The Annals of Applied Probability
1994, Vol. 4, No. 3, 859–901

EVOLUTIONARY FORMALISM FOR PRODUCTS OF POSITIVE RANDOM MATRICES

BY LUDWIG ARNOLD, VOLKER MATTHIAS GUNDLACH AND LLOYD DEMETRIUS

Universität Bremen, Universität Bremen and Harvard University

We present a formalism to investigate directionality principles in evolution theory for populations, the dynamics of which can be described by a positive matrix cocycle (product of random positive matrices). For the latter, we establish a random version of the Perron–Frobenius theory which extends all known results and enables us to characterize the equilibrium state of a corresponding abstract symbolic dynamical system by an extremal principle. We develop a thermodynamic formalism for random dynamical systems, and in this framework prove that the top Lyapunov exponent is an analytic function of the generator of the cocycle. On this basis a fluctuation theory for products of positive random matrices can be developed which leads to an inequality in dynamical entropy that can be interpreted as a directionality principle for the mutation and selection process in evolutionary dynamics.

1. Introduction. Evolutionary theory is concerned with understanding the dynamical behavior of replicating entities—molecules, cells or higher organisms—subject to two main forces: *mutation*, which introduces new variability within the population, and *selection*, which organizes this variability through competition for available resources. The latter ultimately leads to the replacement of one population by another of different genetic structure. Studies of such evolutionary processes indicate that changes can be determined in terms of two population parameters, the *growth rate* λ and the *population entropy* H, and a quantity Φ called *reproductive potential* that describes the effect of ecological conditions on the population. In particular, the reproductive potential determines evolutionary trends for the population entropy (cf. [19]): When certain constraints on the resources corresponding to a reproductive potential $\Phi \leq 0$ are satisfied, a directionality principle, which is an analogue of the *second law of thermodynamics*, holds, namely,

$$\tilde{\Delta} H \geq 0. \tag{1.1}$$

Here $\tilde{\Delta} H$ denotes the change in entropy as the population moves from one stationary state to another under the concerted action of mutation and selection. Population energy as introduced by Demetrius [17] is a special case

Received November 1992; revised October 1993.
AMS 1991 *subject classifications*. Primary 28D99, 58F11, 92D15; secondary 60J10, 54H20, 92D25.
Key words and phrases. Evolutionary theory, random dynamical system, products of random matrices, Perron–Frobenius theory, Markov chain in a random environment, thermodynamic formalism, Gibbs measures, variational principle, equilibrium states.

of the Kolmogorov–Sinai invariant and presents a mathematical concept which characterizes the heterogeneity in the birth and death rates of the individuals in the population. Analytically, H describes the rate of decay of fluctuations in population numbers, induced by sampling effects. Populations with nonoverlapping generations are described by $H = 0$; when generations overlap, $H > 0$. The relation (1.1) thus embodies a fundamental observation concerning evolutionary processes: Evolution by mutation and natural selection results in an increased stability of population numbers, if constraints on resources obtain.

The derivation of (1.1) has been carried out for a class of nonlinear dynamical systems—*Mendelian models of age-dependent population growth* (cf. [18], [19]). Three main elements are involved:

1. The characterization of the equilibrium state of the dynamical system in terms of an extremal principle. A consequence of this principle is the relation

$$\lambda = H + \Phi. \tag{1.2}$$

2. A fluctuation theory for the dynamical state at the equilibrium. This represents the mathematical description of the mutation event. If $\Delta\lambda$ and ΔH denote the change in growth rate and entropy due to mutation, then we have

$$\Phi > 0 \Rightarrow \Delta\lambda\,\Delta H < 0, \tag{1.3}$$

$$\Phi < 0 \Rightarrow \Delta\lambda\,\Delta H > 0. \tag{1.4}$$

3. An analysis of competition between dynamic systems. This yields the mathematical description of the selective event, a process which drives the system to a new state. If $\tilde{\Delta} H$ denotes the change in entropy as the population moves from one state to the next and ΔH the change in entropy which characterizes invading mutants, then

$$\Delta H \tilde{\Delta} H > 0. \tag{1.5}$$

The program which this paper presents aims to elucidate the mathematical structure which underlies (1.2)–(1.5) for random and deterministic systems in both discrete and continuous time, and consequently to identify a general class of dynamical systems for which the directionality principle (1.1) holds.

The ideas that underlie (1.2) have their origins in models of equilibrium statistical mechanics whose mathematical structure is described by the term *thermodynamic formalism* (cf. [37]). The notions that generate (1.3)–(1.5) have their origins in models of mutation and selection in evolutionary dynamics. The directionality principle (1.1) rests centrally on (1.3)–(1.5) and hence we use the term *evolutionary formalism* to describe its mathematical basis.

This article derives an analogue of (1.1) for dynamical systems described by products of nonnegative random matrices that satisfy a strong primitivity condition such that one can consider, without loss of generality, products of positive random matrices. The largest Lyapunov exponent and the (fiber) entropy for random matrix products represent the analogues of the growth rate λ and the population entropy H in age-structured populations.

The selective process in evolutionary models may involve competitive interactions leading to the replacement of one population type by another, as in models of asexual populations, or cooperative interactions resulting in a mixed type which replaces the ancestral type, as in Mendelian populations. In cooperative interactions, (1.5) is a consequence of a perturbation analysis of the new stationary state, whereas in competitive interactions, (1.5) can be shown to hold trivially. As the main thrust of this article is to investigate in a canonical example the effect of randomness on the evolutionary process, we will restrict our study to competitive interactions. Consequently our analysis addresses mainly (1.2)–(1.4).

In our derivations of (1.2)–(1.4) for random matrix products a crucial role is played by a random version of an abstract dynamical system $(\mathscr{S}, \mu, \sigma)$, where $\mathscr{S}$ denotes the symbol space, the space of genealogies, σ the shift operator, and μ a shift invariant probability measure on $\mathscr{S}$. Stationary states of dynamical systems described by positive matrix products can be canonically represented in terms of such abstract dynamical systems. We exploit this connection in the variational characterization of λ which leads to (1.2) and in the fluctuation theory for λ and H leading to (1.3) and (1.4).

The class of dynamical systems we consider is the natural generalization of discrete time models which were studied in the context of population dynamics and are defined via *Leslie matrices*. These population models represent a class of dynamical systems where the expressions in (1.2)–(1.4) can be explicitly computed and consequently we will invoke these models to introduce for non-Mendelian interactions the central methods and results of the evolutionary formalism. We develop this in Section 2.

The general mathematical setup for this article is described in Section 3. There we also study products of random positive matrices and derive a random version of the *Perron–Frobenius theorem* which extends earlier results in the literature. In Section 4, we develop a thermodynamic formalism for random matrix products by exploiting the connection between the phase space on which the products of random matrices are defined and the space of genealogies. We derive a variational principle for the maximal Lyapunov exponent and use this to derive (1.2), which also yields a new characterization of the stationary distribution. We show that the latter can be represented as a Gibbs distribution by investigating analyticity properties of the exponent. Based on the ideas of Gundlach and Rand [24] and on the notions and notations of Crauel [15], we weaken the condition of Ruelle [38] for the analytic dependence of the top Lyapunov exponent on the random variables defining the matrices. We investigate first and second directional derivatives of the exponent to establish a fluctuation theory for the products of random positive matrices. In Section 5, we finally apply this theory to special perturbations of the matrices in order to derive the mutation relations manifested in (1.3) and (1.4) and to obtain a directionality principle for the case of competitive interactions.

2. Matrix models in population growth. Let us consider a simple time-discrete model which has its origin in demography and is well known in population biology and economics. It can be obtained by considering a population divided into d age classes. Thus a vector $z(n) = (z_1(n), \dots, z_d(n))$ represents an age distribution of the population at time n. Changes in this age distribution are described by the discrete dynamical system given by

$$z(n+1) = Az(n), \tag{2.1}$$

where $A = (a_{ij})$ is a so-called *Leslie matrix* given by

$$A = \begin{pmatrix} m_1 & m_2 & \cdots & \cdots & m_d \\ b_1 & 0 & \cdots & \cdots & 0 \\ 0 & b_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{d-1} & 0 \end{pmatrix}.$$

The entries m_j represent the number of offspring an individual in age group j at time n contributes to the first age group at time $n+1$, while the quantities b_j denote the proportion of individuals of age j at time n surviving to age $j+1$ at time $n+1$. The Leslie matrix A can be represented by a directed graph, the so-called *life-cycle graph* (see Figure 1), where the nodes (i) represent the age classes, the transitions $(i) \to (i+1)$ represent the aging process and the transitions $(i) \to (1)$ represent the reproductive event. We assume

$$m_j \geq 0, \qquad 0 < b_j \leq 1, \qquad m_d > 0. \tag{2.2}$$

These conditions ensure that the matrix A is *irreducible*, that is, for all $1 \leq i, j \leq d$ there exists $n = n(i,j)$ such that $a_{ij}^{(n)} > 0$ if $A^n = (a_{ij}^{(n)})$. If such n does not depend on i, j, so that $A^n > 0$, we call A *primitive*. This is the case if and only if we require, in addition to (2.2), that there exists $i \leq d-1$ such that $m_i > 0$ and greatest common divisor of $\{j\colon m_j > 0\} = 1$ (cf., e.g., [16]).

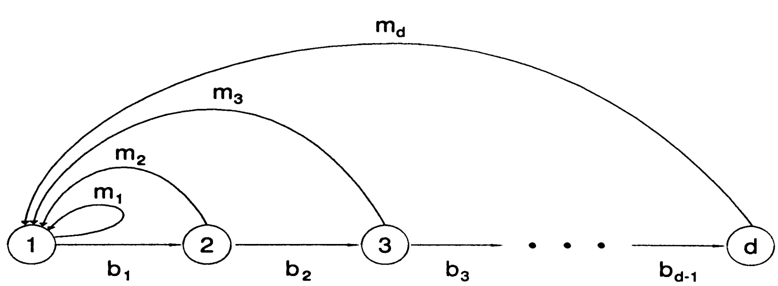

FIG. 1. *Life-cycle graph for a Leslie matrix.*

We assume from now on that A is primitive. Anyway, we conclude from the *Perron-Frobenius theorem* that:

1. A has a simple positive real dominant eigenvalue γ.
2. There exist positive unit vectors $u, v \in \mathbb{R}^d$ such that $Au = \gamma u$ and $A^* v = \gamma v$, where the asterisk (*) denotes the adjoint of a matrix.

The vector $u = (u_i)$ corresponds to the stationary age distribution, while $v = (v_i)$ is a measure of the relative contribution made to the stationary population in the future by the individual age groups;

$$\lambda := \log \gamma$$

is the intrinsic rate of natural increase, which we also call the *population growth rate*. It describes the growth rate of the population number $N_n = \sum_{j=1}^d z_j(n)$, that is, $\lambda = \lim_{n\to\infty}(1/n)\log N_n$ if $N_0 \neq 0$. The elements u_i and v_i can be explicitly expressed in terms of the quantities m_i and b_i. Namely, let us write

$$l_j := \begin{cases} 1, & \text{for } j = 1, \\ \prod_{k=1}^{j-1} b_k, & \text{for } j > 1. \end{cases}$$

The number l_j represents the proportion of individuals surviving to age j, and the dominant eigenvalue γ is the unique positive real root of the equation

$$1 - \sum_{j=1}^{d} \frac{l_j m_j}{\gamma^j} = 0 \tag{2.3}$$

from which we can deduce that the

$$p_j := \frac{l_j m_j}{\gamma^j} \tag{2.4}$$

define a probability distribution for the age of parents. Now put

$$\tau := \sum_{j=1}^{d} j m_j u_j.$$

Then the expressions for u and v are given by

$$u_i = \frac{l_i}{\gamma^i}, \qquad v_i = \frac{\sum_{j=i}^d m_j u_j}{\tau u_i}.$$

Thus we have $m_i u_i = p_i$; hence, $\tau = \sum_{j=1}^d j p_j$, which consequently describes the mean age of parents of all newborns when the equilibrium age distribution is attained. It is also called the *generation time*.

In the ergodic theory of populations we consider the system at steady state characterized by the parameter γ and the vectors u and v. This steady state can be represented in terms of an abstract symbolic dynamical system.

Namely, let

$$X = \prod_{k=0}^{\infty} \{1, \dots, d\}, \qquad X_A = \{x \in X : a_{x_{k+1}x_k} > 0 \text{ for all } k \in \mathbb{N}\}.$$

The set X_A represents the set of all paths x of the life-cycle graph (Figure 1) of the form

$$x = (\dots, x_{-1}, x_0), \qquad x_i \in \{1, \dots, d\}.$$

Such a path is also called *genealogy* as it represents a recording of successive ancestors of a particular individual which at time 0 is in age group x_0.

Let X be equipped with the product of the discrete topology for $\{1, \dots, d\}$ such that X becomes a compact Hausdorff space. The set X_A is a closed subset of X and hence compact. We consider the shift $\sigma: X_A \to X_A$ defined by $(\sigma x)_k = x_{k+1}$ for $k \in \mathbb{N}$, which is a continuous surjection. Note that the dynamics for σ is related to the dynamics given by (2.1), but the two are not conjugated in any way. Considering the meaning of the configuration spaces and the evolution of the states, this is rather obvious: While (2.1) describes the evolution of the age distribution of the population induced by birth and death processes, the shift on X_A is only concerned with the genealogical history of living individuals in the population and consequently corresponds to the dynamics defined by the adjoint of A. Motivated by this fact, we define a potential function φ_A by

$$\varphi_A(x) := \log a_{x_1 x_0}. \tag{2.5}$$

Let M denote the set of all σ-invariant probability measures on X_A and let $H(\mu)$ be the (metric) entropy for the shift σ with respect to $\mu \in M$. As we do not consider any mappings other than σ on X_A, we do not indicate the σ-dependence of $H(\mu)$ in our notations. It follows from Ruelle's thermodynamic formalism [37] that the population growth rate λ satisfies an extremal principle:

$$\lambda = \sup\left\{H(\nu) + \int \varphi_A \, d\nu : \nu \in M\right\} \tag{2.6}$$

and the condition

$$\lambda = H(\mu) + \int \varphi_A \, d\mu \tag{2.7}$$

is satisfied for a unique measure $\mu \in M$. Moreover, this measure is ergodic and a *Gibbs measure* for the potential φ_A. This so-called *equilibrium state* μ, and consequently the quantities $H(\mu)$ and $\int \varphi_A \, d\mu$, can be explicitly computed. We proceed as follows. Consider the transition probability matrix $P = (p_{ij})$ given by

$$p_{ij} := \frac{a_{ji} v_j}{\gamma v_i}.$$

With the help of the probability distribution given by the p_j from above, we can write the matrix P as

$$p_{ij} = \begin{cases} p_i/\sum_{j=i}^d p_j, & \text{for } j = 1, \\ \sum_{j=i+1}^d p_j/\sum_{j=i}^d p_j, & \text{for } j = i+1, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\pi = (\pi_i)$ be the stationary distribution of P defined by $\pi^* P = \pi^*$, which turns out to be

$$\pi_i = \frac{1}{\tau} \sum_{j=i}^{d} p_j.$$

Note that $\pi_i > 0$ for all i, as it follows from (2.4) and (2.2) that $p_d > 0$ and $p_i \geq 0$ for $i \leq d$. Then we can define a Markov measure μ on X_A with the corresponding σ-algebra $\mathscr{B}$ by declaring for all cylinders of any length k,

$$\mu\{x_n = i_n, \ldots, x_{n+k} = i_{n+k}\} := \pi_{i_n} p_{i_n i_{n+1}} \cdots p_{i_{n+k-1} i_{n+k}} \quad \text{for all } n \in \mathbb{N}.$$

This measure is shift-variant and thus we have arrived at a Markov shift $(X_A, \mathscr{B}, \mu, \sigma)$ for which we can compute $H(\mu)$ (cf. 4.27 in [43]) and $\int \varphi_A \, d\mu$ as follows:

$$(2.8) \qquad H(\mu) = -\sum_{i,j} \pi_i p_{ij} \log p_{ij} = -\frac{\sum_{j=1}^d p_j \log p_j}{\sum_{j=1}^d j p_j} = -\frac{1}{\tau} \sum_{j=1}^{d} p_j \log p_j.$$

$$(2.9) \qquad \begin{aligned} \Phi := \int \varphi_A \, d\mu &= \sum_{i,j} \pi_i p_{ij} \log a_{ji} = \frac{\sum_{j=1}^d p_j \log l_j m_j}{\sum_{j=1}^d j p_j} \\ &= \frac{1}{\tau} \sum_{j=1}^{d} p_j \log l_j m_j. \end{aligned}$$

The expression on the right-hand side of (2.8) is called the *population entropy* H, while we refer to the one on the right-hand side of (2.9) as the *reproductive potential* Φ. It follows from (2.8), (2.9) and (2.4) that μ satisfies

$$(2.10) \qquad \lambda = H + \Phi.$$

Hence it must indeed be the unique equilibrium state defined by (2.7).

