Tel-Aviv University
School of Mathematical Sciences

Thermodynamic Formalism for
Countable Markov Shifts

Thesis submitted for the degree ‘Doctor of Philosophy’

by

Omri M. Sarig

Submitted to the Senate of the Tel-Aviv University
April 2000
Tel-Aviv University
School of Mathematical Sciences

Thermodynamic Formalism for Countable Markov Shifts

Thesis submitted for the degree ‘Doctor of Philosophy’
by
Omri M. Sarig

Submitted to the senate of the Tel-Aviv University
April 2000
This thesis was written under the supervision of

**Professor Jon Aaronson**
Acknowledgments

I would like to express my deep gratitude to my teacher Jon Aaronson for accompanying my steps in research from their very beginning, for supporting and assisting me along the way, and for his constant active and highly valuable interest in this work. All that I know of mathematical research, I owe to him.

I owe special thanks to my dear parents Yigal and Gissi Sarig, for their constant support, encouragement and help, without which this work would almost surely never have been finished.
Contents

0 Notational conventions 3

1 Introduction and Summary of Main Results 4
  1.1 Topological Markov shifts: Notation and first definitions . . . . . 4
  1.2 Thermodynamic formalism .................................... 5
    1.2.1 Motivation .............................................. 5
    1.2.2 Basic questions ....................................... 7
    1.2.3 Results for finite Markov shifts ....................... 8
  1.3 Summary of main results ................................... 9
    1.3.1 The Gurevich pressure ................................. 10
    1.3.2 A generalized Ruelle’s Perron-Frobenius theorem ........ 11
    1.3.3 Gibbs measures and equilibrium measures ............... 12
    1.3.4 Phase transitions .................................. 13
    1.3.5 Examples ............................................ 14
  1.4 Summary .................................................. 16

2 The Gurevich Pressure 18
  2.1 The Gurevich pressure .................................... 18
  2.2 The variational principle .................................. 20
  2.3 Notes .................................................... 22

3 The Generalized Ruelle’s Perron-Frobenius Theorem 24
  3.1 A generalized Ruelle’s Perron-Frobenius theorem ............ 24
  3.2 Construction of eigenmeasure ................................ 26
  3.3 Construction of eigenfunction and (3.1) ...................... 31
    3.3.1 Preliminaries: the Schweiger property .................. 31
    3.3.2 Construction of $h$ and $a_n$ .......................... 33
  3.4 Positive recurrence and Null recurrence ..................... 36
  3.5 Notes ................................................ 38

4 Gibbs Measures and Equilibrium Measures 40
  4.1 Gibbs measures ........................................... 40
    4.1.1 The BIP property .................................... 40
    4.1.2 The rate of convergence in the RPF theorem ......... 43
4.2 Equilibrium measures ........................................................................... 44
  4.2.1 Entropy for infinite measures ......................................................... 45
  4.2.2 The induced system ....................................................................... 45
  4.2.3 Two variational principles ............................................................. 47
4.3 Notes .................................................................................................. 51

5 Phase Transitions .................................................................................... 52
  5.1 The discriminant: statement of main results ....................................... 52
  5.2 Proof of the discriminant theorem ...................................................... 55
    5.2.1 A renewal sequence of operators ................................................ 55
    5.2.2 Proof of the discriminant theorem .............................................. 57
  5.3 Proof of the strong positive recurrence theorem .................................. 59
    5.3.1 Preparations ................................................................................ 59
    5.3.2 Proof of theorem 10 .................................................................... 61
  5.4 Notes .................................................................................................. 62

6 Examples .................................................................................................. 63
  6.1 Systems with BIP ............................................................................ 63
  6.2 The renewal shift ............................................................................... 64
    6.2.1 general results ........................................................................... 64
    6.2.2 Application to the Pomeau-Manneville model ............................ 66
  6.3 Pathological examples ...................................................................... 69
    6.3.1 Tool for constructing examples .................................................. 69
    6.3.2 Infinite number of changes in recurrence .................................... 70
    6.3.3 Null recurrence in an interval ..................................................... 71
    6.3.4 Coexistence of all modes of recurrence .................................... 71
    6.3.5 Positive measure set of critical points ....................................... 72
  6.4 Notes .................................................................................................. 74
Notational Conventions

We use the following notational conventions.

- \( a = B^{\pm n} b \) means that \( B^{-n} b \leq a \leq B^n b \)
- \( a_n \asymp b_n \) means that \( \exists B, N \) such that \( \forall n > N, a_n = B^{\pm 1} b_n \)
- \( a_n \propto b_n \) means that \( \exists c \neq 0 \) such that \( a_n/b_n \to c \)
- \( a_n \sim b_n \) means that \( a_n/b_n \to 1 \)
- \( B \) - the Borel \( \sigma \)-field
- \( C_B(X) \) - the space of all bounded continuous complex valued functions on \( X \), equipped with the \( \| \cdot \|_{\infty} \) norm
- \( C_B[a] \) - the space of all bounded continuous complex valued functions on \( X \) which are supported inside \([a]\), equipped with the \( \| \cdot \|_{\infty} \) norm
- \( L_{\mu} \) - Information function (see section 4.2.3)
- \( L_{\phi} \) - Ruelle’s operator, \( L_{\phi} f(x) = \sum_{y=x} e^{\phi(y)} f(y) \)
- \( \log \) - the same as \( \ln \)
- \( \mu(f) \) - another notation for \( \int f \, d\mu \)
- \( \mu \circ T \) - see section 3.3.1
- \( 1_A \) - the indicator of \( A \subseteq X \):
  \[
  1_A(x) = \begin{cases} 
  1 & x \in A \\
  0 & x \notin A 
  \end{cases}
  \]
- \( P_T(X) \) - the collection of all \( T \)-invariant Borel probability measures.
- \( \phi_n \) - the \( n \)-th ergodic sum, \( \phi + \phi \circ T + \ldots + \phi \circ T^{n-1} \)
- \( t_{ij} \) - elements of the transition matrix
- \( \varphi_n(x) \) - first return time function: \( \varphi_n(x) = 1_{[a]}(x) \inf\{n \geq 1 : T^n x \in [a]\} \)
- \( Z_n(\phi, a) \) - local partition function: \( \sum_{x=x} e^{\phi_n(x)} 1_{[a]}(x) \)
- \( Z_n(\phi, a) \) - reduced local partition function: \( \sum_{x=x} e^{\phi_n(x)} 1_{[a]}(x) \)

Equations and sections are numbered according to chapters: equation (3.2) is the second equation in chapter 3, section 5.2.1 is the first subsection in the second section in chapter 5.

An attempt has been made to formulate the main results of each chapter as early as possible. Notes, references and historical remarks are compiled in the last section of each chapter.
Chapter 1

Introduction and Summary of Main Results

The purpose of this work is to study the thermodynamic formalism for countable Markov shifts. This chapter contains an introduction to this subject, and a summary of the main results of the work. Most of these results are contained in [S2],[S3], [S4] and [S5].

1.1 Topological Markov shifts: Notation and first definitions

Topological Markov shifts appear as symbolic models for dynamical systems with Markov partitions (see e.g. [Si], [B1], [KH], [B2], [T1], [PS]). These models are relatively well understood when the Markov partition is finite ([B1], [Ru1], [PP]). This work treats the case when the Markov partition is infinite. This section collects the basic terminology and notation we need.

Let $S$ be a countable set and $A = (t_{ij})_{S \times S}$ a matrix of zeroes and ones with no columns or rows which are all zeroes. Let $X$ be the set

$$X := \left\{ x \in S^{\mathbb{N} \cup \{0\}} : t_{x_i x_{i+1}} = 1, \forall i \geq 0 \right\}$$

endowed with the relative product topology, which is also given by the metric $d(x,y) := 2^{-\min\{n: x_n \neq y_n \}}$ or by the base of cylinders

$$[a_0, \ldots, a_{n-1}] := \{ x \in X : x_i = a_i, 0 \leq i \leq n - 1 \}$$

where $n \in \mathbb{N}$ and $a_0, \ldots, a_{n-1} \in S$. An admissible word is an element $a \in S^n$ such that $[a] \neq \emptyset$. The length of an admissible word is defined as the number of its letters.
1.2. THERMODYNAMIC FORMALISM

Definition 1 Let $T : X \to X$ be the left shift $(T x)_i := x_{i+1}$. The topological dynamical system $(X, T)$ is called a topological Markov shift.\footnote{Sometimes this is called a one sided topological Markov shift, to emphasize that $X$ is a subset of $\mathbb{Z}^\mathbb{N}$, not a subset of $\mathbb{N}^\mathbb{N}$.} The members of $S$ are called the states of the shift, and the matrix $A$ is called the transition matrix. The sets $[a]$ where $a \in S$ are called partition sets.

We also use the terminology finite Markov shift and countable Markov shift to distinguish between the cases $|S| < \infty$ and $|S| \leq \aleph_0$. We say that $X$ is topologically mixing if $(X, T)$ is topologically mixing. This is equivalent to the following property: $\forall a, b \in S, \exists N$ such that for all $n \geq N$ there is an admissible word $a$ of length $n$ such that $a_0 = a$ and $a_{n-1} = b$.

Let $\phi : X \to \mathbb{R}$ be some real function (also called potential). The variations of $\phi$ are $V_n(\phi) := \sup \{ \phi(x) - \phi(y) : x, y \in X, x_i = y_i, 0 \leq i \leq n - 1 \}$. We say that $\phi$ has summable variations if

$$\sum_{n=2}^{\infty} V_n(\phi) < \infty$$

We say that $\phi$ is weakly Hölder continuous (with parameter $\theta$) if there exist $A > 0$ and $\theta \in (0, 1)$ such that for all $n \geq 2$,

$$V_n(\phi) \leq A \theta^n$$

Note that in both cases the quantification begins with $n = 2$ so $\phi$ may be unbounded or even satisfy $V_1(\phi) = \infty$ (This is impossible when $|S| < \infty$. In this case weak Hölder continuity is the same as usual Hölder continuity.)

The Ruelle operator (Ruelle [Ru2]) is given by

$$\begin{equation}
(L_\phi f)(x) := \sum_{T y = x} e^{\phi(y)} f(y)
\end{equation}$$

If $\|L_\phi 1\|_\infty < \infty$ this is a bounded operator on the Banach space $C_B(X) = \{ f \in C(X) : \|f\|_\infty < \infty \}$ (complex valued functions). One checks that $(L_\phi^n f)(x) = \sum_{T^n y = x} e^{\phi_n(y)} f(y)$ where $\phi_n := \sum_{k=0}^{n-1} \phi \circ T^k$.

1.2 Thermodynamic formalism

1.2.1 Motivation

Topological Markov shifts typically admit many invariant measures. This raises the problem of choosing the ‘right’ measure to work with. The thermodynamic formalism is the study of certain procedures for the choice of measures,
motivated by equilibrium statistical physics. This approach has its roots in the
works of Sinai [Si] and Ruelle [Ru3].

The basic motivation is the following. Think of \((x_0, x_1, \ldots) \in X\) as of a
possible configuration of a one-sided one-dimensional lattice \(\mathbb{N} \cup \{0\}\), whose
\(i\)-th site is in ‘state’ \(x_i \in S\). If all the entries of the transition matrix are equal
to 1, all configurations \(S^{\mathbb{N} \cup \{0\}}\) are possible. If not, only configurations which
satisfy some nearest neighbor constraints are considered.

Now assume that the different sites are interacting among themselves, and
with an external potential field. Let

\[
H(x) = H(x_0|x_1, x_2, x_3, \ldots)
\]

be the interaction energy between the first site and the rest of the sites plus
its interaction energy with the external force field. The energy content of
the first \(n\) sites can be calculated by considering the energy released when
we ‘bring from infinity’ the segment \((x_0, \ldots, x_{n-1})\) and place it to the left
of \((x_n, x_{n+1}, \ldots)\). We do this step by step, first adding \(x_{n+1}\), then adding
\(x_{n+2}\) and so forth. The first stage yields energy \(H(x_{n+1} x_{n+2}, \ldots) = H(T^{n+1} x)\),
the second \(H(x_{n+2} x_{n+3}, \ldots) = H(T^n x)\) etc. The total is the ergodic sum
\(H_n(x) = \sum_{k=0}^{n-1} H(T^k x)\). This is where dynamics enters.

Next, assume that the system is allowed to exchange energy with its sur-
rroundings (so energy is not conserved), and that we constrain the temperature
of the surroundings to be fixed.

We are interested in the measure that gives the distribution of the config-
urations in equilibrium. Physics dictates that when there is a finite number of
possible configurations \(\{\xi_i\}\), the equilibrium probability of each of them is

\[
P(\xi = \xi_i) = \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}},
\]

where \(\beta\) is proportional to the inverse of the temperature, and \(E_i\) is the energy
of the configuration \(\xi_i\). We cannot use this law directly, since in our case there is
a continuum of configurations and since \(E = \sum_{k=0}^{\infty} H(T^k x)\) may diverge. Still,
we shall be interested in measures with properties which are weak versions of
the impossible relation \(P(x) \propto e^{-\beta H(x)}\). This can be done either by suitable
limit procedures, or by using other formulations of the physical law which are
meaningful in our situation. An alternative formulation is the variational prin-
ciple: the equilibrium distribution is that which minimizes the mean free energy
per site

\[
\lim_{n \to \infty} \frac{1}{n} \left( \int H_n(\xi) \, dm - \beta^{-1} \sum_{[a]=n} m[a] \log \frac{1}{m[a]} \right)
\]

Using these analogies, the constructions of the thermodynamic formalism
are the same as the calculation of the equilibrium distribution given

\[
\phi(x) := -\beta H(x_0|x_1, x_2, \ldots)
\]

Using this notation, \(e^{-\beta H_n(x)}\) becomes \(e^{\phi_n(x)}\) and minimizing the free energy
is the same as maximizing \(\mu_T = \int \phi \, d\mu\) (as long as \(\mu = \mu \circ T^{-1}\)).
1.2.2 Basic questions

The thermodynamic formalism treats several methods for choosing a measure as a function of $\phi$. These methods are equivalent for finite Markov shifts, but turn out to be non-equivalent in the countable case. We therefore formulate them separately.

