On the Prize Work of O. Sarig on Infinite Markov Chains and Thermodynamic Formalism

Yakov Pesin

Pennsylvania State University

October 19, 2013

List of Papers

- Thermodynamic Formalism for Countable Markov shifts. Erg. Th. Dyn. Sys. 19, 1565-1593 (1999).
- Phase Transitions for Countable Topological Markov Shifts. Commun. Math. Phys. 217, 555-577 (2001).
- Characterization of existence of Gibbs measures for Countable Markov shifts. Proc. of AMS. 131 (no. 6), 1751-1758 (2003).

Lecture notes: *Thermodynamic Formalism for countable Markov shifts*, Penn State, Spring 2009.

The Gibbs Distribution

Thermodynamic formalism, i.e., the formalism of equilibrium statistical physics, originated in the work of Boltzman and Gibbs and was later adapted to the theory of dynamical systems in the classical works of Ruelle, Sinai and Bowen.

Consider a system $\mathcal{A}$ of n-particles. Each particle is characterized by its position and velocity. A given collection of such positions and velocities over all particles is called a state. We assume somewhat unrealistically that the set of all states is a finite set $X = \{1, \ldots, n\}$ and we denote by E_i the energy of the state i. We assume that particles interact with a *heat bath* $\mathcal{B}$ so that

- $\mathcal{A}$ and $\mathcal{B}$ can exchange energy, but not particles;
- $\mathcal{B}$ is at equilibrium and has temperature T;
- $\mathcal{B}$ is much larger than $\mathcal{A}$, so that its contact with $\mathcal{A}$ does not affect its equilibrium state.

Since the energy of the system is not fixed every state can be realized with a probability p_i given by the Gibbs distribution

$$p_i = \frac{1}{Z(\beta)} e^{-\beta E_i}, \text{ where } Z(\beta) = \sum_{i=1}^{N} e^{-\beta E_i},$$

$\beta = \frac{1}{\kappa T}$ is inverse temperature and κ is Boltzman's constant. It is easy to show that the Gibbs distribution maximizes the quantity $H - \beta E = H - \frac{1}{\kappa T} E$, where

$$H = -\sum_{i=1}^{N} p_i \log p_i$$

is the entropy of the Gibbs distribution and

$$E = \sum_{i=1}^{N} (\beta E_i) p_i = \int_X \varphi \, d(p_1, \dots, p_n)$$

is the average energy, where $\varphi(i) = \beta E_i$ is the potential.

In other words, the Gibbs distribution minimizes the quantity $E - \kappa TH$ called the free energy of the system.
The principle that nature maximizes entropy is applicable when energy is fixed, otherwise nature minimizes the free energy.

A substantial generalization of this example, a *"far cry"*, is the following result of Parry and Bowen.

Let (Σ_A^+, σ) be a (one-sided) subshift of finite type. Here $A = (a_{ij})$ is a transition matrix ($a_{ij} = 0$ or 1, no zero columns or rows),

$$\Sigma_A^+ = \{x = (x_n) \colon a_{x_n x_{n+1}} = 1 \text{ for all } n \geq 0\}$$

and σ is the shift. We assume that A is irreducible (i.e., $A^N > 0$ for some $N > 0$ and all $n \geq N$) implying σ is topologically transitive. Consider a Hölder continuous function (potential) φ on Σ_A^+.

Theorem

There exist a unique σ-invariant Borel probability measure μ on Σ_A^+ and constants $C_1 > 0$, $C_2 > 0$ and P such that for every $x = (x_i) \in \Sigma_A^+$ and $m \geq 0$,

$$C_1 \leq \frac{\mu(\{y = (y_i) \colon y_i = x_i, i = 0, \ldots, m\})}{\exp\left(-Pm + \sum_{k=0}^{m-1} \varphi(\sigma^k(x))\right)} \leq C_2.$$

$\mu = \mu_\varphi$ is a Gibbs measure and $P = P(\varphi)$ the topological pressure.

Ruelle's Perron-Frobenius Theorem

The proof of this theorem is based on Ruelle's version of the classical Perron–Frobenius theorem. Given a continuous function φ on Σ_A^+, define a linear operator $\mathcal{L} = \mathcal{L}_\varphi$ on the space $C(\Sigma_A^+)$ by

$$(\mathcal{L}_\varphi f)(x) = \sum_{\sigma(y)=x} e^{\varphi(y)} f(y).$$

$\mathcal{L}_\varphi$ is called the Ruelle operator and it is a great tool in constructing and studying Gibbs measures. Note that for all $n > 0$

$$(\mathcal{L}_\varphi^n f)(x) = \sum_{\sigma^n(y)=x} e^{\Phi_n(y)} f(y).$$

Theorem

Let φ be a Hölder continuous function on Σ_A^+. Then there exist $\lambda > 0$, a continuous positive function h and a Borel measure ν s.t.