The parameters λ and Φ may assume both positive and negative values, while H is always nonnegative:

$$\Phi > 0 \quad \Rightarrow \quad \lambda > H, \qquad \Phi < 0 \quad \Rightarrow \quad \lambda < H.$$

The condition $\Phi < 0$, for which the directionality principle holds, refers to populations whose growth rate is constrained, a situation which will occur when resources are limited.

We should note that the extremal pricniple described by (2.6) is formally identical to the principle of the minimization of the free energy in classical statistical mechanics. A consequence of this variational principle is the relation

$$(2.11) \qquad P = E - ST.$$

Here P denotes the free energy, E the mean energy, S the Gibbs–Boltzmann entropy and T the absolute temperature. The relations (2.10) and (2.11) imply the following formal correspondence between the parameters in population theory and thermodynamics:

Population Theory	Thermodynamics
Growth rate λ	Free energy (pressure) P
Reproductive potential Φ	Mean energy E
Generation time τ	Inverse of absolute temperature T^{-1}
Population entropy $-\tau H$	Gibbs–Boltzmann entropy S

Mutation. *Mutations* describe small changes at the genetic level. Such variations will induce changes in the life cycles of the population as described by the population matrix A. In order to understand the evolutionary dynamics of a population, it is central to understand the effect of these mutations on the population growth rate λ and the entropy H. We call the population described by the matrix A the *ancestral type*. Its macroscopic properties are completely characterized by the potential φ_A defined by (2.5), that is, by A. The mutant population will be represented by a perturbation of φ_A, namely,

$$\varphi(\delta) = \varphi_A + \delta f$$

for some $f \in C(X_A)$ and some $\delta \in \mathbb{R}$ of small absolute value. It will turn out that for the formalism we want to develop, such an f has to be cohomologous to φ_A (see Corollary 5.4), whence in particular

$$\int f \, d\mu = \int \varphi_A \, d\mu = \Phi.$$

All possible f's cohomologous to a potential $\varphi_A(x) = \log a_{x_1 x_0}$ are determined in Appendix A. For definiteness, assume here $f = \varphi_A$. Then $\varphi(\delta)$ is just the potential corresponding to a matrix $A(\delta) = (a_{ij}^{1+\delta})$, and considering all the quantities l_j, m_j, p_j, etc. from above as functions of δ, we find that

$$m_j'(0) = m_j \log m_j, \qquad l_j'(0) = l_j \log l_j,$$

and by differentiating (2.3), we obtain

$$\sum_{j=1}^{d} \gamma^{-j}\left[l_j m_j \log m_j + l_j m_j \log l_j\right] = \sum_{j=1}^{d} j\gamma^{-j-1} l_j m_j \gamma'(0)$$

$$= \frac{\gamma'(0)}{\gamma} \sum_{j=1}^{d} j\gamma^{-j} l_j m_j.$$

Consequently, we deduce from the definitions of p_j, τ and Φ that

$$\lambda'(0) = \frac{\gamma'(0)}{\gamma} = \frac{1}{\tau} \sum_{i=1}^{d} p_j \log l_j m_j = \Phi. \tag{2.12}$$

Now it is clear that

$$p_j'(0) = p_j(\log l_j m_j - j\Phi), \qquad \sum_{j=1}^{d} p_j'(0) = 0,$$

$$\tau'(0) = \sum_{j=1}^{d} j p_j(\log l_j m_j - j\Phi).$$

Therefore,

$$\begin{aligned}
H'(0) &= \frac{\tau'(0)}{\tau^2}\sum_{j=1}^{d} p_j \log p_j - \frac{1}{\tau}\sum_{j=1}^{d} p_j'(0)\log p_j \\
&= \frac{\tau'(0)}{\tau^2}\sum_{j=1}^{d} p_j(\log l_j m_j - j\log\gamma) - \sum_{j=1}^{d}\frac{p_j}{\tau}\log l_j m_j \log p_j \\
&\quad + \frac{\Phi}{\tau}\sum_{j=1}^{d} p_j j \log p_j \\
&= \frac{\tau'(0)}{\tau}(\Phi - \log\gamma) \\
&\quad + \sum_{j=1}^{d}\frac{p_j}{\tau}\left[-(\log l_j m_j)^2 + j(\log\gamma + \Phi)\log l_j m_j - j^2\Phi\log\gamma\right] \\
&= \frac{\tau'(0)}{\tau}\Phi - \frac{1}{\tau}\sum_{j=1}^{d} p_j(\log l_j m_j)^2 + \frac{1}{\tau}\sum_{j=1}^{d} p_j j\Phi\log l_j m_j \\
&= \frac{1}{\tau}\sum_{j=1}^{d} p_j\left[2\Phi_j\log l_j m_j - \Phi^2 j^2 - (\log l_j m_j)^2\right]
\end{aligned}$$

and thus

$$H'(0) = -\frac{1}{\tau}\sum_{j=1}^{d}\left[-j\Phi + \log l_j m_j\right]^2 p_j =: -\sigma^2 \le 0. \tag{2.13}$$

Note that $H'(0) = 0$ if and only if $l_j m_j = (\exp\Phi)^j =: c^j$ for all j or, equivalently,

$$m_1 = c, \qquad b_1 m_2 = c^2, \dots, b_1 \dots b_{d-1} m_d = c^d.$$

If we set $m_1 = c$, $m_i = c\gamma_1/\gamma_i$ for $2 \le i \le d$ and some parameter γ_1, then $b_i = c\gamma_{i+1}/\gamma_i$ for $1 \le i \le d-1$ and the matrix A becomes

$$A = c\begin{pmatrix} 1 & \gamma_1/\gamma_2 & \cdots & \cdots & \gamma_1/\gamma_d \\ \gamma_2/\gamma_1 & 0 & \cdots & \cdots & 0 \\ 0 & \gamma_3/\gamma_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \gamma_d/\gamma_{d-1} & 0 \end{pmatrix},$$

which is satisfied, for instance, if all γ_i coincide.

If δ is small enough, we can, on the basis of (2.12) and (2.13), approximate $\Delta\lambda = \lambda((1+\delta)\varphi_A) - \lambda$ by $\delta \cdot \Phi$, $\Delta H = H(\mu_\delta) - H$ by $-\delta\sigma^2$ and, hence, $\Delta\lambda\,\Delta H$ by $-\delta^2\Phi\sigma^2$, from which we conclude that in the case $\sigma^2 > 0$,

$$\begin{aligned} \Phi > 0 \quad &\Rightarrow \quad \Delta\lambda\,\Delta H < 0, \\ \Phi < 0 \quad &\Rightarrow \quad \Delta\lambda\,\Delta H > 0. \end{aligned}$$

In this article we will give a proof for the two so-called *fluctuation relations* in a context more general than the population model. The above can be taken as an introduction to the methods we will invoke in our studies of the fluctuation relations for products of random matrices. For an elementary proof of $H'(0) \le 0$ for a general primitive matrix A, see Appendix A.

Selection. The mutuant type will invade the population if $\Delta\lambda > 0$ and extinction will obtain if $\Delta\lambda < 0$. In order to investigate the possible directionality of evolution, an interaction between the invading mutants and the ancestral types of the population leading to a new dynamical system has to be considered.

In sexual populations, the mutant will mate with the ancestral type and generate a new type (cooperative interaction). The coupling of the three genotypes is determined by the Mendelian laws. In systems described by a single locus with two alleles, the new dynamical system will involve three types A_1A_1, A_1A_2, and A_2A_2. The relative proportion of the three types at equilibrium is determined by the growth rate of each type and the mating laws.

In asexual populations, the ancestral and mutant types evolve independently and the relative proportion of the two types at equilibrium will be determined uniquely by the growth rate (competitive interaction).

In this paper, we restrict our study to linear models of asexual populations. The result of our analysis will be a directionality principle; see (5.7).

3. Random Perron–Frobenius theory. The inclusion of randomness in dynamical systems has lead to the notion of *random dynamical systems* (for an overview, see [1]). Among these systems, products of random maps play a central role, and their long-term behavior is one of the main topics of the subject. One of the earliest results in this field is due to Furstenberg and Kesten (cf. [23]), who considered products of (positive) random matrices. Later their theorem was improved and extended by Oseledets in his celebrated multiplicative ergodic theorem (cf. [35]). In this section we will present a version for products of *positive* random matrices which could be called a random Perron–Frobenius theorem. It contains and extends all the results known to date.

Let $(\Omega, \mathscr{F}, \mathbb{P}, \theta)$ be a dynamical system in the sense of ergodic theory, that is, let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $\theta\colon \Omega \to \Omega$ be a bijection such that θ and θ^{-1} are measurable, which leaves $\mathbb{P}$ invariant. Assume that θ is ergodic. A *random dynamical system* on a measurable space $(X, \mathscr{B})$ over θ is a family $\{\varphi(n, \omega)\colon n \in \mathbb{N}\}$, $\mathbb{N} = \{1, 2, \dots\}$, of measurable transformations on X

satisfying for $\mathbb{P}$-almost all $\omega \in \Omega$ the *cocycle property*

$$\varphi(n+k, \omega) = \varphi(n, \theta^k\omega) \circ \varphi(k, \omega) \quad \text{for all } n, k \in \mathbb{N}.$$

It follows that, with $\varphi(\omega) := \varphi(1, \omega)$,

$$\varphi(n, \omega) = \varphi(\theta^{n-1}\omega) \circ \cdots \circ \varphi(\omega) \quad \text{for all } n \in \mathbb{N},$$

that is, the time-one mapping $\varphi(\omega)$ generates the cocycle.

A φ-invariant measure is a probability measure μ on $X \times \Omega$ with marginal $\mathbb{P}$ on Ω which is invariant under the induced *skew-product transformation*

$$\Theta\colon X \times \Omega \to X \times \Omega, \qquad (x, \omega) \mapsto (\varphi(\omega)x, \theta\omega).$$

If X is a Polish space and since θ is invertible, the φ-invariance of μ is equivalent to

$$\varphi(\omega)\mu_\omega = \mu_{\theta\omega} \quad \text{for } \mathbb{P}\text{-almost all } \omega \in \Omega,$$

where $\{\mu_\omega\colon \omega \in \Omega\}$ is the disintegration of μ with respect to $\mathbb{P}$ given by $\mu(dx, d\omega) = \mu_\omega(dx)\mathbb{P}(d\omega)$ (cf. [15]).

It will turn out to be important that θ is invertible. This assumption is basically without loss of generality due to the existence of a natural extension (for Lebesgue spaces, see [14], page 239). Invertibility of θ will enable us to consider $\varphi(n, \theta^k\omega)$ for $k \in \mathbb{Z}$ (but $n \in \mathbb{N}$), the mapping which goes from time $k \in \mathbb{Z}$ n steps forward to $n + k \in \mathbb{Z}$. In particular, $\varphi(n, \theta^{-n}\omega)$, which goes from $-n$ to 0, will be studied for $n \to \infty$.

We will now study the matrix cocycle in $\mathbb{R}^d$,

$$\phi_n(\omega) := A(\theta^{n-1}\omega) \cdots A(\omega), \qquad n \in \mathbb{N}, \tag{3.1}$$

over θ generated by the random variable $A\colon \Omega \to \mathscr{M}_+$, where $\mathscr{M}_+$ is the semigroup of positive $d \times d$ matrices [a matrix $A = (a_{ij})$ is positive, $A > 0$, if $a_{ij} > 0$ for all $1 \le i, j \le d$, and analogously for vectors]. Put for $A > 0$,

$$m = \min_{1 \le i, j \le d} a_{ij}, \qquad M = \max_{1 \le i, j \le d} a_{ij}.$$

THEOREM 3.1 (Random Perron–Frobenius theorem). *Consider the product of positive random matrices* (3.1) *and assume*

$$\log^+ \frac{1}{m} \in L^1(\mathbb{P}), \qquad \log^+ M \in L^1(\mathbb{P}).$$

Then there is a θ-invariant set $\tilde{\Omega} \subset \Omega$ of full $\mathbb{P}$-measure on which the following holds:

(i) *There exist a unique positive random unit vector u and a positive random scalar q with* $\log q \in L^1(\mathbb{P})$ *such that*

$$A(\omega)u(\omega) = q(\omega)u(\theta\omega).$$

Further, there is a unique invariant splitting

$$\mathbb{R}^d = W(\omega) \oplus \mathbb{R}u(\omega),$$

that is, $A(\omega)W(\omega) \subset W(\theta\omega)$.

(ii) *If* $x \notin W(\omega)$, *then*

$$\lim_{n\to\infty} \frac{1}{n}\log|\phi_n(\omega)x| = \lambda = \mathbb{E}\log q,$$

where λ *is the Furstenberg constant* (*top Lyapunov exponent*). *If* $x \in W(\omega)$, *then*

$$\limsup_{n\to\infty} \frac{1}{n}\log|\phi_n(\omega)x| < \lambda.$$

The statements of Theorem 3.1, which go beyond Oseledets' multiplicative ergodic theorem, are:

1. The top exponent λ is always *simple*.
2. There is an *invariant* complement $\mathbb{R}u(\omega)$ of $W(\omega)$ (the latter can, in general, only be obtained for invertible matrices).

In the proof of Theorem 3.1, we will establish a group of results one could call random Perron–Frobenius theory. Our principal technique will be to study the action of $A \in \mathscr{M}_+$ on S_+^{d-1},

$$S_+^{d-1} := \mathbb{R}_+^d \cap S^{d-1}, \qquad \mathbb{R}_+^d := \{x \in \mathbb{R}^d: x > 0\},$$

where $S^{d-1} := \{x: |x| = 1\}$ is the unit sphere, and to measure distances on S_+^{d-1} by the *Hilbert metric* (projective distance) d, which is defined by

$$d(x, y) := \log \frac{\max_{1\le i\le d}(x_i/y_i)}{\min_{1\le i\le d}(x_i/y_i)}, \qquad x, y \in \mathbb{R}_+^d.$$

LEMMA 3.2. (i) *On* $\mathbb{R}_+^d$, *we have*

$$\begin{aligned} d(x,y) \ge 0, \qquad & d(x,y) = 0 \quad \Leftrightarrow \quad x = \lambda y \quad \textit{for some } \lambda \in \mathbb{R}_+, \\ & d(x,y) = d(y,x), \\ & d(x,y) \le d(x,z) + d(z,y). \end{aligned}$$

(ii) *On* S_+^{d-1}, d *is a finite metric which makes* (S_+^{d-1}, d) *a complete metric space. We have*

$$|x - y| \le \exp(d(x,y)) - 1.$$

In particular, every Cauchy sequence in (S_+^{d-1}, d) *is also a Cauchy sequence in* $(S_+^{d-1}, |\cdot|)$.

For a proof of the above and further facts, we refer to [4], [5] and [40].

The Hilbert metric is important to us because of the following facts.

LEMMA 3.3. *Let*

$$\tau(A) := \sup_{\substack{x, y \in S_+^{d-1} \\ x \neq y}} \frac{d(Ax, Ay)}{d(x, y)}, \qquad A \in \mathscr{M}_+,$$

be Birkhoff's contraction coefficient of A. Then:

(i) $0 \le \tau(A) < 1$.
(ii) $\tau(AB) \le \tau(A)\tau(B)$ *if* $A, B \in \mathscr{M}_+$.

A positive matrix thus automatically contracts all d distances in S_+^{d-1}:

$$d(Ax, Ay) \le \tau(A) d(x, y) < d(x, y).$$

As shown in [46], the Hilbert metric is not a Riemannian metric, but is the only, up to a scalar factor, Finsler metric which makes the action of any positive matrix on S_+^{d-1} a contraction.

Now consider the matrix cocycle (3.1). Using polar coordinates $s = x/|x|$, $r = |x|$ the action of $\phi_n(\omega)$ on $\mathbb{R}^d \setminus \{0\}$, described by the random difference equation $x_{n+1} = A(\theta^n\omega)x_n$, can be split into

$$s_{n+1} = \bar{A}(\theta^n\omega) s_n, \qquad r_{n+1} = |A(\theta^n\omega) s_n| r_n,$$

where $\bar{A}$: $S^{d-1} \to S^{d-1}$ is defined by $s \mapsto \bar{A}s = As/|As|$. In particular, the linear cocycle ϕ_n induces via the canonical projection p: $\mathbb{R}^d \setminus \{0\} \to S^{d-1}$ a nonlinear cocycle $\bar{\phi}_n$ on S^{d-1} given by

$$\bar{\phi}_n(\omega) = \bar{A}(\theta^{n-1}\omega) \circ \cdots \circ \bar{A}(\omega)$$

with $p \circ \phi_n = \bar{\phi}_n \circ p$, which leaves S_+^{d-1} invariant.

