions of Q to finite subsets $V' \subset V$; $\lambda(Q')$ being the maximal eigenvalue of Q'.

We shall call Q *admissible* if Q^n, the n-th iterate of Q, is finite for all $n \geqq 1$. The entries of Q^n will be denoted by $Q^{(n)}(\cdot, \cdot)$.

By a non-negative matrix Q one can construct a directed graph $G(Q)$ whose vertices are all $v \in V$ and whose edges are those ordered pairs $v, v' \in V$ for which $Q(v, v') > 0$. It is clear that Q is indecomposable if and only if $G(Q)$ is connected.

Let $\gamma = (v_0, v_1, \ldots, v_n)$ be a path of length n in $G(Q)$, i.e., a sequence of vertices such that v_i, v_{i+1} is an edge of $G(Q)$ for $i = 0, \ldots, n-1$. Let

$$Q(\gamma) = \prod_{i=0}^{n-1} Q(v_i, v_{i+1}). \tag{1.8}$$

A path $(v_0, \ldots, v_n)$ in $G(Q)$ will be called a *v-cycle*, $v \in V$, if $v_0 = v_n = v$ and $v_i \neq v$ when $1 \leqq i \leqq n-1$. Let $\Gamma(Q, v)$ denote the set of all v-cycles. Consider the series

$$\varphi_{Q,v}(t) = \sum_{\gamma \in \Gamma(Q,v)} Q(\gamma)\, t^{l(\gamma)}, \tag{1.9}$$

$l(\gamma)$ being the length of γ. If Q is admissible, (1.9) can be regarded as a power series because the sum of coefficients of t^n is finite for each $n \geqq 1$. We shall deal with (1.9) only for $t \geqq 0$.

Proposition. *Let Q be a non-negative matrix. Suppose Q is admissible and indecomposable. Then*

(i) *The radius of convergence of the power series $\sum_n Q^{(n)}(v, v') t^n$ does not depend on $v, v' \in V$ (we denote it by $R(Q)$).*

(ii) *If $\lim_{n\to\infty} [R(Q)]^n Q^{(n)}(v, v') = 0$ for some pair $v, v' \in V$ then this is true for all such pairs (otherwise Q is called $R(Q)$-positive).*

(iii) *The $R(Q)$-positivity of Q is equivalent to each of the following properties:* a) $\varphi_{Q,v}(R(Q))=1$ *and* $\frac{d}{dt}\varphi_{Q,v}(t)|_{t=R(Q)}<\infty$ *for some (and then for any)* $v\in V$; b) *there exist vectors* $\xi=\xi_Q\colon V\to R$ *and* $\eta=\eta_Q\colon V\to R$ *with positive components such that*

$$\sum_{v'\in V} Q(v,v')\xi(v')=\xi(v)/R(Q), \qquad \sum_{v'\in V}\eta(v')Q(v',v)=\eta(v)/R(Q), \qquad v\in V, \quad (1.10)$$

$$\sum_{v\in V}\xi(v)\eta(v)=1; \quad (1.11)$$

ξ, η *being uniquely defined to within a factor.*

(iv) $R(Q)=1/\lambda(Q)$.

Assertions (i)-(iii) can be found in [14], (iv) will be proved below (see Remark 3.3).

Let U be a potential and $Q_U=\exp(-U)$. The indecomposability of U is clearly the same as that of Q_U.

1.7. The main results of the paper are as follows.

Theorem A. *Let U be an indecomposable potential with $\lambda(Q_U)<\infty$ and let $v\in V$, $\mu\in\mathscr{E}(v)$. Then $\mathscr{P}(U,\mu,v)$ does not dedend on the numbering on $\Gamma(v)$. Moreover,* lim sup *can be replaced by* lim *in* (1.5).

Theorem B. *Let U be an indecomposable potential. Then*

(i) $\sup_{\mu\in\mathscr{E}(v)}\mathscr{P}(U,\mu,v)=\ln\lambda(Q_U)$ *for every* $v\in V$;

(ii) $\sup_{\mu\in\mathscr{E}}\mathscr{P}(U,\mu)=\ln\lambda(Q_U)$.

We shall call a measure $\mu\in\mathscr{E}$ *maximal* if $\mathscr{P}(U,\mu)=\ln\lambda(Q_U)$.

Theorem C. *Let U be an indecomposable potential with $\lambda(Q_U)<\infty$. Then*

(i) *Q_U is an admissible matrix;*

(ii) *If there exists a maximal measure $\mu\in\mathscr{E}$, then it is unique, belongs to $\mathscr{E}(v)$ for each $v\in V$, and is a Markov measure with stationary probabilities $\pi(v)$, $v\in V$, and transition probabilities $p(v,v')$, $v,v'\in V$, defined by*

$$\pi(v)=\xi(v)\eta(v), \qquad p(v,v')=Q(v,v')\xi(v')/\lambda(Q)\xi(v),$$

where $Q=Q_U$ and ξ,η are the vectors indicated in (1.10), (1.11).

(iii) *One can find a maximal measure $\mu\in\mathscr{E}$ if and only if Q_U is $1/\lambda(Q_U)$-positive.*

Theorem D. (i) *If for an indecomposable potential U with $\lambda(Q_U)<\infty$ there exists a maximal measure $\mu\in\mathscr{E}$, then μ is a homogeneous Gibbs field with potential U.*

(ii) *If there exists a homogeneous Gibbs field $\mu\in\mathscr{I}$ with an indecomposable potential U, then $\lambda(Q_U)<\infty$, $\mu\in\mathscr{E}$ and μ is a maximal measure for U.*

Theorem D contains the variational principle as stated in 1.1. Together with Theorem C it gives necessary and sufficient conditions for the existence and uniqueness of a homogeneous Gibbs field with a given potential. These conditions coincide with those of Kesten [6] stated for a finite potential.

Theorems A and B are proved in Sect. 2, Theorems C and D in Sects. 3 and 4 respectively. In Sect. 4 we also give an example of a potential for which neither formula (1.2) nor a natural way of its regularization [15] can be applied.

All the results of this paper can be automatically extended to a potential of any finite range.

1.8. We finish this Section with a short comparision between the approach of this paper and that of [2, 10]. The latter is based on the notion of specific information gain which is well known (in the one-dimensional case) in Information theory under the title of the entropy creation rate of one stationary process with respect to another such a process [8]. We recall its definition using the above notation.

Let ξ denote the partition of X into the sets $\{x \in X\colon x_0 = v\}$, $v \in V$, and let $\mu, \nu \in \mathscr{I}$. If there exists the limit

$$h(\mu, \nu) = \lim_{n \to \infty} \frac{1}{n} \sum_C \mu(C) \ln \frac{\mu(C)}{\nu(C)},$$

where the sum is over all the atoms of the partition $\bigvee_{i=1}^{n} S^{-i}\zeta$, then $h(\mu, \nu)$ is called the specific information gain of μ with respect to ν. Assume $\nu \in \mathscr{I}$ is a Gibbs field with potential U. The variational principle as stated in [2, 10] reads as follows: $\mu \in \mathscr{I}$ is also a Gibbs field with potential U if and only if $h(\mu, \nu) = 0$ (generally, $h(\mu, \nu) \geqq 0$).

The following theorem establishes the relation between $\mathscr{P}(U, \cdot)$ introduced above and the specific information gain.

Theorem E. *Let $\nu \in \mathscr{I}$ be a Gibbs field with potential U. Then for every $\mu \in \mathscr{E}$*

$$h(\mu, \nu) = \ln \lambda(Q_U) - \mathscr{P}(U, \mu). \tag{1.12}$$

Thus the right side of (1.12) can be regarded as a generalization of the specific information gain to the case where there are no homogeneous Gibbs fields with potential U. If, however, such a field does exist, the variational principles stated by means of $\mathscr{P}(U, \mu)$ and $h(\mu, \nu)$ are equivalent. But even in this case one could not immediately use the results of [2, 6, 9, 10] combined with Theorem E in order to prove Theorem D (ii) because these results (in spite of that they relate to a more general situation than our one) are obtained under some additional assumptions not necessarily satisfied in the case under consideration (in particular, U is assumed to be bounded in [10]). So we prefer a unified approach regardless of whether a Gibbs field with potential U exists or not.

One can prove Theorem E using the explicit form of the Gibbs field indicated in Theorem C. We shall not give this proof here.

2. Some Properties of $\mathscr{P}(U,\cdot)$

2.1. Let U be an indecomposable potential and let $Q=Q_U$. We denote by $X(Q)$ the set of all doubly infinite paths in the graph $G(Q)$, i.e., the set of sequences $x=(x_i)_{i\in\mathbb{Z}}\in X$ such that $Q(x_i,x_{i+1})>0$ for all $i\in\mathbb{Z}$. We fix an arbitrary $v\in V$ and denote by X_v the set of those $x\in X$ for which $x_i=v$ for infinitely many $i>0$ and for infinitely many $i<0$. The sets $X(Q)$, X_v and $X_v(Q)=X_v\cap X(Q)$ are measurable and shift invariant. Moreover, $X_v(Q)\subset A(v)$ (see (1.4)). Let

$$\mathscr{E}(Q,v)=\{\mu\in\mathscr{E}:\ \mu(X_v(Q))=1\}.$$

It immediately follows that $\mathscr{E}(Q,v)\subseteq\mathscr{E}(v)$.

2.2. **Proposition.** (i) *If $\mu\in\mathscr{E}(v)\setminus\mathscr{E}(Q,v)$, then $\mathscr{P}(U,\mu,v)=-\infty$.*

(ii) *There exists a $\mu\in\mathscr{E}(Q,v)$ such that $\mathscr{P}(U,\mu,v)>-\infty$.*

Proof. (i) Let $\mu\in\mathscr{E}(v)\setminus\mathscr{E}(Q,v)$. Since μ is ergodic and $X(Q)$ is S-invariant we get that either $\mu(X(Q))=0$, or $\mu(X(Q))=1$. In the former case there are $k\in\mathbb{Z}$ and $v',v''\in V$ such that $U(v',v'')=+\infty$ and $\mu(x_k=v',\ x_{k+1}=v'')=\mu(x_0=v',\ x_1=v'')>0$. Then for n sufficiently large the content of the square brackets in (1.5) turns into $-\infty$ and hence $\mathscr{P}(U,\mu,v)=-\infty$. In the latter case $\mu(X_v(Q))=1$ because due to the fact that $\mu\in\mathscr{E}(v)$ we have $\mu(X_v)=1$. Thus $\mu\in\mathscr{E}(Q,v)$ which contradicts the assumption.

(ii) Since Q is indecomposable, one can find a finite subset $V^0\subset V$ such that $v\in V^0$ and the restriction of Q to $V^0\times V^0$ is again an indecomposable matrix, say, Q^0. The set $X(Q^0)$ of all doubly infinite paths in $G(Q^0)$ is a subset of $X(Q)$. Moreover, every measure $\mu\in\mathscr{E}(v)$ concentrated on $X(Q^0)$ belongs to $\mathscr{E}(Q,v)$ because $U(x_0,x_1)$ is bounded on $X(Q^0)$. It is clear that such a measure does exist.

2.3. From 2.2 it follows that in order to discover both the supremum of $\mathscr{P}(U,\cdot,v)$ on $\mathscr{E}(v)$ and the set where this supremum is attained we can restrict ourselves to the subset $\mathscr{E}(Q,v)\subseteq\mathscr{E}(v)$.

We introduce new "coordinates" on $X_v(Q)$ and express $\mathscr{P}(U,\mu,v)$ by these coordinates. For an arbitrary $x=(x_i)_{i\in\mathbb{Z}}\in X_v(Q)$ we represent the set $\{n\in\mathbb{Z}:\ x_n=v\}$ in the form of an increasing sequence $(n_i)_{i\in\mathbb{Z}}$, where n_0 is the largest non-positive number in this set. Let $t(x)=-n_0$ and $y(x)=(y_i)_{i\in\mathbb{Z}}$, where $y_i=(x_{n_i},x_{n_i+1},\ldots,x_{n_{i+1}})$ (so y_i is a v-cycle). It follows that letting $x\mapsto(y(x),\,t(x))$ we obtain a mapping, Φ_v: $X_v(Q)\to Z_v$, where $Z_v=Z$ is the set of pairs (y,t), $y\in(\Gamma(Q,v))^{\mathbb{Z}}$, $0\leqq t\leqq l(y_0)-1$. Obviously Φ_v is one-to-one and it is measurable together with Φ_v^{-1} (Z is provided with a measurable structure being a subset of the product $Y\times\mathbb{Z}^+$, where $Y=(\Gamma(Q,v))^{\mathbb{Z}}$ and $\mathbb{Z}^+$ denotes the non-negative integers).

Let T' denote the shift transformation (by one step to the left) defined on Y and let

$$T(y,t)=\begin{cases}(y,t+1), & \text{when } 0\leqq t\leqq l(y_0)-2,\\ (T'y,0), & \text{when } t=l(y_0)-1,\end{cases}\qquad (y,t)\in Z.$$

So T is the integral (or, special) transformation determined by T' and the function $f: Y \to \mathbb{Z}^+$, where $f(y) = l(y_0)$, $y \in Y$.

From the definition of Φ_v and T it follows that

$$\Phi_v S x = T \Phi_v x, \qquad x \in X_v(Q), \tag{2.1}$$

i.e., Φ_v transfers S into T. Due to (2.1) Φ_v^* transfers each S-invariant measure on $X_v(Q)$ into a T-invariant measure on Z and transfers $\mathscr{E}(Q, v)$ into $\mathscr{E}(Z)$, the family of all T-invariant ergodic probability measures on Z.

Let $\mathscr{I}(Z)$ be the family of all T-invariant probability measures on Z. Normalizing the restriction of every $\nu \in \mathscr{I}(Z)$ to the set

$$Z'_v = Z' = \{(y, t) \in Z : t = 0\}$$

(it follows from the T-invariance of ν that $\nu(Z') > 0$) gives a probability measure ν' on Z' which in its turn uniquely determines ν. We call ν the lifting of ν'.

Taking into account that the v-cycles are numbered we introduce the following notation. Let $\gamma_n = (v_0^n, \ldots, v_k^n) \in \Gamma(Q, v)$. We denote

$$l(n) = l(\gamma_n), \qquad e(n) = \sum_{j=0}^{l(n)-1} U(v_j^n, v_{j+1}^n), \qquad n \geqq 1, \tag{2.2}$$

$$B^i(n) = \{(y, t) \in Z : y_0 = \gamma_n, t = i\}, \qquad n \geqq 1, \quad 0 \leqq i \leqq l(n) - 1, \tag{2.3}$$

$$p_n(\nu) = \nu'(B^0(n)), \qquad n \geqq 1. \tag{2.4}$$

Let β_n denote the partition of Z into the sets $B^i(k)$, $1 \leqq k \leqq n$, $0 \leqq i \leqq l(k) - 1$, and $B^c(n) = Z \setminus \bigcup B^i(k)$, where the union is over i, k indicated just now. Finally, let

$$L(\nu) = \sum_{n=1}^{\infty} p_n(\nu) l(n). \tag{2.5}$$

Obviously, $\sum_{n=1}^{\infty} l(n) \nu(B^0(n)) = 1$ which implies that

$$\nu(Z') L(\nu) = 1. \tag{2.6}$$

2.4. **Proposition.** *Let* $v \in V$, $\mu \in \mathscr{E}(Q, v)$, *and* $\nu = \Phi_v^* \mu$. *Then*

$$h_\mu(S, \alpha_n(v)) = h_\nu(T, \beta_n), \qquad n = 1, 2, \ldots, \tag{2.7}$$

$$\int_{A(v,n)} U(x_0, x_1) d\mu = (1/L(\nu)) \sum_{i=1}^{n} p_i(\nu) e(i), \qquad n = 1, 2, \ldots, \tag{2.8}$$

so that (*see* (1.5))

$$P(U, \mu, v) = \limsup_{n \to \infty} \left[h_\nu(T, \beta_n) - (1/L(\nu)) \sum_{i=1}^{n} p_i(\nu) e(i) \right]. \tag{2.9}$$

Proof. Due to the definition of Φ_v

$$\Phi_v(A(v, \gamma_j, i) \cap X_v(Q)) = B^i(j), \qquad 0 \leqq i \leqq l(j) - 1, \qquad j = 1, 2, \ldots$$

Hence Φ_v transfers $\alpha_n(v)$ into β_n, $n=1, 2, \ldots$, which gives (2.7).