1. $\mathcal{L}_\varphi h = \lambda h$ *and* $\int_{\Sigma_A^+} h\, d\nu = 1$ *(i.e., h is a normalized eigenfunction for the Ruelle operator);*
2. $\mathcal{L}_\varphi^* \nu = \lambda \nu$*;*
3. *for every $f \in C(\Sigma_A^+)$*

$$\lambda^{-n} \mathcal{L}_\varphi^n(f)(x) \to h(x) \int f\, d\nu \ \text{ as } n \to \infty \tag{1}$$

   uniformly in x.

4. *the measure $\mu_\varphi = h\, d\nu$ is a σ-invariant Gibbs measure for φ, which is ergodic (in fact, it is Bernoulli).*

One can show that the rate of convergence in (**??**) is exponential implying that μ_φ has exponential decay of correlations (with respect to the class of Hölder continuous functions on Σ_A^+) and satisfies the Central Limit Theorem.

he measure ν in the above theorem has an important property of being conformal. Given a potential φ on Σ_A^+, we call a Borel probability measure μ on Σ_A^+ (which is not necessarily invariant under the shift) conformal (with respect to φ) if for some constant λ and almost every $x \in \Sigma_A^+$

$$\frac{d\mu}{d\mu \circ \sigma}(x) = \lambda^{-1} \exp \varphi(x).$$

In other words, log Jacobian of μ is $\log \lambda - \varphi$.

One can show that the relation $\mathcal{L}_\varphi^* \nu = \lambda \nu$ is equivalent to the fact that ν is a conformal measure for φ.

The Topological Pressure

The topological pressure $P(\varphi)$ of a continuous potential φ is given by

$$P(\varphi) = \lim_{m\to\infty} \frac{1}{m} \log Z_m(\varphi),$$

where

$$Z_m(\varphi) = \sum_{[x_0x_1\ldots x_{m-1}]} \exp\Big(\sup_{x\in[x_0x_1\ldots x_{m-1}]} \Phi_n(x)\Big),$$

$[x_0x_1 \ldots x_{m-1}]$ is a cylinder and

$$\Phi_n(x) = \sum_{k=0}^{m-1} \varphi(\sigma^k(x))$$

is the n-th ergodic sum of φ.

The Variational Principle

One of the fundamental results in thermodynamics of dynamical system is the Variational Principle for the topological pressure:

Theorem

For every continuous potential φ

$$P(\varphi) = \sup\left\{ h_\mu(f) + \int_{\Sigma_A^+} \varphi \, d\mu \right\},$$

where the supremum is taken over all σ*-invariant Borel probability measures on* Σ_A^+.

Equilibrium Measures

Given a continuous potential φ, a σ-invariant measure $\mu = \mu_\varphi$ on Σ_A^+ is called an equilibrium measure if

$$P(\varphi) = h_{\mu_\varphi} + \int_{\Sigma_A^+} \varphi \, d\mu_\varphi.$$

Theorem

If φ is Hölder continuous, then the Gibbs measure μ_φ in the Ruelle's Perron–Frobenius theorem is the unique equilibrium measure for φ. Moreover, $\log \lambda = P(\varphi)$.

Two-sided subshifts

Many results in thermodynamics of one-sided shubshifts can be extended to two-sided subshifts (Σ_A, σ) where

$$\Sigma_A = \{x = (x_n) \colon a_{x_n x_{n+1}} = 1 \text{ for all } n \in \mathbb{Z}\}$$

and σ is the shift. This is based on results by Sinai and Bowen.

Theorem

Given a Hölder continuous potential φ on Σ_A there are Hölder continuous functions h on Σ_A and ψ on Σ_A^+ such that for every $x = (x_n) \in \Sigma_A$

$$\varphi(x) + h(x) - h(\sigma(x)) = \psi(x_0 x_1 \dots).$$

This equation means that the potentials φ and ψ are cohomologous. Two cohomologous potentials have the a same set of Gibss measures.

Subshifts of countable type

We now move from subshifts of finite type to subshifts of countable type $(X = \Sigma_A^+, \sigma)$ where A is a transition matrix on a countable set S of states and σ is the shift. The Borel σ-algebra $\mathcal{B}$ is generated by all cylinders. The main obstacle in constructing equilibrium measures in this case is that the space X is not compact and hence, the space of probability measures on X is not compact either and new methods are needed.