LEMMA 3.4. *Let* A: $\Omega \to \mathscr{M}_+$. *Then there exists a* θ-*invariant set* $\tilde{\Omega} \subset \Omega$ *of full* $\mathbb{P}$-*measure such that for all* $\omega \in \tilde{\Omega}$, $x, y \in S_+^{d-1}$,

$$\limsup_{n \to \infty} \frac{1}{n} \log |\bar{\phi}_n(\omega) x - \bar{\phi}_n(\omega) y| \le \limsup_{n \to \infty} \frac{1}{n} \log d(\phi_n(\omega) x, \phi_n(\omega) y)$$
$$\le \mathbb{E} \log \tau(A) \in [-\infty, 0).$$

PROOF. By Lemma 3.2(ii) it suffices to prove the r.h.s. inequality. We have

$$d(\phi_n(\omega) x, \phi_n(\omega) y) \le \tau(\phi_n(\omega)) d(x, y),$$

and by Lemma 3.3(ii),

$$\tau(\phi_n(\omega)) \le \prod_{i=0}^{n-1} \tau(A(\theta^i\omega)).$$

By the ergodic theorem there is a θ-invariant set $\tilde{\Omega} \subset \Omega$ of full $\mathbb{P}$ measure on which

$$\limsup_{n \to \infty} \frac{1}{n} \log d(\phi_n(\omega) x, \phi_n(\omega) y) \le \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \log \tau(A(\theta^i\omega)) = \mathbb{E} \log \tau(A),$$

where $\mathbb{E} \log \tau(A) < 0$ by Lemma 3.3(i). □

Lemma 3.4 says that any two orbits of $\overline{\phi}_n(\omega)$ on S_+^{d-1} cluster exponentially fast forward in time. However, they do not, in general, converge to a common point, but approach a stationary process.

LEMMA 3.5. *Let* A: $\Omega \to \mathscr{M}_+$ *with* $\log^+ d(A(\omega)x_0, x_0) \in L^1(\mathbb{P})$ *for some* $x_0 \in S_+^{d-1}$. *Then there is a* θ-*invariant set* $\tilde{\Omega} \subset \Omega$ *of full* $\mathbb{P}$-*measure* (*not depending on* x_0) *such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log d\big(\overline{\phi}_n(\theta^{-n}\omega)x_0, \overline{\phi}_{n+1}(\theta^{-(n+1)}\omega)x_0\big) \le \mathbb{E}\log\tau(A).$$

In particular, $(\overline{\phi}_n(\theta^{-n}\omega)x_0)$ *is exponentially Cauchy with respect to the projective metric and hence also with respect to the Euclidean metric.*

PROOF. Note first that $\log^+ d(A(\omega)x_0, x_0) \in L^1(\mathbb{P})$ if and only if $\log^+ d(A(\omega)x, x) \in L^1(\mathbb{P})$ for any other (hence for all) $x \in S_+^{d-1}$, since

$$\begin{aligned} d(x, A(\omega)x) &\le d(x, x_0) + d(x_0, A(\omega)x_0) + d(A(\omega)x_0, A(\omega)x) \\ &\le 2d(x, x_0) + d(x_0, A(\omega)x_0). \end{aligned} \tag{3.2}$$

Now

$$\begin{aligned} &d\big(\overline{\phi}_n(\theta^{-n}\omega)x_0, \overline{\phi}_{n+1}(\theta^{-(n+1)}\omega)x_0\big) \\ &\qquad \le \prod_{i=1}^{n} \tau\big(A(\theta^{-i}\omega)\big) d\big(A(\theta^{-(n+1)}\omega)x_0, x_0\big) \end{aligned}$$

so that the lemma follows again from the ergodic theorem and the fact that $\log^+ d(A(\omega)x_0, x_0) \in L^1(\mathbb{P})$ guarantees that on a θ-invariant set of full measure [independent of x_0 by (3.2)],

$$\limsup_{n\to\infty} \frac{1}{n} \log d\big(A(\theta^{-(n+1)}\omega)x_0, x_0\big) = 0. \qquad \square$$

REMARK. We can choose $x_0 = (1/\sqrt{d})(1, \ldots, 1)^* = \mathbb{1}/\sqrt{d}$ and observe that

$$d\left(A\frac{\mathbb{1}}{\sqrt{d}}, \frac{\mathbb{1}}{\sqrt{d}}\right) = \log\frac{\max_i \sum_j a_{ij}}{\min_i \sum_j a_{ij}} \le \log\frac{M}{m}.$$

The integrability condition $\log^+ d(A(\omega)\mathbb{1}/\sqrt{d}, \mathbb{1}/\sqrt{d}) \in L^1(\mathbb{P})$ is certainly fulfilled if $\log^+(M/m) \in L^1(\mathbb{P})$, which is implied by the assumptions of Theorem 3.1.

PROPOSITION 3.6. *Let* A: $\Omega \to \mathscr{M}_+$ *and* $\log^+ d(A(\omega)x_0, x_0) \in L^1(\mathbb{P})$ *for some* $x_0 \in S_+^{d-1}$. *Let*

$$u(\omega) := \lim_{n\to\infty} \overline{\phi}_n(\theta^{-n}\omega)x_0$$

be the S_+^{d-1}-valued random variable which is for $\omega \in \tilde{\Omega}$ (cf. Lemma 3.5) equal to the limit in S_+^{d-1} of the Cauchy sequence $(\overline{\phi}_n(\theta^{-n}\omega)x_0)$. Then:

(i) *u: $\Omega \to S_+^{d-1}$ is measurable with respect to the "past" σ-algebra $\mathscr{F}^-$ generated by the random variables $(A(\theta^{-n}\omega))_{n \in \mathbb{N}}$.*

(ii) *For all $x \in S_+^{d-1}$ and $\omega \in \tilde{\Omega}$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \left| \overline{\phi}_n(\theta^{-n}\omega)x - u(\omega) \right| \le \mathbb{E} \log \tau(A) < 0.$$

(iii) *u is the a.s. unique stationary solution of $\overline{\phi}_n$ in S_+^{d-1} (or $\mu_\omega = \delta_{u(\omega)}$ is the unique $\overline{\phi}$-invariant measure on S_+^{d-1}), that is, for $\omega \in \tilde{\Omega}$,*

$$\overline{\phi}_n(\theta^k\omega)u(\theta^k\omega) = u(\theta^{k+n}\omega) \quad \textit{for all } k \in \mathbb{Z},\ n \in \mathbb{N}.$$

Equivalently,

$$\overline{A}(\omega)u(\omega) = u(\theta\omega)$$

or

$$A(\omega)u(\omega) = q(\omega)u(\theta\omega), \qquad q(\omega) = |A(\omega)u(\omega)| > 0.$$

(iv) *For $\omega \in \tilde{\Omega}$ and all $x \in S_+^{d-1}$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \left| \overline{\phi}_n(\omega)x - u(\theta^n\omega) \right| \le \mathbb{E} \log \tau(A).$$

PROOF. (i) The measurability of u is obvious from the definition and $\overline{\phi}_n(\theta^{-n}\omega) = \overline{A}(\theta^{-1}\omega) \circ \cdots \circ \overline{A}(\theta^{-n}\omega)$.

(ii) It suffices to prove the assertion for the Hilbert distance. For $x = x_0$, Lemma 3.5 gives

$$d\left(\overline{\phi}_n(\theta^{-n}\omega)x_0, u(\omega)\right) \le \sum_{i=n}^{\infty} d\left(\overline{\phi}_i(\theta^{-i}\omega)x_0, \overline{\phi}_{i+1}(\theta^{-(i+1)}\omega)x_0\right).$$

Choosing $\varepsilon > 0$ with $\mathbb{E}\log\tau(A) + \varepsilon < 0$, there is a constant $C(\varepsilon, \omega)$ such that for all $i \in \mathbb{N}$,

$$d\left(\overline{\phi}_i(\theta^{-i}\omega)x_0, \overline{\phi}_{i+1}(\theta^{-(i+1)}\omega)x_0\right) \le C(\varepsilon, \omega)\exp((\mathbb{E}\log\tau(A) + \varepsilon)i).$$

Thus

$$\limsup_{n\to\infty} \frac{1}{n} \log d\left(\overline{\phi}_n(\theta^{-n}\omega)x_0, u(\omega)\right) \le \mathbb{E}\log\tau(A).$$

For arbitrary x, use the triangle inequality and Lemma 3.4.

(iii) We have, as $\overline{A}(\omega)$: $S_+^{d-1} \to S_+^{d-1}$ is continuous, for $\omega \in \tilde{\Omega}$,

$$\begin{aligned}\overline{A}(\omega)u(\omega) &= \overline{A}(\omega)\lim_{n\to\infty}\overline{\phi}_n(\theta^{-n}\omega)x_0 = \lim_{n\to\infty}\overline{\phi}_n(\theta^{-n}\theta\omega)\overline{A}(\theta^{-n}\omega)x_0 \\ &= \lim_{n\to\infty}\overline{\phi}_n(\theta^{-n}\theta\omega)x_0 = u(\theta\omega),\end{aligned}$$

since, putting $s_n := \overline{A}(\theta^{-n}\omega)x_0$,

$$d\left(\overline{\phi}_n(\theta^{-n}\theta\omega)s_n, \overline{\phi}_n(\theta^{-n}\theta\omega)x_0\right) \le \tau\left(\overline{\phi}_n(\theta^{-n+1}\omega)\right)d\left(\overline{A}(\theta^{-n+1}\omega)x_0, x_0\right),$$

and the r.h.s. converges to zero exponentially fast.

(iv) By (iii), $u(\theta^n\omega) = \overline{\phi}_n(\omega)u(\omega)$ on $\tilde{\Omega}$ and

$$d\big(\overline{\phi}_n(\omega)x, \overline{\phi}_n(\omega)\big)u(\omega) \le \tau\big(\overline{\phi}_n(\omega)\big)d(x, u(\omega)).$$

Note that (iv) also gives the a.s. uniqueness of the stationary solution. □

We can now apply everything to the cocycle $\Psi_n(\omega) = A^*(\theta^{-n+1}\omega)\cdots A^*(\theta^{-1}\omega)A^*(\omega) = \phi_n^*(\theta^{-n+1}\omega)$ over θ^{-1} generated by $A^*\colon \Omega \to \mathscr{M}_+$ and obtain the following.

COROLLARY 3.7. *Let* $\log^+ d(A^*(\omega)x_0, x_0) \in L^1(\mathbb{P})$ *for some* $x_0 \in S_+^{d-1}$. *Then there is an a.s. unique random variable* $v\colon \Omega \to S_+^{d-1}$, *which is measurable with respect to the "future" σ-algebra* $\mathscr{F}^+$ *generated by the random variables* $(A(\theta^n\omega))_{n\ge 0}$ *and satisfies on a θ-invariant set of full* $\mathbb{P}$*-measure*

$$A^*(\omega)v(\theta\omega) = q^*(\omega)v(\omega), \qquad q^*(\omega) = |A^*(\omega)v(\theta\omega)| > 0.$$

REMARK. (i) The integrability condition of the corollary is again implied by the one of Theorem 3.1.

(ii) If $(A(\theta^n\omega))_{n\in\mathbb{Z}}$ is i.i.d., then $\mathscr{F}^+$ and $\mathscr{F}^-$ and thus u and v are independent. The latter was also discovered by Kesten and Spitzer ([27], (2.18)).

We can now establish the existence of an invariant splitting.

PROPOSITION 3.8. *Assume that for some* $x_0 \in S_+^{d-1}$,

$$\log^+ d(A(\omega)x_0, x_0) \in L^1(\mathbb{P}), \qquad \log^+ d(A^*(\omega)x_0, x_0) \in L^1(\mathbb{P}).$$

Then there exists a θ-invariant set $\tilde{\Omega} \subset \Omega$ *of full* $\mathbb{P}$*-measure on which the following hold*:

(i) *We have*

$$\frac{q^*(\omega)}{q(\omega)} = \frac{\langle u(\theta\omega), v(\theta\omega)\rangle}{\langle u(\omega), v(\omega)\rangle}. \tag{3.3}$$

(ii) *The* $(d-1)$*-dimensional subspace*

$$W(\omega) := v(\omega)^\perp = \{x \in \mathbb{R}^d\colon \langle x, v(\omega)\rangle = 0\}$$

and the half-spaces

$$W^+(\omega) = \{x \in \mathbb{R}^d\colon \langle x, v(\omega)\rangle > 0\}, \qquad W^-(\omega) = \{x \in \mathbb{R}^d\colon \langle x, v(\omega)\rangle < 0\}$$

are ϕ_n*-invariant,*

$$A(\omega)W^{(\pm)}(\omega) \subset W^{(\pm)}(\theta\omega),$$

and $\ker A(\omega) \subset W(\omega)$. *In particular,*

$$\mathbb{R}^d = W(\omega) \oplus \mathbb{R}u(\omega)$$

is an invariant splitting.

(iii) *The unique invariant measure of* $\overline{\phi}_n$ *in* $W^{\pm}(\omega) \cap S^{d-1}$ *is* $\pm u(\omega)$, *and* $\pm u(\omega)$ *is globally attracting with*

$$\limsup_{n\to\infty} \frac{1}{n} \log \left| \overline{\phi}_n(\omega)x - (\pm u(\theta^n\omega)) \right| \le \mathbb{E} \log \tau(A) < 0$$

for all $x \in W^{\pm}(\omega) \cap S^{d-1}$.

(iv) *For each* $x \in \mathbb{R}^d$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \left| \frac{\phi_n(\omega)x}{q_n(\omega)} - \frac{\langle x, v(\omega)\rangle}{\langle u(\omega), v(\omega)\rangle} u(\theta^n\omega) \right| \le \kappa < 0,$$

where $q_n(\omega) = q(\theta^{n-1}\omega) \cdots q(\omega)$ *is a scalar cocycle and* $\kappa \in [-\infty, 0)$ *is a constant which only depends on* A.

PROOF. We work on the intersection Ω_1 of the two invariant sets of Proposition 3.6 and Corollary 3.7.

(i) Take the inner product of $A(\omega)u(\omega) = q(\omega)u(\theta\omega)$ and $v(\theta\omega)$ to obtain

$$\begin{aligned} q(\omega)\langle u(\theta\omega), v(\theta\omega)\rangle &= \langle A(\omega)u(\omega), v(\theta\omega)\rangle = \langle u(\omega), A^*(\omega)v(\theta\omega)\rangle \\ &= q^*(\omega)\langle u(\omega), v(\omega)\rangle. \end{aligned}$$

(ii) Just use $q^*(\omega) > 0$ and

$$\langle v(\theta\omega), A(\omega)x\rangle = \langle A^*(\omega)v(\theta\omega), x\rangle = q^*(\omega)\langle v(\omega), x\rangle.$$

(iii) Any compact set in $W^+(\omega) \cap S^{d-1}$ is mapped by $\overline{\phi}_{n(\omega)}$ into S_+^{d-1} for a finite $n(\omega)$. After that Proposition 3.6(iv) takes over, showing in particular that there can be no other invariant measure in $W^+(\omega)$.