From (1.3), (1.6), (2.6), and the fact that μ, v are S-invariant and T-invariant respectively we get

$$\begin{aligned}\int_{A(v,n)} U(x_0, x_1)\,d\mu &= \sum_{j=1}^{n} \sum_{i=0}^{l(j)-1} \int_{A(v,\gamma_j,i)} U(x_0, x_1)\,d\mu \\ &= \sum_{j=1}^{n} \sum_{i=0}^{l(j)-1} \int_{A(v,\gamma_j,0)} U(x_i, x_{i+1})\,d\mu = \sum_{j=1}^{n} \int_{A(v,\gamma_j,0)} \sum_{i=0}^{l(j)-1} U(x_i, x_{i+1})\,d\mu \\ &= \sum_{j=1}^{n} \mu(A(v,\gamma_j,0))e(j) = \sum_{j=1}^{n} v(B^0(j))e(j) = v(Z') \sum_{j=1}^{n} p_j(v)e(j) \\ &= (1/L(v)) \sum_{j=1}^{n} p_j(v)e(j),\end{aligned}$$

i.e., (2.8) holds, Q.E.D.

2.5 Due to the definition of Z' and Y these two sets can be identified. So we can regard T' as acting on Z'. Let β'_n denote the partition of Z' into the subsets $B^0(k)$, $1 \leqq k \leqq n$, and $B'^c(n) = Z' \setminus B'(n)$, where $B'(n) = \bigcup_{k=1}^{n} B^0(k)$.

Proposition. *If* $v \in \mathscr{I}(Z)$, *then*

$$H_v(T\beta_n/\beta_n) = (1/L(v))H_{v'}(T'\beta'_n/\beta'_n) + \varepsilon_n(v), \tag{2.10}$$

where $|\varepsilon_n(v)| \leqq \rho_n(p(v))$, $p(v) = (p_i(v))_{i=1,2,\ldots}$, *and*

$$\lim_{n\to\infty} \rho_n(p(v)) = 0 \tag{2.11}$$

($H_v(\cdot/\cdot)$ *being the conditional entropy*).

Proof. For short we denote:

$$L = L(v), \quad p_n = p_n(v), \quad q_{nk} = v(T'B^0(k)/B^0(n)) \quad (\text{when } v(B^0(n) > 0)),$$
$$L^{(n)} = \sum_{i \geqq n+1} p_n(v)\,l(n), \quad p^{(n)} = \sum_{i \geqq n+1} p_n(v), \quad n, k \geqq 1. \tag{2.12}$$

We first evaluate $H_v(T\beta_n/\beta_n)$. By definition

$$H_v(T\beta_n/B^j(k)) = 0, \quad j > 0, \tag{2.13}$$

$$v(TB^i(k)/B^0(m)) = \delta(i+1, l(k))\,v(T'B^0(k)/B^0(m)) = \delta(i+1, l(k))q_{mk}, \quad 0 \leqq i \leqq l(k)-1, \tag{2.14}$$

$$v(TB^c(n)/B^0(m)) = \sum_{k \geqq n+1} v(T'B^0(k)/B^0(m)) = \sum_{k \geqq n+1} q_{mk}, \tag{2.15}$$

$$\begin{aligned} v(B^i(k)/B^c(n)) &= \delta(i+1, l(k))\,v(T'B^0(k)/B^c(n)) \\ &= \delta(i+1, l(k)) \sum_{m \geqq n+1} v(T'B^0(k)/B^0(m))\,v(B^0(m)/B^c(n)) \\ &= \delta(i+1, l(k)) \sum_{m \geqq n+1} q_{mk}\,v(B^0(m))/v(B^c(n)) \\ &= \delta(i+1, l(k)) \sum_{m \geqq n+1} q_{mk}p_m/L^{(n)}, \quad 0 \leqq i \leqq l(k)-1, \end{aligned} \tag{2.16}$$

where $\delta(i,j)$ is Kronecker's symbol. (We consider only those n, m, k for which the corresponding conditional probability makes sense.) From (2.13)–(2.16) we obtain

$$
\begin{aligned}
H_\nu(T\beta_n/\beta_n) &= \nu(B^c(n))H_\nu(T\beta_n/B^c(n)) + \sum_{m=1}^{n} \nu(B^0(m))H_\nu(T\beta_n/B^0(m)) \\
&= (-L^{(n)}/L)\nu(TB^c(n))\ln \nu(TB^c(n)/B^c(n)) - (1/L)\sum_{k=1}^{n}\sum_{m\geqq n+1} p_m q_{mk} \\
&\quad \cdot \ln \sum_{m\geqq n+1} p_m q_{mk} + (1/L)(\ln L^{(n)})\sum_{k=1}^{n}\sum_{m\geqq n+1} p_m q_{mk} \\
&\quad - (1/L)\left[\sum_{m=1}^{n} p_m \sum_{i\geqq n+1} q_{mi}\ln \sum_{i\geqq n+1} q_{mi} + \sum_{k=1}^{n} q_{mk}\ln q_{mk}\right], \qquad (2.17)
\end{aligned}
$$

where each term including a conditional probability which makes no sense should be replaced by zero.

To evaluate $H_{\nu'}(T'\beta'_n/\beta'_n)$ let us remark that

$$
\begin{aligned}
\nu'(T'B^0(k)/B'^c(n)) &= \sum_{m\geqq n+1} \nu'(T'B^0(k)/B^0(m))\,\nu'(B^0(m))/\nu'(B'^c(n)) \\
&= \sum_{m\geqq n+1} p_m q_{mk}/p^{(n)}.
\end{aligned}
$$

From this we obtain

$$
\begin{aligned}
H_{\nu'}(T'\beta'_n/\beta'_n) &= -p^{(n)}\Big[\nu'(T'B'^c(n)/B'^c(n))\ln \nu'(T'B'^c(n)/B'^c(n)) \\
&\quad - \sum_{k=1}^{n} \nu'(T'B^0(k)/B'^c(n))\ln \nu'(T'B^0(k)/B'^c(n))\Big] \\
&\quad + \sum_{m=1}^{n} \nu'(B^0(m))\Big[-\nu'(T'B'^c(n)/B^0(m))\ln \nu'(T'B'^c(n)/B^0(m)) \\
&\quad - \sum_{k=1}^{n} \nu'(T'B^0(k)/B^0(m))\ln \nu'(T'B^0(k)/B^0(m))\Big] \\
&= -p^{(n)}\nu'(T'B'^c(n)/B'^c(n))\ln \nu'(T'B'^c(n)/B'^c(n)) \\
&\quad - \sum_{k=1}^{n}\sum_{m\geqq n+1} p_m q_{mk}\ln \sum_{m\geqq n+1} p_m q_{mk} + \ln p^{(n)}\sum_{k=1}^{n}\sum_{m\geqq n+1} p_m q_{mk} \\
&\quad - \sum_{m=1}^{n} p_m\left[\sum_{k\geqq n+1} q_{mk}\ln \sum_{k\geqq n+1} q_{mk} + \sum_{k=1}^{n} q_{mk}\ln q_{mk}\right]. \qquad (2.18)
\end{aligned}
$$

Due to (2.17), (2.18)

$$H_\nu(\beta_n/\beta_n) = (1/L)H_{\nu'}(T'\beta'_n/\beta'_n) + \varepsilon_n(\nu),$$

where

$$
\begin{aligned}
\varepsilon_n(\nu) &= (p^{(n)}/L)\nu'(T'B'^c(n))\ln \nu'(T'B'^c(n)/B'^c(n)) \\
&\quad - (L^{(n)}/L)\nu(TB^c(n)/B^c(n))\ln \nu(TB^c(n)/B^c(n)) \\
&\quad + (1/L)(\ln p^{(n)} + \ln L^{(n)})\sum_{k=1}^{n}\sum_{m\geqq n+1} p_m q_{mk}.
\end{aligned}
$$

Since

$$\sum_{k=1}^{n} \sum_{m \geqq n+1} p_m q_{mk} = \sum_{m \geqq n+1} p_m \sum_{k=1}^{n} q_{mk} \leqq \sum_{m \geqq n+1} p_m = p^{(n)},$$

$p^{(n)} \leqq L^{(n)} \leqq L (n \geqq 1)$, and $|u \ln u| \leqq 1/2$ $(0 \leqq u \leqq 1)$ it follows that

$$\begin{aligned} |\varepsilon_n(v)| &\leqq (1/2L)(p^{(n)} + L^{(n)}) + (1/L) p^{(n)} (|\ln p^{(n)}| + |\ln L^{(n)}|) \\ &\leqq (1/L)[L^{(n)} + p^{(n)}(2|\ln p^{(n)}| + |\ln L|)]. \end{aligned}$$

The last expression can be taken for $\rho_n(p(v))$, because it depends only on $p(v)$ and tends to zero as $n \to \infty$. This finishes the proof.

2.6. Let β' denote the partition of Z' into the sets $B^0(n)$, $n = 1, 2, \ldots$. A probability measure v' on Z' is called a Bernoulli measure (B-measure) if the partitions $(T')^i \beta'$, $i \in \mathbb{Z}$, are independent with respect to v'. Clearly, every B-measure v' is uniquely determined by the probability vector $p(v)$, where v is the lifting of v', and given a probability vector $p = (p_1, p_2, \ldots)$ there is a B-measure v' with $p(v) = p$.

Proposition. *Let $v, \tilde{v} \in \mathscr{I}(Z)$ be such that $\tilde{v}'$ is a B-measure and $p(v) = p(\tilde{v})$. Then*

$$H_{\tilde{v}}(T\beta_n/\beta_n) \geqq H_v(T\beta_n/\beta_n) - 2\rho_n(p(v)), \qquad n \geqq 1,$$

where ρ_n is indicated in Proposition 2.5.

Proof. By assumption the partitions $(T')^i \beta'$, $i \in \mathbb{Z}$, are independent with respect to $\tilde{v}'$. The same holds for $(T')^i \beta'_n$, $i \in \mathbb{Z}$, because β'_n is obtained from β' by joining the elements. From this fact and Proposition 2.5 it follows that

$$H_{\tilde{v}}(T\beta_n/\beta_n) = (1/L(\tilde{v})) H_{\tilde{v}'}(T'\beta'_n/\beta'_n) + \varepsilon_n(\tilde{v}) = (1/L(\tilde{v})) H_{\tilde{v}'}(T'\beta'_n) + \varepsilon_n(\tilde{v}).$$

As $p(v) = p(\tilde{v})$, we have

$$L(v) = L(\tilde{v}), \qquad H_{\tilde{v}'}(T'\beta'_n) = H_{\tilde{v}'}(\beta'_n) = H_{v'}(\beta'_n) = H_{v'}(T'\beta'_n).$$

Again using Proposition 2.5 we get

$$\begin{aligned} H_{\tilde{v}}(T\beta_n/\beta_n) &= (1/L(v)) H_{v'}(T'\beta'_n) + \varepsilon_n(\tilde{v}) \geqq (1/L(v)) H_{v'}(T'\beta'_n/\beta'_n) + \varepsilon_n(\tilde{v}) \\ &= H_v(T\beta_n/\beta_n) - \varepsilon_n(v) + \varepsilon_n(\tilde{v}) \geqq H_v(T\beta_n/\beta_n) - 2\rho_n(p(v)), \quad \text{Q.E.D.} \end{aligned}$$

2.7. **Proposition.** *Let $v \in \mathscr{I}(Z)$ be such that v' is a B-measure. Then*

$$H_v(T\beta_n/\beta_n) = -(1/L(v)) \sum_{i=1}^{n} p_i(v) \ln p_i(v) + \rho_n^{(1)}(p(v)), \qquad n \geqq 1,$$

where $\lim_{n \to \infty} \rho_n^{(1)}(p(v)) = 0$.

Proof. From (2.10) and the B-property of v' it follows that

$$H_\nu(T\beta_n/\beta_n)=(1/L(\nu))H_{\nu'}(T'\beta'_n/\beta'_n)+\varepsilon_n(\nu)=(1/L(\nu))H_{\nu'}(T'\beta'_n)+\varepsilon_n(\nu)$$
$$=(1/L)\left[-p^{(n)}\ln p^{(n)}-\sum_{i=1}^n p_i\ln p_i\right]+\varepsilon_n(\nu)=(1/L)\sum_{i=1}^n p_i\ln p_i+\rho_n^{(1)},$$

where p_i, L, $p^{(n)}$ are defined by (2.12), $p=(p_1, p_2, \ldots)$, and

$$\rho_n^{(1)}=\varepsilon_n(\nu)-(1/L)p^{(n)}\ln p^{(n)}.$$

Obviously, $\rho_n^{(1)}$ depends only on $p(\nu)$ and goes to zero as $n\to\infty$, Q.E.D.

2.8. **Proposition.** *If* $\nu\in\mathscr{I}(Z)$, *then*

$$h_\nu(T,\beta_n)\leqq H_\nu(T\beta_n/\beta_n)\leqq -(1/L(\nu))\sum_{i=1}^n p_i(\nu)\ln p_i(\nu)+\rho_n^{(2)}(p(\nu)),\qquad n\leqq 1,$$

where $\rho_n^{(2)}(p(\nu))$ *goes to zero as* $n\to\infty$.

Proof. Let $\tilde{\nu}$ denote the measure in $\mathscr{I}(Z)$ for which $\tilde{\nu}'$ is a B-measure and $p(\tilde{\nu})=p(\nu)$. Propositions 2.6 and 2.7 imply that

$$h_\nu(T,\beta_n)\leqq H_\nu(T\beta_n/\beta_n)\leqq H_{\tilde{\nu}}(T\beta_n/\beta_n)+2\rho_n(p(\nu))$$
$$-(1/L(\tilde{\nu}))\sum_{i=1}^n p_i(\tilde{\nu})\ln p_i(\tilde{\nu})+\rho_n^{(1)}(p(\tilde{\nu}))+2\rho_n(p(\nu))$$
$$=-(1/L(\nu))\sum_{i=1}^n p_i(\nu)\ln p_i(\nu)+\rho_n^{(1)}(p(\nu))+2\rho_n(p(\nu)).$$

It remains to set $\rho_n^{(2)}(p(\nu))=\rho_n^{(1)}(p(\nu))+2\rho_n(p(\nu))$ which finishes the proof.

2.9. For each $\nu\in\mathscr{I}(Z_v)$ we denote:

$$P^0(U,\nu,v)=\limsup_{n\to\infty}\left[h_\nu(T,\beta_n)-(1/L(\nu))\sum_{i=1}^n p_i(\nu)e(i)\right],\tag{2.19}$$

$$P^1(U,\nu,v)=\limsup_{n\to\infty}(1/L(\nu))\left[-\sum_{i=1}^n p_i(\nu)\ln p_i(\nu)-\sum_{i=1}^n p_i(\nu)e(i)\right].\tag{2.20}$$

Proposition. *For every* $\nu\in\mathscr{I}(Z_v)$

$$P^0(U,\nu,v)\leqq P^1(U,\nu,v).$$

Proof. The inequality claimed follows immediately from the definition of P^0, P^1 and Propositions 2.4, 2.8.