A Bit of History

- Dobrushin, Landford and Ruelle (mid 1960th) who introduced what is now called DLR measures which characterize Gibbsian distributions in terms of families of conditional probabilities.
- Gurevic (early 1970th) who studied the topological entropy (the case $\varphi = 0$) and obtained the variational principle for the topological entropy; Vere-Jones (1960th) who studied recurrence properties that are central for constructing Gibbs measures; both Gurevic and Vere-Jones assumed that the potential function depend on finitely many coordinates which allowed them to use some ideas from the renewal theory.
- Yuri (mid 1990th) who proved convergence in (**??**) requiring the finite images property.
- Aaronson, Denker and Urbanski (mid 1990th) who studied ergodic properties of conformal measures and Aaronson and Denker (2001) who established convergence in (**??**) requiring the big images property.

Dobrushin-Landford-Ruelle (DLR) Measures

We begin by describing DLR measures. Given a probability measure μ on X, consider the conditional measures on cylinders $[a_0, \ldots, a_{n-1}]$ generated by μ, i.e., the conditional distribution of the configuration of the first n sites $(a_0, \ldots, a_{n-1})$ given that site n is in state x_n, site $(n+1)$ is in state x_{n+1} etc. More precisely, for almost all $x \in X$,

$$\mu(a_0, \ldots, a_{n-1} | x_n, x_{n+1}, \ldots)(x) = \mathbb{E}_\mu(1_{[a_0, \ldots, a_{n-1}]} | \sigma^{-n}\mathcal{B})(x).$$

Given $\beta > 0$ and a measurable function $U \colon X \to \mathbb{R}$, we call a probability measure μ on X a Dobrushin-Lanford-Ruelle (DLR) measure for the potential $\varphi = -\beta U$ if for all $N \geq 1$ and a.e. $x \in X$ the conditional measures of μ satisfies the DLR equation:

$$\mu(x_0, \ldots, x_{n-1} | x_n, x_{n+1}, \ldots)(x) = \frac{\exp \Phi_n(x)}{\sum_{\sigma^n(y)=\sigma^n(x)} \exp \Phi_n(y)}.$$

The problem now is to recover μ from its conditional probabilities.

Conformal Measures

In the particular case $\varphi(x) = f(x_0, x_1)$ recovering measure μ from its conditional probabilities is the well known Kolmogorov's theorem in the theory of classical Markov chains where the stochastic matrix $P = (p_{ij})$ is given by $p_{ij} = \exp f(i,j)$ if $a_{ij} = 1$ and $p_{ij} = 0$ otherwise.
For general potentials φ DLR measures can be recovered using conformal measures. More precisely, the following statement holds.

Theorem

Let φ be a Borel function and μ a non-singular conformal probability measure for φ on X. Then μ is a DLR measure for φ.

Note that this result is quite general as it imposes essentially no restrictions on the potential φ. Indeed, one can obtain much stronger statements assuming certain level of regularity of the potential.

Regularity requirements: Summable Variations and Locally Hölder ontinuity

Let $\varphi\colon X \to \mathbb{R}$ be a potential. We denote by

$$\text{var}_n(\varphi) = \sup\{|\varphi(x) - \varphi(y)|\colon x_i = y_i,\ 0 \le i \le n-1\}$$

the n-th variation of φ. We say that φ has summable variations if

$$\sum_{n=2}^{\infty} \text{var}_n(\varphi) < \infty$$

and φ is locally Hölder continuous if the exist $C > 0$ and $0 < \theta < 1$ such that for all $n \ge 2$

$$\text{var}_n(\varphi) \le C\theta^n.$$

The Gurevic-Sarig Pressure

We shall always assume that σ is topologically mixing (i.e., given $i, j \in S$ there is $N = N(i,j)$ s.t., for any $n \geq N$ there is an admissible word of length n connecting i and j). For $i \in S$ let

$$Z_n(\varphi, i) = \sum_{\sigma_n(x)=x, x_0=i} \exp(\Phi_n(x)).$$

The Gurevic-Sarig pressure of φ is the number

$$P_G(\varphi) = \lim_{n\to\infty} \frac{1}{n} \log Z_n(\varphi, i).$$

Theorem (O. Sarig)

Assume that φ has summable variations. Then

1. *The limit exists for all $i \in S$ and is independent of i.*
2. $-\infty < P_G(\varphi) \leq \infty$.
3. $P_G(\varphi) = \sup\{P(\varphi|K) \colon K \subset X$ *compact and* $\sigma^{-1}(K) = K\}$.