(iv) First let us fix a vector $x > 0$ and choose $\lambda_n(x, \omega)$ the largest, $\mu_n(x, \omega)$ the smallest positive number with

$$\lambda_n(x,\omega)q_n(\omega)u(\theta^n\omega) \le \phi_n(\omega)x \le \mu_n(x,\omega)q_n(\omega)u(\theta^n\omega). \tag{3.4}$$

Since $\phi_n(\omega) > 0$, those λ_n, μ_n exist and because of Proposition 3.6(iii) they satisfy

$$\lambda_n(x,\omega) \le \lambda_{n+1}(x,\omega) \le \cdots \le \mu_{n+1}(x,\omega) \le \mu_n(x,\omega).$$

Along the lines of Birkhoff ([4], part 7), from which we adapted the method of this proof, one can show that there is a mapping $\rho\colon \Omega \to [0,1)$ [which can be estimated by $\rho(\omega) \le 1 - (m(\omega)/M(\omega))^2$] such that

$$\begin{aligned} \mu_{n+1}(x,\omega) - \lambda_{n+1}(x,\omega) &\le \rho(\theta^n\omega)(\mu_n(x,\omega) - \lambda_n(x,\omega)) \\ &\le \prod_{k=1}^{n} \rho(\theta^k\omega)(\mu_1(x,\omega) - \lambda_1(x,\omega)) \end{aligned}$$

and thus we may deduce from the ergodic theorem that on a θ-invariant set $\tilde{\Omega} \subset \Omega_1$ of full $\mathbb{P}$-measure which is independent of x,

$$\limsup_{n\to\infty} \frac{1}{n} \log(\mu_n(x,\omega) - \lambda_n(x,\omega)) \le \mathbb{E} \log \rho := \kappa < 0.$$

Consequently the limits of $\mu_n(x, \omega)$ and $\lambda_n(x, \omega)$ for $n \to \infty$ exist, are positive and coincide. We denote this limit by $M(x, \omega) > 0$. By (3.4),

$$(\lambda_n(x, \omega) - M(x, \omega))q_n(\omega)u(\theta^n\omega) \le \phi_n(\omega)x - M(x, \omega)q_n(\omega)u(\theta^n\omega) \le (\mu_n(x, \omega) - M(x, \omega))q_n(\omega)u(\theta^n\omega),$$

and thus we have for all $\omega \in \tilde{\Omega}$ and all $x > 0$,

$$\limsup_{n\to\infty} \frac{1}{n}\log\left|\frac{\phi_n(\omega)x}{q_n(\omega)} - M(x, \omega)u(\theta^n\omega)\right| \le \kappa. \tag{3.5}$$

If we give up the restriction $x > 0$, we consider instead $x = x_+ + x_-$ with $(x_+)_i = \max\{x_i, 0\}$ and $(x_-)_i = \min\{x_i, 0\}$ and if we write $\phi_n(\omega)x = \phi_n(\omega)x_+ + \phi_n(\omega)x_-$, we obtain for every $x \in \mathbb{R}^d$ a mapping $M(x, \cdot)$: $\tilde{\Omega} \to \mathbb{R}_+$ that satisfies (3.5) and is positive and linear in x. In particular, we could find such mappings for every x in a basis of $\mathbb{R}^d$, which, by linear extension, yields a mapping $M: \mathbb{R}^d \times \tilde{\Omega} \to \mathbb{R}$ that satisfies (3.5) for all $x \in \mathbb{R}^d$ and for all $\omega \in \tilde{\Omega}$. We will now further determine that mapping M. By multiplying the exponentially fast decaying sequence

$$\frac{\phi_n(\omega)x}{q_n(\omega)} - M(x, \omega)u(\theta^n\omega)$$

by $\langle \cdot, v(\theta^n\omega)\rangle$, we obtain, with the help of part (i) of this proposition, the exponentially fast decaying

$$\left(\frac{\langle x, v(\omega)\rangle}{\langle u(\omega), v(\omega)\rangle} - M(x, \omega)\right)\langle u(\theta^n\omega), v(\theta^n\omega)\rangle,$$

which is only possible if we have almost surely

$$M(x, \omega) = \frac{\langle x, v(\omega)\rangle}{\langle u(\omega), v(\omega)\rangle}. \qquad \square$$

REMARK. (i) An equivalent characterization of the invariance of the splitting $\mathbb{R}^d = W(\omega) \oplus \mathbb{R}u(\omega)$ is

$$A(\omega) \circ \Pi(\omega) = \Pi(\theta\omega) \circ A(\omega),$$

where

$$x \mapsto \Pi(\omega)x = \frac{\langle x, v(\omega)\rangle}{\langle u(\omega), v(\omega)\rangle}u(\omega)$$

is the projection with range $\mathbb{R}u(\omega)$ and kernel $W(\omega)$.

(ii) Proposition 3.8(iv) says that x and $\Pi(\omega)x$ approach each other exponentially fast under the application of $(\phi_n(\omega)/q_n(\omega))$, as

$$\frac{\phi_n(\omega)}{q_n(\omega)}\Pi(\omega)x = \frac{\phi_n(\omega)}{q_n(\omega)}\frac{\langle x, v(\omega)\rangle}{\langle u(\omega), v(\omega)\rangle}u(\omega) = \frac{\langle x, v(\omega)\rangle}{\langle u(\omega), v(\omega)\rangle}u(\theta^n(\omega)).$$

If u is constant (as in the case of a random stochastic matrix), then ϕ_n/q_n converges exponentially fast to Π.

So far we have established the behavior of ϕ_n on S_+^{d-1} and of ϕ_n/q_n. Our final task is the investigation of the growth of orbits $\phi_n(\omega)x$ in $\mathbb{R}^d$. For this we need the stronger integrability conditions of Theorem 3.1.

PROPOSITION 3.9. *Let* $A\colon \Omega \to \mathscr{M}_+$ *and assume* $\log^+ M, \log^+(1/m) \in L^1(\mathbb{P})$. *Then Propositions* 3.6 *and* 3.8 *hold. Moreover,*

(i) $\log q, \log q^*, \log\langle u,v\rangle \in L^1(\mathbb{P})$ *and* $\log q, \log q^*$ *are cohomologous,*

$$\log q^* = \log q + \theta \log\langle u,v\rangle - \log\langle u,v\rangle \qquad \mathbb{P}\text{-}a.s.$$

(ii) *If* $\mathbb{R}^d = W(\omega) \oplus \mathbb{R}u(\omega)$, *we have*

$$x \notin W(\omega) \quad \Rightarrow \quad \lim_{n\to\infty} \frac{1}{n}\log|\phi_n(\omega)x| = \lambda,$$

$$x \in W(\omega) \quad \Rightarrow \quad \limsup_{n\to\infty} \frac{1}{n}\log|\phi_n(\omega)x| < \lambda,$$

where $\lambda = \mathbb{E}\log q$ *is the Furstenberg constant (top Lyapunov exponent).*

PROOF. (i) An elementary estimate gives $m \le q \le \sqrt{d}M$ and

$$-\log^- m = \log^+\frac{1}{m} \le \log q \le \log\sqrt{d} + \log^+ M,$$

which yields $\log q \in L^1(\mathbb{P})$, analogously $\log q^* \in L^1(\mathbb{P})$. Further, $1 \ge \langle u,v\rangle \ge \min_i u_i \ge m/q$, implying $\log\langle u,v\rangle \in L^1(\mathbb{P})$.

(ii) From Proposition 3.8(iv) for each $x \in \mathbb{R}^d$ and $\omega \in \tilde{\Omega}$,

$$\limsup_{n\to\infty}\frac{1}{n}\log\left|\phi_n(\omega)x - \frac{\langle x, v(\omega)\rangle}{\langle u(\omega), v(\omega)\rangle}q_n(\omega)u(\theta^n\omega)\right| \le \kappa + \lambda < \lambda, \tag{3.6}$$

where $\lambda = \mathbb{E}\log q$. Now write any $x \in \mathbb{R}^d$ as

$$x = r(\omega)w(\omega) + \frac{\langle x, v(\omega)\rangle}{\langle u(\omega), v(\omega)\rangle}u(\omega) \quad \text{with } w(\omega) \in W(\omega) \cap S^{d-1}$$

so that

$$\phi_n(\omega)x = r(\omega)\phi_n(\omega)w(\omega) + \frac{\langle x, v(\omega)\rangle}{\langle u(\omega), v(\omega)\rangle}q_n(\omega)u(\theta^n\omega).$$

The first term has growth rate less than λ by (3.6). The second term (and hence the sum) has *exact* growth rate λ provided $\langle x, v(\omega)\rangle \ne 0$, that is $x \notin W(\omega)$. It remains to prove that λ is equal to the Furstenberg constant

$$\Lambda = \lim_{n\to\infty}\frac{1}{n}\log\|\phi_n(\omega)\|$$

[which exists by the subadditive ergodic theorem and the fact that $\log^+ M \in L^1(\mathbb{P})$ is equivalent to $\log^+\|A\| \in L^1(\mathbb{P})$]. Clearly $\lambda \le \Lambda$. For the converse, use

$$\|\phi_n(\omega)\| \le R(\omega) \sum_{i=1}^{d-1} |\phi_n(\omega) w_i(\omega)| + \frac{1}{\langle u(\omega), v(\omega)\rangle} q_n(\omega),$$

where $(w_i(\omega))_{i=1,\dots,d-1}$ is a basis in $W(\omega)$, so $\Lambda \le \max\{\kappa + \lambda, \lambda\} = \lambda$. □

REMARK. (i) For all statements *before* Proposition 3.9, it suffices to assume $\log^+\log(M/m) \in L^1(\mathbb{P})$. For part (ii) of Proposition 3.9, $\log^+\log(M/m) \in L^1(\mathbb{P})$ and $\log^+\log M \in L^1(\mathbb{P})$ would suffice, with $\lambda = \mathbb{E}\log q$, which is possibly $-\infty$. However, we will later need part (i) of Proposition 3.9, which also gives a finite λ. Note that $\log^+ M$, $\log^+(1/m) \in L^1(\mathbb{P})$ imply $\log M, \log m, \log(M/m) \in L^1(\mathbb{P})$ and thus $\log a_{ij} \in L^1(\mathbb{P})$.

(ii) If $A(\omega) \ge 0$, but for some fixed $r \in \mathbb{N}$, $\phi_r(\omega) > 0$, then all the results remain true.

Discussion of the literature. There has been and still is intensive research on products of (positive) random matrices. Cohen [12] has compiled a bibliography up to 1986 which contains several hundred papers. We only comment on those papers which have contributed to a random Perron–Frobenius theory.

Furstenberg and Kesten worked in their classical paper [23] with the assumptions

$$1 \le \frac{M(\omega)}{m(\omega)} \le C < \infty, \qquad \log^+\|A\| \in L^1,$$

and proved (see the Corollary on page 462) that for all $i, j = 1, \dots, d$,

$$\frac{1}{n}\log \phi_n(\omega)_{ij} \to \lambda. \tag{3.7}$$

This follows from our Proposition 3.8(iv) by noting that

$$\begin{aligned} \phi_n(\omega)_{ij} = \langle e_i, \phi_n(\omega) e_j\rangle &= r_j(\omega)\langle e_i, \phi_n(\omega) w(\omega)\rangle \\ &\quad + \frac{v_j(\omega)}{\langle u(\omega), v(\omega)\rangle} q_n(\omega) u_i(\theta^n\omega). \end{aligned} \tag{3.8}$$

The first term has growth rate less than $\lambda = \mathbb{E}\log q < \infty$. The second term has *exact* growth rate λ since $\lim_{n\to\infty} \log u_i(\theta^n\omega) = 0$, which is true because $\log u_i \in L^1$, as

$$1 \ge u_i(\theta\omega) \ge \frac{m(\omega)}{q(\omega)} \ge c_1 \frac{m(\omega)}{M(\omega)} \ge c_2 > 0$$

[$\log(M/m) \in L^1$ would suffice for this step].

Kingman [28] has also obtained (3.7) under the condition $\log a_{ij} \in L^1$ [which implies $\log(M/m) \in L^1$] by utilizing his subadditive ergodic theorem.

The first one whose work deserves the name random Perron–Frobenius theory was Evstigneev [20]. He proved our Proposition 3.6(ii) [except for the rate $\mathbb{E}\log\tau(A)$] and (iii) under the stronger assumption $\log(M/m) \in L^1$. Evstigneev worked with a metric on S_+^{d-1} similar to the Hilbert metric and also used the pull-back to $n = -\infty$.

Kesten and Spitzer [27] investigated convergence in distribution of sequences of products of i.i.d. nonnegative random matrices given by $\Lambda_n(\omega) = A(\omega)\cdots A(\theta^{n-1}\omega)$. In their Lemma 2 they state that u and v exist as limits in distribution and are independent. The link to our results is provided by the observation that in the i.i.d. case (and if we assume w.l.o.g. that θ is invertible) the law of $(\Lambda_n)_{n\in\mathbb{N}}$ is equal to the law of $(\phi_n(\theta^{-n}\cdot))_{n\in\mathbb{N}}$, for which we have Proposition 3.6, which also implies the independence of u and v.

The book by Bougerol and Lacroix [9] contains a few remarks and exercises (on pages 59, 60 and 68) using the Hilbert metric for products of i.i.d. positive matrices. They present the statement of Lemma 3.4 and the fact that $\phi_n(\omega)^*x$ converges in probability to a random variable $\tilde{v}\colon \Omega \to S_+^{d-1}$ which is a weak version of our Corollary 3.7 since $\phi_n(\omega)^* = \Psi_n(\tilde{\theta}^{-n}\tilde{\theta}\omega)$ for $\tilde{\theta} = \theta^{-1}$.

Heyde [25] has observed that, under $1 \le M/m \le C$ and $\log M \in L^1$, $\log \phi_n(\omega)_{ij}$ is well approximated by a sum of stationary random variables [namely, $\sum_{i=0}^{n}\log q(\theta^i\omega)$], which also follows from Proposition 3.8(iv); see (3.8).

Cohn, Nerman and Peligrad [13] proved various limit theorems for $\log \phi_n(\omega)_{ij}$. Their method is based on a representation of nonnegative matrices via stochastic matrices [see also (4.1)].

Ferrero and Schmitt [21] investigated the more general case of products of random Ruelle–Perron–Frobenius operators. They derive generalizations of all our statements, but under conditions where are, if specialized to the matrix case $\psi_\omega(x) = \log a_{x_1x_0}(\omega)$, much stronger than ours. To establish the results concerning the invariant splitting (up to and including our Proposition 3.8), they need $\log M$, $\log m \in L^1$ (condition H3), while $\log^+\log(M/m) \in L^1$ would do. We thus believe that our "elementary" approach is not obsolete, as it also clearly reveals the minimal integrability conditions.

The approach of Ferrero and Schmitt has recently been extended by Bogenschütz and Gundlach [8]. Under the same conditions they show that the results remain valid if the assumption that the random matrices are quadratic is dropped.

Baccelli [2] investigated products of nonnegative random matrices in the semifield in which "addition" is max and "multiplication" is $+$. He stated in his Corollary 5 the existence of $q, u \in L^1$ with $A(\omega)u(\omega) = q(\omega)u(\theta\omega)$, provided $a_{ij} \in L^1$ and a certain event has positive probability.

Wojtkowski [45] gives conditions under which the top Lyapunov exponent of a product of nonnegative random matrices with $|\det A(\omega)| = 1$ $\mathbb{P}$-a.s. is positive.

There is a huge body of literature about products of random *stochastic* matrices (i.e., with row sums 1) as they appear naturally as n-step transition matrices of Markov chains in a stationary random environment. See the survey paper by Orey [34] in which systematic use is made of the Hilbert metric, and an asymptotic theory of products of the form $A(\omega) \cdots A(\theta^{n-1}\omega)$ and $A(\theta^{-n}\omega) \cdots A(\theta^{-1}\omega)$ is developed. Orey reviews in particular the work of Cogburn (see also Cogburn [11]) and Nawrotzki.

In case $A(\omega)$ is stochastic a.s., $u(\omega) \equiv \mathbb{1}/d$ and $q(\omega) = |A(\omega)u(\omega)| \equiv 1$ are nonrandom, and $\lambda = \mathbb{E} \log q = 0$. The splitting is $\mathbb{R}^d = W(\omega) \oplus \mathbb{R}\mathbb{1}$, the projection is $\Pi(\omega) = \langle \cdot , v(\omega) \rangle / \langle \mathbb{1}, v(\omega) \rangle \mathbb{1}$ and Proposition 3.8(iv) gives $\phi_n(\omega) \to \Pi(\omega)$ a.s. exponentially fast with speed $\kappa < 0$. Since $\log^+ M = 0$ and $\log q$, $\log\langle u, v \rangle$ and $\log q^* \in L^1$ automatically, the whole theory including Theorem 3.1 is valid in case of products of random stochastic matrices under the sole condition that $\log^+ d(A^*(\omega)\mathbb{1}, \mathbb{1}) \in L^1$. Berger [3] has given a beautiful proof of $\phi_n(\omega) \to \Pi(\omega)$ which is modeled after the deterministic proof of $A^n \to \Pi$, but relies on the fact that the matrices are i.i.d.

4. Thermodynamic formalism. From Propositions 3.6 and 3.8 we know that under certain conditions there exist globally attracting stationary states for the cocycle $\overline{\phi}_n$. Of course these states are of particular interest from the viewpoint of dynamics. Therefore, we will in the following only be concerned with these states and their changes due to small perturbations of the cocycle.

We know from the population model of Section 2 that the stationary states for the cocycle can in a canonical way be carried over to a related abstract system on a symbol space. The new symbolic system is not a coded version of the original system, though it can be seen as the description of the historical evolution of the system in graph-theoretical terms. In particular it can be used to simulate the multiplication of the adjoint matrices. The configuration space for our symbolic system is known in statistical mechanics as the one for random spin models. Since the works of [41], [10] and [37] it is popular to exploit such connections between dynamical systems and statistical mechanics by adopting methods from the latter to dynamics. This approach is known as *thermodynamic formalism*.