2.10. **Proposition.** *For every* $\nu\in\mathscr{I}(Z_v)$

$$P^1(U,\nu,v)\leqq\ln\lambda(Q).$$

Proof. We develop the proof by contradiction. Suppose there are $\nu\in\mathscr{I}(Z_v)$, $\varepsilon>0$, and an infinite set $\mathscr{N}\subset\mathbb{Z}^+$ such that

$$-(1/L(v))\sum_{i=1}^{n} p_i(v)(\ln p_i(v)+e(i))\geqq \ln\lambda(Q)+\varepsilon, \tag{2.21}$$

if $n\in\mathcal{N}$ (clearly, only the case $\lambda(Q)<\infty$ should be treated). For each $n\in\mathcal{N}$ we denote

$$p_i=p_i(v),\qquad i=1,2,\ldots,\qquad q_n=\sum_{i=1}^{n} p_i,\qquad p_i'=p_i/q_n,\qquad 1\leqq i\leqq n.$$

Let ν_n be the measure on Z_v for which $p(\nu_n)=(p_1',p_2',\ldots,p_n',0,0,\ldots)$ and ν_n' is a B-measure. Let $\mu_n=(\Phi_v^{-1})^*\nu_n$. One can find a finite subset $V_n\subset V$ such that $v\in V_n$ and μ_n is concentrated on the set of doubly-infinite paths in the graph $G(Q_n)$, where Q_n, the restriction of Q to $V_n\times V_n$, is an indecomposable matrix. Since V_n is finite, the potential $U_n=-\ln Q_n$ satisfies Theorem B (iii) so that

$$P(U_n,\mu_n)\leqq\ln\lambda(Q_n)\leqq\ln\lambda(Q). \tag{2.22}$$

On the other hand, using Proposition 2.4 and the explicit form of the measure $\nu_n=\Phi_v^*\mu_n$ we have

$$\begin{aligned}P(U_n,\mu_n)&=-\Big(\sum_{i=1}^{n} p_i'l(i)\Big)^{-1}\sum_{i=1}^{n} p_i'(\ln p_i'+e(i))\\&=-\Big(\sum_{i=1}^{n} p_il(i)\Big)^{-1}\sum_{i=1}^{n} p_i(\ln p_i+e(i))+\Big(\sum_{i=1}^{n} p_il(i)\Big)^{-1}q_n\ln q_n.\end{aligned} \tag{2.23}$$

As $n\to\infty$, the last term in (2.23) goes to zero. In combination with (2.21), (2.22) this implies that for $n\in\mathcal{N}$ the sum $\sum_{i=1}^{n} p_i(v)(\ln p_i(v)+e(i))$ is bounded. From this fact and the equality $L(v)=\sum_{i=1}^{\infty} p_il(i)$ it follows that (2.22) together with (2.23) contradict (2.21), Q.E.D.

2.11. **Proposition.** *If $\lambda(Q)<\infty$, then for each $\nu\in\mathcal{I}(Z_v)$ there are only two possibilities: either the series*

$$\sum_i p_i(\nu)(-\ln p_i(\nu)-e(i)) \tag{2.24}$$

converges absolutely, or its partial sum goes to $-\infty$ and the same holds for every series obtained from (2.24) *by permutation.*

Proof. The sum of the positive terms in (2.24) is finite, for otherwise there would be a permutation of terms making the sum of (2.24) equal to $+\infty$ which contradicts Proposition 2.10 (note that Proposition 2.10 does not depend on ordering).

Now move on to the series consisting of the negative terms in (2.24). If it converges, we have the former of the possibilities indicated above. If it diverges, we have the latter one. This finishes the proof.

2.12. **Proof of Theorem A.** Due to (2.20) and Proposition 2.4 it suffices to show that no change of the ordering on $\Gamma(Q,v)$ effects $P^0(U,\nu,v)$ and that the upper limit in (2.9) can be replaced by the limit.

If for $v \in \mathscr{E}(Z_v)$ the series (2.24) diverges, then by (2.19), (2.20) $P^1(U, v, v) = -\infty$ and the assertion claimed follows from Propositions 2.9 and 2.11. It remains to consider the case where (2.24) converges absolutely. In that case for an arbitrary permutation $n(1), n(2), \dots$ of the positive integers we have

$$\lim_{m\to\infty} (1/L(v)) \sum_{i=1}^{m} p_{n(i)}(v)[-\ln p_{n(i)}(v) - e(n(i))]$$
$$= \lim_{m\to\infty} (1/L(v)) \sum_{i=1}^{m} p_i(v)[-\ln p_i(v) - e(i)] = P^1(U, v, v). \tag{2.25}$$

Arrange the v-cycles in the following order: $\gamma_{n(1)}, \gamma_{n(2)}, \dots$. Let $\tilde{\beta}_n$ and $\tilde{\beta}'_n$ denote the partitions defined with respect to the new ordering in the same way as β_n and β'_n were defined with respect to the initial one, $n = 1, 2, \dots$. For every $n \geqq 1$ one can find a k_n such that each $i \leqq n$ coincides with $n(j)$ for some $j \leqq k_n$. For every $k \geqq k_n$ let ζ_k (respectively, ζ'_k) denote the partition of Z_v (respectively, Z'_v) whose elements are $B^j(n(i))$, $i \leqq k$, $n(i) > n$, $0 \leqq j \leqq l(n(i)) - 1$ (respectively, $B^0(n(i))$, $i \leqq k$, $n(i) > n$) and the complement $C(k)$ (respectively, $C'^c(k)$) of the union of these sets. By definition

$$\tilde{\beta}'_k = \beta'_n \vee \zeta'_k, \qquad \tilde{\beta}_k = \beta_k \vee \zeta_k, \qquad n \geqq 1, \ k \geqq k_n.$$

Let

$$\tilde{P}^0(U, v, v) = \limsup_{n\to\infty} \left[h_v(T, \tilde{\beta}_n) - \sum_{i=1}^{n} p_{n(i)}(v)\, e(n(i)) \right],$$

so that $\tilde{P}^0(U, v, v)$ is defined with respect to the new ordering just similarly as $P^0(U, v, v)$ was defined with respect to the initial one.

Due to (2.26) for every $n \geqq 1$ and every $k \geqq k_n$ we have the following:

$$\left[h_v(T, \beta_n) - (1/L(v)) \sum_{i=1}^{n} p_i(v)\, e(i) \right]$$
$$- \left[h_v(T, \tilde{\beta}_k) - (1/L(v)) \sum_{i=1}^{k} p_{n(i)}(v)\, e(n(i)) \right]$$
$$= h_v(T, \beta_n) - (1/L(v)) H_{v'}(\beta'_n) + (1/L(v)) H_{v'}(\beta'_n)$$
$$- (1/L(v)) \sum_{i=1}^{n} p_i(v)\, e(i) - h_v(T, \tilde{\beta}_k) + (1/L(v)) H_{v'}(\tilde{\beta}'_k)$$
$$- (1/L(v)) H_{v'}(\tilde{\beta}'_k) + (1/L(v)) \sum_{i=1}^{k} p_{n(i)}(v)\, e(n(i))$$
$$= [h_v(T, \beta_n) - h_v(T, \beta_n \vee \zeta_k)] + (1/L(v))[H_{v'}(\beta'_n \vee \zeta'_k) - H_{v'}(\beta'_n)]$$
$$+ (1/L(v)) \left[H_{v'}(\beta'_n) - \sum_{i=1}^{n} p_i(v)\, e(i) \right]$$
$$- (1/L(v)) \left[H_{v'}(\tilde{\beta}'_k) - \sum_{i=1}^{k} p_{n(i)}(v)\, e(n(i)) \right].$$

The content of the first brackets in the resulting expression can be transformed by the formula

$$h_\nu(T, \zeta \vee \vartheta) - h_\nu(T, \zeta) = H_\nu(\vartheta / T^{-1} \vartheta_T^- \vee \zeta_T),$$

where ζ, ϑ are finite measurable partitions and $\vartheta_T^- = \bigvee_{i=0}^{\infty} T^{-i}\vartheta$, $\zeta_T = \bigvee_{i=-\infty}^{\infty} T^i \zeta$ (see [7], 7.7). The content of the second brackets is $H_{\nu'}(\zeta_k'/\beta_n')$. Finally,

$$H_{\nu'}(\beta_n') = -\sum_{i=1}^{n} p_i(\nu) \ln p_i(\nu) - \left(1 - \sum_{i=1}^{n} p_i(\nu)\right) \ln \left(1 - \sum_{i=1}^{n} p_i(\nu)\right),$$

$$H_{\nu'}(\tilde{\beta}_k') = -\sum_{i=1}^{k} p_{n(i)}(\nu) \ln p_{n(i)}(\nu) + \left(1 - \sum_{i=1}^{k} p_{n(i)}(\nu)\right) \ln \left(1 - \sum_{i=1}^{k} p_{n(i)}(\nu)\right).$$

Summarizing we get

$$\begin{aligned}
&\left[h_\nu(T, \beta_n) - (1/L(\nu)) \sum_{i=1}^{n} p_i(\nu)\, e(i)\right] - \left[h_\nu(T, \tilde{\beta}_k) - (1/L(\nu)) \sum_{i=1}^{k} p_{n(i)}(\nu)\, e(n(i))\right] \\
&= -H_\nu(\zeta_k / T^{-1}(\zeta_k)_T^- \vee (\beta_n)_T) + (1/L(\nu))\, H_{\nu'}(\zeta_k'/\beta_n') + (1/L(\nu)) \\
&\quad \cdot \sum_{i=1}^{n} p_i(\nu)(-\ln p_i(\nu) - e(i)) - (1/L(\nu)) \sum_{i=1}^{k} p_{n(i)}(\nu)(-\ln p_{n(i)}(\nu) \\
&\quad - e(n(i)) + (1/L(\nu)) \left(1 - \sum_{i=1}^{n} p_i(\nu)\right) \ln \left(1 - \sum_{i=1}^{n} p_i(\nu)\right) \\
&\quad - (1/L(\nu)) \left(1 - \sum_{i=1}^{k} p_{n(i)}(\nu)\right) \ln \left(1 - \sum_{i=1}^{k} p_{n(i)}(\nu)\right) \\
&= (1/L(\nu))\, H_{\nu'}(\zeta_k'/\beta_n') - H_\nu(\zeta_k / T^{-1}(\zeta_k)_T^- \vee (\beta_n)_T) + \varepsilon(\nu, n, k),
\end{aligned} \tag{2.27}$$

where $\varepsilon(\nu, n, k)$ goes to zero uniformly in $k > k_n$ as $n \to \infty$ (see (2.25)).

Introduce an auxiliary partition ϑ_k with elements $B^j(n(i))$, $i \leqq k$, $n(i) > n$, $0 \leqq j < l(n(i)) - 1$, and $Z_v \setminus D_k$, where D_k is the union of $B^j(n(i))$ over all i, j indicated above. Every $B^j(n(i))$ just mentioned is an element of $T^{-1}\zeta_k$. So $Z_v \setminus D_k$ consists of entire elements of $T^{-1}\zeta_k$. Hence $T^{-1}\zeta_k \geqq \vartheta_k$, i.e. $T^{-1}\zeta_k$ refines ϑ_k. With this in mind and using well known properties of conditional entropy we have

$$H_\nu(\zeta_k / T^{-1}(\zeta_k)_T^- \vee (\beta_n)_T) \leqq H_\nu(\zeta_k / T^{-1}\zeta_k \vee \beta_n) \leqq H_\nu(\zeta_k / \vartheta_k \vee \beta_n). \tag{2.28}$$

To evaluate $H_\nu(\zeta_k / \vartheta_k \vee \beta_n)$ we note that ζ_k decomposes in a non-trivial way only one element of $\vartheta_k \vee \beta_n$, namely the element $D_k \cup E_k$, where $E_k = \bigcup_{i>k} \bigcup_{j=0}^{l(n(i))-1} B^j(n(i))$. By the definition of conditional entropy

$$\begin{aligned}
H_\nu(\zeta_k / \vartheta_k \vee \beta_n) = &- \nu(D_k \cup E_k) \frac{\nu(E_k)}{\nu(D_k \cup E_k)} \ln \frac{\nu(E_k)}{\nu(D_k \cup E_k)} \\
&- \nu(D_k \cup E_k) \sum_{i:\, i \leqq k,\, n(i) > n} \frac{\nu(B^{l(n(i))-1}(n(i))}{\nu(D_k \cup E_k)} \ln \frac{\nu(B^{l(n(i))-1}(n(i))}{\nu(D_k \cup E_k)}
\end{aligned}$$

$$= -\nu(E_k)\ln\nu(E_k)+\nu(E_k)\ln\nu(D_k\cup E_k)-\sum_{i:\,i\leq k,\,n(i)>n}\nu(B^0(n(i))\ln\nu(B^0(n(i))$$
$$+\nu(D_k)\ln\nu(D_k\cup E_k)= -\nu(E_k)\ln\nu(E_k)+\nu(D_k\cup E_k)\ln\nu(D_k\cup E_k)$$
$$-(1/L(\nu))\sum_{i:\,i\leq k,\,n(i)>n}p_{n(i)}(\nu)\ln p_{n(i)}(\nu)+\nu(D_k)\ln L(\nu)$$
$$= -(1/L(\nu))\sum_{i:\,i\leq k,\,n(i)>n}p_{n(i)}(\nu)\ln p_{n(i)}(\nu)+\varepsilon_1(\nu,n,k),$$

where $\varepsilon_1(\nu,n,k)\to 0$ uniformly in $k>k_n$ as $n\to\infty$ (here we as usually set $0\ln 0 = 0$).

Similarly, letting $D'_k = \bigcup_{i:\,i\leq k,\,n(i)>n} B^0(n(i))$ gives

$$H_{\nu'}(\zeta'_k/\beta'_n)= -\nu'(D'_k\cup E'_k)\frac{\nu'(E'_k)}{\nu'(D'_k\cup E'_k)}\ln\frac{\nu'(E'_k)}{\nu'(D'_k\cup E'_k)}$$
$$-\nu'(D'_k\cup E'_k)\sum_{i:\,i\leq k,\,n(i)>n}\frac{\nu'(B^0(n(i))}{\nu'(D'_k\cup E'_k)}\ln\frac{\nu'(B^0(n(i))}{\nu'(D'_k\cup E'_k)}$$
$$= -\nu'(E'_k)\ln\nu'(E'_k)+\nu'(E'_k)\ln\nu'(D'_k\cup E'_k)-\sum_{i:\,i\leq k,\,n(i)>n}p_{n(i)}(\nu)\ln p_{n(i)}(\nu)$$
$$+\nu'(D'_k)\ln\nu'(D'_k\cup E'_k)=\nu'(D'_k\cup E'_k)\ln\nu'(D'_k\cup E'_k)-\nu'(E'_k)\ln\nu'(E'_k)$$
$$-\sum_{i:\,i\leq k,\,n(i)>n}p_{n(i)}(\nu)\ln p_{n(i)}(\nu), \tag{2.29}$$

whence

$$(1/L(\nu))\,H_{\nu'}(\zeta'_k/\beta'_n)= -(1/L(\nu))\sum_{i:\,i\leq k,\,n(i)>n}p_{n(i)}(\nu)\ln p_{n(i)}(\nu)+\varepsilon_2(\nu,n,k),$$

where $\varepsilon_2(\nu,n,k)\to 0$ uniformly in $k>k_n$ as $n\to\infty$. Joining (2.27)–(2.29) we conclude that for $k>k_n$

$$\left[h_\nu(T,\beta_n)-(1/L(\nu))\sum_{i=1}^{n}p_i(\nu)\,e(i)\right]$$
$$-\left[h_\nu(T,\tilde\beta_n)-(1/L(\nu))\sum_{i=1}^{k}p_{n(i)}(\nu)\,e(n(i))\right]\geq\varepsilon_3(\nu,n,k),$$

where $\varepsilon_3(\nu,n,k)\to 0$ uniformly in $k>k_n$ as $n\to\infty$. This yields both assertions required.

2.13. **Proof of Theorem B.** (i) We shall first show that for any $v\in V$

$$\sup_{\mu\in\mathscr{E}(v)}P(U,\mu,v)\geq\ln\lambda(Q). \tag{2.30}$$

Due to the indecomposability of Q there exists an increasing sequence of finite subsets $V_n\subset V$ such that their union is V and the restriction of Q to $V_n\times V_n$ is an indecomposable matrix, say Q_n. Clearly $v\in V_n$ for n large enough. Let $X(Q_n)$ be the set of all doubly-infinite paths in the graph $G(Q_n)$. For every n

there is an ergodic shift-invariant probability measure μ^n on $X(Q_n)$ such that

$$h_{\mu^n}(S) - \int_{X(Q_n)} U(x_0, x_1)\, d\mu^n = \ln \lambda(Q_n).$$

The above measure is unique and has the property: $\mu^n(x_0 = \tilde{v}) > 0$ for each $\tilde{v} \in V_n$ (see [4]). Since $X(Q_n) \subset X(Q)$, μ^n can be regarded as a measure on $X(Q)$. With this in mind and arguing as in the proof of Proposition 1.4 we get

$$P(U, \mu^n, v) = \ln \lambda(Q_n),$$

which in view of the definition of $\lambda(Q)$ yields (2.30). Now (i) follows from (2.30) and Propositions 2.4, 2.9, 2.10.