1. Look for Gibbs measures (in the sense of Bowen): these are invariant probability measures $m$ which satisfy the Gibbs property: there are constants $P$ and $M$ such that for all cylinders $[a] = [a_0, \ldots, a_{n-1}]$,

$$\forall x \in [a], \quad m[a] = M \pm 1 e^{\phi(x)} n^P$$

2. Solve a variational problem: find the invariant measure which maximizes the quantity $h_\mu(T) + \int \phi d\mu$. A maximizing measure is called an equilibrium measure. The maximum attained (or the supremum when the maximum is not attained) is called the topological pressure of $\phi$.

3. Thermodynamic limit: we approximate the variational problem of the previous step by simpler discrete problems. Fix some ‘boundary condition’ $x \in X$, let $\alpha_0^{n-1}$ be the partition into cylinders of length $n$, and maximize $\frac{1}{n} (H_\mu(\alpha_0^{n-1}) + \int \phi d\mu)$ where $\mu$ is supported on $\{y : T^n y = x\}$. The solution is

$$\nu_n = \frac{1}{Z_n} \sum_{T^n y = x} e^{\phi_n(y)} \delta_y$$

where $\delta_y$ is the probability measure concentrated on $\{y\}$ and $Z_n$ are normalization constants. The $w^*$-Limit points as $n$ tends to infinity are called Gibbs states. The idea is to find a Gibbs state. We will also be interested in the existence of an equivalent invariant measure.

Note that for every function $f$, $\nu_n^\ast (f) = (L^n \phi)(f)/(L^n \phi)^\ast (f)$ where $L_\phi$ is Ruelle’s operator given by (1.3), so studying the thermodynamic limit can be done by analyzing the asymptotics of $L^n_\phi$.

4. Look for measures with given conditionals: find a measure $\nu$ such that $\nu(x_0|x_1, x_2, \ldots) := \mathbb{E}_\nu(1_{x_0} | T^1 B)(x)$ is proportional to $e^{\phi(x)}$. For non-singular measures, this is the same as finding a measure $\nu$ and a proportionality constant $\lambda$ such that

$$\frac{d\nu}{d\nu \circ T} = \lambda e^{\phi}$$

Solutions of this equation are called conformal measures. Equation (1.5) is equivalent to $\lambda L_\phi$ being equal to the transfer operator of $\nu$. Thus solutions of (1.5) satisfy $L_\phi^2 \nu = \lambda \nu$, and invariant densities of $\nu$ are functions $h$ such that $L_\phi h = \lambda h$. This relates this procedure to finding eigenmeasures and eigenfunctions for the Ruelle operator.

---

3We are assuming here for simplicity that for all $n$, $\sum_{y} e^{\phi_n(y)} < \infty$. When this is not the case, it is usually possible to work with $\nu_n^\ast_{[a]}$ where $a$ is some fixed partition set.

4Also called by some authors ‘Gibbs states’ or ‘Gibbsian fields’.
The following questions arise: When do Gibbs measures, equilibrium measures and conformal measures exist, and how can they be calculated? How can the topological pressure be calculated? What happens in the thermodynamic limit, and do the asymptotics depend on the choice of the boundary condition \( x \)? Are these methods equivalent?

Another basic question is the determination of ‘thermodynamic’ quantities arising in the thermodynamic limit. By this we mean the asymptotics as \( n \) tends to infinity of the \( \nu^n_n \)-distributions of \( \psi_n / n \), where \( \psi(x) \) is some fixed function. For example, when \( X = \{-1, +1\}^N \) and \( \psi(x) = x_0 \), \( \psi_n / n \) measures the average charge of the \( n \)-th prefix, and it is natural to ask whether this converges to a uniquely determined charge density.

The basic tool to study these questions is the theory of large deviations (see Ellis [E] and Martin-Löf [ML]). Passing to moment generating functions, set

\[
(1.6) \quad c(t) := \lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}_{\nu^n_n} \left( e^{t\psi_n} \right)
\]

when the limit exists. A well-known theorem in large deviations theory (see Ellis [E], theorem II.6.3) says that if \( c(t) \) exists and is finite for all \( t \) then the \( \nu^n_n \)-distributions of \( \psi_n / n \) tend exponentially to a unique value if \( c(t) \) is differentiable at zero.\(^5\) This value is interpreted as the corresponding thermodynamic quantity, and the exponential convergence is used as a justification for describing a system which is random on the microscopic level by one number on the macroscopic level. When \( c(t) \) is not differentiable at zero, there is no exponential convergence to a unique value. This situation is often called a phase transition (of the first order) (see [E]). Higher order phase transitions correspond to non-differentiability of higher derivatives of \( c(t) \). These are also important, since for certain \( \psi \) the higher derivatives of \( c(t) \) have a thermodynamic description.

The study of phase transitions is also strongly connected to the properties of the Ruelle operator. This is because

\[
\mathbb{E}_{\nu^n_n}(e^{t\psi_n}) = (L^n_{\psi+t\psi}(x))/(L^n_\psi(x))
\]

whence \( c(t) = P(\phi + t\psi) - P(\phi) \) where \( P(\cdot) \) is defined by

\[
(1.7) \quad P(\varphi) := \lim_{n \to \infty} \frac{1}{n} \log \sum_{x} e^{\varphi_n(x)}
\]

1.2.3 Results for finite Markov shifts

We now give a brief summary of known results on the thermodynamic formalism for finite Markov shifts (see [Rul], [B1], [Wall], [PP]).

Let \( X \) be a topologically mixing finite Markov shift and let \( \phi \) be some fixed real function on \( X \) with summable variations. The topological pressure is given by the Variational Principle:

\[
P_{\text{top}}(\phi) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{T^nx=x} e^{\phi_n(x)}
\]

\(^5\)by Exponential convergence to a limit \( L \) we mean that \( \forall \epsilon > 0, \exists N(\epsilon) > 0 \) such that for \( n \) large enough, \( \nu^n_n(\psi_n \geq \epsilon) \leq e^{-N(\epsilon)} \).
This value is attained by a unique equilibrium measure, which is also the unique Gibbs measure.

The thermodynamic limit is described by the following theorem of Ruelle [Ru2] (given in the form found in [Wal2])

**Ruelle’s Perron-Frobenius theorem.** Let $X$ be a finite Markov shift which is topologically mixing, and let $\phi : X \to \mathbb{R}$ have summable variations. Then $\exists \lambda > 0$, $\exists h$ positive continuous function and $\exists \nu$ Borel probability measure such that $L_\phi \nu = \lambda \nu, \int h d\nu = 1$ and such that for all $f \in C(X),$

\[
\| \lambda^{-n} L_\phi^n f - h \int f d\nu \| \to 0 \quad \text{as} \quad n \to \infty
\]

This theorem implies by the discussion in the previous section, that $\nu^n \to \nu$ where $\nu$ is the conformal measure, and that $h d\nu$ is an invariant probability. It can also be proved that $h d\nu$ is the unique Gibbs measure and the unique equilibrium measure. In other words, for finite Markov shifts there is no difference between Gibbs measures, equilibrium measures, and the invariant probability measures equivalent to the Gibbs state and to the conformal measure.

The convergence part in RPF theorem also implies that $P_{\text{top}}(\phi) = \log \lambda = P(\phi)$ where $\lambda$ is given by the RPF theorem and $P(\phi)$ is given by (1.7). It follows that $c(t) = P_{\text{top}}(\phi + t \psi) - P_{\text{top}}(\phi)$ so the existence of phase transitions is closely related to the differentiability of the topological pressure functional.

For Hölder continuous functions, it can be shown that $\lambda$ is a simple isolated eigenvalue of $L_\phi$ when the action of this operator is restricted to the space of Lipschitz functions. By perturbation theory, $\lambda$ depends analytically on $\phi$, whence by the previous remarks, $c(t)$ is differentiable. Thus for finite Markov shifts and Hölder continuous functions, there are no phase transitions (see Ruelle [Ru2]).

### 1.3 Summary of main results

The subject of this work is the thermodynamic formalism for topological Markov shifts with a countable number of states. The main difference between this case and the finite state case is that the shift space $X$ is no longer compact. This difference gives rise to new qualitative phenomena. The purpose of this study is to understand these new phenomena, and classify them.

This section contains a summary of our main results. In what follows, $X$ is a topologically mixing countable Markov shift with set of states $S$ and transition matrix $A = (a_{ij})_{S 	imes S}, T : X \to X$ is the left shift, and $\phi : X \to \mathbb{R}$ is a some function with summable variations.
1.3.1 The Gurevich pressure

When $|S| < \infty$ the topological pressure is given by

$$P_{\text{top}}(\phi) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{T^n x = x} e^{\phi_n(x)}.$$  

When $|S| = \infty$ the left side may be strictly smaller than the right side (Gurevich [G2, G1], Salama [Sal]).

The topological entropy for a countable Markov shift was calculated by Gurevich [G1]. This is the same as calculating the pressure of $\phi = 0$. Gurevich showed that the supremum of metric entropies is equal to the supremum of topological entropies of all finite Markov shifts $Y \subseteq X$ (for the calculation of the topological entropy of a finite shift see Parry [P]). His calculation is based on finding a metric compactification for $X$ whose topological entropy is equal to the supremum of the topological entropies of all compact invariant subsets of $X$. The result then follows from the variational principle for compact metric spaces. This argument was generalized by Zargaryan [Z] for some bounded functions $\phi$ which can be uniformly extended to a certain compactification of $X$, and by Gurevich and Savchenko in [GS] to all functions which are uniformly continuous and bounded.

The assumption that $\phi$ is bounded is a severe restriction, since for bounded potentials, $\sup \{h_\mu(T) + \int \phi d\mu\} \geq \sup h_\mu(T) + \inf \phi$ and this is infinite for all countable Markov shifts with infinite topological entropy (e.g., $X = \mathbb{N}^\omega \{0\}$). For such shifts, functions with finite pressure must be unbounded from below. Obviously, such potentials have no continuous extension to a compactification, and cannot be treated by the compactification approach.

In chapter 2 (section 2.2, theorem 2) we prove that for every $\phi$ with summable variations such that $\sup \phi < \infty$,  

$$\sup \left\{ h_\mu(T) + \int \phi d\mu : \mu \in \mathcal{P}_T(X), \quad -\int \phi d\mu < \infty \right\} = P_G(\phi)$$

where $\mathcal{P}_T(X)$ is the collection of all $T$-invariant Borel probability measures and

$$P_G(\phi) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{T^n x = x} e^{\phi_n(x)} 1_{a}(x)$$

where $a \in S$ is arbitrary. This is the Variational Principle for countable Markov shifts.

The restriction $-\int \phi d\mu < \infty$ in needed to prevent the case $h_\mu(T) = \infty$ and $\int \phi d\mu = -\infty$ (when the sum $h_\mu(T) + \int \phi d\mu$ is meaningless). The restriction $\sup \phi < \infty$ is needed to ensure that $\int \phi d\mu$ is well-defined with respect to every probability measure.\footnote{Actually, we prove later (section 3.1, corollary 1) that if $P_G(\phi) < \infty$ then $\phi$ is cohomologous to $\phi'$ such that $\sup \phi' < \infty$ via a continuous transfer function $\xi$. Unfortunately, $\int (\xi - \xi \circ T) d\mu$ may not be defined for some $\mu \in \mathcal{P}_T(X)$.}
1.3. SUMMARY OF MAIN RESULTS

Theorem 1 in section 2.1 shows that

\[ P_G(\phi) = \sup \{ P_{\nu|\phi} (\phi|_K) : K \subseteq X \text{ is compact and invariant } \} \]

This shows that \( P_G(\phi) \) is indeed a generalization of Gurevich’s topological entropy. We therefore call it the Gurevich pressure of \( \phi \).

### 1.3.2 A generalized Ruelle’s Perron-Frobenius theorem

It is well-known that Ruelle’s Perron-Frobenius theorem is not true for countable Markov shifts (see [G1] or chapter 6 below). One of the main problems is the existence of the conformal measure \( \nu \).

This problem appears only in the non-compact case \( |S| = \infty \). In the compact case, \( \nu \) can be constructed by applying a fixed-point theorem for the action of \( G(\mu) := L^*_\mu \mu(L_01) \) on the \( w^* \)-compact set of Borel probability measures on \( X \) (see Ruelle [Ru2]). This fails in the non-compact situation, since the set of probability measures is not \( w^* \)-compact. Indeed, without further assumptions \( \nu \) may not exist.

Versions of the RPF theorem in non-compact situations were proved in certain special cases. Walters [Wal2] used assumptions which imply that \( \phi \) can be extended to a compactification \( X \) of \( X \) in a way that \( L_\phi \) acts continuously on \( C_B(X) \). This allowed the fixed-point argument to hold on the set of probability measures on \( X \). A similar approach was adopted by Yuri [Yu2]. Other authors were only interested in functions \( \phi \) for which \( L_\phi \) is the transfer operator of a given measure \( \nu \). In this case it follows automatically (see section 3.3.1, lemma 1) that \( L_\phi^* \nu = \nu \) and that \( \nu \) is a conformal measure (see, e.g., [ADU], [T3], [Yu1]).

We define the following notions of recurrence for functions \( \phi \) with a finite Gurevich pressure (section 1.1, definition 3): Let \( \lambda := e^{P_\nu(\phi)} \) and fix some \( a \in S \). Set \( \varphi_n(x) := 1_{[a]}(x) \inf \{ n \geq 1 : T^n x \in [a] \} \) and

\[
Z_n(\phi,a) := \sum_{T^n x = x} e^{\phi(x)} 1_{[a]}(x) \quad \text{and} \quad Z_n^*(\phi,a) := \sum_{T^n x = x} e^{\phi(x)} 1_{[a]}(x)
\]

We say that \( \phi \) is **recurrent** if \( \sum \lambda^n Z_n(\phi,a) = \infty \) and **transient** otherwise.

We say that \( \phi \) is **positive recurrent** if it is recurrent and \( \sum n \lambda^n Z_n^*(\phi,a) < \infty \).

We say it is **null recurrent** if it is recurrent and \( \sum n \lambda^n Z_n^*(\phi,a) = \infty \). (These notions turn out to be independent of the choice of \( a \in S \).)

If \( \phi(x) = \log p_{xx} \), where \( (p_{ij})_{S \times S} \) is a stochastic matrix, these notions are the same as the usual notions of recurrence for the corresponding Markov chain.