The Variational Principle For The Gurevic-Sarig Pressure

The Gurevic-Sarig pressure is a generalization of the notion of topological entropy introduced by Gurevic, so that $P_G(0) = h_G(\sigma)$.

Theorem (O. Sarig)

Assume that φ has summable variations and $\sup \varphi < \infty$. Then

$$P_G(\varphi) = \sup\left\{ h_\mu(\sigma) + \int \varphi \, d\mu \right\} < \infty,$$

where the supremum is taken over all σ-invariant Borel probability measures on Σ_A^+ such that $-\int \varphi \, d\mu < \infty$.

Our goal is to construct an equilibrium measure μ_φ for φ, find conditions under which it is unique and study its ergodic properties. We will achieve this by first constructing a Gibbs measure for φ and then showing that it is an equilibrium measure for φ providing it has finite entropy.

Gibbs Measures For Subshifts of Countable Type

The construction of Gibbs measures is based on the study of the Ruelle operator $\mathcal{L}_\varphi$ and on establishing a generalized version of the Ruelle's Perron-Frobenius theorem. The role of the Ruelle operator in the study of Gibbs measures can be seen from the following result that connects this operator with the Gurevic-Sarig pressure. We say that a non-zero function f is a test function if it is bounded continuous non-negative and is supported inside a finite union of cylinders.

Theorem (O. Sarig)

Assume that φ has summable variations. Then for every test function f and all $x \in X$

$$\lim_{n\to\infty} \frac{1}{n} \log(\mathcal{L}_\varphi^n f)(x) = P_G(\varphi).$$

This results implies that if $P_G(\varphi)$ is finite then for every $x \in$ the asymptotic growth of $\mathcal{L}_\varphi^n f(x)$ is λ^n where $\lambda = \exp P_G(\varphi)$.

Recurrence Properties of the Potential

We now wish to obtain a more refined information on the asymptotic behavior of $\lambda^{-n}\mathcal{L}_{\varphi}^{n}$. To this end given a state $i \in S$, let

$$Z_n(\varphi, i) = \sum_{\sigma^n(x)=x} \exp(\Phi_n(x)) 1_{[i]}(x)$$

and

$$Z_n^*(\varphi, i) = \sum_{\sigma^n(x)=x} \exp(\Phi_n(x)) 1_{[\varphi_i = n]}(x),$$

where φ_i is the first return time to the cylinder $[i]$.
We say that φ is

- recurrent if $\sum \lambda^{-n} Z_n(\varphi, i) = \infty$;
- positive recurrent if it is recurrent and $\sum n\lambda^{-n} Z_n^*(\varphi, i) < \infty$;
- null recurrent if it is recurrent and $\sum n\lambda^{-n} Z_n^*(\varphi, i) = +\infty$;
- transient if $\sum \lambda^{-n} Z_n(\varphi, i) < \infty$.

Genralized Ruelle's Perron–Frobenius (GRPF) Theorem (O. Sarig)

Assume that the potential φ has summable variations and $P_G(\varphi) < \infty$. Then

1. φ is recurrent if and only if there are $\lambda > 0$, a positive continuous function h and a conservative measure ν (i.e., a measure that allows no nontrivial wandering sets) which is finite and positive on cylinders such that

$$\mathcal{L}_\varphi h = \lambda h, \quad \mathcal{L}_\varphi^* \nu = \lambda \nu$$

In this case $\lambda = \exp P_G(\varphi)$.

2. φ is positive recurrent if and only if $\int h\,d\nu = 1$; in this case for every cylinder $[a]$,

$$\lambda^{-n}\mathcal{L}_{\varphi}^{n}(1_{[a]})(x) \to h(x)\frac{\nu([a])}{\int h\,d\nu} \text{ as } n \to \infty$$

uniformly in x on compact sets. Furthermore, ν is a conformal measure for φ (and hence, a DLR measure) and $\mu_{\varphi} = h\nu$ is the unique σ-invariant Gibbs measure for φ.

3. φ is null recurrent if and only if $\int h\,d\nu = \infty$; in this case for every cylinder $[a]$,

$$\lambda^{-n}\mathcal{L}_{\varphi}^{n}(1_{[a]})(x) \to 0 \text{ as } n \to \infty$$

uniformly in x on compact sets.

1. φ is transient if and only if there are no conservative measures ν which are finite on cylinders and such that $\mathcal{L}_{\varphi}^{*}\nu = \lambda\nu$ for some $\lambda > 0$.