(ii) For every $\mu \in \mathscr{E}$ there is a $v \in V$ such that $\mu \in \mathscr{E}(v)$. The definition of $P(U, \mu)$ and assertion (i) imply that

$$P(U, \mu) \leqq P(U, \mu, v) \leqq \ln \lambda(Q).$$

Hence

$$\sup_{\mu \in \mathscr{E}} P(U, \mu) \leqq \ln \lambda(Q).$$

In order to prove the converse inequality we find for an arbitrary $\varepsilon > 0$ a finite subset $V_\varepsilon \subset V$ such that Q_ε, the restriction of Q to $V_\varepsilon \times V_\varepsilon$, is an indecomposable matrix and $\ln \lambda(Q_\varepsilon) > \ln \lambda(Q) - \varepsilon$. As in (i) there is a probability measure μ_ε concentrated on $X(Q_\varepsilon)$ and such that $P(U, \mu_\varepsilon, \tilde{v}) = \ln \lambda(Q_\varepsilon)$ for every $\tilde{v} \in V_\varepsilon$. Moreover, μ_ε is a Markov measure and $\mu_\varepsilon(x_0 = \tilde{v}) > 0$ for every $\tilde{v} \in V_\varepsilon$. Therefore, $\{\tilde{v} \in V \colon \mu_\varepsilon \in \mathscr{E}(\tilde{v})\} = V_\varepsilon$ and hence

$$\mathscr{P}(U, \mu_\varepsilon) = \ln \lambda(Q_\varepsilon) > \ln \lambda(Q) - \varepsilon.$$

Since this is true for any $\varepsilon > 0$, we have

$$\sup_{\mu \in \mathscr{E}} \mathscr{P}(U, \mu) \geqq \ln \lambda(Q).$$

This finishes the proof.

3. The Structure of a Maximal Measure

3.1. We shall first establish some properties of the series in (1.9).

Proposition. *Assume V is finite and Q is indecomposable. Then for any $v \in V$*

$$\varphi_{Q,v}(1/\lambda(Q)) = 1.$$

This is actually well known [14] and can be, for example, derived from the fact that any indecomposable finite state Markov chain is recurrent. We now turn to an infinite V.

3.2. **Proposition.** *Let Q be an indecomposable admissible matrix and let $r(Q, v)$, $v \in V$, denote the radius of convergence of the series in* (1.9). *Then $1/\lambda(Q)$ is a* *unique positive solution of the equation* $\varphi_{Q,v}(t)=1$, *provided that* $\varphi_{Q,v}(r(Q,v))\geqq 1$, *and* $1/\lambda(Q)=r(Q,v)$ *otherwise.*

Proof. Take a sequence of finite subsets $V_n \subset V$ such that: a) $\bigcup_n V_n = V$, b) $V_{n+1} \supset V_n$, and c) Q_n, the restriction of Q to $V_n \times V_n$, is an indecomposable matrix, $n=1,2,\ldots$. Clearly, $\lambda(Q_n)\uparrow\lambda(Q)$ as $n\to\infty$. Let for short

$$\tau=1/\lambda(Q), \quad \tau_n=1/\lambda(Q_n), \quad \varphi=\varphi_{Q,v}, \quad r=r(Q,v), \quad \varphi_n=\varphi_{Q_n,v}$$

(in the last equality we assume that $v\in V_n$ which is the case when n is large enough).

We first assume that $\varphi(r)>1$. In this case $r>0$ and since φ strictly increases within the set $\mathbb{R}^+\cap D$, where D is the domain of definition of φ, it suffices to show that $\varphi(\tau)=1$. If $\varphi(\tau)>1$ (e.g., $\varphi(\tau)=+\infty$), there is a finite number of terms in (1.9) whose sum is greater than 1. By the definition of φ_n these terms are also terms of $\varphi_n(\tau)$ for n large enough. Since $\varphi_n(\tau)$ is a series with no negative terms we conclude that $\varphi_n(\tau)>1$ when n is large enough. By Proposition 3.1 $\varphi(\tau_n)=1$ and hence $\tau>\tau_n$ for n large enough. But this contradicts the inequality $\lambda(Q_n)\leqq\lambda(Q_{n+1})$, $n\geqq 1$.

If $\varphi(\tau)<1$ we have $\tau<r$ because $\varphi(r)>1$. Using the continuity of the sum of a power series we can find $\varepsilon>0$ such that $\tau+\varepsilon<r$, $\varphi(\tau+\varepsilon)<1$. On the other hand, $\tau_n<\tau+\varepsilon$ for n large enough and $\varphi(\tau+\varepsilon)\geqq\varphi_n(\tau+\varepsilon)$ for every n. From the fact that φ_n is monotonic within the positive semi-axis combined with Proposition 3.1 we conclude that when n is large enough, $\varphi(\tau+\varepsilon)\geqq\varphi_n(\tau+\varepsilon)\geqq\varphi_n(\tau_n)=1$, so that $\varphi(\tau+\varepsilon)\geqq 1$ which leads to a contradiction. Thus $\varphi(\tau)=1$ when $\varphi(r)>1$.

Now consider the case where $\varphi(r)\leqq 1$. Since $\varphi_n(r)\leqq\varphi(r)$ and φ_n is monotonic, Proposition 3.1 implies that $\tau_n\geqq r$ for $n\geqq 1$ and hence $\tau\geqq r$. Suppose $\tau>r$ and pick an arbitrary $\varepsilon>0$ such that $r+\varepsilon<\tau$. Clearly, $\varphi_n(r+\varepsilon)\leqq\varphi_n(\tau)\leqq\varphi_n(\tau_n)=1$ for every n. On the other hand, $\varphi(r+\varepsilon)=+\infty$ because r is the radius of convergence of (1.9). But each term in the series for $\varphi(r+\varepsilon)$ is also a term in the series for $\varphi_n(r+\varepsilon)$ when n is large enough which contradicts the above mentioned inequality $\varphi_n(r+\varepsilon)\leqq 1$. It remains to accept that $\tau=r$ when $\varphi(r)\leqq 1$.

3.3. *Remark.* From Proposition 3.2 and some results by Vere-Jones [14, Theorem C, Lemma 2.1] one can easily derive that if Q is an indecomposable admissible matrix, $\lambda(Q)=1/R(Q)$ as stated in Proposition 1.6(iv). But if Q is inadmissible, $R(Q)=0$ and $\lambda(Q)=+\infty$.

3.4. *Definition.* A measure $\mu\in\mathscr{E}(v)$ will be referred to as *v-maximal* if it maximizes $\mathscr{P}(U,\cdot,v)$ on $\mathscr{E}(v)$.

Theorem. *Let* $\lambda(Q_U)<\infty$ *and assume that for some* $v\in V$ *there exists a v-maximal measure* $\mu\in\mathscr{E}(v)$. *Then the measure* $\nu=\Phi_v^*\mu$ *is such that* ν' (see 2.3) *is a B-measure with*

$$p_n(\nu)=(\lambda(Q_U))^{-l(n)}\exp(-e(n)), \qquad n=1,2,\ldots \tag{3.1}$$

(the notation used here is introduced in Sects. 1, 2).

For the proof of this theorem we need some more auxiliary facts.

3.5. **Proposition.** *Under the conditions of Theorem* 3.4 *the following equalities hold:*

$$\lim_{n\to\infty}[H_\nu(T\beta_n/\beta_n)-h_\nu(T,\beta_n)]=0, \tag{3.2}$$

$$\lim_{n\to\infty}[H_{\nu'}(\beta'_n)-H_{\nu'}(\beta'_n/(T')^{-1}\beta'_n)]=0. \tag{3.3}$$

Proof. Propositions 2.9 and 2.10 imply that

$$\mathscr{P}^1(U,\nu,v)=\mathscr{P}^0(U,\nu,v)=\ln\lambda(Q_U). \tag{3.4}$$

By Proposition 2.8

$$h_\nu(T,\beta_n)-(1/L(\nu))\sum_{i=1}^n p_i(\nu)\,e(i)\leqq H_\nu(T\beta_n/\beta_n)-(1/L(\nu))\sum_{i=1}^n p_i(\nu)\,e(i)$$

$$-(1/L(\nu))\sum_{i=1}^n p_i(\nu)[\ln p_i(\nu)+e(i)]+\rho_n^{(2)}(p(\nu)), \tag{3.5}$$

where $\rho_n^{(2)}(p(\nu))$ goes to zero when $n\to\infty$. Using Proposition 2.11 we can direct n to infinity in (3.5). Due to (3.4) the limits of both sides are the same and finite. Hence

$$\lim_{n\to\infty}[H_\nu(T\beta_n/\beta_n)-h_\nu(T,\beta_n)]=0,$$

$$\lim_{n\to\infty}\Big[-(1/L(\nu))\sum_{i=1}^n p_i(\nu)\ln p_i(\nu)-H_\nu(T\beta_n/\beta_n)\Big]=0.$$

The first of these equalities is just (3.2). The second one combined with Proposition 2.5 yield

$$\lim_{n\to\infty}(1/L(\nu))\Big[-\sum_{i=1}^n p_i(\nu)\ln p_i(\nu)-H_{\nu'}(T'\beta'_n/\beta'_n)\Big]=0. \tag{3.6}$$

Since

$$\lim_{n\to\infty}\Big[H_{\nu'}(\beta'_n)+\sum_{i=1}^n p_i(\nu)\ln p_i(\nu)\Big]=0,$$

(3.6) implies (3.3). This finishes the proof.

3.6. **Lemma** (Smorodinsky [12]). *Given* $\varepsilon>0$ *there is* $\delta(\varepsilon)>0$ *such that if* $\zeta=(A_1,\dots,A_n)$ *and* $\vartheta=(B_1,\dots,B_m)$ *are finite measurable partitions of a probability space* (Ω,A,μ) *for which* $H_\mu(\zeta)-H_\mu(\zeta/\vartheta)<\delta(\varepsilon)$, *then the total measure of those* B_j *for which*

$$\sum_{i=1}^n|\mu(A_i/B_j)-\mu(A_i)|\geqq\varepsilon \tag{3.7}$$

does not exceed ε (in such a case ζ is called ε-independent of ϑ).

Those B_j which satisfy (3.7) will be referred to as irregular.

3.7. **Proposition.** *If a measure* ν' *on* Z' *satisfies* (3.3), *then the partitions* β' *and* $T'\beta'$ *are independent relative to* ν'.

Proof. We take Z', ν', β'_n, and $(T')^{-1}\beta'_n$ as Ω, μ, ζ, and ϑ, respectively, in Lemma 3.6. Let B be an arbitrary element of $(T')^{-1}\beta'$ with $\nu'(B)>0$ and let an arbitrary positive $\varepsilon<\nu'(B)$ be fixed. We can find an $n(B)$ such that if $n\geqq n(B)$, then B is also an element of $(T')^{-1}\beta'_n$. Using (3.3) we can find an n_ε such that if $n\geqq n_\varepsilon$, then

$$H_{\nu'}(\beta'_n)-H_{\nu'}(\beta'_n/(T')^{-1}\beta'_n)\leqq\delta(\varepsilon).$$

Suppose that $n\geqq\max(n(B),n_\varepsilon)$. By Lemma 3.6 β_n is ε-independent of $(T')^{-1}\beta'_n$, i.e., the total measure of the irregular elements of $(T')^{-1}\beta'_n$ does not exceed ε. Since $\nu'(B)>\varepsilon$, the element B must be regular. It follows that

$$|\nu'(A_n/B)-\nu'(A_n)|\leqq\varepsilon \tag{3.8}$$

for any element A_n of β'_n. But every element A of β' is also an element of β'_n for $n\geqq n(A)$. Thus, for arbitrary elements A and B of β' and $(T')^{-1}\beta'$, respectively, we have (3.8) as n is large enough. Since ε is arbitrary, it follows that $\nu'(A/B)=\nu'(A)$, as required.

3.8. **Proposition.** *If a measure $\nu\in\mathscr{I}(Z_v)$ satisfies* (3.2), *then $(T^i\beta)_{i\in\mathbb{Z}}$ is a Markov sequence of partitions, that is, given an integer $m\geqq 1$, an element C of $\bigvee_{i=1}^{m}T^{-i}\beta$ with $\nu(C)>0$, and elements A and B of $T\beta$ and β respectively* (*with* $\nu(B)>0$) *the following equality holds*

$$\nu(A/B\cap C)=\nu(A/B).$$

Proof. For each element B of β_n with $\nu(B)>0$ we define a measure ν^B on Z_v via $\nu^B(\cdot)=\nu(\cdot/B)$. Due to the properties of conditional entropy

$$H_\nu(T\beta_n/\beta_n)-H_\nu\left(T\beta_n/\beta_n\vee\bigvee_{i=1}^{m}T^{-i}\beta_n\right)$$
$$=\sum_{B\in\beta_n}\nu(B)\left[H_{\nu^B}(T\beta_n)-H_{\nu^B}\left(T\beta_n/\bigvee_{i=1}^{m}T^{-i}\beta_n\right)\right],\quad n\geqq 1, \tag{3.9}$$

where $B\in\beta_n$ means that B is an element of β_n and the sum is over all such B with $\nu(B)>0$.

Let A_1, B_1, C_1 be arbitrary elements of $T\beta$, β, and $\bigvee_{i=1}^{m}T^{-i}\beta$, respectively, where $\nu^{B_1}(C_1)>0$. Let ε be a positive number such that

$$\varepsilon<\nu^{B_1}(C_1), \tag{3.10}$$

$$\delta(\varepsilon)<\nu(B_1), \tag{3.11}$$

where $\delta(\varepsilon)$ is the constant from Lemma 3.6. There are $n(A_1, B_1, C_1)$ and n_ε such that if $n\geqq n(A_1,B_1,C_1)$, then A_1,B_1,C_1 are also elements of $T\beta_n$, β_n, $\bigvee_{i=1}^{m}T^{-i}\beta_n$, respectively, and if $n\geqq n_\varepsilon$, then (see (3.2))

$$H_\nu(T\beta_n/\beta_n)-h_\nu(T,\beta_n)\leqq(\delta(\varepsilon))^2. \tag{3.12}$$

Since $h_\nu(T,\beta_n) \leqq H_\nu(T\beta_n / \bigvee_{i=0}^{m} T^{-i}\beta_n)$ for every $m \geqq 0$, we see from (3.12) that

$$H_\nu(T\beta_n/\beta_n) - H_\nu\left(T\beta_n / \bigvee_{i=0}^{m} T^{-i}\beta_n\right) \leqq (\delta(\varepsilon))^2. \tag{3.13}$$

We now suppose that $n \geqq \max(n(A_1, B_1, C_1), n_\varepsilon)$. Due to (3.9), (3.13), and the Chebyshev inequality the total measure ν of those $B \in \beta$ for which

$$H_{\nu^B}(T\beta_n) - H_{\nu^B}\left(T\beta_n / \bigvee_{i=1}^{m} T^{-i}\beta_n\right) > \delta(\varepsilon)$$

does not exceed $\delta(\varepsilon)$. Due to (3.11) they don't include B_1 and hence

$$H_{\nu^{B_1}}(T\beta_n) - H_{\nu^{B_1}}\left(T\beta_n / \bigvee_{i=1}^{m} T^{-i}\beta_n\right) \leqq \delta(\varepsilon).$$

By Lemma 3.6 it follows that $T\beta_n$ is ε-independent of $\bigvee_{i=1}^{m} T^{-i}\beta_n$ relative to ν^{B_1}, i.e., the total measure ν^{B_1} of those $C \in \bigvee_{i=1}^{m} T^{-i}\beta_n$ for which

$$\sum_{A \in T\beta_n} |\nu^{B_1}(A/C) - \nu^{B_1}(A)| > \varepsilon$$

does not exceed ε. Due to (3.10) they don't include C_1 and hence

$$|\nu^{B_1}(A_1) - \nu^{B_1}(A_1/C_1)| \leqq \varepsilon.$$

This inequality holds for all $A_1 \in T\beta$, $B_1 \in \beta$, $C_1 \in \bigvee_{i=1}^{m} T^{-i}\beta$ such that $\nu^{B_1}(C_1) > 0$. It clearly implies the assertion claimed.