---

The Gurevich pressure was introduced in my M.Sc. thesis [S1], where it was shown to satisfy a variational principle in the special case when \( \phi \) has a conformal probability measure with invariant density bounded away from zero and infinity.
12 CHAPTER 1. INTRODUCTION AND SUMMARY OF MAIN RESULTS

In chapter 3 we prove the following generalization of the RPF theorem (section 3.1, theorem 3): Assume $P_G(\phi) < \infty$. Then

- $\phi$ is recurrent iff there exists a $\lambda > 0$ and a possibly infinite conservative Borel measure $\nu$, which is finite and positive on cylinders, such that $L_\phi^* \nu = \lambda \nu$. In this case $\lambda = e^{P_G(\phi)}$ and there exists a positive continuous function $h$, such that $L_\phi h = \lambda h$.

- if $\phi$ is positive recurrent, then $\int h d\nu < \infty$ and for all cylinders $[a]$, $\lambda \in L_\phi^* \nu$ $(x) = h(x) \nu [a]$, where $h$ is normalized so that $\int h d\nu = 1$.

- if $\phi$ is null recurrent, then $\int h d\nu = \infty$ and $\lambda \in L_\phi^* \nu$ $(x) \to 0$ for every cylinder. Nevertheless, there is a sequence $a_n = o(n)$, such that for every cylinder $[a]$,

$$\frac{1}{a_n} \sum_{k=0}^{a_n - 1} \lambda^k (L_\phi^k \nu) (x) \to h(x) \nu [a]$$

The sequence $a_n$ is given asymptotically by $a_n \propto \sum_{k=1}^{a_n} \lambda^k Z_k(\phi, a)$.

This theorem generalizes D. Vere-Jones’ Perron-Frobenius theorem for positive countable matrices [VJ1, VJ2], which in itself, is a generalization of known results for probabilistic countable Markov chains (see [Fe], [Se],[Ki]). The main point in the proof is the construction of $\nu$. This is done in section 3.2 using a tightness argument.

1.3.3 Gibbs measures and equilibrium measures

Let $h$ and $\nu$ be as in the RPF theorem. When $|S| < \infty$ the measure $h d\nu$ is an invariant Gibbs measure in the sense of Bowen. This is not necessarily true when $|S| = \infty$, even if $\phi$ is positive recurrent.

It was shown in [S2] that $\phi$ has an invariant Gibbs probability measure iff $V_1(\phi) < \infty$, $\phi$ is positive recurrent, $\|\log h\|_\infty < \infty$ and $A$ satisfies a certain property, called ‘big images’ (see also [ADU]). In a recent preprint [MU], Mauldin and Urbański give a condition, which we call the BIP property, which is sufficient for the existence of Gibbs measures for all Hölder continuous functions such that $V_1(\phi) < \infty$. We show that this condition is actually necessary and sufficient (section 4.1, theorem 4). The heart of the proof is Mauldin and Urbański’s result, to which we give a new proof, based on the notion of positive recurrence.

In section 4.1.2 we use results from [AD] to show that if $X$ has the BIP property and $\phi$ is weakly Hölder continuous with finite pressure, such that $V_1(\phi) < \infty$, then $\lambda \in L_\phi^* f \to h \nu (f)$ exponentially fast whenever $f \in C_B(X)$ is sufficiently smooth (see theorem 5 for details). For such functions $f, g$, $\int (f - \nu (f))(g - \nu(g)) \circ T^n d\nu$ decays exponentially, a property which is sometimes called exponential decay of correlations. This is one more example of the similarity between countable decay of Markov shifts with the BIP property and finite Markov shifts.
1.3. SUMMARY OF MAIN RESULTS

Another topic we treat in chapter 4 is that of equilibrium measures. When $|S| < \infty$, the measure $h \, dv$ maximizes the quantity $h_m(T) + \int \phi \, dm$. When $|S| = \infty$ and $\phi$ is recurrent, we still have a measure $h \, dv$. This, however, cannot be directly viewed as an equilibrium measure, because it may happen that $h_m(T) = \infty$ and $\int \phi = -\infty$ in which case $h_m(T) + \int \phi \, dm$ is not defined. Furthermore, if $\phi$ is null recurrent, $m$ is infinite and it is not clear how to define its entropy. Theorems 6 and 7 show how to interpret $dm$ as an equilibrium measure in this case, using Krenkel’s notion of entropy for a conservative infinite invariant measure [Kr] and Rokhlin’s formula (see also [Wal2], [Le], [T1] and [Yu1]).

1.3.4 Phase transitions

Ruelle showed in [Ru2] that for finite Markov shifts, whenever $\phi$ and $\psi$ are Hölder continuous, $t \mapsto P_{\phi + t \psi}$ is real analytic. Since, for finite Markov shifts $P_{\phi + t \psi} = P^t(\phi)$, where $P^t(\phi)$ is given by (1.7), this implies that $c(t)$ in (1.6) is differentiable at zero, so there are no phase transitions.

The situation for countable Markov shifts is known to be different. Specific examples appear in Hofbauer [Hof], Wang [W1, W2], Pomeau and Manneville [PM], Prellberg and Slawny [PS] and Lopes [Lo].

Our results on phase transitions are formulated in terms of differentiability results for functions of the form $t \mapsto P_\phi(t \psi)$. Such results describe the behavior of $P_{\phi + t \psi}$, as well as the large deviations behavior of the $\nu^x_n[a]$-distributions of $\psi_n/n$, where $[a]$ is some arbitrary partition set.8

These results relate non-differentiability of $P_\phi(\psi)$ to changes in the recurrence properties of the one-parameter family $\{\phi + t \psi\}_t$. To make this precise we need the notion of discriminant which we now proceed to define (section 5.1, definition 4): Let $a \in S$ be some fixed state, and let $\overline{S}$ denote the set of all non-empty cylinders $[a, \xi_1, \ldots, \xi_n]$ where $n$ is arbitrary, $\xi_i \neq a$ for all $i$ and such that $[a, \xi_1, \ldots, \xi_n, a]$ is not empty. Set $\overline{X} := \overline{S}^{\mathbb{N} \cup \{0\}}$ and let $\overline{T} : \overline{X} \to \overline{X}$ be the left shift. For every function $f$ on $X$, let $\overline{T} := \left( \sum_{k=0}^{n-1} f \circ T^k \right) \circ \pi$ where $\pi : \overline{X} \to [a]$ is the natural embedding. The passage from $(X, f)$ to $(\overline{X}, \overline{T})$ is the symbolic counterpart of the inducing process (see section 4.2.2 for details). Now consider

$$p^* = p^*([\phi] := \sup \{ p \in \mathbb{R} : P_\phi(\phi + p) < \infty \} \leq \infty$$

The a-discriminant of $\phi$ is defined as

$$\Delta = \Delta_a([\phi] := \sup \{ P_\phi(\phi + p) : p < p^* \} \leq \infty$$

Equations (5.1)-(5.4) in section 5.1 show how these numbers can be estimated, and in certain situations completely calculated.

---

8The reason we restrict ourselves to partition sets is that for general countable Markov shifts, $(L^*_\phi)^\mathcal{S}$ may have infinite mass, in which case $\nu^x_n$ is not well-defined. However, when $P_\phi(\phi) < \infty$, $L^*_\phi [a]$ is always finite on $[a]$ (see section 3.1), so $\nu^x_n[a] := (L^*_\phi)^\mathcal{S}_n[a] / L^*_\phi [a](x)$ is a well-defined probability measure.
The relevance of the discriminant to phase transitions is explained by the following theorem (section 5.1, theorem 8): Assume $P_G(\phi) < \infty$. Then

- If $\Delta_n[\phi] > 0$, then $\phi$ is positive recurrent. If $\Delta_n[\phi] = 0$ then $\phi$ is either positive recurrent or null recurrent. If $\Delta_n[\phi] < 0$, then $\phi$ is transient.

- If $\Delta_n[\phi] \geq 0$ then there exists a unique solution $p(\phi)$ for the equation $P_G(\phi + p) = 0$ and if $\Delta_n[\phi] < 0$, no such solution exists.

- If $\Delta_n[\phi] \geq 0$, then $P_G(\phi) = -p(\phi)$. If $\Delta_n[\phi] < 0$, then $P_G(\phi) = -p^*$.

This theorem shows that if the discriminant of a one parameter family $\Delta_n[\phi + t\psi]$ changes sign as $t$ varies, the potential $\phi + t\psi$ changes its mode of recurrence, and the formula for $P_G(\phi + t\psi)$ changes from $-p(\phi + t\psi)$ to $-p^*(\phi + t\psi)$. This implies a change in the qualitative behavior of the equilibrium measure, as well as a possible non-smooth behavior for the pressure.

Our next result discusses the case when $\Delta$ does not change sign, and remains positive. We say that a potential $\phi$ is strongly positive recurrent if its discriminant is strictly positive (this turns out to be independent of the choice of $\phi$). We say that $\psi \in Dir(\phi)$ if $\psi$ has summable variations and $\exists \delta > 0$ such that $P_G(\phi + t\psi)$ is finite for all $|t| < \delta$. The set $Dir(\phi)$ is the set of directions in which $P_G(\psi)$ can be derived at $\phi$.

We prove the following result (section 5.1, theorem 9): if $\phi$ and $\psi$ are weakly Hölder continuous, and $\psi \in Dir(\phi)$, then $\exists \varepsilon > 0$ such that $\phi + t\psi$ is positive recurrent for all $|t| < \varepsilon$ and $t \mapsto P_G(\phi + t\psi)$ is real analytic in $(-\varepsilon, \varepsilon)$.

This shows that if the discriminant remains positive, there are no phase transitions. We remark that it is not necessarily true that if the discriminant is negative then there are no phase transitions. A counter-example is given in chapter 6, section 6.3.5.

The discriminant theorem and the strong positive recurrence theorem suggest the following methodology for proving the existence of phase transitions: estimate $\Delta_n[\phi + t\psi]$ and prove that it must change sign. We use this method in chapter 6.

### 1.3.5 Examples

Chapter 6 is dedicated to the study of one-parameter families of the form $\{\beta\phi\}_{\beta \geq 1}$ where $\phi$ is some fixed function with finite Gurevic pressure. We focus on the analyticity of the function $\beta \mapsto P_G(\beta\phi)$.

In section 6.1 we show that if $X$ satisfies the BIP property and if $V_1(\phi) < \infty$, then there are no phase transitions: $\beta\phi$ remains positive recurrent, and $\beta \mapsto P_G(\beta\phi)$ is real analytic. This is another example of the similarity between systems admitting the BIP property and finite Markov shifts. It can be also be viewed as a generalization of Ruelle's result for the compact case, since every finite Markov shift admits the BIP property.

---

9 Such families appear as models for systems whose temperature is changed. The parameter $\beta$ corresponds to the inverse of the 'temperature'.
1.3. **SUMMARY OF MAIN RESULTS**

Various examples which do admit phase transitions were considered in the context of interval maps by Pomeau and Manneville [PM], Lopes [Lo], Prellberg and Slawny [PS] and for a certain countable Markov chain by Wang [W1, W2]. These examples exhibit the same critical phenomena: for certain $\phi$, $\exists \beta_c$ such that the pressure of $\beta \phi$ is analytic in $\beta$ for $\beta < \beta_c$, non-analytic in $\beta_c$, and linear for $\beta > \beta_c$. These examples were treated using different methods, and it was not clear why they exhibit the same critical behavior, and what other kinds of behavior are possible.

It turns out that these examples can all be modeled by the same countable Markov shift, the renewal shift. In section 6.2 we consider this shift and prove the following general result (theorem 12): If $\phi$ is weakly Hölder continuous and $\sup \phi < \infty$, then there exists $0 < \beta_c < \infty$ such that

- in $(0, \beta_c)$, $\beta \phi$ is positive recurrent and $P_\beta(\beta \phi)$ is real analytic in $\beta$;
- in $(\beta_c, \infty)$, $\beta \phi$ is transient, and $P_\beta(\beta \phi)$ is a linear function;
- $\beta_c$ is a point of non-analyticity for $P_\beta(\beta \phi)$.

We also give sufficient conditions for $\{\beta \phi\}$ to have a critical point (i.e. $\beta_c < \infty$) as well as not to have a critical point (i.e. $\beta_c = \infty$).

In section 6.2.2 we apply this theorem to obtain some new results for the Pomeau-Manneville model, concerning potentials with singular conformal measures (see theorem 13 for details).

Next we discuss the possibility of different critical behavior. Section 6.3 presents some pathological examples. Theorem 14 in section 6.3.1 provides the basic tool for constructing examples, by reducing the problem of constructing the pair $(X, \phi)$ to a problem of constructing certain power series with appropriate behavior at the radius of convergence. We then use this technique to construct the following examples:

- **Infinite number of changes in recurrence.** Gurevich and Savchenko have constructed an example which changes its mode of recurrence an arbitrarily large but finite number of times, but it was unknown whether an infinite number of changes is also possible (private communication). In section 6.3.2 we construct a one parameter family $\{\beta \phi\}_\beta$ which changes from recurrent to transient an infinite number of times.

- **Null recurrence in an interval.** For potentials on the renewal shift, null recurrence can appear only for a single value of $\beta$. In section 6.3.3 we construct a $\phi$ which is not cohomologous to a constant, such that $\beta \phi$ is null recurrent for all $\beta$ in an interval.

- **Coexistence of all modes of recurrence in intervals.** In section 6.3.4 we construct a potential $\phi$, such that $\beta \phi$ is null recurrent in an interval, then becomes positive recurrent in an interval, and then becomes transient in an interval. This shows that all modes of recurrence can ‘coexist’ in intervals of temperatures, and that null recurrence is not necessarily an interface phenomena between positive recurrence and transience.
• *Positive measure set of critical points.* In section 6.3.5 we construct a \( \phi \), such that \( \beta \mapsto P_G(\beta \phi) \) has a positive Lebesgue measure Cantor set of critical points. This one-parameter family is always transient. Thus, this example also shows that unlike strict positivity, strict negativity of the discriminant does not imply lack of phase transitions.

### 1.4 Summary

These results may be summarized in the following way.