In this section we develop such a theory for products of random positive matrices. In particular, we introduce parameters like *pressure* and *entropy* for random dynamical systems in order to give useful characterizations of the system in terms of a few parameters. The crucial step for this scheme is the construction of a unique *equilibrium state* satisfying a variational principle for the top Lyapunov exponent. We can deduce the existence of this probability measure from our Perron–Frobenius theory, so that we do not need to follow the usual route of statistical mechanics and investigate the related problem of the differentiability properties of the pressure function. Due to an observation of ours that the evolutionary dynamics due to mutation and selection can be mathematically expressed by directional derivatives of the population growth rate, which coincides with the pressure function for random dynamical systems, we do nevertheless examine the analyticity of the pressure. For this purpose we adapt ideas of [24] to the setup of [15]. This leads to analyticity properties of the top Lyapunov exponent under weaker conditions than in [38] and to explicit formulas for directional derivatives of the pressure and the entropy which we need in Section 5.

Variational principle. In the following we consider a fixed random matrix $A\colon \Omega \to \mathscr{M}_+$ for which we assume that the results of Section 3 are valid; that is, we require for such A that $\log^+ M$, $\log^+(1/m) \in L^1(\mathbb{P})$ and consequently also $\log a_{ij} \in L^1(\mathbb{P})$ by a remark from Section 3. Then let us define a random matrix $P(\omega) = (p_{ij}(\omega))$ and a random vector $p(\omega) = (p_i(\omega))$ by

$$p_{ij}(\omega) := \frac{a_{ij}^*(\omega) v_j(\theta\omega)}{q^*(\omega) v_i(\omega)}, \qquad p_i(\omega) := \frac{u_i(\omega) v_i(\omega)}{\langle u(\omega), v(\omega)\rangle}. \tag{4.1}$$

PROPOSITION 4.1. *For almost all $\omega \in \Omega$, $P(\omega)$ is a stochastic matrix and $p(\omega)$ is a probability vector. Furthermore the random variable $p\colon \Omega \to \mathbb{R}_+^d$ is a stationary solution of the cocycle defined by P^*, that is,*

$$P^*(\omega) p(\omega) = p(\theta\omega). \tag{4.2}$$

PROOF. The assertions follow by straightforward calculations and with the help of (3.3). □

Hence the stochastic matrices $P(\omega)$, $\omega \in \Omega$, define a random Markov chain and $(P(\omega), p(\omega))$, $\omega \in \Omega$, induces a random Markov shift, also known as a *Markov chain in a random stationary environment*, on $X = \{1, \dots, d\}^{\mathbb{N}}$. Namely, let us denote by σ the shift map on X, by $\mathscr{B}$ the natural Borel σ-algebra of X and by μ the measure defined via

$$\begin{aligned} \mu_\omega(x_i = y_0, \dots, x_{i+n} = y_n) &= p_{y_0}(\theta^i\omega) \prod_{j=1}^{n} p_{y_{j-1}y_j}(\theta^{i+1}\omega), \\ \mu(dx, d\omega) &= d\mu_\omega(x)\, d\mathbb{P}(\omega). \end{aligned} \tag{4.3}$$

A measure ν on $(X \times \Omega, \mathscr{B} \otimes \mathscr{F})$ with marginal $\mathbb{P}$ on $(\Omega, \mathscr{F})$ is called invariant if it is invariant under the (skew) product transformation $\Theta = (\sigma, \theta)$. From Section 3 we know that this is the case if and only if $\sigma\nu_\omega = \nu_{\theta\omega}$ holds $\mathbb{P}$-a.s. Now the next result is obvious.

LEMMA 4.2. *The probability measure μ on $(X \times \Omega, \mathscr{B} \otimes \mathscr{F})$ is invariant and $(X \times \Omega, \mathscr{B} \otimes \mathscr{F}, \mu, \sigma)$ is a random Markov shift.*

Though the nature of the shift σ is deterministic, in our setup we will refer to it as the random Markov shift. The randomness enters the problem via the random initial probabilities $p(\omega)$ and the random transition probabilities $P(\omega)$. Let us also remark that we could have defined a two-sided shift space $X = \{1, \dots, d\}^{\mathbb{Z}}$, but for the following investigations it will be important to work on the one-sided space. Considering that the original system defined by (3.1) was one-sided, too, this should not cause any headaches. Incidentally, in the same canonical way as for the original system one could extend the shift to one on the two-sided sequence space by using pull-backs.

Let $\mathscr{P}$ be a finite partition of X and denote by $H_\nu(\mathscr{P})$ the usual entropy for a measure ν on $(X, \mathscr{B})$. The entropy of a random shift system can be introduced as the fiber entropy of the skew product of θ and σ (cf. Bogenschütz [6]): $h(\nu) = \sup h(\nu; \mathscr{P})$, where ν is a probability measure invariant under the skew-product transformation, the supremum is taken over all finite partitions of X and

$$h(\nu; \mathscr{P}) = \lim_{n \to \infty} \frac{1}{n} H_{\nu_\omega}\left(\bigvee_{i=0}^{n-1} \sigma^{-i}\mathscr{P} \right) \quad \text{a.s.}$$

We will refer to $h(\nu)$ just as the entropy of σ with respect to the measure ν. We have omitted indicating the dependence of $h(\nu)$ on σ as throughout this paper we do not consider any transformation apart from the shift. Using the standard partition of X into 1-cylinders and the version of the Kolmogorov–Sinai theorem from Bogenschütz [6], one can show analogously to the deterministic case (see, e.g., 4.27 in [43]) that for the measure μ constructed from A by (4.3),

$$h(\mu) = -\int \sum_{i,j} p_i(\omega) p_{ij}(\omega) \log p_{ij}(\omega) \, d\mathbb{P}(\omega).$$

Let us remark that in applications $h(\mu)$ is also called *adaptive value* (cf. 7.1 in [16]) in order to give a physical interpretation to the entropy concept and underline the fact that it is concerned with both the structure of the dynamical system and the uncertainty of the environment. Namely, on the one hand, $h(\mu)$ describes the asymptotic rate of increase of the number of symbol sequences with increasing number of coordinates on the basis of a Shannon–McMillan–Breiman theorem (cf. Theorem 4.2 in [6]); on the other hand, it describes the stability of the equilibrium state due to a large deviation characterization (cf. [29]).

Let us consider the space $C(X)$ of real-valued continuous functions on X equipped with the topology induced by the sup-norm $\|\cdot\|$ and furthermore the space $\mathbb{L}^1(\Omega, C(X))$ of $\mathbb{P}$-integrable random continuous functions on X. On $L^1(\Omega, C(X))$ we will use the norm defined by

$$\|f\|_1 := \int \|f(\omega)\| \, d\mathbb{P}(\omega)$$

and the induced topology which makes $\mathbb{L}^1(\Omega, C(X))$ a Banach space. Here we have used notions and notations from Crauel [15]. Now let us denote

$$\varphi_A(x, \omega) := \log a^*_{x_0 x_1}(\omega) = \log a_{x_1 x_0}(\omega)$$

and call the function $\varphi_A \in \mathbb{L}^1(\Omega, C(X))$ a *potential*. If we define the so-called *reproductive potential* Φ_ν for the probability measure ν on $(X \otimes \Omega, \mathscr{B} \otimes \mathscr{F})$ by

$$\Phi_\nu(A) := \int \varphi_A(x, \omega) \, d\nu(\omega, x),$$

then we obtain

$$\Phi_\mu(A) = \int \sum_{i,j} p_i(\omega) p_{ij}(\omega) \log a^*_{ij}(\omega)\, d\mathbb{P}(\omega).$$

The reproductive potential and the entropy are related via an extremal principle which can be found for general random dynamical systems in [32] and for even more intrinsic shift systems in [7]. For our setup this *variational principle* is as follows.

THEOREM 4.3. *Let μ be the measure constructed above for A. Then $h(\mu) + \Phi_\mu(A) = \lambda$ and μ is the unique invariant probability measure on $(X \otimes \Omega, \mathscr{B} \otimes \mathscr{F})$ satisfying*

$$h(\mu) + \Phi_\mu(A) = \sup\{h(\nu) + \Phi_\nu(A) | \nu \text{ is invariant}\} = \lambda.$$

PROOF. The first assertion is a consequence of the random Perron–Frobenius theorem, namely,

$$\begin{aligned} h(\mu) &+ \Phi_\mu(A) \\ &= \int \sum_{i,j} p_i(\omega) p_{ij}(\omega) \big(\log a^*_{ij}(\omega) - \log p_{ij}(\omega)\big)\, d\mathbb{P}(\omega) \\ &= \int \sum_{i,j} p_i(\omega) p_{ij}(\omega) \big(\log q^*(\omega) + \log v_i(\omega) - \log v_j(\theta\omega)\big)\, d\mathbb{P}(\omega) \\ &= \lambda + \int \Big(\sum_i p_i(\omega) \log v_i(\omega) - \sum_j p_j(\theta\omega) \log v_j(\theta\omega)\Big) d\mathbb{P}(\omega) = \lambda. \end{aligned}$$

The proof of the second assertion is a consequence of the following well known information theory result.

LEMMA. *If $(p_1, \dots, p_d)$ and $(q_1, \dots, q_d)$ are two probability vectors and $p_i > 0$ for $1 \le i \le d$, then*

$$\sum_{i=1}^d q_i \log p_i - \sum_{i=1}^d q_i \log q_i \le 0$$

with equality if and only if $p_i = q_i$ for $1 \le i \le d$.

Let $\alpha = \{C_1, \dots, C_d\}$ be the partition of the shift space X in 1-cylinders and let ν be a shift-invariant probability measure. Then we choose

$$p_i(x, \omega) = \frac{p_{ix_1}(\omega) p_i(\omega)}{p_{x_1}(\theta\omega)}, \qquad q_i(x, \omega) = E_{\nu_\omega}\big(\chi_{C_i} | \sigma^{-1}\mathscr{B}\big)(x),$$

where $E_{\nu_\omega}(\cdot | \sigma^{-1}\mathscr{B})$ denotes the conditional expectation with respect to $\sigma^{-1}\mathscr{B}$ and χ_C denotes the characteristic function corresponding to a set C. Using the conditional entropy $H_{\nu_\omega}(\alpha | \sigma^{-1}\mathscr{B})$ of α with respect to $\sigma^{-1}\mathscr{B}$ we deduce from the lemma by integration that

$$0 \geq \sum_{i=1}^{d} \int_{C_i} \log \frac{p_{x_0 x_1}(\omega) p_{x_0}(\omega)}{p_{x_1}(\theta\omega)} \, d\nu_\omega(x) \, d\mathbb{P}(\omega) + \int H_{\nu_\omega}(\alpha | \sigma^{-1}\mathscr{B}) \, d\mathbb{P}(\omega)$$
$$= \phi_\nu(P^*) + h(\nu)$$

with equality if and only if $p_i(x, \omega) = q_i(x, \omega)$ ν-a.s. This condition is equivalent to

$$\int g(\sigma x) \frac{p_{i x_1}(\omega) p_i(\omega)}{p_{x_1}(\theta\omega)} \, d\nu_\omega(x) = \int g(\sigma x) p_i(x, \omega) \, d\nu_\omega(x)$$
$$= \int g(\sigma x) q_i(x, \omega) \, d\nu_\omega(x)$$
$$= \int_{C_i} g(\sigma x) \, d\nu_\omega(x)$$

for all $\mathscr{B}$-measurable functions g. If we choose in particular g to be the characteristic function of an n-cylinder $C_{a_1, \dots, a_n} = \{x \in X: x_1 = a_1, \dots, x_n = a_n\}$, then we obtain the condition

$$\nu_\omega(C_{i, a_1, \dots, a_n}) = \frac{p_{i a_1}(\omega) p_i(\omega)}{p_{a_1}(\theta\omega)} \nu_{\theta\omega}(C_{a_1, \dots, a_n}).$$

This shows that the measure ν must be Markov and that the transition probabilities for ν must be given by $p_i(x, \omega)$. The corresponding probability vector has to satisfy (4.2), but from Proposition 3.8 we know that such a vector is unique. Hence the resulting measure is unique, that is, equal to μ. Now the assertion of the theorem finally follows from

$$\sup\{h(\nu) + \Phi_\nu(A) | \nu \text{ is invariant}\}$$
$$= \sup\{h(\nu) + \Phi_\nu(P^*) + \lambda | \nu \text{ is invariant}\}$$
$$= h(\mu) + \Phi_\mu(P^*) + \lambda = \lambda. \qquad \square$$

The pressure function. For $A: \Omega \to \mathscr{M}_+(d)$ with $\varphi_A \in \mathbb{L}^1(\Omega, C(X))$ let us put

$$\tilde{\pi}_\sigma(A) := \pi_\sigma(\varphi_A) := \sup\left\{h(\nu) + \int \varphi_A \, d\nu | \nu \text{ is invariant}\right\};$$

$\tilde{\pi}_\sigma(A)$ and $\pi_\sigma(\varphi_A)$ are called the *pressure* or *free energy* of A or φ_A, respectively, for the random shift σ. In order to study perturbations of A and φ_A we have to use the extension of this definition as given in [7].

DEFINITION 4.4. For $f \in L^1(\Omega, C(X))$ we define the pressure (free energy) of f for the shift σ as

$$\pi_\sigma(f) := \sup\left\{h(\nu) + \int f d\nu | \text{is invariant}\right\}.$$

Any probability measure μ satisfying $\pi_\sigma(f) := h(\mu) + \int f\,d\mu$ is called an equilibrium state for f.

In our situation we can combine Proposition 3.9(ii) and Theorem 4.3 to state the following result according to the last definition.

PROPOSITION 4.5. *For the cocycle ϕ_n induced by A we have*:

(i) $\pi_\sigma(\varphi_A) = \lambda$, *where λ is the top Lyapunov exponent of ϕ_n.*
(ii) $\pi_\sigma(\varphi_A) = \lim_{n\to\infty}(1/n)\log\|\phi_n(\omega)\|$ *for almost all $\omega \in \Omega$.*
(iii) *μ from Lemma* 4.2 *is a unique equilibrium state for φ_A.*

The uniqueness of equilibrium states does not hold in general. In [7], [30] and [21] it is shown under some Hölder assumptions for f that one can find such unique probability measures. Originating from statistical mechanics and, in particular, from the theory of phase transitions it is common to reduce the investigation of uniqueness of equilibrium states to the differentiability of the pressure. Though we already know about the uniqueness of our equilibrium state, we need the connection between it and the differentiability of the pressure to find further characterizations of this probability measure. We start with some basic properties of π_σ.

PROPOSITION 4.6. *The pressure π_σ: $L^1(\Omega, C(X)) \to \mathbb{R}$ is Lipschitz continuous with constant* 1, *convex and increasing.*

PROOF. Note that it follows from

$$\left|\sup_{i\in I} a_i - \sup_{i\in I} b_i\right| \le \sup_{i\in I}|a_i - b_i|$$

for any families $(a_i)_{i\in I}, (b_i)_{i\in I}$ of real numbers that

$$|\pi_\sigma(f) - \pi_\sigma(g)| \le |f-g|_1$$

for any $f, g \in \mathbb{L}^1(\Omega, C(X))$. It is also clear that $f(\omega) \le g(\omega)$ a.s. implies

$$\pi_\sigma(f) = \int f\,d\mu_f + h(\mu_f) \le \int g\,d\mu_f + h(\mu_f) \le \int g\,d\mu_g + h(\mu_g) = \pi_\sigma(g)$$

if μ_f, μ_g denote any equilibrium states for f and g, respectively. Moreover, it is obvious for the same reasons that $\pi_\sigma(\lambda f + (1-\lambda)g) \le \lambda\pi_\sigma(f) + (1-\lambda)\pi_\sigma(g)$ for any $\lambda \in [0,1]$. □

As a consequence of this result, one can consider *tangent functionals* to the pressure. Usually one defines for a convex function g: $X \to \mathbb{R}$ a tangent functional ζ: $X \to \mathbb{R}$ at $x_0 \in X$ as a continuous linear mapping that satisfies

$$g(x_0 + y) \ge g(x_0) + \zeta(y) \quad \text{for all } y \in X.$$

Due to the Riesz representation theorem, we prefer the following definition (cf. Section 9.5 in [43] and Section 1 in [44]).