3.9. **Proof of Theorem 3.4.** We first prove that ν' is a B-measure. By a standard argument one can deduce from Proposition 3.8 that for any positive integers $m, k_1, k_2, \ldots, k_{m+1}$ such that $k_{i+1} > k_i$, $i = 1, 2, \ldots, m$, and for any elements A, B, C of $T^{k_{m+1}}\beta$, $T^{k_m}\beta$, and $\bigvee_{i=1}^{m} T^{k_i}\beta$, respectively, such that $\nu(B \cap C) > 0$ the following holds:

$$\nu(A/B \cap C) = \nu(A/B). \tag{3.14}$$

Recall that to each element $B^0(i)$ of β' there corresponds a v-cycle $\gamma_i \in \Gamma(Q, v)$ of length $l(i)$. By definition

$$T' B^0(i) = T^{l(i)} B^0(i) \tag{3.15}$$

which implies that for any $n \geqq 1$ and any $i_1, \ldots, i_n$

$$\begin{aligned} &B^0(i_n) \cap T'B^0(i_{n-1}) \cap \ldots \cap (T')^{n-1} B^0(i_1) \\ &= B^0(i_n) \cap T^{l(i_{n-1})} B^0(i_{n-1}) \cap \ldots \cap T^{l(i_{n-1}) + \ldots + l(i_1)} B^0(i_1). \end{aligned} \tag{3.16}$$

Indeed, for $n=2$ (3.16) is a direct consequence of (3.15). With this in mind (3.16) can be easily checked by induction for any n.

In order to prove that ν' is a B-measure it suffices to check that

$$\nu'((T')^m B^0(i_0)/(T')^{m-1}B^0(i_1)\cap\ldots\cap T'B^0(i_{m-1})\cap B^0(i_m))=\nu'(B^0(i_0)) \qquad (3.17)$$

for any $i_0,\ldots,i_m$ provided that the condition on the left side of (3.17) has positive measure. Using successively (3.16), (3.14), and (3.15) we have

$$\begin{aligned}
&\nu'((T')^m B^0(i_0)/(T')^{m-1}B^0(i_1)\cap\ldots\cap T'B^0(i_{m-1})\cap B^0(i_m))\\
&=\frac{\nu(B^0(i_m)\cap T'B^0(i_{m-1})\cap\ldots\cap(T')^m B^0(i_0))}{\nu(B^0(i_m)\cap T'B^0(i_{m-1})\cap\ldots\cap(T')^{m-1}B^0(i_1))}\\
&=\frac{\nu(B^0(i_m)\cap T^{l(i_{m-1})}B^0(i_{m-1})\cap\ldots\cap T^{l(i_{m-1})+\ldots+l(i_0)}B^0(i_0))}{\nu(B^0(i_m)\cap T^{l(i_{m-1})}B^0(i_{m-1})\cap\ldots\cap T^{l(i_{m-1})+\ldots+l(i_1)}B^0(i_1))}\\
&=\nu(T^{l(i_{m-1})+\ldots+l(i_0)}B^0(i_0)/B^0(i_m)\cap\ldots\cap T^{l(i_{m-1})+\ldots+l(i_1)}B^0(i_1))\\
&=\nu(T^{l(i_{m-1})+\ldots+l(i_0)}B^0(i_1)/T^{l(i_{m-1})+\ldots+l(i_1)}B^0(i_1))\\
&=\nu(T^{l(i_0)}B^0(i_0)/B^0(i_1))=\nu(T'B^0(i_1)/B^0(i_1))=\nu'(T'B^0(i_0)/B^0(i_1)).
\end{aligned}$$

Due to Proposition 3.7 the last expression equals $\nu'(T'B^0(i_0))=\nu'(B^0(i_0))$. Thus (3.17) is true.

We now wish to prove (3.1). Let

$$p_i^0=p_i(\nu),\qquad i=1,2,\ldots,\qquad p^0=(p_1^0,p_2^0,\ldots), \qquad (3.18)$$

and let P denote the set of probability vectors $p=(p_1,p_2,\ldots)$ such that $\sum_{i=1} p_i l(i)<\infty$. Every $p\in P$ determines (in a natural fashion) a T'-invariant B-measure on Z_v' which, in its turn, determines a measure $\nu(p)\in\mathscr{E}(Z_v)$. Clearly, $p^0\in P$, $\nu(p^0)=\nu$. By our assumption $\nu(p^0)$ maximizes $\mathscr{P}^0(U,\cdot,v)$ on $\mathscr{E}(Z_v)$. Due to Propositions 2.9, 2.10 and Theorem $B(i)$ it also maximizes $\mathscr{P}^1(U,\cdot,v)$ on $\mathscr{E}(Z_v)$. Moreover, $\mathscr{P}^1(U,\nu(p^0),v)=\ln\lambda(Q_U)$. In view of (2.20) and Proposition 2.11 it follows that

$$\begin{aligned}
&-(1/\sum_{i=1} p_i^0 l(i))\sum_{i=1} p_i^0(\ln p_i^0+e(i))\\
&=\max_{p\in P}[(-1/\sum_{i=1} p_i l(i))\sum_{i=1} p_i(\ln p_i+e(i))]\\
&=\ln\lambda(Q_U),
\end{aligned} \qquad (3.19)$$

where, as usual, $0\ln 0=0$.

It follows that $p_i^0>0$ for all i. Indeed, it can be immediately checked that if $p_i^0=0$, $p_j^0>0$ for some i,j, one can increase the left side of (3.19) by substituting ε and $p_j^0-\varepsilon$ for p_i^0 and p_j^0 respectively and by choosing $\varepsilon>0$ sufficiently small.

Setting $p_1=1-\sum_{i=2}^{\infty} p_i$ we can regard the content of the square brackets in (3.19) as a function, call it F, of $p_2,p_3,\ldots$, where $p_i\geqq 0$ for $i>1$, $\sum_{i=2}^{\infty} p_i\leqq 1$ and $\sum_{i=2}^{\infty} p_i l(i)<\infty$. As we have seen above, all of these inequalities are strict when p_i

$=p_i^0$, $i>1$. Hence F is differentiable with respect to each p_i at this point and moreover, $\partial F/\partial p_i=0$, $i>1$. It follows that

$$(l(i)-l(1))\sum_{j=1}^{\infty} p_j^0(\ln p_j^0+e(j))$$
$$-(e(i)-e(1)+\ln p_i^0-\ln p_1^0)\sum_{j=1}^{\infty} p_j^0 l(j), \qquad i>1, \tag{3.20}$$

which yields

$$p_i^0=cu^{l(i)}\exp(-e(i)), \qquad i\geqq 1, \tag{3.21}$$

where u is a constant and

$$c=1/\left[\sum_{i=1}^{\infty} u^{l(i)}\exp(-e(i))\right]. \tag{3.22}$$

It remains to show that

$$u=1/\lambda(Q_U), \qquad c=1. \tag{3.23}$$

We set

$$\varphi(t)=\sum_{i=1}^{\infty} t^{l(i)}\exp(-e(i)), \qquad 0\leqq t\leqq r, \tag{3.24}$$

$$p_i(t)=(1/\varphi(t))\,t^{l(i)}\exp(-e(i)), \qquad i\geqq 1, \quad 0<t<r, \tag{3.25}$$

$$f(t)=-\left[\sum_{i=1}^{\infty} p_i(t)\,l(i)\right]^{-1}\sum_{i=1}^{\infty} p_i(t)[\ln p_i(t)+e(i)]$$
$$=-\ln t+[t\varphi'(t)]^{-1}\varphi(t)\ln\varphi(t), \qquad 0<t<r, \tag{3.26}$$

where r is the radius of convergence of the series in (3.24) (it is worth observing that this is the same series as in (1.9)). By (3.26)

$$f'(t)=[\varphi(t)/t\varphi'(t)]'\ln\varphi(t), \qquad 0<t<r. \tag{3.27}$$

From the fact that φ is the sum of a power series with non-negative coefficients of which at least two are positive, it follows that $\varphi(t)/t\varphi'(t)>\varphi(t+\delta)/(t+\delta)\,\varphi'(t+\delta)$ when $0<t<t+\delta<r$. Therefore, when $\varphi(r)\leqq 1$, f monotonically increases on $(0,r)$, and when $\varphi(r)>1$, it has a unique absolute maximum at the point $r_1\in(0,r)$ for which $\varphi(r_1)=1$.

From (3.21), (3.22) we see that $u\leqq r$. If $u<r$, then by (3.19), (3.21), (3.22), (3.24)–(3.26) $f(u)=\max f(t)$, $0<t<r$. Taking into account the above mentioned properties of f we conclude that $\varphi(r)>1$, $\varphi(u)=1$. Proposition 3.2 now implies that $u=1/\lambda(Q_U)$. So we obtain (3.23). If $u=r$, the fact that $p^0\in P$ implies that the series for $\varphi(t)$ and $\varphi'(t)$ both converge when $t=u=r$ and moreover, that f can be defined by (3.26) to be a continuous function on $(0,r]$. Moreover,

$$f(r)=\max_{0\leqq t\leqq r} f(t)=\lambda(Q_U). \tag{3.28}$$

From (3.26) and from the properties of f it follows that $\varphi(r) \leqq 1$. Hence (see Proposition 3.2) $r = 1/\lambda(Q_U)$. By substituting this into (3.26) we see that (3.28) can be satisfied only when $\varphi(r) = 1$. So we arrive at (3.23) again.

3.10. **Proof of Theorem C.** (i) Suppose Q_U is indecomposable. There are $v, w \in V$ and a positive integer k such that $Q_U^{(k)}(v, w) = +\infty$. Due to the indecomposability of Q_U there is a path leading from w to v in $G(Q_U)$. Let l denote the length of this path. It follows that $Q_U^{(k+1)}(v, v) = +\infty$. By definition

$$Q_U^{(k+1)}(v, v) = \sum_\gamma Q_U(\gamma),$$

where the sum is over the set $\Gamma(Q_U, v, k+l)$ of all the paths γ in $G(Q_U)$ of length $k+l$ leading from v to v. Let c be an arbitrary positive number. There exists a finite subset $\Gamma_c \subset \Gamma(Q_U, v, k+l)$ such that $\sum_{\gamma \in \Gamma_c} Q_U(\gamma) \geqq c^{k+l}$. Since Q_U is indecomposable, there exists a finite subset $V_c \subset V$ such that the restriction $Q_{U,c}$ of Q_U to $V_c \times V_c$ is an indecomposable matrix and V_c contains every $w \in V$ visited by a path $\gamma \in \Gamma_c$. Obviously, $Q_{U,c}^{(k+l)}(v, v) \geqq c^{k+l}$ which implies that $Q_{U,c}^{((k+l)n)}(v, v) \geqq c^{(k+l)n}$ for any positive integer n. Hence the radius of convergence of the series $\sum_{n=1}^{\infty} Q_{U,c}^{(n)}(v, v)\, t^n$ does not exceed c. Due to Proposition 1.6(i) $R(Q_{U,c}) \leqq 1/c$, whence $\lambda(Q_{U,c}) \geqq c$. If $c > \lambda(Q_U)$ we come to a contradiction.

(ii) Let $\mu \in \mathscr{E}$ be a maximal measure, i.e., $P(U, \mu) = \ln \lambda(Q_U)$. There is a $v \in V$ such that $\mu \in \mathscr{E}(v)$. Moreover, μ is a v-maximal measure. By Theorem 3.4 $\nu = \Phi_v^* \mu$ is determined by (3.1). Taking into account that ν and ν' are probability measures we have

$$\sum_{i=1}^{\infty} p_i(\nu) = 1, \qquad \sum_{i=1}^{\infty} p_i(\nu)\, l(i) = L(\nu) < \infty. \tag{3.29}$$

Let w be an arbitrary vertex of $G(Q_U)$ and let $\Gamma^-(v, w)$ (respectively, $\Gamma^+(w, v)$) denote the family of paths in $G(Q_U)$ leading from v to w (respectively, from w to v) and containing v only as the initial (respectively, terminal) vertex. We let

$$\xi(w) = \sum_{\gamma \in \Gamma^+(w, v)} \lambda^{1 - l(\gamma)} \exp(-e(\gamma)), \tag{3.30}$$

$$\eta(w) = \sum_{\gamma \in \Gamma^-(v, w)} \lambda^{-l(\gamma)} \exp(-e(\gamma)), \tag{3.31}$$

where $\lambda = \lambda(Q_U)$ and $e(\gamma) = \sum_{i=1}^{n-1} U(v_i, v_{i+1})$ as $\gamma = (v_1, \ldots, v_n)$. Due to (3.1), (3.29)

$$\xi(v) = \lambda, \quad \eta(v) = 1. \tag{3.32}$$

By the same reason and since $\Gamma^+(v, v) = \Gamma^-(v, v)$ we have

$$\sum_{w} \xi(w)\,\eta(w) = \xi(v)\,\eta(v) + \sum_{w \neq v} \xi(w)\,\eta(w) = \lambda\Big[\sum_{\gamma\in\Gamma^+(v,v)} \lambda^{-l(\gamma)} \exp(-e(\gamma))\Big]^2$$
$$+\lambda \sum_{w\neq v}\ \sum_{\gamma\in\Gamma^-(v,w)}\ \sum_{\gamma_1\in\Gamma^+(w,v)} \lambda^{-l(\gamma)-l(\gamma_1)} \exp(-e(\gamma)-e(\gamma_1))$$
$$=\lambda\left[\sum_{i=1}^{\infty} \lambda^{-l(i)} \exp(-e(i))\right]^2 + \lambda \sum_{\gamma\in\Gamma^+(v,v)} \lambda^{(l(\gamma)-1)} \lambda^{-l(\gamma)} \exp(-e(\gamma))$$
$$=\lambda+\lambda(L(v)-1)=\lambda L(v)<\infty. \tag{3.33}$$

It follows in particular that $\xi(w)\,\eta(w)<\infty$ for each $w\in V$. Furthermore,

$$\sum_{w_1\in V} Q_U(w, w_1)\,\xi(w_1)=\lambda\,\xi(w), \quad \sum_{w_1\in V} \eta(w_1)\,Q_U(w_1, w)=\lambda\,\eta(w), \qquad w\in V. \tag{3.34}$$

Indeed, due to (3.1), (3.29)

$$\sum_{w_1\in V} Q_U(w, w_1)\,\xi(w_1)=Q_U(w, v)\,\xi(v)+\sum_{w_1\neq v} Q_U(w, w_1)\,\xi(w_1)$$
$$=\lambda \exp(-U(w, v)) \sum_{\gamma\in\Gamma^+(v,v)} \lambda^{-l(\gamma)} \exp(-e(\gamma))$$
$$+\lambda \sum_{w_1\neq v} \exp(-U(w, w_1)) \sum_{\gamma\in\Gamma^+(w_1,v)} \lambda^{-l(\gamma)} \exp(-e(\gamma))$$
$$=\lambda \exp(-U(w, v))+\lambda^2 \sum_{\substack{\gamma\in\Gamma^+(w,v),\\ l(\gamma)>1}} \lambda^{-l(\gamma)} \exp(-e(\gamma))$$
$$=\lambda \sum_{\gamma\in\Gamma^+(w,v)} \lambda^{1-l(\gamma)} \exp(-e(\gamma))=\lambda\,\xi(w).$$

So the former of the two equalities in (3.34) is true. The latter one can be checked in the same way.