1. The ‘correct’ notion of pressure for countable Markov shifts is that of the Gurevich pressure, in the sense that a variational principle holds:

\[
P_G(\phi) = \sup \{ h_\mu(T) + \int \phi \, d\mu : \mu \in \mathcal{P}_T(X) \text{ and } -\int \phi \, d\mu < \infty \}
\]

2. For \( X \) topologically mixing and \( \phi \) with summable variations and finite pressure the following trichotomy holds:

   (a) *Positive recurrence.* There is a conservative conformal measure \( \nu \), equivalent to an invariant finite equilibrium measure (in the sense of theorem 7), and this \( \nu \) is the thermodynamic limit.

   (b) *Null recurrence.* There is a conservative conformal measure, but the invariant measure it is equivalent to is infinite. The thermodynamic limit exists in the following weak sense:

   \[
   \lim_{n \to \infty} \frac{1}{Z_n(a)} \sum_{k=0}^{n-1} \lambda^k \langle L^n \rangle \delta_y \to \nu
   \]

   where \( Z_n(a) \) is a normalization constant which gives the partition set \([a]\) measure one.

   (c) *Transience.* There is no conservative conformal measure.

3. Unlike finite Markov shifts, countable Markov shifts admit phase transitions even for potentials which only depend on a finite number of coordinates. These transitions are closely related to changes in the recurrence properties of the system, and can be analyzed by studying the discriminant:

   (a) when the discriminant of a one-parameter family of potentials changes sign, the system changes its mode of recurrence, and the pressure may lose its differentiability.

   (b) when the discriminant remains positive, the one-parameter family remains positive recurrent and the pressure varies real analytically.

   (c) when the discriminant remains negative, the one-parameter family remains transient, but nothing can be said about the differentiability of the pressure function.

4. Without additional assumptions on the transition matrix, one-parameter families \( \{ \beta \phi \}_{\beta > 1} \) can exhibit a variety of pathological behavior, even if \( \phi \) depends only on a finite number of coordinates.
5. Finite Markov shifts are particular cases of topological Markov shifts with the BIP property. The behavior of such shifts is highly regular: if \( \phi \) is weakly Hölder continuous, such that \( V_1(\phi) \) and \( P_0(\phi) \) are finite, then \( \phi \) admits a Gibbs measure, there are no phase transitions and there is exponential decay of correlations.
Chapter 2

The Gurevich Pressure

The purpose of this chapter is to review the definition of the Gurevich pressure from [S1] and to prove that it satisfies a variational principle (theorem 2). The results of this chapter appeared in [S2].

2.1 The Gurevich pressure

Throughout this section let $X$ be topologically mixing and let $\phi$ be a function with summable variations. For every $a \in S$ set

$$Z_n(\phi, a) = \sum_{T^n x = x} e^{\phi(x)} 1_{[a]}(x).$$

Definition 2 Let $X$ be topologically mixing and $\phi$ be some function with summable variations. The Gurevich Pressure of $\phi$ is the number

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

The following proposition shows that the Gurevich pressure is well-defined.

Proposition 1 Let $X$ be topologically mixing and $\phi$ have summable variations. For every $a \in S$, the limit $\lim_{n \to \infty} \frac{1}{n} \log Z_n(\phi, a)$ exists and is independent of $a$. This limit is never $-\infty$. If $\|L_\phi 1\|_\infty < \infty$, it is not $+\infty$.

Proof. Fix $a$ and set $a_n = \log Z_n(\phi, a)$. By the summable variations property $\{a_n\}_{n \geq 1}$ is almost sub-additive, in the sense that for some constant $c$, and all $m, n \in \mathbb{N}$

$$a_n + a_m < a_{n+m} + c.$$

It follows that the series $\{\frac{a_n}{n}\}$ converges to a limit.

This limit is independent of $a$ because of the topological mixing of $X$. To see that it is never $-\infty$, note that by topological mixing $\exists x \in [a]$ which is periodic.
2.1. **THE GUREVIC PRESSURE**

If its period is $n_c$, then $Z_{knc}(\phi,a) \geq e^{k\phi_{nc}(x)}$, so $a_{knc}/kn_c$ is bounded from below.

To see that the limit is finite when $\|L_\phi 1\|_\infty < \infty$, use summable variations to show that for all $x \in [a]$, $Z_n(\phi,a) = O \left( \left( L_{\phi}^1 [a] \right)(x) \right) = O \left( \|L_\phi 1\|_\infty \right)$. □

The following proposition describes three basic properties of $P_G(\cdot)$. These are known for finite Markov shifts (see e.g. [PP]).

**Proposition 2** Let $X$ be topologically mixing and assume $\phi$ and $\psi$ have summable variations. Then,

1. Addition of constants. For every $c \in \mathbb{R}$, $P_G(\phi + c) = P_G(\phi) + c$.
2. Convexity. For every $t \in [0,1]$, $P_G(t\phi + (1-t)\psi) \leq tP_G(\phi) + (1-t)P_G(\psi)$.
3. Invariance under cohomology. If for some $f$, $\phi - \psi = f - f \circ T$, then $P_G(\phi) = P_G(\psi)$.

**Proof.** The first property is proved by direct calculation. The second follows from $Z_n(t\phi + (1-t)\psi,a) \leq Z_n(\phi,a)^t Z_n(\psi,a)^{1-t}$ (Hölder’s inequality). The third follows from the identity $\phi_n(x) = \psi_n(x) + f(x) - f(T^n x)$, which implies that if $T^n x = x$ then $\phi_n(x) = \psi_n(x)$. □

The Gurevic pressure is a generalization of the Gurevic entropy $h_G(T)$, in the sense that $h_G(T) = P_G(\phi)$ for $\phi \equiv 0$. We omit this straightforward calculation, as it appeared in [S1].

The following theorem gives an alternative formula for $P_G(\phi)$, which relates the Gurevic pressure to the usual definition of topological pressure for finite Markov shifts (see [Rui1], [B1]). This was proved for the case $\phi \equiv 0$ by Gurevic in [G2].

**Theorem 1** If $X$ is topologically mixing and $\phi$ has summable variations, then

$$P_G(\phi) = \sup \{ P_{top}(\phi|_Y) : Y \subseteq X \text{ top. mixing finite Markov shift} \}$$

$$= \sup \{ P_{top}(\phi|_K) : K \subseteq X \text{ compact ; } T^{-1}K = K \}$$

where $P_{top}(\phi|_Y)$, respectively $P_{top}(\phi|_K)$, is the topological pressure of the restriction of $\phi$ to the compact metric spaces $Y$, respectively $K$.

**Proof.** The second equality is trivial since every finite Markov shift is a compact invariant set, and every compact invariant set is a subset of a topologically mixing finite Markov shift.

We prove the first equality. Let $Y \subseteq X$ be an arbitrary topologically mixing finite Markov shift, and set $Z_n(Y,\phi,a) = \sum_{T^n x = x} e^{\phi_n(x)} 1_Y(x) 1_{[a]}(x)$. It is known that for finite Markov shifts (see e.g. [PP]),

$$P_{top}(\phi|_Y) = \lim_{n \to \infty} \frac{1}{n} \log Z_n(Y,\phi,a)$$

whence, since $Z_n(Y,\phi,a) \leq Z_n(\phi,a)$, $P_{top}(\phi|_Y) \leq P_G(\phi)$. Thus,

$$P_G(\phi) \geq \sup \{ P_{top}(\phi|_Y) : Y \subseteq X \text{ top. mixing finite Markov shift} \}.$$
We show the reverse inequality, under the assumption that \( P_G(\phi) < \infty \) (the case \( P_G(\phi) = \infty \) is similar). Set \( B := \sum_{n \geq 2} V_n(\phi) \), fix \( \varepsilon > 0 \) and let \( m > \frac{2B}{\varepsilon} \) be large enough so that

\[
P_G(\phi) < \frac{1}{m} \log Z_m(\phi, a) + \varepsilon.
\]

Identifying \( S \) with \( N \), we choose \( M \) large enough so that

\[
\frac{1}{m} \log Z_m(\phi, a) \leq \frac{1}{m} \log Z_m\left(\{1, \ldots, M\}^{\mathbb{N}^+} \cap X, \phi, a\right) + \varepsilon.
\]

Adding a finite number of states to \( \{1, \ldots, M\} \), one can construct a topologically mixing finite Markov shift \( Y \subseteq X \), such that

\[
\frac{1}{m} \log Z_m(\phi, a) < \frac{1}{m} \log Z_m(Y, \phi, a) + \varepsilon.
\]

Set \( a_n = \log Z_n(Y, \phi, a) \). By the definition of \( B \), \( a_n + a_m \leq a_{n+m} + 2B \), whence for \( n = km + r \) \( (r = 0, \ldots, k - 1) \)

\[
\frac{ka_m + ar}{km + r} \leq \frac{a_{km+r} + 2(k+1)B}{km + r} \leq \frac{a_n}{n} + \frac{k+1}{k} \varepsilon.
\]

Fixing \( m \) and passing to the limit as \( n \to \infty \), we have

\[
\frac{1}{m} \log Z_m(Y, \phi, a) \leq P_{\text{op}}(\phi|_Y) + \varepsilon.
\]

By (2.1), (2.2), (2.3), \( P_G(\phi) \leq P_{\text{op}}(\phi|_Y) + 3\varepsilon \). Since \( \varepsilon \) was arbitrarily small, the first equality is proved.

\[
\square
\]

### 2.2 The variational principle

Denote by \( \mathcal{P}_T(X) \) the set of all invariant Borel probability measures. The variational principle for finite Markov shifts states that the topological pressure of a continuous function \( \phi \) is equal to \( \sup \{ h_\mu(T) + \int \phi \, d\mu : \mu \in \mathcal{P}_T(X) \} \). Such a theorem was proved for countable Markov shifts by Gurevic [G2] in the case \( \phi \equiv 0 \) (in other words, the Gurevic entropy is equal to the supremum of all metric entropies). The following theorem extends this result.

We restrict ourselves to the case sup \( \phi < \infty \) and consider only probabilities \( \mu \), such that \( -\int \phi \, d\mu < \infty \) to avoid problems with the integrability of \( \phi \) and the definition of \( h_\mu(T) \) for general \( \mu \in \mathcal{P}_T(X) \). (The latter is not defined when \( h_\mu(T) = \infty \) and \( \int \phi \, d\mu = -\infty \).)

**Theorem 2** Let \( X \) be a topologically mixing countable Markov shift and \( \phi \) have summable variations. If sup \( \phi < \infty \) then

\[
P_G(\phi) = \sup \left\{ h_\mu(T) + \int \phi \, d\mu \mid \mu \in \mathcal{P}_T(X) ; -\int \phi \, d\mu < \infty \right\}.
\]
2.2. THE VARIATIONAL PRINCIPLE

Proof. Let $P_\mathcal{G} (\phi) \leq \sup \{ h_\mu (T) + \int \phi \, d \mu \}$ by theorem 1 and the variational principle for finite Markov shifts $Y \subseteq X$ (see [B1]). Note that this inequality is actually an equality for the case $P_\mathcal{G} (\phi) = \infty$. It is therefore enough to assume $P_\mathcal{G} (\phi) < \infty$ and prove the reverse inequality.

Fix some $\mu \in \mathcal{P} (X)$ such that $- \int \phi \, d \mu < \infty$. Ordering the states, we assume without loss of generality that $S = \mathbb{N}$. Set $[\geq m] = \{ x : x_0 \geq m \}$, $\alpha_m = \{ [1], \ldots, [m - 1], [\geq m] \}$ and $\mathcal{B}_m = \sigma (\alpha_m)$. As $m \uparrow \infty$, $\mathcal{B}_m \uparrow \bigcup_m \mathcal{B}_m \subseteq \sigma (\bigcup_m \mathcal{B}_m) = \mathcal{B}$, whence (see e.g. [DGS])

$$h_\mu (T, \alpha_m) + \int \phi \, d \mu \rightarrow_{m \rightarrow \infty} h_\mu (T) + \int \phi \, d \mu.$$

Fix $m$ and set $\beta = \alpha_m$. For every $\mathbf{a} = (a_0, \ldots, a_n)$ where $\forall i$, $a_i \in \beta$ set

$$\langle \mathbf{a} \rangle = \{ a_0, \ldots, a_n \} := \bigcap_{k=0}^{n} T^{-k} a_i.$$

Set $\phi_n (\mathbf{a}) := \sup \{ \phi_n (x) : x \in \langle \mathbf{a} \rangle \}$. Then,

$$\frac{1}{n} H_\mu (\beta^n_0) + \int \phi \, d \mu = \frac{1}{n} \left( H_\mu (\beta^n_0) + \int \phi_n \, d \mu \right) \leq$$

$$\leq \frac{1}{n} \sum_{a, b \in \beta} \mu (a \cap T^{-n} b) \sum_{\langle \mathbf{a} \rangle \subseteq a \cap T^{-n} b} \mu (\langle \mathbf{a} \rangle) \log \frac{e^{\phi_n (\langle \mathbf{a} \rangle)}}{\mu (\langle \mathbf{a} \rangle)}$$

$$\leq \frac{1}{n} \sum_{a, b \in \beta} \mu (a \cap T^{-n} b) \log \sum_{\langle \mathbf{a} \rangle \subseteq a \cap T^{-n} b} e^{\phi_n (\langle \mathbf{a} \rangle)} + \frac{1}{n} H_\mu (\beta \cap T^{-n} \beta)$$

where the last inequality follows from Jensen's inequality for sums. Set

$$P_n (a, b) := \frac{1}{n} \log \sum_{\langle \mathbf{a} \rangle \subseteq a \cap T^{-n} b} e^{\phi_n (\langle \mathbf{a} \rangle)}.$$

Since $\beta$ is a finite partition, $H_\mu (\beta \cap T^{-n} \beta) = O(1)$. Therefore, since $\beta = \alpha_m$,

$$h_\mu (T, \alpha_m) + \int \phi \, d \mu \leq \limsup_{n \rightarrow \infty} \left\{ \sum_{a, b \in \beta} \mu (a \cap T^{-n} b) P_n (a, b) \right\}$$

We estimate $P_n (a, b)$. Assume, first, that $a, b \notin [\geq m]$. In this case, since $\alpha_0^n$ is finer than $\beta^n_0$,

$$P_n (a, b) \leq \frac{1}{n} \log \sum_{\langle \mathbf{a} \rangle \subseteq [a] \cap T^{-n} [a]} e^{\sup \{ \phi_n (x) : x \in \langle \mathbf{a} \rangle \}}$$

$$\phi_n (x) = \sup \{ \phi_n (x) : x \in \langle \mathbf{a} \rangle \}.$$
By the summable variations of \( \phi \) and the topological mixing of \( X \), the right side tends to \( P_G(\phi) \), whence

\[
(2.5) \quad a, b \notin [\geq m] \implies \limsup_{n \to \infty} P_n(a, b) \leq P_G(\phi)
\]

Assume, now, that \( a = [\geq m] \) or \( b = [\geq m] \). For every \( \langle a \rangle \in \beta^n \) such that \( \langle a \rangle \subseteq a \cap T^{-n} b \), \( \exists i, j, k \geq 0 \) such that \( i + j + k = n + 1 \) and

\[
\langle a \rangle = \langle \underbrace{m, \ldots, m}_i, \ldots, \underbrace{m, \ldots, m}_j \rangle
\]

where \( \xi_i, \xi_j \notin [\geq m] \). For such \( i, j, k \), \( \phi_n(\langle a \rangle) \leq (i + k)\sup \phi + \phi_j(\xi_i, \ldots, \xi_j) \).