DEFINITION 4.7. A finite signed meausre ν on $(X \otimes \Omega, \mathscr{B} \otimes \mathscr{F})$ is called a tangent functional to π_σ at $f \in \mathbb{L}^1(\Omega, C(X))$ if $\pi_\sigma(f+g) - \pi_\sigma(f) \geq \int g\, d\nu$ for all $g \in \mathbb{L}^1(\Omega, C(X))$.

Since any equilibrium state μ for $f \in \mathbb{L}^1(\Omega, C(X))$ satisfies, for all $g \in \mathbb{L}^1(\Omega, C(X))$,

$$\begin{aligned}\pi_\sigma(f+g) &= \sup\left\{h(\nu) + \int f\, d\nu + \int g\, d\nu \,|\, \nu \text{ is invariant}\right\}\\ &\geq h(\mu) + \int f\, d\mu + \int g\, d\mu = \pi_\sigma(f) + \int g\, d\mu,\end{aligned}$$

it is clear that any equilibrium state for f is also a tangent functional to π_σ at f. Using exactly the same arguments as Walters ([44], Corollary 2 and Theorem 5, and 9.15 in [43]) for the deterministic case and the fact that the entropy map h is an upper semicontinuous function on the space of all shift-invariant probability measures, one can even show a closer connection between tangent functionals to π_σ and equilibrium states.

PROPOSITION 4.8. *For $f \in \mathbb{L}^1(\Omega, C(X))$ the pressure π_σ is Gateaux differentiable at f [i.e., the limit*

$$\lim_{t \to 0} \frac{1}{t}(\pi_\sigma(f+tg) - \pi_\sigma(f)) =: \frac{d}{dt}\pi_\sigma(f+tg)\bigg|_{t=0}$$

exists for all $g \in \mathbb{L}^1(\Omega, C(X))$] if and only if there is a unique tangent functional μ to π_σ at f, and then

$$\frac{d}{dt}\pi_\sigma(f+tg)\bigg|_{t=0} = \int g\, d\mu \quad \text{for all } g \in \mathbb{L}^1(\Omega, C(X)).$$

If, in particular, π_σ is Gateaux differentiable at f and there is a unique equilibrium state μ for f, then μ is the unique tangent functional to π_σ at f.

Thus we have established a connection between equilibrium states and the derivatives of the pressure. In order to calculate the latter explicitly, let us go back now to Proposition 4.5 and present assertion (ii) of that proposition as

$$\pi_\sigma(\varphi_A) = \lim_{n\to\infty} \frac{1}{n} \log Z_n(\varphi_A, \omega) \quad \text{a.s.,} \tag{4.4}$$

where $Z_n(\varphi_A, \omega)$ is defined by

$$Z_n(\varphi_A, \omega) := \sum_{i_0, \ldots, i_{n-1}} \exp\left(\sum_{j=0}^{n-1} \varphi_A(\sigma^j i^*, \theta^j \omega)\right) \tag{4.5}$$

with i^* being the n-periodic member of X starting with $i_0, \dots, i_{n-1}$. That the r.h.s. of (4.4) indeed yields λ can be seen as follows. We have

$$\begin{aligned} Z_n(\varphi_A, \omega) &= \sum_{i_0, \dots, i_{n-1}} \prod_{j=0}^{n-1} a_{i_{j+1} i_j}(\theta^j \omega) = \sum_{i_0} \left(A^*(\omega) \cdots A^*(\theta^{n-1}\omega) \right)_{i_0 i_0} \\ &= \operatorname{tr}(\phi_n^*(\omega)) = \operatorname{tr}(\phi_n(\omega)) = \sum_{i=1}^{d} \langle e_i, \phi_n(\omega) e_i \rangle, \end{aligned}$$

and (as shown in our discussion of the literature) by Proposition 3.9(i) each term on the r.h.s., and hence the sum, has exponential growth rate λ.

It is easy to see that the sequence $\log Z_n$ satisfies almost everywhere the inequality

$$\begin{aligned} \log Z_{n_1 + \cdots + n_k}(\varphi_A, \omega) \le{}& \log Z_{n_k}(\varphi_A, \theta^{n_{k-1} + \cdots + n_1}\omega) + \cdots \\ &+ \log Z_{n_1}(\varphi_A, \omega) \\ &+ c(\theta^{n_1 + \cdots + n_k - 1}\omega) + \cdots + c(\theta^{n_1 - 1}\omega) \\ &\text{for all } n_1, \dots, n_k \in \mathbb{N},\ k \ge 2, \end{aligned} \tag{4.6}$$

where $c\colon \Omega \to \mathbb{R}$, $c \ge 0$ is integrable. Such a "weak" subadditivity property can be used to derive the existence of the limit in (4.4) (cf. [39]). It also holds if we replace φ_A by a mapping f in the space $\mathbb{L}^1_{EHC}(\Omega, C(X))$ of mappings which have equi-Hölder-continuous support. Such mappings f satisfy almost surely

$$|f(\omega)x - f(\omega)y| \le c\alpha^n \quad \text{if } x_i = y_i \text{ for } i \le n \tag{4.7}$$

for some $\alpha = \alpha(f) \in (0,1)$ and $c = c(f) > 0$. They play a major role in a topological introduction of the pressure function which are going to present now. For this purpose, let us define Z_n by (4.5) for all $f \in \mathbb{L}^1(\Omega, \tilde{C}(X))$, where $\tilde{C}(X)$ denotes the space of complex-valued continuous functions on X, and call it the *n-partition function* for f. For each $n \in \mathbb{N}$ it is clear that $\pi_n(f, \omega) := (1/n)\log Z_n(f, \omega)$ is an analytic function in f. We have to examine the limit points of this sequence and their analyticity.

PROPOSITION 4.9. *For every $f \in \mathbb{L}^1(\Omega, C(X))$ the limit of the sequence $(\pi_n(f, \omega))_{n \in \mathbb{N}}$ of analytic functions given by $\pi_n(f, \omega) = (1/n)\log Z_n(f, \omega)$ exists almost surely and equals $\pi_\sigma(f)$.*

PROOF. It follows from (4.7) that for $f \in \mathbb{L}^1_{EHC}(\Omega, C(X))$,

$$\log Z_{n+k}(f, \omega) \le \log Z_n(f, \theta^k\omega) + \log Z_k(f, \omega) + \sum_{j=0}^{n-1} c\alpha^{n-j} + \sum_{j=0}^{k-1} c\alpha^{k-j}$$

a.s. for all $n, k \in \mathbb{N}$ and hence

$$\log Z_{n+k}(f, \omega) \le \log Z_n(f, \theta^k\omega) + \log Z_k(f, \omega) + c_1$$

for some constant $c_1 > 0$. Thus if we define $G_n(f, \omega) = \log Z_n(f, \omega) + c_1$, we can apply Kingman's subadditive ergodic theorem to G_n to deduce the existence of the limit $\lim_{n \to \infty}(1/n)\log Z_n(f, \omega)$ a.s. for all $f \in \mathbb{L}^1_{EHC}(\Omega, C(X))$. Note that for general $f, g \in \mathbb{L}^1(\Omega, C(X))$ one can show analogously to Ruelle ([3.7], 3.3) that $\pi_n(f, \omega)$ is almost everywhere convex and satisfies

$$|\pi_n(f, \cdot) - \pi_n(g, \cdot)|_1 = \int |\pi_n(f, \omega) - \pi_n(g, \omega)|\, d\mathbb{P}(\omega) \le |f - g|_1,$$

where $|\cdot|_1$ denotes the L^1-norm. This inequality and the fact that $\mathbb{L}^1_{EHC}(\Omega, C(X))$ is dense in $\mathbb{L}^1(\Omega, C(X))$ (cf. Lemma 6.2 in [6]) enables us to deduce the existence of $\lim_{n \to \infty} \pi_n(f, \omega)$ for *all* $f \in \mathbb{L}^1(\Omega, C(X))$. This limit coincides with the so-called *topological pressure* for random dynamical systems (cf. 5.2 in [6]). In the last reference this quantity is introduced on the basis of so-called spanning or separated sets. As shown by Ruelle ([37], 7.19) and using the results of Bogenschütz ([6], Section 5), the two ways of introducing topological pressure for random dynamical systems are equivalent for the random shift we are dealing with. This enables us to have recourse to further results of Bogenschütz ([6], Theorem 6.1) like the variational principle for the topological pressure which gives the equality of $\lim_{n \to \infty} \pi_n(f, \omega)$ and $\pi_\sigma(f)$ for $f \in \mathbb{L}^1(\Omega, C(X))$. □

Analyticity of the pressure function. With this topological definition of pressure it is easy to investigate the analyticity of π_σ using a method of Dobrushin, which was adopted from statistical mechanics to dynamical systems in [24].

THEOREM 4.10. *For φ_A there exists a neighborhood V in $\mathbb{L}^1(\Omega, \tilde{C}(X))$ on which the pressure is an analytic function.*

PROOF. We know from the weak subadditivity property (4.6) for φ_A that

$$\left|\log Z_{n_1 + \cdots + n_k}(\varphi_A, \cdot)\right|_1 \le \left|\log Z_{n_1}(\varphi_A, \cdot)\right|_1 + \cdots + \left|\log Z_{n_k}(\varphi_A, \cdot)\right|_1 + k|c|_1$$

for all $n_1, \dots, n_k \in \mathbb{N}$, $k \ge 2$, which guarantees the existence of a constant C_1 such that

$$|\log Z_n(\varphi_A, \cdot)|_1 \le nC_1.$$

Let us now consider $f \in \mathbb{L}^1(\Omega, \tilde{C}(X))$. Then we write

$$Z_n(\varphi_A + f, \omega) = Z_n(\varphi_A, \omega)\tilde{Z}_n^{\varphi_A}(f, \omega),$$

where

$$\tilde{Z}_n^{\varphi_A}(f, \omega) = \sum_{i_0, \dots, i_{n-1}} \exp\left(\sum_{j=0}^{n-1} f(\sigma^j i^*, \theta^j \omega)\right) \exp\left(\frac{\sum_{k=0}^{n-1} f(\sigma^k i^*, \theta^k \omega)}{Z_n(\varphi_A, \omega)}\right),$$

$$\left|\log \tilde{Z}_n^{\varphi_A}(f, \cdot)\right|_1 \le \|f\|_1 n.$$

Thus if we choose a neighborhood U of 0 in $\mathbb{L}^1(\Omega, \tilde{C}(X))$ such that $\|f\|_1 \le C_1$ for all $f \in U$, then

$$\frac{1}{n}|\log Z_n(\varphi_A + f, \cdot)|_1 \le 2C_1. \tag{4.8}$$

Note that $\pi_n(g, \omega)$ is trivially an analytic function of $g \in \mathbb{L}^1(\Omega, \tilde{C}(X))$. It follows from (4.8) and the Cauchy formula for the derivatives of analytic functions that there exist a neighborhood V of φ_A in $\mathbb{L}^1(\Omega, \tilde{C}(X))$ and a constant $C_2 > 0$ such that

$$\left|\frac{d}{dg}\pi_n(g, \cdot)\right|_1 \le C_2.$$

In particular, the $\pi_n(g, \omega)$ form almost surely an equicontinuous family of analytic functions on V. Thus by the theorem of Arzela-Ascoli, $(\pi_n(g, \omega))_{n \in \mathbb{N}}$ has a subsequence which almost surely converges uniformly on every compact subset of V. In fact, all possible limit functions must agree on the set $V \cap \mathbb{L}^1(\Omega, C(X))$ (there we know the existence of the limit); hence, they are all continuations of the same real analytic function and consequently equal. □

This theorem together with Proposition 4.5 gives us an immediate corollary.

COROLLARY 4.11. *The top Lyapunov exponent* $\lambda = \lambda(A)$ *is an analytic function in a neighborhood of* A *in* $\mathbb{L}^1(\Omega, \mathbb{C}^{d \times d})$.

A similar result had been obtained by Ruelle (cf. 3.1 in [38]) for the case that Ω is compact and the matrix A satisfies some general positivity conditions: Then there exists an open subset of $C(\Omega, \mathbb{R}^{d \times d})$ on which the pressure is real and analytic. Our conditions on A concerning the dependence on ω are less restrictive and assure the analyticity of the top exponent in a wider domain. For further regularity results see [33] for the i.i.d. case and A: $\Omega \to GL(d, \mathbb{R})$ (analytic dependence of A on a parameter implies C^∞ dependence of λ) and [26] for i.i.d. positive matrices (C^{k+2} dependence of A on a parameter implies C^k dependence of λ).

THEOREM 4.12. *Let* μ *be the unique equilibrium state for* φ_A. *Then*

(i) μ *is the probability measure which takes the average of mappings in* $\mathbb{L}^1(\Omega, C(X))$ *using almost surely the weights*

$$W_n(i^*, \omega) = W_n^{\varphi_A}(i^*, \omega) = \frac{\exp\left(\sum_{j=0}^{n-1}\varphi_A(\sigma^j i^*, \theta^j\omega)\right)}{Z_n(\varphi_A, \omega)}, \qquad \text{as } n \to \infty;$$

(ii) $$\left.\frac{\partial^2}{\partial s\, \partial t}\pi_\sigma(\varphi_A + sf + tg)\right|_{s=t=0} \ge 0 \quad \text{for all } f, g \in \mathbb{L}^1(\Omega, C(X))$$

and

$$\frac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A+sf+tf)\Big|_{s=t=0} = \mu(f^2)-\mu(f)^2+2\sum_{j=1}^{\infty}\left\{\mu(f\cdot f\circ\sigma^j)-\mu(f)^2\right\};$$

(iii) *for* $f\in\mathbb{L}^1_{EHC}(\Omega,C(X))$, *we have*

$$\frac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A+sf+tf)\Big|_{s=t=0}=0 \quad\Leftrightarrow\quad f=C+H\circ\sigma-H,$$

for some $C\in\mathbb{R}$, $H\in\mathbb{L}^1_{EHC}(\Omega,C(X))$.

PROOF. From Proposition 4.8 we know that the first derivative of π_σ at φ_A defines the unique equilibrium state for φ_A. Since

$$\frac{d}{dt}\pi_\sigma(\varphi_A+tf)\Big|_{t=0} = \lim_{n\to\infty}\frac{1}{n}\frac{1}{Z_n(\varphi_A,\omega)}\sum_{i_0,\dots,i_{n-1}}\exp\left(\sum_{j=0}^{n-1}\varphi_A(\sigma^j i^*,\theta^j\omega)\right)\sum_{k=0}^{n-1}f(\sigma^k i^*,\theta^k\omega)$$

for all $f\in\mathbb{L}^1(\Omega,C(X))$ and almost all $\omega\in\Omega$, the first assertion of the corollary is proven. The second one is a consequence of the convexity of the pressure and the observation that for $f,g\in\mathbb{L}^1(\Omega,C(X))$,

$$\begin{aligned}\frac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A+sf+tg)\Big|_{s=t=0} = \lim_{n\to\infty}\frac{1}{n}\Bigg[&\sum_{i_0,\dots,i_{n-1}}W_n(i^*,\omega)\sum_{p=0}^{n-1}f(\sigma^p i^*,\theta^p\omega)\sum_{k=0}^{n-1}g(\sigma^k i^*,\theta^k\omega)\\ &-\sum_{i_0,\dots,i_{n-1}}W_n(i^*,\omega)\sum_{p=0}^{n-1}f(\sigma^p i^*,\theta^p\omega)\\ &\times\sum_{l_0,\dots,l_{n-1}}W_n(l^*,\omega)\sum_{r=0}^{n-1}g(\sigma^r l^*,\theta^r\omega)\Bigg]\end{aligned}$$

holds almost surely. Note that for each $n\in\mathbb{N}$ and almost all $\omega\in\Omega$ the $W_n(\cdot,\omega)$ define a probability distribution on $\{1,\dots,d\}^{\{0,\dots,n\}}$. So if we denote the corresponding probability measure and the expectation by μ_n and $\mathbb{E}_n$, respectively, and if we introduce random variables X_n, Y_n by

$$X_n(i^*,\omega)=\sum_{j=0}^{n-1}f(\sigma^j i^*,\theta^j\omega),\qquad Y_n(i^*,\omega)=\sum_{j=0}^{n-1}g(\sigma^j i^*,\theta^j\omega),$$

then $\mathbb{P}$-a.s.,

$$\left.\frac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A + sf + tg)\right|_{s=t=0} = \lim_{n\to\infty}\frac{1}{n}[\mathbb{E}_n X_n Y_n - (\mathbb{E}_n X_n)(\mathbb{E}_n Y_n)]$$
$$= \lim_{n\to\infty}\frac{1}{n}\operatorname{cov}_n(X_n, Y_n)$$
$$= \lim_{n\to\infty}\frac{1}{n}[\mathbb{E}_n(X_n - \mathbb{E}_n X_n)(Y_n - \mathbb{E}_n Y_n)]$$

where cov_n denotes the covariance with respect to $\mathbb{E}_n$. Let us point out that the indicated limits exist by Theorem 4.10 and that

$$X_n(i^*, \omega) - \mathbb{E}_n(X_n) = \sum_{j=0}^{n-1}\left[f(\sigma^j i^*, \theta^j\omega) - \frac{1}{n}\sum_{l_0,\dots,l_{n-1}} W_n(l^*, \omega)\sum_{k=0}^{n-1} f(\sigma^k l^*, \theta^k\omega)\right].$$