This theorem is a generalization of earlier results by Vere-Jones, by Aaronson and Denker and by Yuri. In particular, the result by Yuri requires

- finite images property (FIP): the set $\{\sigma([i]) \colon i \in S\}$ is finite;

and the result by Aaronson and Denker requires

- big images and pre-images property (BIP): there exist $i_1, \ldots i_m \in S$ such that for all $j \in S$ there are $1 \leq k, \ell \leq m$ for which $a_{i_k j} a_{j i_\ell} = 1$.

In fact, the BIP property can be used to characterize existence of σ-invariant Gibbs measures.

Theorem (O. Sarig)

Assume that the potential φ has summable variations. Then φ admits a unique σ-invariant Gibbs measure μ_φ if and only if

1. *X satisfies the BIP property;*
2. *$P_G(\varphi) < \infty$ and $var_1 \varphi < \infty$ (i.e., $\sum_{n \geq 1} var_n(\varphi) < \infty$).*

In this case φ is positive recurrent and $\mu_\varphi = h\nu$, where ν is the conformal measure for φ in the GRPF theorem.

Existence and Uniqueness of Equilibrium Measures

Theorem (Existence: O. Sarig)

Assume that the potential φ has summable variations and $P_G(\varphi) < \infty$. Assume also that φ is positive recurrent and $\sup \varphi < \infty$. If the measure ν in GRPF theorem has finite entropy then the measure $\mu_\varphi = h\nu$ is an equilibrium measure for φ.

Theorem (Uniqueness: J. Buzzi, O. Sarig)

Assume that the potential φ has summable variations and $P_G(\varphi) < \infty$. Assume also that $\sup \varphi < \infty$. Then φ has at most one equilibrium measure. In addition, if such a measure exists then φ is positive recurrent and this measure coincides with the measure ν in GRPF theorem and has finite entropy.

Ergodic Properties

Theorem (O. Sarig)

Assume that the potential φ has summable variations and $P_G(\varphi) < \infty$. Assume also that $\sup \varphi < \infty$. If $\mu = \mu_\varphi$ is an equilibrium measure for φ then μ is strongly mixing (Bernoulli) and

$$h_\mu(\sigma) = \int \log \frac{d\mu}{d\mu \circ \sigma}\, d\mu \text{ (the entropy formula).}$$

The strong mixing property is a corollary of a general result by Aaronson, Denker and Urbanski that claims that if ν is a non-singular σ-invariant measure, which is finite on cylinders, conservative and whose log of the Jacobian has summable variations, then ν is strongly mixing.

Decay of Correlations and CLT

Theorem (O. Sarig)

Assume that the potential φ is locally Hölder continuous and $P_G(\varphi) < \infty$. Assume also that $\sup \varphi < \infty$. Then the equilibrium measure μ_φ for φ has exponential decay of correlations (with respect to the class of Hölder continuous functions on X) and satisfies the Central Limit Theorem.

The proof of this results is based on the crucial spectral gap property (SGP) of the Ruelle operator that claims that in an appropriate (sufficiently "large") Banach space $\mathcal{B}$ of continuous functions $\mathcal{L}_\varphi = \lambda\mathcal{P} + \mathcal{N}$ where $\lambda = P_G(\varphi)$ and

$$\mathcal{P}\mathcal{N} = \mathcal{N}\mathcal{P} = 0, \quad \mathcal{P}^2 = \mathcal{P}, \quad \dim(\mathrm{Im}\mathcal{P}) = 1.$$

Furthermore, the spectral radius of $\mathcal{N}$ is less than λ. The SGP implies the exponential rate of convergence in (**??**) leading to the exponential decay of correlations and the Central Limit Theorem.

For subshifts of finite type there is a subspace $\mathcal{B}$ on which the Ruelle operator has the SGP (due to Ruelle and Doeblin-Fortet) but this may not be true for subshifts of countable type due to the presence of phase transitions. Indeed, the SGP guarantees that the function $t \to p(t) = P_G(\varphi + t\psi)$ (where φ and ψ are locally Hölder continuous) is real-analytic meaning absence of phase transitions. However, for subshifts of countable type as t varies the function $\varphi + t\psi$ can change its mode of recurrence (e.g., move from being positive recurrent to null recurrent or to transient) resulting in non-analyticity of the function $p(t)$ and hence, the appearance of phase transitions. Given a subshift of countable type, Cyr and Sarig found a necessary and sufficient condition for the existence of a space $\mathcal{B}$ on which the Ruelle operator has the SGP. Using this condition they showed that absence of phase transitions is open and dense in the space of locally Hölder continuous potentials.