Summing over all possibilities we have

\[
P_n(a, b) \leq \frac{1}{n} \log \left( \sum_{a', b'=1}^n e^{iP_{\gamma}(a', b')} \right) + \sup \phi
\]

Thus, by the preceding discussion,

\[
(2.6) \quad a = [\geq m] \text{ or } b = [\geq m] \implies \limsup_{n \to \infty} P_n(a, b) \leq P_G(\phi) + [\sup \phi]
\]

We now use (2.5) and (2.6) to finish the proof. Summing over all \( a, b \in \beta \), we have

\[
\limsup_{n \to \infty} \left\{ \sum_{a, b \in \beta} \mu(a \cap T^{-n} b)P_n(a, b) \right\}
\]

\[
\leq P_G(\phi) \sum_{a, b \in [\geq m]} \mu(a \cap T^{-n} b) + (P_G(\phi) + [\sup \phi]) \sum_{\neg ([a, b \in [\geq m]}) \mu(a \cap T^{-n} b)
\]

\[
\leq P_G(\phi) + 2\mu[\geq m]\sup \phi = P_G(\phi) + o(1), \quad \text{as } m \to \infty.
\]

Combining this with (2.4), we have that

\[
h_\mu(T) + \int \phi d\mu = \lim_{m \to \infty} \left( h_\mu(T, \alpha_m) + \int \phi d\mu \right) \leq P_G(\phi)
\]

as required. \( \square \)

### 2.3 Notes

The topological pressure for \( \phi \equiv 0 \) on a general countable Markov shift (i.e. the topological entropy) was calculated by B. Gurevic [G1], who showed that it is equal to the supremum of topological entropies of all finite Markov shifts \( Y \subseteq X \). Gurevic’s calculation is based on passing to a certain compactification of \( X \), using the variational principle for compact metric spaces to obtain an
inequality, and showing that the topological entropy on the compactification can be approximated by the topological entropy on compact invariant subsets of $X$ to obtain the reverse inequality. This argument was generalized by Zargaryan [Z] for some bounded functions $\phi$ which can be uniformly extended to a certain compactification of $X$, and by Gurevic and Savchenko [GS] to all functions which are uniformly continuous and bounded.

Definition 2 and proposition 1 previously appeared in [S1], where it was also shown that $P_G(0)$ is equal to Gurevic’s entropy. Theorem 1 extends this by showing that $P_G(\phi)$ coincides with the quantities considered by Zargaryan [Z], and by Gurevic and Savchenko [GS] for non-constant $\phi$. Theorem 2 improves the variational principle of Gurevic and Savchenko in the sense that it also covers functions which are not bounded from below. The motivation for considering such functions is that for bounded potentials, $\sup \{ h_\mu(T) + \int \phi d\mu \} \geq \sup h_\mu(T) + \inf \phi$ and this is infinite for most countable Markov shifts. Thus, functions $\phi$ with non-trivial pressure are typically unbounded from below. Such potentials have no continuous extension to a compactification, and cannot thus be treated using the approach of [Z] and [GS].
Chapter 3

The Generalized Ruelle’s Perron-Frobenius Theorem

The purpose of this chapter is to formulate and prove a generalization of Ruelle’s Perron-Frobenius theorem for countable Markov shifts. The results of this chapter will appear in [S3].

3.1 A generalized Ruelle’s Perron-Frobenius theorem

Let \((X, T)\) be a topologically mixing countable Markov shift with a set of states \(S\) and a transition matrix \(A = (t_{ij})_{S \times S}\). For every \(a \in S\) set

\[ \varphi_a(x) := 1_{[a]}(x) \inf \{n \geq 1 : T^n x \in [a]\} \]

where \(\inf \emptyset := \infty\) and let

\[ Z_n(\phi, a) := \sum_{T^n x = x} e^{\phi_n(x)} 1_{[a]}(x) \quad \text{and} \quad Z^*_n(\phi, a) := \sum_{T^n x = x} e^{\phi_n(x)} 1_{[\varphi_x = n]}(x) \]

where \(\phi_n := \phi + \phi \circ T + \ldots + \phi \circ T^{n-1}\).

**Definition 3** Let \(X\) be topologically mixing and \(\phi\) have summable variations and finite Gurevich pressure \(P_G(\phi) = \log \lambda\). Fix some \(a \in S\). We call \(\phi\):

1. **recurrent** if \(\sum_{n \geq 1} \lambda^{-n} Z_n(\phi, a) = \infty\), and **transient** otherwise.

2. **positive recurrent** if it is recurrent and \(\sum_{n \geq 1} n \lambda^{-n} Z^*_n(\phi, a) < \infty\).

3. **null recurrent** if it is recurrent and \(\sum_{n \geq 1} n \lambda^{-n} Z^*_n(\phi, a) = \infty\).
3.1. A GENERALIZED RUELLE’S PERRON-FROBENIUS THEOREM

By topological mixing and summable variations for all $a, b \in S$ there is a constant $C$ and a number $n_0$, such that $Z_n(\phi, a) \leq CZ_{n+n_0}(\phi, b)$ for all $n$. This shows that $\sum k Z_n(\phi, a)$ converges and diverges together with $\sum k Z_n(\phi, b)$, whence recurrence and transience are independent of the choice of $a$. It turns out that the notions of positive recurrence and null recurrence are also independent of $a$. This follows from theorem 3 below.

Recall that Ruelle’s operator is given by $L_\phi f(x) = \sum_{y=x} e^{\phi(y)} f(y)$. The main result of this chapter is the following theorem, which is a generalization of Ruelle’s Perron-Frobenius theorem [Ru1] to countable Markov shifts.

**Theorem 3 (Generalized RPF theorem)** Let $X$ be topologically mixing and assume $\phi$ has summable variations and finite Gurevich pressure. $\phi$ is recurrent iff there exist $\lambda > 0$, a conservative $\sigma$-finite measure $\nu$, finite and positive on cylinders, and a positive continuous function $h$, such that $L_\lambda^\nu = \nu$ and $L_\phi h = \lambda h$. In this case $\lambda = \exp P_G(\phi)$ and $\exists \alpha_n \uparrow \infty$, such that for every cylinder $[a]$ and $x \in X$

$$\frac{1}{\alpha_n} \sum_{k=1}^{n} k (F^{k+1}_{\phi 1}[a]) (x) \to \nu [a]$$

where $\{\alpha_n\}$ satisfies $\alpha_n \sim \left( \int_{[a]} h d\nu \right)^{\frac{1}{\lambda}} \sum_{k=1}^{n} k Z_k(\phi, a)$ for every $a \in S$. Furthermore,

1. if $\phi$ is positive recurrent then $\int h d\nu < \infty$, $\alpha_n \propto n$ and for every cylinder $[a]$, $\lambda \left( n L_{\phi}^n[1][a] \to \nu [a] / \nu [\Omega] \right)$ uniformly on compacts.

2. if $\phi$ is null recurrent then $\int h d\nu = \infty$, $\alpha_n = o(n)$, and $\lambda \left( n L_{\phi}^n[1][a] \to 0 \right)$ uniformly on partition sets.

It follows from the proof that $\nu$ is exact, that $h$ is bounded away from zero and infinity on partition sets, and that

$$V_m(\log h) < \log \sum_{i=m+1}^\infty V_i(\phi)$$

In particular, if $\phi$ is weakly Hölder continuous, then so are $\log h$ and $\log h \circ T$.

The generalized RPF theorem implies the following useful result (compare with Walters [Wal1] and Keane [Ke]):

**Corollary 1** Let $X$ be topologically mixing and let $\phi$ be a function with summable variations and finite Gurevich pressure. Then there exist two continuous functions $\phi'$ and $\varphi$, such that $\phi' \leq 0$, $P_G(\phi') = 0$ and $\phi' = \phi + \varphi - \varphi \circ T - P_G(\phi')$. The function $\varphi$ is bounded on partition sets. If $\phi$ is recurrent, then $L_{\phi'} 1 = 1$, and if $\phi$ is transient, then $L_{\phi'} 1 \leq 1$. If $\phi$ is weakly Hölder continuous, then so are $\phi'$ and $\varphi$. 
Proof. Set $\lambda := \exp P_{\mathcal{G}}(\phi)$. Assume that $\phi$ is transient. Fix some state $a \in S$ and set $h := \sum_{n \geq 0} \lambda^n L^n \varphi_1[|a|]$. By transience, topological mixing and summable variations $h$ is finite. By (3.1), if $\phi$ is weakly Hölder, then so are $\log h$ and $\log h \circ T$. It is easy to check that $\lambda^1 L_\phi h \leq h$. Set $\varphi := \log h$ and $\varphi' := \phi + \varphi - \varphi \circ T - P_{\mathcal{G}}(\phi)$. Clearly, $L_{\varphi'} 1 \leq 1$ whence $\varphi' \leq 0$ as required. The case when $\phi$ is recurrent is handled by replacing $h$ in the last argument by the $h$ given by theorem 3. \qed

The rest of the chapter is devoted to the proof of theorem 3. Throughout the proof we assume that $X$ is a topologically mixing countable Markov shift and that $\phi$ has summable variations and finite pressure $P_{\mathcal{G}}(\phi) = \log \lambda$. Set

$$B_k = \exp \sum_{n=k+1}^{\infty} V_n(\phi) \quad (k = 1, 2, \ldots)$$

By summable variations $\forall n \geq 1 \ B_n < \infty$ and $B_n \downarrow 1$. The following inequality will be used constantly:

$$(3.2) \ x_0 = y_0, \ldots, x_{n-1} = y_{n-1} \Rightarrow \forall m \leq n - 1 \quad \left( e^{\phi_m(x)} = B_n^{\pm 1} e^{\phi_m(y)} \right)$$

A frequently used corollary is that $\forall x_a \in [a],$

$$Z_n(\phi, a) = B_n^{\pm 1} \left( L^n \varphi_1[|a|] \right)(x_a).$$

The reader should note that the assumption that the Gurevic pressure is finite implies that all of the $Z_n(\phi, a)$ are finite (because by summable variations $\exists C > 1$, such that $\forall m, n \ C \ m Z_n(\phi, a)^m < Z_{mn}(\phi, a)$). This assumption also implies that the $L^n_\phi$ are all defined on bounded functions supported inside a finite union of partition sets.

We break the proof into several steps. These appear in the following sections.

## 3.2 Construction of eigenmeasure

Recall that for every measure $\mu$ such that $\mu \sim \mu \circ T^{-1}$ there exists an operator $\hat{T} : L^1(\mu) \rightarrow L^1(\mu)$ uniquely defined by the condition

$$\forall f \in L^1(\mu) \forall \varphi \in L^\infty(\mu) \int \varphi \cdot \hat{T} f \ d\mu = \int \varphi \circ T \cdot f \ d\mu$$

This operator is called the transfer operator, or the dual operator of $\mu$ (see e.g. [A1], section 1.3).

### Proposition 3
If there exists $\lambda > 0$ and a conservative $\sigma$-finite measure $\nu$ which is finite and positive on some cylinder, such that $L^*_\phi \nu = \lambda \nu$, then $\phi$ is recurrent and $\lambda = e^{P_{\mathcal{G}}(\phi)}$. 

3.2. CONSTRUCTION OF EIGENMEASURE

Proof. Choose a cylinder $[b]$ with a finite positive measure. It easy to verify that $\lambda^1 L_\phi$ acts as the transfer operator of $\nu$ whence by conservativity

$$\sum_{n \geq 1} \lambda^n Z_n (\phi, b_c) \geq B_1 \sum_{n \geq 1} \lambda^n (L_\phi^n [b]) (x) = \infty.$$  

We show that $\lambda = e^{P_\phi (\phi)}$. It follows from what we just proved that $\lambda \leq e^{P_\phi (\phi)}$ because the radius of convergence of the series $\sum_{k \geq 1} Z_k (\phi, b_c) x^k$ is $e^{P_\phi (\phi)}$. Consider $Z_n (\phi, b) = \sum_{T^n x = x} e^{\phi_n (x)} 1_{[b]} (x)$. By summable variations,

$$\lambda^n Z_n (\phi, b) \leq B_1 \left[ \frac{1}{\nu (b)} \int \left( \lambda^n L_\phi^n [b] \right) d\nu \right] \leq B_1.$$

By topological mixing and summable variations $\frac{1}{n} \log Z_n (\phi, b) \to P_\phi (\phi)$ whence $\lambda \geq e^{P_\phi (\phi)}$.  

Proposition 4 If $\phi$ is recurrent, then there exist $\lambda > 0$ and a conservative measure $\nu$, finite and positive on cylinders, such that $L_\nu \nu = \lambda \nu$.