Thus if we choose $f = g$ and take into consideration that

$$\left(\sum_{j=0}^{n-1} f(\sigma^j i^*, \theta^j\omega)\right)\left(\sum_{j=0}^{n-1} f(\sigma^j i^*, \theta^j\omega)\right) = \sum_{j=0}^{n-1} f(\sigma^j i^*, \theta^j\omega)^2 + 2\sum_{j=0}^{n-1}\sum_{k=1}^{n-1-j} f(\sigma^j i^*, \theta^j\omega) f(\sigma^{j+k} i^*, \theta^{j+k}\omega),$$

then we obtain

$$\left.\frac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A + sf + tf)\right|_{s=t=0}$$
$$= \lim_{n\to\infty}\frac{1}{n}\sum_{i_0,\dots,i_{n-1}} W_n(i^*, \omega)$$
$$\times \sum_{j=0}^{n-1}\left(f(\sigma^j i^*, \theta^j\omega) - \frac{1}{n}\sum_{l_0,\dots,l_{n-1}} W_n(l^*, \omega)\sum_{k=0}^{n-1} f(\sigma^k l^*, \theta^k\omega)\right)$$
$$\times \left(f(\sigma^j i^*, \theta^j\omega) - \frac{1}{n}\sum_{l_0,\dots,l_{n-1}} W_n(l^*, \omega)\sum_{k=0}^{n-1} f(\sigma^k l^*, \theta^k\omega)\right)$$
$$+ \lim_{n\to\infty}\frac{1}{n}2\sum_{i_0,\dots,i_{n-1}} W_n(i^*, \omega)$$
$$\times \sum_{j=0}^{n-1}\sum_{p=1}^{n-1-j}\left(f(\sigma^j i^*, \theta^j\omega) - \frac{1}{n}\sum_{l_0,\dots,l_{n-1}} W_n(l^*, \omega)\sum_{k=0}^{n-1} f(\sigma^k l^*, \theta^k\omega)\right)$$
$$\times \left(f(\sigma^{j+p} i^*, \theta^{j+p}\omega) - \frac{1}{n}\sum_{l_0,\dots,l_{n-1}} W_n(l^*, \omega)\sum_{k=0}^{n-1} f(\sigma^k l^*, \theta^k\omega)\right).$$

So we can deduce from part (i) that

$$\frac{\partial^2}{\partial s\,\partial t}\Pi_\sigma(\varphi_A + sf + tf)\Big|_{s=t=0} = \operatorname{cov}(f,f) + 2\sum_{p=1}^{\infty}\operatorname{cov}(f, f\circ\sigma^p),$$

where cov is the covariance with respect to the expectation defined by μ. Thus part (ii) is also proven. For the proof of the final assertion of this corollary we adopt an idea of Ruelle ([37], Chapter 5, Exercise 5) and note that the convexity of π_σ implies not only $(\partial^2/\partial s\,\partial t)\pi_\sigma(\varphi_A + sf + tg)|_{s=t=0} \geq 0$ for all $f, g \in \mathbb{L}^1(\Omega, C(X))$, but also that the matrix

$$\begin{pmatrix} \dfrac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A + sf + tf)\Big|_{s=t=0} & \dfrac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A + sf + tg)\Big|_{s=t=0} \\ \dfrac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A + sf + tg)\Big|_{s=t=0} & \dfrac{\partial^2}{\partial s\,\partial t}\pi_s(\varphi_A + sg + tg)\Big|_{s=t=0} \end{pmatrix}$$

is positive semidefinite. Consequently,

$$\frac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A + sf + tf)\Big|_{s=t=0} = 0 \quad\Rightarrow\quad \frac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A + sf + tg)\Big|_{s=t=0} = 0$$
$$\text{for all } g \in \mathbb{L}^1(\Omega, C(X)),$$

and in this case one can show with a bit of work that for all $h \in \mathbb{L}^1(\Omega, C(X))$,

$$\begin{aligned} &\frac{\partial^3}{\partial s\,\partial t\,\partial u}\pi_\sigma(\varphi_A + sf + tg + uh)\Big|_{s=t=u=0} \\ &\quad = \frac{d}{du}\pi_\sigma(\varphi_A + uh)\Big|_{u=0}\frac{\partial^2}{\partial s\,\partial t}\pi_\sigma(\varphi_A + sf + tg)\Big|_{s=t=0} \\ &\qquad + \frac{d}{dt}\pi_\sigma(\varphi_A + tg)\Big|_{t=0}\frac{\partial^2}{\partial s\,\partial u}\pi_\sigma(\varphi_A + sf + uh)\Big|_{s=u=0} \\ &\qquad + \frac{d}{ds}\pi_\sigma(\varphi_A + sf)\Big|_{s=0}\frac{\partial^2}{\partial t\,\partial u}\pi_\sigma(\varphi_A + tg + uh)\Big|_{t=u=0} = 0. \end{aligned}$$

Using induction one deduces that for $l \geq 3$ all directional derivatives of the form

$$\frac{\partial^l}{\partial s_1\,\partial s_2\cdots\partial s_l}\pi_\sigma(\varphi_A + s_1 f_1 + s_2 f_2 + \cdots + s_l f_l)\Big|_{s_1 = s_2 = \cdots = s_l = 0}$$

with $f_i \in \mathbb{L}^1(\Omega, C(X))$ can be represented as linear combinations of

$$\frac{\partial^k}{\partial s_1\cdots\partial s_k}\pi_\sigma\big(\varphi_A + s_1 f_{p(1)} + \cdots + s_k f_{p(k)}\big)\Big|_{s_1 = \cdots = s_k = 0}, \qquad 2 \leq k \leq l-1,$$

where p is a permutation of $\{1,\dots,l\}$, and hence must vanish. So we conclude by the analyticity of π_σ that its first derivative at φ_A in direction of f is a constant, if $(\partial^2/\partial s\, \partial t)\pi_\sigma(\varphi_A + sf + tf)|_{s=t=0} = 0$. Then, in particular,

$$s \mapsto \frac{d}{dt}\pi_\sigma(\varphi_A + sf + tf)\bigg|_{t=0} = \mu_{\varphi_A+sf}(f)$$

is a constant. Here $\mu_{\varphi_A+sf}(f)$ denotes the equilibrium state for $\varphi_A + sf$, which is constant in s. Along the lines of Parry and Pollicott ([36], Proposition 3.6) in the deterministic case, one can show in the random case (cf. [8]) that two functions in $\mathbb{L}^1_{EHC}(\Omega, C(X))$ have the same equilibrium state if and only if they are cohomologous via a function in $\mathbb{L}^1_{EHC}(\Omega, C(X))$. Thus φ_A, $\varphi_A + sf$ and $\varphi_A + c$ (where c is a constant) all must be cohomologous and we deduce the existence of a function $H \in \mathbb{L}^1_{EHC}(\Omega, C(X))$ and a constant C such that

$$\frac{\partial^2}{\partial s\, \partial t}\pi_\sigma(\varphi_A + sf + tf)\bigg|_{s=t=0} = 0 \quad \Rightarrow \quad f = C + H \circ \sigma - H,$$

which completes the proof of this corollary. □

The assertions of the corollary can be written down in a shorter and more elegant way if one uses Lanford's characterization of *Gibbs states* (cf. C1 in [31]) and the definition of the *spectral density* S_f for the random stationary sequence $(Z_k^f)_{k\in\mathbb{N}}$ defined by $Z_k^f(x, \omega) = f(\sigma^k x, \theta^k\omega)$ corresponding to a function $f \in \mathbb{L}^1(\Omega, C(X))$ and a Markov chain $(X_n)_{n\in\mathbb{N}}$ *in the random environment given by* P, π, *namely*

$$S_f(y) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\exp(-iyk)\rho_k, \quad \text{where } \rho_{-k} = \rho_k = \operatorname{cov}\big(Z_0^f, Z_k^f\big).$$

COROLLARY 4.13. *If μ is the unique equilibrium state for φ_A, then it is the unique invariant Gibbs state for φ_A and*

$$\frac{\partial^2}{\partial s\, \partial t}\pi_\sigma(\varphi_A + sf + tf)\bigg|_{s=t=0} = 2\pi S_f(0)$$

for any $f \in \mathbb{L}^1(\Omega, C(X))$.

5. An evolutionary principle. We have seen that the macroscopic behavior of dynamical systems described by products of random positive matrices can be represented in terms of the top Lyapunov exponent λ, the entropy H and the reproductive potential Φ, parameters related by $\lambda = H + \Phi$. Small directed perturbations of that dynamical system at equilibrium have an analogy to mutations in evolutionary theory, as seen in Section 2. Such perturbations generate a new system with different macroscopic parameters. Interaction, *competitive* and *cooperative*, between the "ancestral" and the "mutant" dynamical system, an analogy to the selective process in evolutionary theory, will drive the system to a new stationary state characterized by new macroscopic parameters. The description of the change of these parameters due to this mutation–selection process is the object of our *evolutionary algorithm*. Our analysis is restricted to random dynamical systems which are linear. As cooperative interactions in biology usually result in nonlinear systems describing the mixed population which does not fit into the framework which we have developed so far, we prefer to restrict our attention only to competitive interactions, though the selection process becomes trivial in that case.

Mutation relations. Corresponding to evolutionary theory, we split our analysis into one for the mutation and one for the selection process. The original system (describing the ancestral population in biology) is given by the cocycle induced by a fixed random matrix $A\colon \Omega \to \mathscr{M}_+$ satisfying $\log^+ M$, $\log^+(1/m) \in L^1(\mathbb{P})$, such that we can use all the results of Sections 3 and 4. We are interested in changes of our parameters λ and H due to small directional changes of the potential. While Proposition 4.8 describes the change of growth rate λ, we still have to derive an analogous result for the entropy H. For solving this problem, it is our intention to make use of Theorem 4.12. This can be done if one considers in that theorem the Gateaux derivatives in the right directions. This observation is manifested in the following result.

THEOREM 5.1. *For any* $f \in \mathbb{L}^1(\Omega, C(X))$ *one has*

$$\frac{d}{d\delta} h(\mu_\delta)\bigg|_{\delta=0} = -\frac{\partial^2}{\partial\varepsilon\,\partial\delta}\pi_\sigma((1+\varepsilon)\varphi_A + \delta f)\bigg|_{\varepsilon=\delta=0} \le 0, \tag{5.1}$$

where μ_δ *denotes the unique equilibrium state for* $\varphi_A + \delta f$ *and* $\delta \in \mathbb{R}$ *is sufficiently small in absolute value.*

PROOF. Let us consider any mapping $f \in \mathbb{L}^1(\Omega, C(X))$. Then we deduce from Proposition 4.8 that with $\mu_0 = \mu$ being the unique equilibrium state for φ_A,

$$\frac{d}{d\delta}\pi_\sigma(\varphi_A + \delta f)\bigg|_{\delta=0} = \int f\,d\mu. \tag{5.2}$$

On the other hand, we know from Proposition 4.8 and Theorem 4.10 that for $\delta \in \mathbb{R}$ having sufficiently small absolute value,

$$\pi_\sigma(\varphi_A + \delta f) = h(\mu_\delta) + \int \varphi_A\,d\mu_\delta + \delta\int f\,d\mu_\delta,$$

where μ_δ denotes the unique equilibrium state for $\varphi_A + \delta f$ and hence

$$\frac{d}{d\delta}\pi_\sigma(\varphi_A + \delta f)\bigg|_{\delta=0} = \frac{d}{d\delta}h(\mu_\delta)\bigg|_{\delta=0} + \frac{d}{d\delta}\left(\int \varphi_A\,d\mu_\delta\right)\bigg|_{\delta=0} + \int f\,d\mu. \tag{5.3}$$

So comparing (5.2) and (5.3) gives

$$\frac{d}{d\delta}h(\mu_\delta)\bigg|_{\delta=0} = -\frac{d}{d\delta}\left(\int \varphi_A\,d\mu_\delta\right)\bigg|_{\delta=0}. \tag{5.4}$$

Going even one step further, we consider now an additional perturbation of the form $\varepsilon\varphi_A$ for small $\varepsilon \in \mathbb{R}$. Under the same conditions as above we obtain from Theorem 4.8 that

$$\frac{\partial}{\partial\varepsilon}\pi_\sigma((1+\varepsilon)\varphi_A + \delta f)\bigg|_{\varepsilon=0} = \int \varphi_A \, d\mu_\delta$$

and, therefore, it follows from (5.4) that we can differentiate π_σ in the directions of φ_A and f to obtain

$$\frac{\partial^2}{\partial\varepsilon\,\partial\delta}\pi_\sigma((1+\varepsilon)\varphi_A + \delta f)\bigg|_{\varepsilon=\delta=0} = \frac{d}{d\delta}\left(\int \varphi_A \, d\mu_\delta\right)\bigg|_{\delta=0} = -\frac{d}{d\delta}h(\mu_\delta)\bigg|_{\delta=0}.$$

The nonpositivity of those expressions is a consequence of Corollary 4.12. □

Note that also the left-hand side of (5.1) depends on the choice of $f \in \mathbb{L}^1(\Omega, C(X))$. This dependence is manifested in μ_δ and hence in the direction of the Gateaux derivative of the entropy. Due to Corollary 4.13, we can give a further characterization of the derivative of the entropy. Namely, we obtain the following.

COROLLARY 5.2. *If μ_δ denotes the unique equilibrium state for $(1+\delta)\varphi_A$ with $\delta \in \mathbb{R}$ having sufficiently small absolute value, then*

$$\frac{d}{d\delta}h(\mu_\delta)\bigg|_{\delta=0} = -2\pi S_{\varphi_A}(0).$$

For $\delta \in \mathbb{R}$ having sufficiently small absolute value and any fixed $f \in \mathbb{L}^1(\Omega, C(X))$, we define

$$\Delta\lambda := \pi_\sigma(\varphi_A + \delta f) - \pi_\sigma(\varphi_A), \qquad \Delta H := h(\mu_\delta) - h(\mu).$$

It follows from (5.2) that the sign of $\Delta\lambda$ is given by the sign of $\delta \times \int f \, d\mu$, and (5.1) implies that the sign of ΔH is destined by the sign of $-\delta$, if $(\partial^2/\partial\varepsilon\,\partial\delta)\pi_\sigma((1+\varepsilon)\varphi_A + \delta f)|_{\varepsilon=\delta=0}$ does not vanish. Though soon we will consider in Corollary 5.4 a case where we can exclude such a possibility easily by a nice and explicit assumption, it is, in general, hard to examine the last condition. Anyway, one obtains immediately the next result.

THEOREM 5.3. *For any $f \in \mathbb{L}^1(\Omega, C(X))$, one has*

$$\int f \, d\mu > 0 \quad \Rightarrow \quad \Delta\lambda\,\Delta H \le 0,$$

$$\int f \, d\mu < 0 \quad \Rightarrow \quad \Delta\lambda\,\Delta H \ge 0.$$

If S_f does not vanish at 0, then the deduced inequalities are strict.

Of particular interest in evolutionary theory is the following case, which is in accordance with our observations in Section 2.

COROLLARY 5.4. *For any $f \in \mathbb{L}^1(\Omega, C(X))$ cohomologous to φ_A, one has*

$$\Phi_\mu(A) > 0 \quad \Rightarrow \quad \Delta\lambda\,\Delta H \le 0,$$
$$\Phi_\mu(A) < 0 \quad \Rightarrow \quad \Delta\lambda\,\Delta H \ge 0.$$

Except for the case that A is almost surely of the form

$$a_{ij}(\omega) = \gamma \frac{b_i(\omega)}{b_j(\omega)} \tag{5.5}$$

for some $\gamma > 0$ and $b \in \mathbb{L}^1(\mathbb{P})$ with $b > 0$, the implied inequalities are even strict.