For any w, $w_1\in V$ we let

$$\pi(w)=\xi(w)\,\eta(w)/\lambda\, L(v), \qquad p(w, w_1)=Q_U(w, w_1)\,\xi(w_1)/\lambda\,\xi(w_1) \tag{3.35}$$

It follows from (3.34) that $\pi(\cdot)$ and $p(\cdot,\cdot)$ being regarded as a vector and a matrix turn out to be stochastic.

In order to clarify the structure of μ it suffices to find $\mu(C)$ for any cylinder set C of the form

$$C=\{x\in X(Q_U)\colon x_0=v_0, \dots, x_n=v_n\}, \ v_i\in V, \quad 0\leqq i\leqq n.$$

There are 4 cases: 1) $v_0=v$, $v_n=v$; 2) $v_0\neq v$, $v_n\neq v$; 3) $v_0=v$, $v_n\neq v$; 4) $v_0\neq v$, $v_n=v$. We start with the first case. The definition of Φ_v, the explicit form of ν, and (3.33) together imply that

$$\mu(C)=\exp\left[-\sum_{i=0}^{n-1} U(v_i, v_{i+1})/\lambda^n L(v)=\lambda^{-n}(1/L(v)) \prod_{i=0}^{n-1} Q_U(v_i, v_{i+1}).\right.$$

On the other hand due to (3.35), (3.32)

$$\pi(v_0)\prod_{i=0}^{n-1} p(v_i, v_{i+1}) = \eta(v_0)(\lambda L(\nu))^{-1}\lambda^{-n}\xi(v_n)\prod_{i=0}^{n-1} Q_U(v_i, v_{i+1})$$
$$= \lambda^{-n}(L(\nu))^{-1}\prod_{i=0}^{n-1} Q_U(v_i, v_{i+1}).$$

Thus in the case under consideration

$$\mu(C) = \pi(v_0)\prod_{i=0}^{n-1} p(v_i, v_{i+1}).$$

We now suppose that $v_0 \neq v, v_n \neq v$. For each pair of paths $\gamma^- = (v_1^-, \ldots, v_k^-) \in \Gamma^-(v, v_0)$, $\gamma^+ = (v_1^+, \ldots, v_m^+) \in \Gamma^+(v_n, v)$ (obviously, $v_1^- = v_m^+ = v$, $v_k^- = v_0$, $v_1^+ = v_n$) we let

$$C(\gamma^-, \gamma^+) = \{x \in X(Q_U): x_{-k+1} = v_1^-, \ldots, x_{-1} = v_{k-1}^-, x_0 = v_0, \ldots, x_n = v_n,$$
$$x_{n+1} = v_1^+, \ldots, x_{n+m} = v_m^+\}.$$

Due to (3.30), (3.31)

$$\mu(C) = \sum_{\gamma^- \in \Gamma^-(v, v_0)}\ \sum_{\gamma^+ \in \Gamma^+(v_n, v)} \mu(C(\gamma^-, \gamma^+))$$
$$= \sum_{\gamma^- \in \Gamma^-(v, v_0)}\ \sum_{\gamma^+ \in \Gamma^+(v_n, v)} \exp\left[-e(\gamma^-) - e(\gamma^+) - \sum_{i=0}^{n-1} U(v_i, v_{i+1})\right] / \lambda^{n+l(\gamma^-)+l(\gamma^+)} L(\nu)$$
$$= \eta(v_0)\,\xi(v_n)\sum_{i=0}^{n-1} Q_U(v_i, v_{i+1})/\lambda^{n+1} L(\nu) = \pi(v_0)\prod_{i=0}^{n-1} p(v_i, v_{i+1}),$$

i.e., (3.36) is true again. The remaining two cases are treated similarly. We see from (3.36) that μ is a Markov measure with initial distribution $\pi(\cdot)$ and transition probability $p(\cdot, \cdot)$. Hence there can be only one ergodic maximal measure as claimed before.

(iii) Necessity. As proved before, if there exists a maximal measure, then one can find non-negative vectors ξ and η with finite inner product satisfying (3.34). By Proposition 1.6(iii) and Remark 3.3 this yields the $1/\lambda$-positivity of Q_U.

Sufficiency. Assume that Q_U is $1/\lambda(Q_U)$-positive and fix both an arbitrary $v \in V$ and an arbitrary enumeration on $\Gamma(Q_U, v)$. Let

$$p_n = \lambda^{-l(n)}\exp(-e(n)), \qquad n = 1, 2, \ldots, \tag{3.37}$$

where $\lambda = \lambda(Q_U)$ and $l(n)$, $e(n)$ are given by (2.2). Due to Proposition 1.6(iii)

$$\sum_{n=1}^{\infty} p_n = 1, \qquad \sum_{n=1}^{\infty} l(n)\,p_n < \infty.$$

So there exists a measure ν_v on Z_v for which ν_v' is a B-measure on Z_v' and $p_n(\nu_v) = p_n$. Let $\mu_v = (\Phi_v^*)^{-1}\nu_v$. In the proof of assertion (ii) we could see that μ_v was an ergodic Markov measure (note that (3.37) is just the same as (3.1)) specified by (3.35), (3.30), (3.31). From this it is easy to deduce that for any $w \in V$ the measure $\nu_w = \Phi_w^* \mu_w = \nu$ on Z_w is the lifting of the B-measure $\nu'_w = \nu'$ on Z'_w specified by (3.1), where now $l(n)$ and $e(n)$ correspond to the n-th w-cycle (this is true for an arbitrary enumeration on $\Gamma(Q_U, w)$).

We shall show that

$$P^0(U, \nu, w) \geqq \ln \lambda(Q_U), \qquad w \in V. \tag{3.38}$$

For this let us bound below the quantity

$$h_\nu(T, \beta_n) = \lim_{m \to \infty} H_\nu \left(T\beta_n \Big/ \bigvee_{i=0}^{m} T^{-i} \beta_n \right), \qquad n \geqq 1.$$

Let m be a positive integer and C an arbitrary element of $\bigvee_{i=0}^{m} T^{-i} \beta_n$ with $\nu(C) > 0$. By definition, $C = \bigcap_{i=0}^{m} T^{-i} C_i$, where C_i is an element of β_n, $i = 0, 1, \ldots, m$. If $C_0 = B^j(k)$, $1 \leqq k \leqq n$, $0 < j \leqq l(k) - 1$, then $H_\nu(T\beta_n/C) = 0$. If $C_0 = B^0(k)$, $1 \leqq k \leqq n$, then since ν' is a B-measure, we have

$$\nu(B/C) = \nu(B/C_0) = \nu(B/B^0(k))$$

for any element B of $T\beta_n$. If, finally, $C_0 = B^c(n) = \bigcup_{i \geqq n+1} \bigcup_{j=0}^{l(i)-1} B^j(i)$, then for the same reason

$$\begin{aligned}
\nu(B/C) &= \sum_{i \geqq n+1} \sum_{j=0}^{l(i)-1} \nu(B/B^j(i) \cap C)\, \nu(B^j(i)/C) \\
&= \sum_{i \geqq n+1} \nu(B/B^0(i) \cap C)\, \nu(B^0(i)/C) = \sum_{i \geqq n+1} \nu(B/B^0(i))\, \nu(B^0(i)/C).
\end{aligned}$$

As a result we have

$$\begin{aligned}
H_\nu &\left(T\beta_n \Big/ \bigvee_{i=0}^{m} T^{-i} \beta_n \right) = \sum_{k=1}^{n} \sum_{C \subset B^0(k)} \nu(C)\, H_\nu(T\beta_n/C) + \sum_{C \subset B^c(n)} \nu(C)\, H_\nu(T\beta_n/C) \\
&= -\sum_{k=1}^{n} \sum_{C \subset B^0(k)} \nu(C) \Big[\nu(TB^c(n)/B^0(k)) \ln \nu(TB^c(n)/B^0(k)) \\
&\quad + \sum_{j=1}^{n} \nu(TB^{l(j)-1}(j)/B^0(k)) \ln \nu(TB^{l(j)-1}(j)/B^0(k)) \Big] \\
&\quad - \sum_{C \subset B^c(n)} \nu(C) \Big[\sum_{k \geqq n+1} \nu(TB^c(n)/B^0(k))\, \nu(B^0(k)/C) \\
&\quad \cdot \ln \sum_{k \geqq n+1} \nu(TB^c(n)/B^0(k))\, \nu(B^0(k)/C) + \sum_{j=1}^{n} \sum_{k \geqq n+1} \nu(B^{l(j)-1}(j)/B^0(k)) \\
&\quad \cdot \nu(B^0(k)/C) \ln \sum_{k \geqq n+1} \nu(B^{l(j)-1}(j)/B^0(k))\, \nu(B^0(k)/C)
\end{aligned}$$

$$= -\sum_{k=1}^{n} \sum_{C \subset B^0(k)} \nu(C) \left(p^{(n)} \ln p^{(n)} + \sum_{i=1}^{n} p_i \ln p_i\right)$$

$$- \sum_{C \subset B^c(n)} \nu(C) \left[p^{(n)} \sum_{k \geqq n+1} \nu(B^0(k)/C) \left(\ln p^{(n)} + \ln \sum_{k \geqq n+1} \nu(B^0(k)/C)\right)\right.$$

$$\left. + \sum_{i=1}^{n} \sum_{k \geqq n+1} p_i \, \nu(B^0(k)/C) (\ln p_i + \ln \sum_{k \geqq n+1} \nu(B^0(k)/C))\right]$$

$$= -(1/L(\nu))(1 - p^{(n)}) \left(p^{(n)} \ln p^{(n)} + \sum_{i=1}^{n} p_i \ln p_i\right)$$

$$-(1/L(\nu))\, p^{(n)} (p^{(n)} \ln p^{(n)}) - p^{(n)} \sum_{C \subset B^c(n)} \nu(C) \sum_{k \geqq n+1} \nu(B^0(k)/C)$$

$$\cdot \ln \sum_{k \geqq n+1} \nu(B^0(k)/C) - (1/L(\nu))\, p^{(n)} \sum_{i=1}^{n} p_i \ln p_i$$

$$- \sum_{C \subset B^c(n)} \nu(C) \sum_{i=1}^{n} p_i \sum_{k \geqq n+1} \nu(B^0(k)/C) \ln \sum_{k \geqq n+1} \nu(B^0(k)/C)$$

$$\geqq -(1/L(\nu)) \left(p^{(n)} \ln p^{(n)} + \sum_{i=1}^{n} p_i \ln p_i\right).$$

It follows that

$$h_\nu(T, \beta_n) \geqq -(1/L(\nu)) \sum_{i=1}^{n} p_i \ln p_i$$

and due to (3.37)

$$h_\nu(T, \beta_n) - (1/L(\nu)) \sum_{i=1}^{n} p_i \, e(i) \geqq -(1/L(\nu)) \sum_{i=1}^{n} p_i (\ln p_i + e(i))$$

$$= (1/L(\nu)) \sum_{i=1}^{n} p_i \, l(i) \ln \lambda.$$

As $n \to \infty$ we obtain (3.38). In view of Proposition 2.4, Theorem A(ii), and the ergodicity of μ we conclude that μ is a maximal measure. This finishes the proof of Theorem C.

4. Proof of the Variational Principle

In this section we shall prove Theorem D. Assertion (i) of this theorem is an immediate consequence of Theorem C(ii). So we restrict attention on the proof of assertion (ii). It will consist of several steps.

4.1 **Proposition.** *If for an indecomposable potential U there exists a Gibbs measure $\mu \in \mathscr{I}$, then $\lambda(Q_U) < \infty$.*

Proof. The existence of a Gibbs measure implies that $\Xi_l(v, w)$ is finite for any l and any $v, w \in V$. Hence Q_U is an admissible matrix and due to Remark 3.3 it suffices to make sure that $R(Q_U) > 0$. This can be done by a small modification of the Kesten argument ([6], Lemma 6) giving the same inequality when $U<\infty$. We shall not repeate here this argument.

4.2. *Notation.* Let D denote an arbitrary directed graph with a countable set of vertices W and let K be a real function defined on the edges of D. For every path $\delta=(w_1, \ldots, w_n)$ in D we set

$$K(\delta)=\sum_{i=1}^{n-1} K(w_i, w_{i+1}).$$

Let $\Delta_n(w, w')$, $w, w' \in W$, denote the set of paths in D of length n leading from w to w' and $\Delta_n^+(w, w')$ denote the set of those $\delta \in \Delta_n(w, w')$ which does not contain w' as an intermediate vertex. Let

$$K_n^+(w, w')=\sum_{\delta \in \Delta_n^+(w, w')} K(\delta).$$

4.3. **Lemma.** *If the graph D is connected, then*
(i) *for any* $n \geqq 1$ *and any* $w \in W$

$$K_n^+(w, w) \leqq \lambda^n \tag{4.1}$$

(ii) *for any* $n \geqq 1$ *and any* $w, w' \in W$, $w' \neq w$,

$$K_n^+(w', w) \leqq \lambda^{n+m(w, w')}/K(\delta(w, w')), \tag{4.2}$$

where $\delta(w, w')$ is any path of minimal length from w to w', $m(w, w')=l(\delta(w, w'))$ the length of $\delta(w, w')$, and $\lambda=\lambda(K)$ (K can be thought of as a matrix).

Proof. (i) It suffices to restrict ourselves to the case where $\lambda<\infty$. By Theorem C(i) the matrix is admissible and for any $w \in W$ we have

$$\sum_{n=1}^{\infty} K_n^+(w, w) R^n \leqq 1,$$

where $R=R(K)$. By Remark 3.3 $1/R=\lambda$ which yields (i).

(ii) Let $\delta(w, w')=(w. w'_1, \ldots, w')$. Obviously, $\delta(w, w')$ does not contain w as an intermediate vertex. Therefore, if $\delta=(w', w_1, \ldots, w_k, w) \in \Delta_n^+(w', w)$, then $\delta(w, w')\, \delta \in \Delta_{n+m(w, w')}^+(w, w)$, where $\delta(w, w')\, \delta=(w, w'_1, \ldots, w', w_1, \ldots, w_k, w)$. It follows from Lemma 4.3(i) that

$$K(\delta(w, w'))\, K_n^+(w', w) \leqq K_{n+m(w, w')}^+(w, w) \leqq \lambda^{n+m(w, w')},$$

which yields (4.2), Q.E.D.

4.4. *Definition.* Let D be a directed graph, W the set of its vertices (W is finite or infinite countable), and K a real function on the set of edges of D. A shift-invariant probability measure μ defined on the space of all doubly-infinite paths in D will be called *K-Gibbsian* if for any $n \in \mathbb{Z}$, $k \in \mathbb{Z}^+$, and any path

$(w_0, w_1, \dots, w_{k+1})$ in D the following equality holds

$$\mu(x_n = w_0, x_{n+1} = w_1, \dots, x_{n+k+1} = w_{k+1}) = c_k(w_0, w_{k+1}) \prod_{i=0}^{k} K(w_i, w_{i+1}),$$

$c_k(w_0, w_{k+1})$ being the normalizing factor.

Obviously, every Gibbs field corresponding to the potential U is a Q_U-Gibbsian measure. Moreover, $c_k(w, w') = 1/Q_U^{(k)}(w, w')$.

4.5. Suppose that μ is a shift-invariant probability measure on $X(Q_U)$, $\mu \in \mathscr{E}(v)$ for some $v \in V$, and let $v = \Phi_v^* \mu$ as in 2.3. What conditions on v follow if (in addition) we suppose that μ is a Gibbs measure corresponding to U?

To answer this question we first note that Z_v (see 2.3) can be regarded as the space of all doubly-infinite paths in the directed graph G_v whose vertices have the form (γ, i), where γ is an arbitrary v-cycle in $G(Q_U)$ (i.e. $\gamma \in \Gamma(Q_U, v)$), $i \in \mathbb{Z}$, $0 \leqq i \leqq 1(\gamma) - 1$, and whose edges are defined as follows: there is an edge from (γ, i) to (γ', i') iff either $\gamma' = \gamma$, $i' = i + 1$, or $i = l(\gamma) - 1$, $i' = 0$. Denote by W the set of vertices of G_v. Thus Z_v is the set of all sequences $(z_i)_{i \in \mathbb{Z}}$ such that $z_i \in W$ and (z_i, z_{i+1}) defines an edge of G_v, $i \in \mathbb{Z}$. The vertices of the form $(\gamma, 0)$ and $(\gamma, l(\gamma) - 1)$ will be called *lower* and *upper* respectively.