Proof. Fix some $a \in S$, set $\lambda := e^{P_\phi (\phi)}$ and let $a_n := \sum_{k=1}^n \lambda^k Z_k (\phi, a)$. Since $\phi$ is recurrent, $a_n \uparrow \infty$. For every $y \in X$ let $\delta_y$ denote the probability measure concentrated on $\{y\}$. Fix a periodic point $x_a \in [a]$ and set for every $b \in S$

$$\nu^b_n := \frac{1}{a_n} \sum_{k=1}^n \lambda^k \sum_{T^n y = x_a} e^{\phi_n (y)} 1_{[b]} (y) \delta_y.$$  

By definition, $\nu^b_n (X) = \nu^b_n [b] = \frac{1}{a_n} \sum_{k=1}^n \lambda^k \left( L_\phi^k [b] \right) (x_a)$. We claim that for every $b \in S$,

$$0 < \lim_{n \to \infty} \nu^b_n (X) \leq \lim_{n \to \infty} \nu^b_n (X) < \infty.$$  

First, note that

(3.3) \quad \forall k \lambda^k Z_k (\phi, a) < 2 B_1.

This follows from the inequality $\lambda^{kn} Z_k (\phi, a) \geq B_1 e^{m \lambda^k Z_k (\phi, a)}$. Next, note using topological mixing and summable variations that $\exists C, n_0$, such that $L_\phi^n (x_a) < C Z_{n+n_0} (\phi, a)$. It follows that for $n$ large

$$\frac{1}{a_n} \sum_{k=1}^n \lambda^k L_\phi^k [b] (x_a) \leq \frac{\lambda^{n_0}}{a_n} \sum_{k=1+n_0} C \lambda^k Z_k (\phi, a) \leq C \lambda^{n_0} \left( 1 + \frac{2 B_1 n_0}{a_n} \right) = O(1)$$

1If this sum is finite at some point in $[b]$, it is finite everywhere in $[b]$ and converges there uniformly (because of summable variations). Thus for some $N$, $[b] \setminus \bigcup_{n \geq N} T^{-n} [b]$ has positive measure in contradiction to conservativity.
as \( n \uparrow \infty \). The opposite estimation is done in the same way.

We show how to choose a subsequence \( \{ m_k \}_{k \geq 1} \), such that for every \( b \in S \), \( \{ \nu_{m_k}^b \} \) is \( w^* \) convergent, and show that the non-trivial measure \( \nu \) given by
\[
\nu_{m_k}^b \xrightarrow{w^*} \nu_{[1]} \text{ satisfies } L_b^* \nu = \lambda \nu.
\]
Since \( X \) is not compact, to do this we have to prove that \( \{ \nu_{m_k}^b \}_{k \geq 1} \) are all tight, i.e.,
\[
\forall b \forall \varepsilon > 0 \exists F = F_{b, \varepsilon} \text{ compact such that } \forall n \quad \nu_n^b(F^c) < \varepsilon
\]
Since \( X \) is topologically mixing and \( \phi \) has summable variations, if \( \{ \nu_n^b \}_{n \geq 1} \) is tight for some \( b \), then it is tight for every \( b \). Therefore, we may restrict ourselves to the case \( b = a \) and set
\[
\nu_n := \nu_n^a.
\]

**Step 1.** We show that \( \sum_{k \geq 1} \lambda^{-k} Z_k^* (\phi, a) < \infty \). To see this, set
\[
t(x) := 1 + \sum_{k=1}^{\infty} Z_k (\phi, a) x^k \quad \text{and} \quad r(x) := \sum_{k=1}^{\infty} Z_k^* (\phi, a) x^k.
\]
It is not difficult to verify\(^2\) that \( \forall x \in (0, \lambda^{-1}) \), \( t(x) - 1 = B_1^{a2} r(x) t(x) \). Therefore \( \forall x \in (0, \lambda^{-1}) \) \( r(x) \leq B_1^2 \) whence \( r(\lambda^{-1}) < \infty \).

**Step 2.** Set \( \tau_n(x) := \tau_a(x) \) and
\[
\tau_n(x) := \begin{cases} 
\varphi_a(T^{\tau_1(x)} + \cdots + \tau_{n-1}(x)x) & \text{if } \tau_{n-1}(x) < \infty \\
+\infty & \text{else}
\end{cases}
\]
Note that \( \tau_n > 0 \) only if \( x_0 = a \). For every sequence of natural numbers \( \{ n_i \}_{i \geq 1} \) set
\[
R\left( \{ n_i \}_{i \geq 1} \right) := \{ x \in [a]: \forall i \quad \tau_i(x) \leq n_i \}.
\]
Fix some \( \varepsilon > 0 \). We construct a sequence \( \{ n_i \} \) such that
\[
(3.4) \quad \forall n \quad \nu_n \left[ R\left( \{ n_i \} \right) \right] < \varepsilon
\]
Set \( \Lambda_x(k_1, \ldots, k_m) := \{ x \in [a]: \forall j \leq m \quad \tau_j(x) = k_j \} \) and
\[
Z_{k_1, \ldots, k_m}^* := \sum_{T^{k_1 + \cdots + k_m} x = x} e^{\phi_{k_1 + \cdots + k_m}} \Lambda_x(k_1, \ldots, k_m)(x)
\]
For every \( \{ n_i \}_{i \geq 1} \), such that \( n_i \) is larger than the period of \( x_a \) for every \( i \),
\[
\nu_n \left[ R\left( \{ n_i \} \right) \right] \leq \sum_{i=1}^{\infty} \nu_n \left[ \tau_i > n_i \right] = \frac{\nu_n \left[ \tau_i > n_i \right]}{\nu_n \left[ \tau_i > n_i \right]}.
\]
\(^2\)This follows from the approximate renewal relation
\[
Z_a(\phi, a) = B_1^{a2} \left( Z_a^* (\phi, a) + Z_{a-1}^* (\phi, a) Z_1 (\phi, a) + \cdots + Z_{a-1}^* (\phi, a) Z_{a-1} (\phi, a) \right).
\]
\[ \sum_{i=1}^{\infty} \frac{1}{\alpha_n} \sum_{k=1}^{n} \lambda^k \sum_{n_{i+1}=1}^{n} e^{\psi_k(y)} 1_{\tau_i > n_i} (y) \]

\[ \leq \sum_{i=1}^{\infty} \frac{1}{\alpha_n} \sum_{k=n_i+1}^{n} \lambda^k \sum_{n_{i+1}=1}^{n} e^{\psi_k(y)} \sum_{k_i > n_i, N \leq k} Z_{k_i, \ldots, k_{i-1}, k_i}^* Z_{k_{i+1}, \ldots, k_N}^* \]

\[ \leq \sum_{i=1}^{\infty} \frac{1}{\alpha_n} \sum_{k=n_i+1}^{n} \lambda^k Z_{k_i}^* \sum_{k_i > n_i, N \leq k} Z_{k_i, \ldots, k_{i-1}, k_i}^* Z_{k_{i+1}, \ldots, k_N}^* \]

\[ \leq \frac{1}{\alpha_n} \sum_{k=n_i+1}^{n} \lambda^k Z_{k_i}^* \left( \sum_{k_i > n_i, N \leq k} \sum_{k=1}^{\infty} \lambda^k Z_{k_i}^* \right) \]

\[ \leq \sum_{i=1}^{\infty} \sum_{k=n_i+1}^{n} \lambda^k Z_{k_i}^* \]

It remains to apply the previous step and choose \( n_i \) such that

\[ \sum_{k=n_i+1}^{\infty} \lambda^k Z_{k_i}^* < \frac{\varepsilon}{2B_1^i}. \]

**Step 3.** Fix \( \{n_i\} \) such that \( \forall n \ \nu_n (R \{n_i\}^c) < \varepsilon. \) For every sequence of natural numbers \( \{k_i\} \) set

\[ S(\{k_i\}_{i \geq 1}) = \{x \in [a]: \forall i \ \tau_i (x) = k_i \} \]

We show that for every \( \varepsilon > 0 \) there exists a compact set \( F \subseteq [a] \) such that

(3.5) \[ \forall i \ k_i \leq n_i \Rightarrow \forall n \ \nu_n (F^c \cap S \{k_i\}) \leq \varepsilon \nu_n (S \{k_i\}). \]

This is enough to prove tightness, because (3.5) implies that for every \( n, \)

\[ \nu_n (F^c) \leq \nu_n (F^c \cap R \{n_i\}^c) + \sum_{\{k_i\}: \forall k_i \leq n_i} \nu_n (F^c \cap S \{k_i\}) \]

\[ \leq \varepsilon + \varepsilon \nu_n (X) \]

whence \( \nu_n (F^c) = O(\varepsilon), \) since we already know that \( \nu_n (X) = O(1). \)

The \( F \) we construct is of the form \( F = \{x \in [a]: \forall i \ x_i \leq N_i\} \) where \( N_i \in S \)

(we are using an order on \( S \) induced by the identification \( S \approx N). \) This is a compact set.
We show how to choose \( \{N_i\} \). Set
\[
\Theta_s(k;N) := \{x \in [a]; \tau_1(x) = k \text{ and } \exists i x_i > N\}
\]
\[
Z^*_k(N) = \sum_{T^x=x} e^{\theta_s(x)1_{\Theta_s(k;N)}(x)}
\]
Obviously, \( Z^*_k(N) \downarrow 0 \) as \( N \uparrow \infty \). For every \( i \), we choose \( N_i \) in a way such that
\[
\forall k \leq n_i, \quad Z^*_k(N_i) \leq \frac{\varepsilon}{2B^1_i} Z^*_k.
\]
We make sure that \( \{N_i\} \) are increasing and that \( N_1 > \sup_{i \geq 0} \{x_n(i)\} \).\(^3\)

Fix \( \{k_i\} \) such that \( \forall i k_i \leq n_i \) and \( \nu_n(S \{k_i\}) > 0 \). Fix \( N = N(n, \{k_i\}) \) such that \( k_1 + \ldots + k_N \geq n \). Then
\[
\nu_n(F^c \cap S \{k_i\}) \leq \sum_{i=1}^{\infty} \nu_n \left\{ x \in S \{k_i\} : \exists j \epsilon \left( \sum_{m=1}^{i} k_m \right) x_j > N_j \right\}
\]
We can replace the half open intervals by open intervals, since \( x_j = a < N_j \) in the edges. We can replace \( \infty \) by \( N \), because for \( j > k_1 + \ldots + k_N = n \), if \( x \in \text{supp}(\nu_n) \) then \( x_j \in \{x_n(i) : i \geq 0\} \) whence \( x_j \leq \text{sup}\{x_n(i)\} < N_j \). Since \( j > k_1 + \ldots + k_{i-1} \geq i-1, N_j \geq N_i \), so if we replace \( N_j \) by \( N_i \) we only increase the quantity on the right. Thus,
\[
\nu_n(F^c \cap S \{k_i\}) \leq \sum_{i=1}^{N} \nu_n \left\{ x \in S \{k_i\} : \exists j \epsilon \left( \sum_{m=1}^{i} k_m \right) x_j > N_i \right\}
\]
To calculate the right side, note that for every \( y \) in the support of \( \nu_n \), \( \exists k \) such that \( T^k y = x_a \). If \( y \in S \{k_i\} \) then \( k \) must be of the form \( k_1 + \ldots + k_i \), and if \( \exists j > \sum_{m=1}^{i} k_m \) such that \( y_j > N_i \) then \( \ell \geq i \), else \( j \geq k \) and \( y_j < N_i \). Thus,
\[
\nu_n(F^c \cap S \{k_i\}) \leq \sum_{i=1}^{N} \left( \frac{1}{a_n} \sum_{\ell=i}^{N} \lambda^{(k_1 + \ldots + k_i)} Z^*_k \ldots Z^*_1 1_{S \{k_m\} m > \ell} (x_a) \right)
\]
\[
\leq \sum_{i=1}^{N} \left( \frac{1}{2B^1_i a_n} \sum_{\ell=i}^{N} \lambda^{(k_1 + \ldots + k_i)} Z^*_k \ldots Z^*_1 1_{S \{k_m\} m > \ell} (x_a) \right)
\]
\[
\leq \sum_{i=1}^{N} \left( \frac{\varepsilon}{2B^1_i a_n} \sum_{\ell=i}^{N} \lambda^{(k_1 + \ldots + k_i)} Z^*_k \ldots Z^*_1 1_{S \{k_m\} m > \ell} (x_a) \right)
\]
\[
\leq \varepsilon \nu_n(S \{k_i\})
\]
Tightness is proved.

\( ^3 \)This supremum is finite because \( x_n \) is periodic.
3.3. CONSTRUCTION OF EIGENFUNCTION AND (3.1)

By tightness, there exists a subsequence \( m_k \) such that \( \forall b \in S, \{\nu^b_{m_k}\}_{k \geq 1} \) is \( \nu^* \)-convergent. We denote its limit by \( \nu^b \) and set \( \nu = \sum_{b \in S} \nu^b \). It is not difficult to check that
\[
\forall [b] \quad 0 < \nu ( [b] ) < \infty
\]

We show that \( L^*_\phi \nu = \lambda \nu \). By recurrence, \( a_n \uparrow \infty \). We have already pointed out that \( \lambda^k Z_k ( \phi, a ) = O(1) \), whence \( a_n \sim a_{n+1} \). A standard calculation shows that for every \( [b] \) and \( N, \nu ( [1_{[x_0 < N]} L^\phi 1_{[b]} ] ) = \lambda \nu ( [1_{[x_0 < N]} 1_{[b]} ] ) \). By the monotone convergence theorem, \( \nu ( L^\phi 1_{[b]} ) = \lambda \nu ( [b] ) \). Since \( [b] \) was arbitrary, we have that \( L^*_\phi \nu = \lambda \nu \).