PROOF. Due to the assumed cohomology and Theorem 4.12(iii), it only remains to prove that the condition of φ_A being cohomologous to a constant C is equivalent to (5.5). So let us assume that $\varphi_A = C + H \circ \sigma - H$ for some $H \in \mathbb{L}^1_{EHC}(\Omega, C(X))$. As A depends only on the first two coordinates of the elements in X, we can deduce that H only depends on the first one. Thus if we put $\gamma = \exp(C)$ and $b_i(\omega) = \exp(H(i, \omega))$, we are done. □

Selection dynamics. As already mentioned above we will only consider competitive interaction between the ancestral and mutant types in the population. This competitive interaction assumption implies that each type can be considered as increasing at the expense of each other. Under such conditions there exists a priori a directionality principle, namely, the types with larger growth rate will obviously soon dominate the population, while the ones with smaller growth rate will eventually become extinct. Thus

$$\tilde{\Delta}\lambda > 0, \tag{5.6}$$

where $\tilde{\Delta}_\lambda$ represents the change in growth rate as the population moves from one stationary state to another due to the mutation and selection process. If $\tilde{\Delta}H$ is the analogous change in population entropy and ΔH is the change in entropy of the invading mutant, then trivially $\Delta H \tilde{\Delta} H \ge 0$; hence, we can deduce on the basis of Corollary 5.4 that for populations with $\Phi_\mu(A) < 0$ the following *directionality principle* holds:

$$\tilde{\Delta}H > 0. \tag{5.7}$$

Note that both (5.6) and (5.7) determine a direction for evolutionary changes under certain constraints defined by the reproductive potential, but (5.7) is far more general. This can be seen, for example, in the common nonlinear models where the stationary states are characterized by $\lambda = 0$. For continuous deterministic systems of that kind, even with cooperative interaction the directionality principle (5.7) has already been proven for some important examples (cf. [18]).

6. Conclusion. Products of random positive matrices represent a class of dynamical systems which describe processes which occur in population biology and economics. The evolutionary algorithm involving mutation, an event generating new types, and selection, a dynamic ordering of the types, represent a canonical mechanism for inducing change within dynamical systems.

The problem of associating this process with some measurable property that characterizes persistence of the process and that also increases in evolutionary time has been central in both biological evolution and classical economic theory. The interest in such a property is partly philosophical, deriving from the idea that progress and stability must be the inevitable outcome of variation and conflict: mutation and selection in biological systems and innovation and competition in economic systems.

Entropy represents one such measurable property: The mathematical concept describes the stability of the dynamical system and in this sense it reflects the intuitive notion of persistence. This article has shown that a unidirectional increase in stability, as described by the directionality principle $\tilde{\Delta} H > 0$, holds for a subclass of the dynamical systems considered, namely, systems whose stationary states at equilibrium satisfy the condition $\Phi < 0$, that is, slowly growing systems.

Two central assumptions underlie the derivation of this principle:

1. Linearity of the dynamical system.
2. Competitive interaction between ancestral and mutant types.

These two conditions are highly restrictive. Random evolutionary processes in biology and economics are typically described by nonlinear dynamics. In population biology, density-dependent population growth represents a classical case of nonlinearity. The growth models of Solow and Samuelson [42] represented by homogeneous operators of degree 1 represent a well-known example of nonlinearity in economics.

Competitive interactions leading to the elimination of one type by another is also atypical. In biological systems involving sexual reproduction, the mutants that are generated will mate with the ancestral type to produce new types. Depending on the viability of the new types, the system may evolve to what is called a *polymorphic* state in which the ancestral and the mutant types are represented. In economics, the analogue of polymorphic states may occur through cooperation of the ancestral type and the mutant. When interactions are cooperative, the dynamical system will be driven to a new state with properties distinct from the ancestral and the mutants.

A directionality principle has been derived for time-continuous systems in which both nonlinear dynamics and cooperative interactions are considered (cf. [18]). These models are deterministic and refer to biological systems where the nonlinearity is of a special kind and the cooperative interaction has its basis in the Mendelian laws.

The issue of extending our theory to general nonlinear systems with general cooperative interactions applicable to both biology and economics is still open both in the deterministic and the random case. This paper thus constitutes a small step in our general program whose resolution will rest on two main technical developments:

1. A general Perron–Frobenius theory for nonlinear random and deterministic operators.
2. A thermodynamic formalism to characterize the stationary states of these operators.

The work of Fujimoto and Krause [22] in the deterministic case and Bogenschütz and Gundlach [8] in the random case provide a basis for the development of these two directions.

APPENDIX

Elementary proof of $H'(0) \leq 0$ in the deterministic case. Let $A(\delta) = (a_{ij}(\delta)) = (a_{ij}^{1+\delta})$ for δ in a neighborhood of $\delta = 0$, with $A(0) = A \geq 0$ a primitive matrix. The functions $u(\delta)$, $v(\delta)$, $\gamma(\delta)$, $\Phi(\delta)$ and $H(\delta)$ given by

(A.1) $$A(\delta)u(\delta) = \gamma(\delta)u(\delta), \qquad A^*(\delta)v(\delta) = \gamma(\delta)v(\delta),$$

where $\|u(\delta)\| = 1$ and $\langle u(\delta), v(\delta)\rangle = 1$,

$$p_{ij}(\delta) = \frac{a_{ji}(\delta)v_j(\delta)}{\gamma(\delta)v_i(\delta)}, \qquad \pi_i(\delta) = u_i(\delta)v_i(\delta),$$

$$\Phi(\delta) = \sum_{i,j=1}^{d} \pi_i(\delta)p_{ij}(\delta)\log a_{ji}(\delta),$$

$$H(\delta) = -\sum_{i,j=1}^{d} \pi_i(\delta)p_{ij}(\delta)\log p_{ij}(\delta)$$

are analytic with respect to δ in a neighborhood of $\delta = 0$. The definitions yield

$$\Phi(\delta) + H(\delta) = \log\gamma(\delta)$$

and

(A.2) $$\gamma(\delta)H(\delta) = -(1+\delta)\langle A'(\delta)u(\delta), v(\delta)\rangle + \gamma(\delta)\log\gamma(\delta),$$

where $A'(\delta) = (a_{ij}(\delta)\log a_{ij})$. Differentiating (A.1) and (A.2) at $\delta = 0$ gives [writing $\gamma, \gamma', A', u, v$ instead of $\gamma(0), \gamma'(0), A'(0), u(0), v(0)$]

(A.3) $$\gamma H'(0) = -\gamma\rho_0 - 2\big\langle (A' - \gamma')(\gamma - A)^{-1}(A' - \gamma')u, v\big\rangle.$$

In (A.3), ρ_0 is the variance of the random variable $Y_0 = \log a_{x_1x_0}$. Since $\tilde{u} := (A' - \gamma')u \in v^{\perp}$, the invariant subspace in which $\gamma - A$ is invertible, the second term in (A.3) makes sense, and the expansion

$$(\gamma - A)^{-1}\tilde{u} = \frac{1}{\gamma}\sum_{k=0}^{\infty}\left(\frac{A}{\gamma}\right)^k\tilde{u}$$

converges geometrically. Thus

$$\gamma H'(0) = -\gamma\rho_0 - 2\sum_{k=0}^{\infty}\left\langle (A'-\gamma')\left(\frac{A}{\gamma}\right)^k (A'-\gamma')u, v\right\rangle.$$

Let now $X_0, X_1, \ldots$ be the stationary Markov chain given by the coordinate variables in $\{1,\ldots,d\}^{\mathbb{N}}$ with measure μ generated by π and P. Then $Y_k := \log a_{X_{k+1}X_k}$, $k = 0, 1, \ldots$, defines a stationary sequence. A lengthy but elementary calculation gives

$$\left\langle (A'-\gamma')\left(\frac{A}{\gamma}\right)^k (A'-\gamma')u, v\right\rangle = \gamma^2\rho_{k+1}, \quad k = 0, 1, \ldots,$$

where

$$\rho_k = \mu(Y_k - \mu(Y_k))(Y_0 - \mu(Y_0))$$

are the covariances of (Y_k). Hence,

$$H'(0) = -\rho_0 - 2\sum_{k=1}^{\infty}\rho_k.$$

Remember now that the *spectral density* S of (Y_k) is given by

$$S(y) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\exp(-iyk)\rho_k, \qquad \rho_{-k} = \rho_k,$$

which is a bounded smooth function due to the geometric convergence of the series $\sum_{k=1}^{\infty}\rho_k$. Thus, finally,

$$H'(0) = -2\pi S(0) \leq 0.$$

In the same way as in Section 4, for the random case we could look for an alternative condition for $H'(0) = 0$ and hence $S(0) = 0$. Namely, we obtain analogously to the random case

$$H'(0) = 0 \quad \Leftrightarrow \quad \varphi_A = C + H\circ\sigma - H \tag{A.4}$$

for a constant C and a continuous function H on X_A which depends only on the first coordinate of the sequences. Thus (A.4) gives a formula for the nonvanishing elements of A. So we can conclude, in accordance with the result obtained for the Leslie matrix in Section 2, that primitive matrices A with vanishing $S(0)$ must be of the form

$$a_{ij} = 0 \quad \text{or} \quad a_{ij} = c\frac{b_i}{b_j}$$

for some positive numbers c, b_i.

Acknowledgment. We would like to thank Thomas Bogenschütz for many helpful discussions. In particular, we are indebted to him for his help in introducing topological pressure by showing us the denseness of $\mathbb{L}^1_{EHC}(\Omega, C(X))$ in $\mathbb{L}^1(\Omega, C(X))$.

REFERENCES

[1] ARNOLD, L. and CRAUEL, H. (1991). Random dynamical systems. *Lyapunov Exponents. Proceedings of a Conference in Oberwolfach 1990. Lecture Notes in Math.* **1486**. Springer, Berlin.

[2] BACCELLI, F. (1992). Ergodic theory of stochastic Petri networks. *Ann. Probab.* **20** 375–396.

[3] BERGER, M. A. (1992). Random affine iterated function systems: Curve generation and wavelets. *SIAM Rev.* **34** 361–385.

[4] BIRKHOFF, G. (1957). Extensions of Jentzsch's theorem. *Trans. Amer. Math. Soc.* **85** 219–227.

[5] BIRKHOFF, G. (1967). *Lattice Theory*, 3rd. ed. Amer. Math. Soc., Providence, RI.

[6] BOGENSCHÜTZ, TH. (1992). Entropy, pressure, and a variational principle for random dynamical systems. *Random Comput. Dynamics* **1** 99–116.

[7] BOGENSCHÜTZ, TH. (1993). Equilibrium States for random dynamical systems. Dissertation, Univ. Bremen.

[8] BOGENSCHÜTZ, TH. and GUNDLACH, V. M. (1993). Ruelle's transfer operator for random subshifts of finite type. Preprint, Univ. Bremen.

[9] BOUGEROL, P. and LACROIX, J. (1985). *Products of Random Matrices with Applications to Schrödinger Operators.* Birkhäuser, Boston.

[10] BOWEN, R. (1975). *Equilibrium states and the Ergodic Theory of Anosov Diffeomorphisms. Lecture Notes in Math.* **470**. Springer, Berlin.

[11] COGBURN, R. (1986). On products of random stochastic matrices. *Contemp. Math.* **50** 199–213.

[12] COHEN, J. (1986). Products of random matrices and related topics in mathematics and science—A bibliography. *Contemp. Math.* **50** 337–358.

[13] COHN, H., NERMAN, O. and PELIGRAD, M. (1993). Weak ergodicity and products of random matrices. *J. Theoret. Probab.* **6** 389–405.

[14] CORNFELD, R. I. P., FOMIN, S. V. and SINAI, YA. G. (1982). *Ergodic Theory.* **245**. Springer, Berlin.

[15] CRAUEL, H. (1986). Lyapunov exponents and invariant measures of stochastic systems on manifolds. *Lyapunov Exponents. Proceedings of a Workshop in Bremen 1984. Lecture Notes in Math.* Springer, Berlin.

[16] DEMETRIUS, L. (1971). Primitivity conditions for growth matrices. *Math. Biosci.* **12** 53–58.

[17] DEMETRIUS, L. (1983). Statistical mechanics and population biology. *J. Statist. Phys.* **30** 709–753.

[18] DEMETRIUS, L. (1992). Growth rate, population entropy and evolutionary dynamics. *Theoret. Population Biol.* **4** 208–239.

[19] DEMETRIUS, L. (1992). Thermodynamics of evolution. *Phys. A* **189** 417–439.

[20] EVSTIGNEEV, I. V. (1974). Positive matrix cocycles over dynamical systems. *Uspekhi Mat. Nauk* **29** 219–220.

[21] FERRERO, P. and SCHMITT, B. (1988). Produits aléatoires d'opérateurs matrices de transfert. *Probab. Theory Related Fields* **79** 227–248.

[22] FUJIMOTO, T. and KRAUSE, U. (1988). Asymptotic properties for inhomogeneous iterations of nonlinear operators. *SIAM J. Math. Anal.* **19** 841–853.

[23] FURSTENBERG, H. and KESTEN, H. (1960). Products of random matrices. *Ann. Math. Statist.* **31** 457–489.

[24] GUNDLACH, V. M. and RAND, D. A. (1993). Spatio-temporal chaos. 2. Unique Gibbs states for higher-dimensional symbolic systems. *Nonlinearity* **6** 201–213.

[25] HEYDE, C. C. (1985). An asymptotic representation for products of random matrices. *Stochastic Process. Appl.* **20** 307–314.

[26] HENNION, H. (1991). Dérivabilité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes à coefficients positifs. *Ann. Inst. H. Poincaré Probab. Statist.* **27** 27–59.

[27] KESTEN, H. and SPITZER, F. (1984). Convergence in distribution of products of random matrices. *Z. Wahrsch. Verw. Gebiete* **67** 363–386.

[28] KINGMAN, J. F. C. (1973). Subadditive ergodic theory. *Ann. Probab.* **1** 883–909.

[29] KIFER, YU. (1991). Large deviations for random expanding maps. *Lyapunov Exponents. Proceedings of a Conference in Oberwolfach 1990. Lecture Notes in Math.* **1486**. Springer, Berlin.

[30] KIFER, YU. (1991). Equilibrium states for random expanding transformations. Random Comput. **1**, 1–31.

[31] LANFORD, O. E. (1973). Entropy and equilbrium states in classical statistical mechanics. *Statistical Mechanics and Mathematical Problems. Lecture Notes in Phys.* **20** 1–113. Springer, Berlin.

[32] LEDRAPPIER, F. (1974). Principe variationnel et systémes dynamiques symboliques. *Z. Wahrsch. verw. Gebiete* **30** 185–202.

[33] LE PAGE, É. (1989). Régularité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes et applications. *Ann. Inst. H. Poincaré Probab. Statist.* **25** 109–142.

[34] OREY, S. (1991). Markov chains with stochastically stationary transition probabilities. *Ann. Probab.* **19** 907–928.

[35] OSELEDETS, V. I. (1968). A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.* **19** 197–231.

[36] PARRY, W. and POLLICOTT, M. (1990). *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics. Astérisque*, **187–188**.

[37] RUELLE, D. (1978). Thermodynamic formalism. *Encyclopedia of Mathematics and Its Applications* **5**. Addison-Wesley, Reading, MA.

[38] RUELLE, D. (1979). Analyticity properties of the characteristic exponents of random matrix products. *Adv. in Math.* **32** 68–80.

[39] SCHÜRGER, K. (1991). Almost subadditive extensions of Kingman's ergodic theorem. *Ann. Probab.* **19** 1575–1586.

[40] SENETA, E. (1981). *Non-negative matrices and Markov chains*, 2nd ed. Springer, New York.

[41] SINAI, YA. G. (1972). Gibbs measures in ergodic theory. *Russian Math. Surveys* **27** 21–69 (translation from *Uspekhi Math. Nauk.* **27** 21–64).

[42] SOLOW, R. SAMUELSON, P. A. (1953). Balanced growth under constant return to scale. *Econometrica* **21**.

[43] WALTERS, P. (1982). *An Introduction to Ergodic Theory*. Springer, New York.

[44] WALTERS, P. (1991). Differentiability properties of the pressure of a continuous transformation on a compact metric space. Warwick Preprint.

[45] WOJTKOWSKI, M. (1985). Invariant families of cones and Lyapunov exponents. *Ergodic Theory Dynamical Systems* **5** 145–161.

[46] WOJTKOWSKI, M. (1986). On uniform contraction generated by positive matrices. *Contemp. Math.* **50** 109–118.

LUDWIG ARNOLD
VOLKER MATTHIAS GUNDLACH
INSTITUT FÜR DYNAMISCHE SYSTEME
UNIVERSITÄT, POSTFACH 330 440
28334 BREMEN
GERMANY

LLOYD DEMETRIUS
DEPARTMENT OF ORGANISMIC AND
EVOLUTIONARY BIOLOGY
HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS 02138