We now define a function Q_U^v on the edges of G_v as follows. Let $\gamma = (v_0, v_1, \dots, v_{l-1}, v_l) \in \Gamma(Q_U, v)$, $\gamma' = (v'_0, v'_1, \dots, v'_{l'-1}, v'_{l'}) \in \Gamma(Q_U, v)$, where $v_0 = v'_0 = v'_{l'} = v_l = v$. We set

$$Q_U^v((\gamma, i), (\gamma', i')) = \begin{cases} Q_U(v_i, v_{i+1}), & \text{when } \gamma' = \gamma, \quad 0 \leqq i < i + 1 \leqq l - 1, \\ Q_U(v_{l-1}, v), & \text{when } i = l - 1, \quad i' = 0. \end{cases} \tag{4.3}$$

We can extend $Q_U^v = Q^v$ to the set of all vertex pairs setting $Q^v = 0$ for those pairs which define no edge. After that Q^v can be regarded as a matrix. From (4.1) combined with Proposition 3.2 and the definition of $\lambda(Q_U)$ it follows that $\lambda(Q^v) = \lambda(Q_U)$.

4.6. **Proposition.** *The measure $v = \Phi_v^* \mu$ is Q^v-Gibbsian.*

Proof. We first consider a path $\delta = (w_0, \dots, w_k)$, $w_i \in W$, where $w_0 = (\gamma, 0)$, $w_k = (\gamma', l(\gamma') - 1)$, $\gamma, \gamma' \in \Gamma(Q_U, v)$. Using the structure of G_v we can divide δ into blocks corresponding to v-cycles, that is represent it in the form

$$((\gamma_0, 0), \dots, (\gamma_0, l(0) - 1, \dots, (\gamma_m, 0), \dots, (\gamma_m, l(m) - 1)),$$

where $\gamma_0 = \gamma, \gamma_m = \gamma'$, $l(i)$ being the length of γ_i. Let us set $B(\delta) = \{z \in Z_v: z_0 = w_0, \dots, z_k = w_k\}$ and find $\Phi_v^{-1} B(\delta)$. Let $\gamma_i = (v_0^i, \dots, v_{l(i)}^i), v_0^i = v_{l(i)}^i = v$, $0 \leqq i \leqq m$. By the definition of Φ_v

$$\begin{aligned} \Phi_v^{-1} B(\delta) = \{ & x \in X(Q_U): x_0 = v, \dots, x_{l(0)-1} = v_{l(0)-1}^0, \dots, \\ & x_{l(0)+\dots+l(i-1)} = v, \dots, x_{l(0)+\dots+l(i-1)+l(i)-1} = v_{l(i)-1}^i, \\ & \dots, x_{l(0)+\dots+l(m-1)} = v, \dots, x_{l(0)+\dots+l(m-1)+l(m)-1} \\ & = v_{l(m)-1}^m, x_{l(0)+\dots+l(m)} = v\}, \ \sum_{i=0}^{m} l(i) = k + 1. \end{aligned}$$

Due to (4.3) and the fact that μ is a Gibbs measure corresponding to U we have

$$\nu(B(\delta))=\mu(\Phi_v^{-1}B(\delta))=[1/Q_U^{(k)}(v,v)]\prod_{i=0}^{m}\prod_{j=0}^{l(i)-1}Q_U(v_j^i,v_{j+1}^i)$$

$$=c_k(w_0,w_k)\prod_{i=1}^{k}Q^v(w_i,w_{i+1}),$$

where $c_k(w_0,w_k)=c_k((\gamma,0),(\gamma',l'-1))=1/Q_U^{(k)}(v,v)$. Thus in our case $\nu(B(\delta))$ has the desired form.

Let, further, $\delta=(w_0,\ldots,w_k)$, where $w_0=(\gamma,i)$, $w_k=(\gamma',i')$, $0<i\leqq l-1$, $0\leqq i'<l'-1=l(\gamma')-1$. We continue δ so as to obtain the path $\delta_1=((\gamma,0),\ldots,(\gamma,i),\ldots,(\gamma',i'),\ldots,(\gamma',l'-1))$. Obviously, $\nu(B(\delta_1))=\nu(B(\delta))$. Due to the above

$$\nu(B(\delta_1))=c_{k+i+l'-i'-1}((\gamma,0),(\gamma',l'-1))\prod_{j=0}^{i-1}Q^v((\gamma,j),(\gamma,j+1))$$

$$\cdot\prod_{j'=i'}^{l'-2}Q^v((\gamma',j'),(\gamma',j'+1))\prod_{n=0}^{k}Q^v(w_n,w_{n+1}),$$

and we set

$$c_k((\gamma,i),(\gamma',i'))=c_{k+i+l'-i'-1}((\gamma,0),(\gamma',l'-1))\prod_{j=0}^{i-1}Q^v((\gamma,j),(\gamma,j+1))$$

$$\cdot\prod_{j'=i'}^{l'-2}Q^v((\gamma',j'),(\gamma',j'+1)).$$

Then $(B(\delta))$ again has the desired form. The cases where $w_0=(\gamma,0)$, $w_k=(\gamma',i')$, $0\leqq i'<l(\gamma')-1$ and where $w_0=(\gamma,i)$, $0\leqq i<l(\gamma)-1$, $w_k=(\gamma',l(\gamma')-1)$ are treated similarly.

4.7. An arbitrary Gibbs measure from $\mathscr{I}$ can be decomposed into ergodic components and it is easy to see that each of these is a Gibbs measure with the same potential. So to complete the proof it suffices to check that if μ is an ergodic Gibbs measure such that $\mu\in\mathscr{E}(v)$ for some $v\in V$, then μ is v-maximal. Letting $\nu=\Phi_v^*\mu$ we see that $\nu\in\mathscr{E}(Z_v)$. By Proposition 4.4 ν is a Q^v-Gibbsian measure and it remains to show that

$$P^0(U,\nu,v)\geqq\ln\lambda(Q_U).$$

4.8. We now turn to the partition β_n, $n\geqq 1$, of the space $Z_v=X(Q^v)$ introduced in 2.3. Its element $B^i(k)$ can be represented in the form

$$B^i(k)=\{z\in X(Q^v)\colon z_0=(\gamma_k,i)\},\qquad 1\leqq k\leqq n,\qquad 0\leqq i\leqq l(\gamma_k)-1$$

(the v-cycles $\gamma\in\Gamma(Q_U,v)$ are assumed to be ordered in an arbitrary way). Thus the vertex (γ_k,i) of the graph G_v is associated with $B^i(k)$. We now introduce a new symbol $\#$ and associate it with the set $B^c(n)$ being also an element of β_n. $\#$ will be referred to as the generalized vertex. Let $\beta_n^k=\bigvee_{i=0}^{k-1}T^{-i}\beta_n$. An arbitrary element C of β_n^k has the form $C=\bigcap_{i=0}^{k-1} T^{-i}A_i$, where A_i, $0\leqq i\leqq k-1$, is an element of β_n. With C we associate the sequence $\delta(C)=w_0(C),\dots,w_{k-1}(C)$, where $w_i(C)$ is the vertex (possibly generalized) associated with A_i, $0\leqq i\leqq k-1$.

For $i=0,\dots,k-2$ we set

$$q(w_i(C),w_{i+1}(C))=\begin{cases}1, & \text{if } w_i(C)=\#,\\ Q^v(w_i(C),w_{i+1}(C)), & \text{if } w_i(C)\neq\#,\ w_{i+1}(C)\neq\#,\\ Q^v(w_i(C),(\gamma,0)), & \text{if } w_{i+1}(C)=\#,\end{cases}\tag{4.4}$$

where $\gamma\in\Gamma(Q_U,v)$ is arbitrary chosen. We note that in the last case $w_i(C)=(\gamma_j,l(j)-1)$ for some $j\leqq n$, and the definition of Q^v shows that $Q^v(w_i(C),(\gamma,0))$ does not depend on γ. Let further

$$q(C)=\prod_{i=0}^{k-2} q(w_i(C),w_{i+1}(C)),\tag{4.5}$$

$$k^-(C)=\min\{i\colon 0\leqq i\leqq k-1,\ w_i(C)\neq\#\},$$

$$k^+(C)=\max\{i\colon 0\leqq i\leqq k-1,\ w_i(C)\neq\#\}.$$

We shall refer to C as regular if $k^-(C)$, $k^+(C)$ exist and they are different. Let $W(n)$ denote the set of vertices (γ_k,i) of G_v with $1\leqq k\leqq n$, $0\leqq i\leqq l(\gamma_k)-1$. For every w^-, $w^+\in W(n)$ and every positive integers k^-, k^+ denote by $M_k(k^-,w^-,k^+,w^+)$ the union of those $C\in\beta_n^k$ for which $k^-(C)=k^-$, $w_{k^-}(C)=w^-$, $k^+(C)=k^+$, $w_{k^+}(C)=w^+$.

Fixing an arbitrary $\gamma\in\Gamma(Q_U,v)$ such that $(\gamma,0)\in W(n)$ we set

$$\varepsilon(\gamma)=\lambda^{l(\gamma)}/Q^v(\sigma(\gamma)),$$

where $\lambda=\lambda(Q_U)$ and $\sigma(\gamma)=((\gamma,0),\dots,(\gamma,l(\gamma)-1))$ is the shortest cycle in G_v going through $(\gamma,0)$.

4.9. **Proposition.** *Let v be a shift-invariant ergodic Q^v-Gibbsian measure on Z_v. Then for every k^-, $k^+\geqq 0$ such that $k^-<k^+$, every w^-, $w^+\in W(n)$ such that $v(M_k(k^-,w^-,k^+,w^+))>0$, and every $C\subset M_k(k^-,w^-,k^+,w^+)$ the following inequality holds*

$$(1/q(C))\,v(C/M_k(k^-,w^-,k^+,w^+))$$
$$\leqq\lambda^{s(C)}\max\{1,(\varepsilon(\gamma))^{s(C)}\}/(Q^v)^{(k^+-k^--1)}(w^-,w^+),$$

$s(C)$ being the number of symbols $\#$ in $\delta(C)$.

Proof. For every path δ in G_v we denote by $\delta^\#$ the sequence obtained from δ by replacing all the vertices outside of $W(n)$ by $\#$. By the definition of Gibbs measure for each $C\subset M_k(k^-,w^-,k^+,w^+)$ we have

$$v(C/M_k(k^-,w^-,k^+,w^+))$$
$$=\sum_{\delta=(w_0,\dots,w_{k-1})\colon\,\delta^\#=\delta(C)} v\{(z_0,\dots,z_{k-1})=\delta/z_{k^-}$$
$$=w^-,\ z_{k^+}=w^+,\ z_i\notin W(n)\text{ for }0\leqq i<k^-\text{ and for }k^+<i\leqq k-1\}$$

$$= \sum_{\delta=(w_0,\ldots,w_{k-1}):\,\delta^\#=\delta(C)} v(z_{k^-}=w_{k^-}, z_{k^-+1}=w_{k^-+1},\ldots,z_{k^+}=w_{k^+}/z_{k^-}$$

$$=w^-, z_{k^+}=w^+) = \sum_{\delta=(w_0,\ldots,w_{k-1}):\,\delta^\#=\delta(C)} Q^v(\delta(k^-,k^+))/(Q^v)^{(k^+-k^--1)}(w^-,w^+), \tag{4.6}$$

where $\delta(k^-,k^+)=(w_{k^-}, w_{k^-+1},\ldots,w_{k^+})$.

For every path δ such that $\delta^\#=\delta(C)$ we single out all the segments in $\delta(k^-,k^+)$ consisting of vertices from $W\setminus W(n)$ (if they exist) and number them from left to right as follows: $\sigma(1),\ldots,\sigma(m)$. Obviously, both m and the location of each $\sigma(i)$ within $\delta(k^-,k^+)$ depend only on C (in particular $m=m(C)$). The vertices immediately followed and immediately preceded by each $\sigma(i)$ belong to $W(n)$. We denote them by w_i^- and w_i^+ respectively. Let $|\sigma(i)|$ denote the number of vertices in $\sigma(i)$, $1\leq i\leq m$.

For every $w', w''\in W(n)$ and every $r\in\mathbb{Z}^+$ let $\Delta_r^n(w',w'')$ denote the set of paths in G_v of length r leading from w' to w'' and going outside of $W(n)$ between the initial and the terminal vertices. If $\delta=(w', w_1,\ldots,w_{r-1}, w'')\in\Delta_r^n(w',w'')$ we denote by $w^-(\delta)$ and $w^+(\delta)$ the vertices w_1 and w_{r-1} respectively. It follows from (4.6), (4.4), and (4.5) that

$$\begin{aligned}(1/q(C))\,v(C/M_k(k^-,w^-,k^+,w^+)) &= [1/(Q^v)^{(k^+-k^--1)}(w^-,w^+)] \\ \cdot \sum_{\delta:\,\delta^\#=\delta(C)} Q^v(\delta(k^-,k^+))/q(C) &= [1/(Q^v)^{(k^+-k^--1)}(w^-,w^+)] \\ \cdot \prod_{i=1}^{m} \sum_{\delta\in\Delta^n_{|\sigma(i)|+1}(w_i^-,w_i^+)} Q^v(\delta)/Q^v(w_i^-, w^-(\delta)).\end{aligned} \tag{4.7}$$

From the structure of G_v and the definition of $W(n)$ we see that w_i^- and $w^+(\delta)$ are upper vertices while w_i^+ and $w^-(\delta)$ are lower ones. We replace in each $\delta\in\Delta^n_{|\sigma(i)|+1}(w_i^-,w_i^+)$ the vertices w_i^- and w_i^+ by $(\gamma, l(\gamma)-1)$ and $(\gamma,0)$ respectively, where γ is an arbitrary v-cycle in G. The sequence obtained, say $\Psi(\delta)$, is also a path in G_v. Moreover, Ψ defines a one-to-one mapping from $\Delta^n_{|\sigma(i)|+1}(w_i^-,w_i^+)$ onto $\Delta^n_{|\sigma(i)|+1}((\gamma,l(\gamma)-1),(\gamma,0))$. From (4.3) we see that if w' is an upper vertex, then $Q^v(w',w'')$ does not depend on w'' within the set of lower vertices w''. With this in mind we have

$$\begin{aligned}Q^v(\delta)/Q^v(w_i^-, w^-(\delta)) &= Q^v(\Psi(\delta))/Q^v((\gamma,l(\gamma)-1), w^-(\delta)) \\ &= Q^v(\Psi(\delta))/Q^v((\gamma,l(\gamma)-1),(\gamma,0))\end{aligned}$$

and hence (see (4.7))

$$\begin{aligned}(1/q(C))\,v(C/M_k(k^-,w^-,k^+,w^+)) &= [1/(Q^v)^{(k^+-k^--1)}(w^-,w^+)] \\ \cdot \prod_{i=1}^{m(C)} \sum_{\delta\in\Delta^n_{|\sigma(i)|+1}((\gamma,l(\gamma)-1),(\gamma,0))} & Q^v(\delta)/Q^v((\gamma,l(\gamma)-1),(\gamma,0)) \\ \leq [(Q^v)^{(k^+-k^--1)}(w^-,w^+)]^{-1} & [Q^v((\gamma,l(\gamma)-1),(\gamma,0))]^{-m(C)} \\ \cdot \prod_{i=1}^{m(C)} (Q^v)^+_{|\sigma(i)|+1}((\gamma,l(\gamma)-1),(\gamma,0)) &\end{aligned} \tag{4.8}$$

(see also 4.2 for notation).