We show that \( \nu \) is conservative. One checks that the transfer operator of \( \nu \) is \( \hat{T} = \lambda^{-1} L^\phi \). To prove conservativity it is enough to show that for some positive integrable function \( f, \sum_{k \geq 1} \hat{T}^k f = \infty \) almost everywhere. Set \( f = \sum_{b \in S} f_b 1_{[b]} \) where \( f_b > 0 \) are chosen so that \( \nu ( f ) < \infty \). For every \( b \in S \) and \( x \in [b] \),
\[
\sum_{k=1}^\infty \lambda^k ( L^k \phi f ) ( x ) \geq B_1 \sum_{k=1}^\infty \lambda^k Z_k ( \phi, b )
\]
Choose some \( C_t, n_t \) such that for \( n \) large, \( Z_n ( \phi, b ) \geq C_t Z_{n_t} ( \phi, b ) \). Then \( \sum \lambda^k L^k \phi f \) dominates a divergent series, and therefore diverges. \( \square \)

3.3.1 Construction of eigenfunction and (3.1)

3.3.1 Preliminaries: the Schweiger property

Let \( \mu \) be some Borel measure which is finite on partition sets. Let \( \mu \circ T \) denote the measure given on cylinders by
\[
(\mu \circ T) ( A ) = \sum_{a \in S} \mu ( T ( A \cap [a] ) )
\]
Integrals with respect to \( \mu \circ T \) are given by
\[
\int f d ( \mu \circ T ) = \sum_{a \in S} \int_{T[a]} f ( ax ) d \mu ( x )
\]
If \( \mu \) is non-singular (i.e. \( \mu \sim \mu \circ T^{-1} \)), then \( \mu \ll \mu \circ T \) and the function \( g_\mu = d\mu/d(\mu \circ T) \) is well-defined \( \mu \circ T \) almost everywhere. We will also make use of the measures \( \mu \circ T^n \) defined by induction by \( \mu \circ T^n = (\mu \circ T^{n-1}) \circ T \). The following lemma, due to Ledrappier [Le], summarizes the information needed on \( \mu \circ T \) and \( g_\mu \).

**Lemma 1** Let \( X \) be a topological Markov shift and let \( \mu \) be a Borel measure, which is finite on partition sets and non-singular. Then
1. \( g = d\mu/d(\mu \circ T) \) mod \( \mu \circ T \) iff \( L^\mu g \) is the transfer operator of \( \mu \). In particular, if \( \mu \) is \( T \)-invariant then \( L^\mu g_\mu 1 = 1 \).
2. If $\mu$ is a $T$-invariant probability measure, $E_{\mu}(f|T^{-1}B) = (L_{\log g_\mu} f) \circ T$ for every $f \in L^1(\mu)$.

3. Let $\alpha := \{[a] : a \in S\}$. If $\mu$ is a $T$-invariant probability measure, $I_\mu(\alpha|T^{-1}B) = -\log g_\mu$.4

Let $X$ be a topological Markov shift and $\mu$ be a measure supported on $X$, such that $\mu \sim \mu \circ T^{-1}$ and $\mu \sim \mu \circ T$. $\mu$ is said to have the Schweiger property (see [ADU]) if there exists a collection of cylinders $\mathcal{R}$, such that:

1. The members of $\mathcal{R}$ have finite positive measures and $\cup \mathcal{R} = X \mod \nu$;
2. For every $[\bar{a}] \in \mathcal{R}$ and arbitrary cylinder $[a]$ if $[a, \bar{a}] \neq \emptyset$ then $[a, \bar{a}] \in \mathcal{R}$;
3. There exists a constant $C > 1$, such that for every $[\bar{a}] \in \mathcal{R}$ of length $n$ and $\mu \times \mu$ almost all $x, y \in [\bar{a}] \times [\bar{a}]$

\[
\frac{d\mu}{d\mu \circ T^n} (x) = C^{\pm 1} \frac{d\mu}{d\mu \circ T^n} (y).
\]

Aaronson, Denker and Urbanski proved in [ADU] that if $\mu$ has the Schweiger property, is supported on a topologically mixing topological Markov shift, and is conservative then:

1. The measure $\mu$ is exact (hence ergodic);
2. There exists a $\sigma$-finite invariant measure $m \sim \mu$ such that $\log (\frac{d\mu}{dm})$ is bounded on every $B \in \mathcal{R}$.
3. The measure $m$ is pointwise dual ergodic: there exist $a_n > 0$ such that for every $f \in L^1(m)$

\[
\frac{1}{a_n} \sum_{k=1}^{n} \hat{T}^k f \underset{n \to \infty}{\rightarrow} m(f) \quad \text{a.e.}
\]

where $\hat{T}$ is the transfer operator of $m$.

By the above remarks, if $L^*_\phi \nu = \lambda \nu$, then

\[
\frac{d\nu}{d\nu \circ T^n} = \lambda^{n \phi_n}
\]

By summable variations, $\nu$ satisfies (3.8) with respect to the partition generated by cylinders of length two. It is not true in general, however, that $\nu$ satisfies this property with respect to all cylinders, including those of length one. In order to obtain information on cylinders of length one as well, we need the following lemma (compare with example 1 in [ADU]). For every $c \in S$ set

$\mathcal{R}_c = \{[b_0, \ldots, b_n] : n \in \mathbb{N}, b_{n-1} = c\}$. Note that $[c] \in \mathcal{R}_c$.

4Recall that the information function is defined by

\[
I_\mu(\alpha|T^{-1}B) := - \sum_{A \in \alpha} 1_A \log \mu(A|T^{-1}B)
\]
Lemma 2 Let $X$ be topologically mixing and $\phi$ have summable variations. Suppose that $\nu$ is a conservative measure, finite and positive on cylinders, such that $L^*_\phi \nu = \lambda \nu$. Then $\forall c \in S$ there exists a density function $q = q(c): X \to (0, \infty)$, such that $d\nu_c = q(c) \, d\nu$ has the Schweiger property with respect to $\mathcal{R}_c$. $q$ can be chosen to be constant on partition sets.

Proof. For every $1 \leq m \leq n - 1$ and $[b]$ of length $n$ set
\[
\phi_m (b) := \inf \{ \phi_m (x) : x \in [b] \}.
\]
By (3.2) $\forall x \in [b] \ \phi_m (x) = \phi_m (x_0, \ldots, x_{n-1}) \pm \log B_n \ \nu_m$. Set $q (x) := q_{x_0}$ where
\[
q_b := \left\{ \begin{array}{ll}
e^{\phi (c, b)} & [b] \subseteq T [c] \\ 1 & \text{else} \end{array} \right.,
\]
and set $d\nu_c = q \, d\nu$. A calculation shows that $d\nu_c \circ T^n = q_c \circ T^n \, d\nu_c \circ T^n$ whence
\[
\frac{d\nu_c}{d\nu_c \circ T^n} (x) = \frac{q_{x_0} \, e^{\phi_m (x)}}{q_{x_0}} = B_1 \lambda^n \, e^{\phi_{n-1} (x)}.
\]
Since $e^{\phi_{n-1} (x)} = B_1 e^{\phi_{n-1} (b_0, \ldots, b_{n-1})}$, (3.8) is proved.

Obviously for every $[b]$ in $\mathcal{R}_c$ and for every $[a]$, $[a, b]$ is either empty or in $\mathcal{R}_c$. We show that $X = \bigcup \mathcal{R}_c (\text{mod } \nu_c)$. Assume this were not the case. Then $\exists a \in S \ \exists A \subseteq [a]$ measurable of positive measure, such that $\nu_c (A \cap \bigcup \mathcal{R}_c) = 0$. By topological mixing there exists a $[c] \subseteq [c]$, such that $[c, a] \neq \emptyset$. Choose such a $c$ of minimal length. Set $[c, A] = [c] \cap T^{-1} [A]$ where $|c|$ denotes the length of $[c]$. Then $[c, A] \neq \emptyset$ and
\[
\int_{[c, A]} \frac{d\nu_c \circ T^n [c]}{d\nu_c} \, d\nu_c = \nu_c (A) > 0
\]
whence $\nu_c [c, A] > 0$. Since $|c|$ is minimal,
\[
[c, A] \subseteq [c] \setminus T^{-1} (\bigcup \mathcal{R}_c) = [c] \setminus \bigcup_{n \geq 1} T^{-n} [c]
\]
so, by conservativity, $\nu_c [c, A] = 0$. \hfill \Box

3.3.2 Construction of $h$ and $a_n$

We are now ready to construct the eigenfunction and prove (3.1).

Proposition 5 If $\phi$ is recurrent, then $\exists h > 0$ and $\{a_n\}_{n=1}^{\infty}$, such that $L_\phi h = \lambda h$, and such that for every cylinder $[b]$ and $x \in X$
\[
\frac{1}{a_n} \sum_{h=1}^{n} \lambda \, h \left( L_\phi^k [b] \right) (x) \underset{n \to \infty}{\longrightarrow} h (x) \nu [b].
\]
Proof. Since $\phi$ is recurrent, there exists a conservative measure $\nu$, finite and positive on cylinders, such that $L^{n}_{\phi} \nu = \lambda \nu$. Fix an arbitrary $c \in S$ and set $\mathcal{R}_{c} = \{b_{0}, \ldots, b_{n-1} \}: n \in \mathbb{N}, b_{n-1} = c$. By lemma 2, $\exists \nu_{c} \sim \nu$ with the Schweiger property with respect to $\mathcal{R}_{c}$, such that $d\nu_{c}/d\nu$ is constant on partition sets. According to the results cited in the last section, there exists an exact invariant measure $m$ which is equivalent to $\nu_{c}$, hence also to $\nu$. Its derivative $dm/d\nu$ is bounded away from zero and infinity on members of $\mathcal{R}_{c}$ (because $d\nu_{c}/d\nu$ is constant on partition sets). This measure is pointwise dual ergodic: there exist $a_{n} > 0$ such that\footnote{We will later see that $a_{n}$ is asymptotically proportional to the $a_{n}$ of the last section.} for every $f \in L^{1}(m)$

\begin{equation}
\frac{1}{a_{n}} \sum_{k=1}^{n} \hat{T}^{k} f \longrightarrow_{n \to \infty} \int f \, dm \quad a.e
\end{equation}

Set $h = dm/d\nu$. Since $\nu$ is equivalent to $m$ and $m$ is exact, $\nu$ is conservative ergodic and can only have one invariant density (up to a constant). Thus $h$ and $m$ are independent of $c$. It also follows from (3.9) that $\{a_{n}\}$ is independent of $c$ (up to asymptotic proportionality). Since $h$ is bounded away from zero and infinity on members of $\mathcal{R}_{c}$ for every $c$, $h$ is bounded away from zero and infinity on all cylinders. Thus, since $\nu$ is positive and finite on cylinders, so is $m$.

We show that $h$ and $\{a_{n}\}$ are the required eigenfunction and sequence. The transfer operator of $dm$ is given by $\hat{T} f = \lambda \ h \ 1^{1} L_{\phi} (hf)$, (because $dm = h \, d\nu$ and the transfer operator of $\nu$ is given by $\lambda \ 1^{1} L_{\phi}$). Thus, for every cylinder $[\mathcal{B}]$

\begin{equation}
\frac{1}{a_{n}} \sum_{k=1}^{n} \lambda \ h \ 1^{k} L_{\phi}^{k} [\mathcal{B}] = \frac{1}{a_{n}} h \sum_{k=1}^{n} \hat{T}^{k} (h \ 1^{k} [\mathcal{B}]).
\end{equation}

For every cylinder $[\mathcal{B}]$ the function $h \ 1^{k} [\mathcal{B}]$ is $m$-integrable (because $h$ is bounded away from zero on cylinders). Thus (3.10) implies that for $m$—almost every $x \in X$ for every cylinder $[\mathcal{B}]$

\begin{equation}
\frac{1}{a_{n}} \sum_{k=1}^{n} \lambda \ h \ 1^{k} (L_{\phi}^{k} [\mathcal{B}]) (x) \longrightarrow_{n \to \infty} h \ (x) \ \nu \ [\mathcal{B}].
\end{equation}

Since $\nu$ is positive on cylinders, and $m \sim \nu$, there is a dense set of points $x \in X$ for which (3.11) is valid for every cylinder $[\mathcal{B}]$. By (3.2)

$$\forall m \geq 1 \ \forall k \ \ V_{m} [\log (L_{\phi}^{k} [\mathcal{B}])] < \log B_{m}$$

and we have that the logarithm of each of the summands in the left hand of (3.11) is uniformly continuous in $x$. It follows that $h$ has a version for which (3.11) holds everywhere for every cylinder $[\mathcal{B}]$. This version must satisfy

\begin{equation}
\forall m \geq 1 \ V_{m} [\log h] < \log B_{m},
\end{equation}

whence $\log h, \ log hcT$ are continuous. We see, again, that $h$ is uniformly bounded away from zero and infinity on partition sets, because the last estimation is also valid for the case $m = 1$. 
3.3. CONSTRUCTION OF EIGENFUNCTION AND (3.1) 35

It is now possible to show that \( h \) is an eigenfunction. Applying \( L_\phi \) on both
hands of (3.11) (and noting that by conservativity \( a_n \to \infty \)), it is easy to see
that \( L_\phi h = \lambda h \). Set \( f = h - \lambda^{-1} L_\phi h \). This is a non-negative function, which
satisfies \( \sum_{k \geq 0} \lambda^k L_\phi^k f \to \infty \). Since \( \nu \) is ergodic conservative with transfer
operator \( \lambda^{-1} L_\phi \), this is impossible unless \( f = 0 \) \( \nu \)-a.e. Since \( f \) is continuous
and since \( \nu \) is supported everywhere, \( f = 0 \) whence \( L_\phi h = \lambda h \). \( \Box \)

**Proposition 6** Let \( h \) and \( \{a_n\}_n \) be as in proposition 5 and let \( dm = h \, d\nu \).
Then for every \( a \in S \),

\[
a_n \sim \frac{1}{m[a]} \sum_{k=1}^{n} \lambda^k Z_n(\phi, a).
\]

**Proof.** Let \( \hat{T} \) denote the transfer operator of \( m \). For every cylinder \( [\mathcal{a}] \) of length
\( N \) set \( Z_n(\phi, \mathcal{a}) = \sum_{\sigma \in \mathbb{Z}^N} \epsilon_{\phi_n(\sigma)} \, 1_{[\mathcal{a}]}(\sigma) \) and choose some \( x_{\mathcal{a}} \in [\mathcal{a}] \). By (3.12)
for every \( N \geq 1 \) and almost all \( x_{\mathcal{a}} \in [\mathcal{a}] \)

\[
(3.13) \quad \lambda^{-n} Z_n(\phi, \mathcal{a}) = B^{+1}_N \left( \lambda^{-n} L_{\phi}^n 1_{[\mathcal{a}]} \right)(x_{\mathcal{a}}) = B^{+2}_N \left( \hat{T}^n 1_{[\mathcal{a}]} \right)(x_{\mathcal{a}})
\]

By (3.9)

\[
(3.14) \quad \lim_{n \to \infty}, \lim_{n \to \infty} \left[ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k Z_k(\phi, \mathcal{a}) \right] = B^{+2}_N \, m[a] \; \sin \theta
\]

The idea is to sum over \( [\mathcal{a}] \subseteq [a] \) where \( |a| = N \) and deduce that

\[
\lim_{n \to \infty}, \lim_{n \to \infty} \left[ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k Z_k(\phi, a) \right] = B^{+2}_N \, m[a] \; \sin \theta
\]

which implies, since \( N \) is arbitrary, that both limits coincide and are equal to
\( m[a] \). We need a regularity argument to deal with the possibility that there
may be an infinite number of \( [\mathcal{a}] \subseteq [a] \) such that \( |\mathcal{a}| = N \).