We shall apply Lemma 4.3 to bound the last expression. Let for definiteness $l(\gamma)>1$, that is $(\gamma, l(\gamma)-1) \neq (\gamma, 0)$. Obviously, the shortest path from $(\gamma, 0)$ to $(\gamma, l(\gamma)-1)$ has the form $((\gamma, 0), (\gamma, 1), \ldots, (\gamma, l(\gamma)-1))$. Due to (4.2)

$$(Q^v)^+_{|\sigma(i)|+1}((\gamma, l(\gamma)-1), (\gamma, 0)) \leqq \lambda^{|\sigma(i)|+1+l(\gamma)-1} \Big/ \prod_{j=0}^{l(\gamma)-2} Q^v((\gamma, j), (\gamma, j+1)).$$

Substituting this bound into (4.8) gives

$$\begin{aligned}
&(1/q(C))\, v(C/M_k(k^-, w^-, k^+, w^+)) \\
&\quad \leqq [(Q^v)^{(k^+-k^--1)}(w^-, w^+)]^{-1} [Q^v((\gamma, l(\gamma)-1), (\gamma, 0)) \\
&\quad \cdot \prod_{j=0}^{l(\gamma)-2} Q^v((\gamma, j), (\gamma, j+1))]^{-m(C)} \exp\left[\left(m(C)\, l(\gamma) + \sum_{i=1}^{m(C)} |\sigma(i)|\right) \ln \lambda\right] \\
&\quad = \lambda^{s(C)} [(Q^v)^{(k^+-k^--1)}(w^-, w^+)]^{-1} [\lambda^{l(\gamma)}/Q^v(\sigma(\gamma))]^{m(C)} \\
&\quad = \lambda^{s(C)} [(Q^v)^{(k^+-k^--1)}(w^-, w^+)]^{-1} (\varepsilon(\gamma))^{m(C)}.
\end{aligned} \tag{4.9}$$

If $\varepsilon(\gamma) \leqq 1$, then the right hand side of (4.9) does not exceed

$$\lambda^{s(C)}/(Q^v)^{(k^+-k^--1)}(w^-, w^+).$$

If $\varepsilon(\gamma)>1$ it does not exceed $(\lambda \varepsilon(\gamma))^{s(C)}/(Q^v)^{(k^+-k^--1)}(w^-, w^+)$, because $m(C) \leqq s(C)$. Thus the inequality claimed is proved.

4.10. For every $n \geqq 1$ we set

$$U_n^v(z) = \begin{cases} \ln Q^v(z_0, z_1), & \text{if } z_0 \in W(n), \\ 0, & \text{if } z_0 \notin W(n), \quad z=(z_i)_{i\in\mathbb{Z}} \in Z_v. \end{cases}$$

Thus the function U_n^v is defined on Z_v and it is constant on each element of β_n. Due to (2.2)–(2.6) and (4.3)

$$-\int_{Z_v} U_n^v \, dv = (1/L(v)) \sum_{i=1}^{n} p_i(v)\, e(i).$$

Due to (4.4), (4.5)

$$(k-1) \int_{Z_v} U_n^v \, dv = \int_{Z_v} \sum_{i=0}^{k-2} U_n^v \circ T^i \, dv = \sum_{C \in \beta_n^k} v(C) \ln q(C), \qquad k \geqq 2,$$

where $(U_n^v \circ T^i)(z) = U_n^v(T^i z)$, $z \in Z_v$. Therefore,

$$(1/L(v)) \sum_{i=1}^{n} p_i(v)\, e(i) = -(k-1)^{-1} \sum_{C \in \beta_n^k} v(C) \ln q(C)$$

and by (2.9)

$$P^0(U, v, v) = \lim_{n\to\infty} \lim_{k\to\infty} (1/k)[H_v(\beta_n^k) + \sum_{C \in \beta_n^k} v(C) \ln q(C)]. \tag{4.10}$$

Our goal now is to bound from below the content of the square brackets on the right hand side of (4.10).

Fixing an arbitrary positive $\varepsilon<1/4$ we pick a number n_ε so that

$$\sum_{j=1}^{n}\sum_{i=0}^{l(j)-1} \nu(B_j(i))=\nu(z_0\in W(n))>1-\varepsilon/2 \tag{4.11}$$

for every $n\geqq n_\varepsilon$. Let $n\geqq n_\varepsilon$, $k\geqq 2$, and let $N_k(\varepsilon)$ denote the union of those $C\in\beta_n^k$ for which $s(C)<k\varepsilon$. If $C\subset N_k(\varepsilon)$, then C is regular (see 4.8) and

$$k^-(C)\leqq k\varepsilon, \quad k^+(C)\geqq k(1-\varepsilon), \quad k^-(C)<k^+(C). \tag{4.12}$$

Consider the partition χ_k of Z_v whose elements are all the sets $M_k(k^-, w^-, k^+, w^+)$ of positive measure and the complement of their union which is to within a set of measure 0 just the element $\underbrace{\#\#\ldots\#}_{k}$. It follows that (for short we denote $M_k(k^-, w^-, k^+, w^+)=M$)

$$\begin{aligned}
H_\nu(\beta_n^k)&+\sum_{C\in\beta_n^k}\nu(C)\ln q(C)\geqq H_\nu(\beta_n^k/\chi_k)+\sum_{C\in\beta_n^k}\nu(C)\ln q(C)\\
&=-\sum_M \nu(M)\sum_{C\subset M}\nu(C/M)\ln\nu(C/M)+\sum_M\nu(M)\sum_{C\subset M}\nu(C/M)\ln q(C)\\
&=\sum_M\nu(M)\sum_{C\subset M}\nu(C/M)[-\ln\nu(C/M)+\ln q(C)],
\end{aligned} \tag{4.13}$$

where the sum is over all the elements C of β_n^k and those M which have positive measure.

If $C\subset M_k(k^-, w^-, k^+, w^+)\cap N_k(\varepsilon)$, then by Proposition 4.9

$$\begin{aligned}
-\ln\nu(C/M_k(k^-, w^-, k^+, w^+))+\ln q(C)&\geqq\ln(Q^v)^{(k^+-k^--1)}(w^-, w^+)\\
-s(C)\ln(\kappa\lambda)&\geqq\ln(Q^v)^{(k^+-k^--1)}(w^-, w^+)-k\varepsilon|\ln(\kappa\lambda)|,
\end{aligned} \tag{4.14}$$

where $\kappa=\max\{1, \varepsilon(\gamma)\}$ (the absolute value is taken to include the case $\kappa\lambda<1$). It follows from (4.12) and the assumption $\varepsilon<1/4$ that $k^+-k^--1\geqq k/4$, when $k>4$. Hence k^+-k^--1 tends to infinity together with k. So we can find a $k(n)$ such that for every $k\geqq k(n)$, every $w^-, w^+\in W(n)$, and every k^-, k^+ satisfying (4.12)

$$\ln(Q^v)^{(k^+-k^--1)}(w^-, w^+)\geqq(k^+-k^-)(\ln\lambda-\varepsilon)\geqq k(1-2\varepsilon)(\ln\lambda-\varepsilon). \tag{4.15}$$

From (4.13)–(4.15) we have

$$\begin{aligned}
(1/k)&[H_\nu(\beta_n^k)+\sum_{C\in\beta_n^k}\nu(C)\ln q(C)]\\
&\geqq\sum_M\nu(M)[(1-2\varepsilon)(\ln\lambda-\varepsilon)-\varepsilon|\ln(\kappa\lambda)|]\sum_{C\subset M\cap N_k(\varepsilon)}\nu(C/M)\\
&+\sum_M(-1/k)\,\nu(M)\sum_{C\subset M\cap N_k^c(\varepsilon)}\ln\nu(C/M)\\
&+\sum_M(-1/k)\,\nu(M)\sum_{C\subset M\cap N_k^c(\varepsilon)}\nu(C/M)\ln q(C),
\end{aligned} \tag{4.16}$$

where $N_k^c(\varepsilon)=Z_v\setminus N_k(\varepsilon)$, $k\geqq k(n)$. We shall bound each of the three sums on the right hand side of (4.16) denoting them by S_1, S_2, and S_3 respectively.

Obviously,

$$S_1 = [\ln\lambda - \varepsilon(2\ln\lambda + 1 - 2\varepsilon + |\ln(\kappa\lambda)|)]\, \nu(N_k^c(\varepsilon)).$$

The ergodic theorem combined with (4.11) and the ergodicity of ν imply that $\nu(N_k(\varepsilon)) \to 1$ as $k \to \infty$. Therefore S_1 can be made arbitrarily close to $\ln\lambda - \varepsilon(2\ln\lambda + 1 - 2\varepsilon + |\ln(\kappa\lambda)|)$ by choosing k sufficiently large. Further, $S_2 \geqq 0$. Finally, by (4.4), (4.5)

$$S_3 = (1/k) \sum_{C \subset N_k^c(\varepsilon)} \nu(C) \ln q(C)$$
$$\geqq k^{-1}(k-1)[\ln \min\{1, \min_{w', w'' \in W(n)} Q^v(w', w'')\}]\, \nu(N_k^c(\varepsilon)).$$

The last expression tends to zero as $k \to \infty$. Since ε can be chosen arbitrarily small we conclude that for any $\varepsilon_1 > 0$ there are $n(\varepsilon_1)$ and $k(\varepsilon_1)$ such that for $n \geqq n(\varepsilon_1)$, $k \geqq k(\varepsilon_1)$ we have

$$(1/k)[H_\nu(\beta_n^k) + \sum_{C \in \beta_n^k} \nu(C) \ln q(C)] \geqq \ln\lambda - \varepsilon_1,$$

which shows that $\mathscr{P}^0(U, \nu, v) \geqq \ln\lambda$. Q.E.D.

4.11. *An Example.* Let $(c_i)_{i=1}^\infty$ be a sequence of positive numbers with

$$\sum_{i=1} c_i = 1, \qquad -\sum_{i=1} c_i \ln c_i = +\infty \tag{4.17}$$

and let $(a_i)_{i=1}^\infty$, $(b_i)_{i=1}^\infty$ be arbitrary sequences of positive numbers such that $a_i b_i = c_i$, $1 \leqq i < \infty$. We take the natural numbers for V and set $U(i,j) = -\ln(a_i b_j)$, $i, j \in V$. It can be immediately checked that Q_U is an indecomposable admissible matrix with $\lambda(Q_U) = 1$ and that the Bernoully measure μ^0 with one-dimensional distribution $\pi(i) = c_i$, $i \in V$, is a Gibbs measure with potential U. By Theorem D, μ^0 maximizes the $\mathscr{P}$ defined by (1.5), (1.6) and moreover, $\mathscr{P}(U, \mu^0) = 0$. We note that $\mathscr{P}$ cannot be defined by (1.2) because $h_{\mu^0} = +\infty$ and also $\int_X (U(x_0, x_1))^+ \, d\mu^0 = +\infty$, where $u^+ \equiv \max\{0, u\}$, $u \in \mathbb{R}$. To prove the latter we use the following assertion which can be easily proved by the reader.

Lemma. *Let $(\Omega_i, \mathscr{A}_i, \nu_i)$, $i = 1, 2$, be two copies of a measure space $(\Omega, \mathscr{A}, \nu)$ and f a non-negative measurable function on $\Omega_1 \times \Omega_2$ such that $\int_{\Omega_1 \times \Omega_2} f \, d\nu_1 \, d\nu_2 = +\infty$. Then for every measurable function g on Ω*

$$\int_{\Omega_1 \times \Omega_2} [f(\omega_1, \omega_2) + g(\omega_2) - g(\omega_1)]^+ \, \nu_1(d\omega_1)\, \nu_2(d\omega_2) = +\infty.$$

There is another way of regularizing the right hand side of (1.2). One replaces it by

$$-\int_X [\ln \sum_{v \in V} I_v \, \mu(x_1 = v/\mathscr{B}^-) + U(x_0, x_1)] \, d\mu, \tag{4.18}$$

where I_v is the indicator function of the set $\{x \in X : x_1 = v\}$ and $\mathscr{B}^-$ the σ-algebra generated by the random variables x_i for $i \leqq 0$ (cf. [9]). It can be shown that if the content of the square brackets in (4.18) is an integrable function, then for every $v \in V$ and every $\mu \in \mathscr{E}(v)$ the right hand side of (1.5) coincides with (4.18). On the other hand, there are cases where (4.18) makes no sense. To show this we specify the sequence $(a_i)_{i=1}^{\infty}$ in the example above as follows. Let

$$a_{i+1} > a_i, \quad \ln a_{i+1} \geqq \left((1/c_{i+1}) + \sum_{j=1}^{i} c_j \ln a_j\right) \Big/ \sum_{j=1}^{i} c_j, \quad i \geqq 1. \tag{4.19}$$

Such a sequence can be easily produced by induction. Due to (4.19)

$$\begin{aligned} -\int_X [\ln \sum_{v\in V} I_v \mu^0(x_1 = v/\mathscr{B}^-) + U(x_0, x_1)]^+ \, d\mu^0 \\ = \sum_{i,j=1}^{\infty} c_i c_j (\ln a_i - \ln a_j)^+ = \sum_{i=2}^{\infty} c_i \sum_{j=1}^{i-1} c_j (\ln a_i - \ln a_j) \\ = \sum_{i=2}^{\infty} c_i \left(\ln a_i \sum_{j=1}^{i-1} c_j - \sum_{j=1}^{i-1} c_j \ln a_j\right) = +\infty \end{aligned}$$

and similarly

$$-\int_X [\ln \sum_{v\in V} I_v \mu^0(x_1 = v/\mathscr{B}^-) + U(x_0, x_1)]^- \, d\mu^0 = -\infty$$

(we set $u^- = \min\{0, u\}$, $u \in \mathbb{R}$) so that one can prescribe no reasonable value to (4.18).

Acknowledgment. The author is grateful to the referees for their critical reading of the manuscript and useful comments.

References

1. Dobrushin, R.L.: Gibbsian random fields. The general case. Funct. Analiz. **3**, 27–35 (1968); engl. transl. in Functional analysis and applications, **3**, 22–28 (1969)
2. Föllmer, H.: On entropy and information gain in random fields. Z. Wahrscheinlichkeitstheorie verw. Gebiete 26, 207–217 (1973)
3. Gurevich, B.M.: A variational principle for one-dimensional lattice Gibbs random fields with a countable number of states. Dokl. Akad. Nauk SSSR **241**, 749–752 (1978); engl. transl. in Soviet Math. Dokl. **19**, 914–917 (1978)
4. Ito, Sh., Mori, M.: Isomorphism of a piecewise linear transformation to a Markov automorphism. Publ. Res. Inst. Math. Sci. **13**, 1–16 (1977)
5. Kesten, H.: Existence and uniqueness of countable one-dimensional Markov random fields. Ann. Probability **4**, 557–569 (1976)
6. Künsch, H.: Almost sure entropy and the variational principle for random fields with unbounded state space. Preprint (1980)
7. Lanford, O.E., Ruelle, D.: Observables at infinity and states with short range correlations in statistical mechanics. Comm. Math. Phys. **13**, 194–215 (1968)
8. Pinsker, M.S.: Information and information stability of random variables and processes. Akad. Nauk SSSR, Moskva (1960) (in Russian)
9. Pirlot, M.: A strong variational principle for continuous spin systems. J. Appl. Probability **17**, 47–58 (1980)
10. Preston, C.: Random fields. Lecture Notes in Mathematics **534**. Berlin-Heidelberg-New York: Springer 1976

11. Rohlin, V.A.: Lectures on the entropy theory of measure-preserving transforations. Uspehi Mat. Nauk **12**, 3–56 (1967); engl. transl. in Russ. Math. Surveys **22**(5), 1–52 (1967)
12. Smorodinsky, M.: Ergodic theory. Entropy. Lect. Notes Math. **214**, 1–64. Berlin-Heidelberg-New York: Springer 1971
13. Spitzer, F.: A variational characterization of finite Markov chains. Ann. Math. Statist. **43**, 303–307 (1972)
14. Vere-Jones, D.: Ergodic properties of nonnegative matrices 1. Pacific J. Math. **22**, 361–386 (1967)
15. Walters, P.: Invariant measures and equilibrium states for some mappings which expand distances. Trans Amer. Math. Soc. **236**, 121–153 (1978)

Received May 30, 1980