Let \( \varepsilon > 0 \) and \( F = F_{\varepsilon} \subseteq [a] \) be a compact such that \( m([a] \setminus F) < \varepsilon \). We
denote by \( [a] \cap \alpha_{\varepsilon} \) the set of all cylinders of length \( N \) that are included in
\( [a] \). Then,

\[
\frac{1}{a_n} \sum_{k=1}^{n} \lambda^k Z_k(\phi, a) =
\]

\[
= \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a})
\]

\[
= \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a}) \quad \sin \theta
\]

\[
+ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}, \mathcal{a} \not\subseteq \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a})
\]

\[
+ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq \mathcal{a} \setminus \mathcal{a}} Z_k(\phi, \mathcal{a})
\]

\[
+ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a})
\]

\[
= \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a})
\]

\[
+ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}, \mathcal{a} \not\subseteq \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a})
\]

\[
+ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq \mathcal{a} \setminus \mathcal{a}} Z_k(\phi, \mathcal{a})
\]

\[
+ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a})
\]

\[
= \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a})
\]

\[
+ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}, \mathcal{a} \not\subseteq \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a})
\]

\[
+ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq \mathcal{a} \setminus \mathcal{a}} Z_k(\phi, \mathcal{a})
\]

\[
+ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\mathcal{a}] \subseteq [a] \cap \alpha_{\varepsilon}^{-1}} Z_k(\phi, \mathcal{a})
\]
Using (3.12), (3.13) and the pointwise dual ergodicity of \( m \) we have that for almost every \( z_a \in [a] \)

\[
\frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\omega] \subseteq [a] \cap \alpha_{-1}^{n}} \mathbb{Z}_k (\phi, \omega) \leq \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\omega] \subseteq [a] \cap \alpha_{-1}^{n}} [h^{-1} L^k_{\phi} (h_1 \omega)] (x_\omega) \leq B_N \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \sum_{[\omega] \subseteq [a] \cap \alpha_{-1}^{n}} [h^{-1} L^k_{\phi} (h_1 \omega)] (z_a) \leq B_N B_1 \frac{1}{a_n} \sum_{k=1}^{n} \bigg( \hat{T}^{-1} \bigg| \bigcup_{[\omega]} \big( [a] \setminus F \big) \bigg) (z_a) \xrightarrow{n \to \infty} B_N B_1 m ([a] \setminus F)
\]

Thus,

\[
\frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \mathbb{Z}_k (\phi, a) = \sum_{[\omega] \subseteq [a] \cap \alpha_{-1}^{n}} \bigg[ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \mathbb{Z}_k (\phi, \omega) \bigg] + O (\varepsilon)
\]

The sum on the right is finite, because \( F \) is compact. It follows from this and (3.14) that

\[
\lim_{n \to \infty} \lim_{m \to \infty} \bigg[ \frac{1}{a_n} \sum_{k=1}^{n} \lambda^k \mathbb{Z}_k (\phi, a) \bigg] = B_N^{\pm 2} m (A_{N, \varepsilon}) + O (\varepsilon)
\]

where \( A_{N, \varepsilon} := \bigcup \{ [\omega] : [\omega] = N, [\omega] \cap F_\varepsilon \neq \emptyset \} \). Clearly, \( A_{N, \varepsilon} \downarrow F_\varepsilon \) as \( N \uparrow \infty \) whence both limits are equal to \( m (F_\varepsilon) + O (\varepsilon) \). Passing to the limit \( \varepsilon \downarrow 0 \), we have that the upper limit and lower limit coincide and are equal to \( m [a] \). \( \square \)

### 3.4 Positive recurrence and Null recurrence

**Proposition 7** Assume \( \phi \) is recurrent, fix some arbitrary state \( a \in S \) and let \( \lambda, \nu \) and \( h \) be as the previous sections. If \( \sum_n n \lambda \mathbb{Z}_n^* (\phi, a) < \infty \) then \( \int h \, d\nu < \infty \).
If \( \sum_n n \lambda \mathbb{Z}_n^* (\phi, a) = \infty \) then \( \int h \, d\nu = \infty \).

**Proof.** Let \( dm = h \, d\nu \). The transfer operator of \( m \) is \( \hat{T} f = \lambda^{-1} h^{-1} L_{\phi} (h f) \).
Set \( \psi_N = 1_{[\varphi_a = N]} \). By (3.12) \( \forall N \forall k \)

\[
(\hat{T}^k \psi_N) 1_{[a]} = B_{\pm 2} \lambda^N \mathbb{Z}_N^* (\phi, a) \left( \hat{T}^k 1_{[a]} \right) 1_{[a]}
\]

Summing over \( k = 1, \ldots, N \), dividing by \( a_n \) and passing to the limit using pointwise dual ergodicity, we see that

\[
\lambda^N \mathbb{Z}_N^* (\phi, a) = B_{\pm 2} m [\varphi_a = N] / m [a].
\]
It follows that
\[ \sum_{n=1}^{\infty} n \lambda^{-n} Z^*_n(\phi, a) = B_1^{-\frac{1}{2}} \int_{[a]} \phi_a \, dm. \]

Since \( m \) is conservative and ergodic, \( \int_{[a]} \phi_a \, dm = m(X) \) (Kac’s formula), whence \( m(X) \) is finite if \( \phi \) is positive recurrent. \( \square \)

**Proposition 8.** Under the above assumptions, for every cylinder \([a]\)

1. If \( \phi \) is null recurrent, then \( \lambda^{-n} L^1_{\phi, [a]} \to 0 \) uniformly on partition sets whence \( a_n = o(n) \).

2. If \( \phi \) is positive recurrent, then \( \lambda^{-n} \left(L^1_{\phi, [a]}(x) \to \frac{h(x)}{\nu(b)} \right) \) uniformly on compacts whence \( a_n \to \infty \).

**Proof.** Assume that \( \phi \) is null recurrent. Since \( L^1_{\phi} \) is positive and \( h \) is uniformly bounded away from zero and infinity on \([a]\), it is enough to show that \( \lambda^{-n} h^{-1} L^1_{\phi, [a]} \to 0 \) uniformly on cylinders. Choose unions of partition sets \( F_n \), such that \( F_n \to X \) and \( 0 < m(F_n) < \infty \). \( \phi \) is null recurrent so \( m(F_n) \to \infty \). Set \( f_n = 1_{[a]} - 1_{F_n} \cdot m[a] / m(F_n) \). For every \( b \in S \) the usual estimations yield (for \( \| \cdot \|_1 = \| \cdot \|_{L^1(m)} \))

\[
\| 1_b \hat{T}^n 1_{[a]} \|_\infty \leq \frac{B_1^3}{m([b])} \left( 1_{[b]} \hat{T}^n 1_{[a]} \right) \|_1 \\
\leq \frac{B_1^3}{m([b])} \left( \| 1_{[b]} \hat{T}^n f_n \|_1 + \frac{m[a]}{m(F_n)} \| 1_{[b]} \hat{T}^n 1_{F_n} \|_1 \right) \\
\leq \frac{B_1^3}{m([b])} \left( \| \hat{T}^n f_n \|_1 + \frac{m[a] m[b]}{m(F_n)} \right)
\]

\( \hat{T} \) is the transfer operator of \( m \). Since \( m(F_n) = 0 \) and \( m \) is exact (it is equivalent to \( \nu \), and \( \nu \) has the Schweiger property), it follows from a theorem of M. Lin (see theorem 1.3.3 in [A1]) that \( \| \hat{T}^n f_n \|_{L^1(m)} \to 0 \). It follows from this and from the fact that \( m(F_n) \to \infty \) that \( \| 1_b \hat{T}^n 1_{[a]} \|_\infty \to 0 \) as required.

Assume now that \( \phi \) is positive recurrent. Without loss of generality, assume that \( \int h \, dv = 1 \). Standard estimations show that for every cylinder \([a]\) the family \( \{ \lambda^{-n} L^1_{\phi, [a]} \}_n \) is equicontinuous and uniformly bounded on partition sets \([b]\). It follows that every subsequence has a subsequence of its own, which converges uniformly on compacts. It is enough to show that the only possible limit point is \( h \nu([a]) \), because it will then immediately follow from the equicontinuity of \( \{ \lambda^{-n} L^1_{\phi, [a]} \}_n \) that this sequence tends uniformly on compacts to \( h \nu([a]) \).

Assume that \( \lambda^{-n} L^1_{\phi, [a]} \) tends to \( \varphi \) pointwise. For every \( k \),

\[ \lambda^{-n} L^1_{\phi, [a]} \leq Ch \]
where $C = 1/\inf \{ h(x) : x \in [a] \}$. Since $Ch$ is integrable, we have by the dominated convergence theorem that

$$
\int |\varphi - h\nu [a]| \, dv = \lim_{k \to \infty} \int \left| \lambda^{m^*} L_{\phi, k}^{n^*} 1_{[a]} - h\nu [a] \right| \, dv
$$

$$
= \lim_{k \to \infty} \int \left| \hat{F}^{n^*} (h^{-1}_{x} [a] - \nu [a]) \right| \, dm.
$$

Since $m$ is exact, the last limit is equal to zero and we have that $\varphi = h\nu [a]$ almost everywhere. Since $\varphi$ must be continuous, it must be equal to $h\nu [a]$ everywhere. (Note that this argument does not work if $\phi$ is null recurrent, because in this case $h^{-1}_{x} [a] - \nu [a]$ is not integrable.)

\[ 3.5 \text{ Notes} \]

The notions of recurrence discussed in this chapter are rooted in the probabilistic theory of countable Markov chains (see e.g. Feller [Fe]). These were generalized to the context of (not necessarily stochastic) positive countable matrices by D. Vere-Jones, who used them to state and prove a generalization of the Perron-Frobenius theorem for infinite positive matrices [VJ1, VJ2]. Definition 3 generalizes the concepts of Vere-Jones in the following sense: for $\phi$ of the form $\phi(x) = \phi(x_0, x_1)$, $\phi$ is positive recurrent (respectively, null recurrent and transient) iff the matrix $(e^{-t}\phi)_{s \times s}$ is positive recurrent (respectively null recurrent and transient) in the sense of [VJ1, VJ2].\(^6\) Theorem 3 is a version of Vere-Jones’ Perron-Frobenius theorem for functions $\phi$ with ‘infinite memory’.

The main difference between the proof of Ruelle’s Perron-Frobenius theorem in the compact case and that in the non-compact case is the construction of $\nu$. In the compact case, the existence of $\nu$ is immediate: Let $G(\mu) := L_{\phi, 1}^* \mu / \mu(L_{\phi, 1})$. This acts $w^*$-continuously on the set of Borel probability measures on $X$. Since this set is a compact convex subset of the locally convex topological vector space $C(X)^*$, $G$ has a fixed point and this is the eigenmeasure (see [Ru2]). This argument fails in the non-compact situation, because then the set of probability measures is not $w^*$-compact.

Versions of the RPF theorem in non-compact situations were proved in special cases by several authors. Walters [Wal2] used assumptions which imply that $\phi$ can be extended to a compactification $\hat{X}$ of $X$ in a way that $L_{\phi}$ acts continuously on $C_{B}(\hat{X})$. This allowed him to apply the previous argument on the set of probability measures on $\hat{X}$ to obtain $\nu$. A similar approach was adopted by Yuri in [Yu2]. Other authors were only interested in functions $\phi$ for which $L_{\phi}$ is the transfer operator of a given measure $\nu$. In this case it is automatic (see section 3.3.1, lemma 1) that $L_{\phi}^* \nu = \nu$ and that $\nu$ is a conformal measure (see, e.g., [ADU], [F3], [Yu1]).

\(^6\)In the terminology of [VJ1, VJ2], $R$-positive, $R$-null and $R$-transient for $R$ is the inverse of the Perron value of the matrix.
3.5. NOTES

The main new ingredient in the proof of theorem 3 is a tightness argument (see proposition 4), which is needed to deal with the lack of compactness of $X$. This is used to construct $\nu$. The construction of $h$ given $\nu$ is done using the techniques of Aaronson, Denker and Urbanski [ADU] (see also [A1], [LY], [T3], [Yu1], [Yu2], [Yu3], [Yu4]).
Chapter 4

Gibbs Measures and Equilibrium Measures

In the first section we characterize the case when a Gibbs measure in the sense Bowen exists, and show that in this case the convergence in The RPF theorem is uniformly exponential for suitable functions. In the second section we describe the (possibly infinite) $h dm$ as an equilibrium measure, using Krenkel’s definition for entropy of an infinite conservative measure. Some of the results of this chapter are contained in [S3].

4.1 Gibbs measures

Let $X$ be a topological Markov shift, which is topologically mixing, and let $\phi$ be some function with summable variations. A Borel probability measure $m$ is called a Gibbs measure for $\phi$ (in the sense of Bowen) if there exist two global constants $M$ and $P$, such that for every cylinder $[a_0, \ldots, a_{-1}]$ and every $x \in [a_0, \ldots, a_{-1}]$

$$\frac{1}{M} \leq \frac{m[a_0, \ldots, a_{-1}]}{e^{\delta \phi(x)} P} \leq M.$$  

This definition appeared in [B1]. As mentioned there, this is a simplification of the more subtle notion of Gibbs state used in statistical mechanics [E].

4.1.1 The BIP property

Denote the transition matrix of $X$ by $A = (t_{ij})_{S \times S}$. We say that $X$ satisfies the BIP property (big images and preimages property) if $\exists b_1, \ldots, b_N \in S$ such that

$$\forall a \in S \exists i, j \text{ such that } t_{b_i a} t_{a b_j} = 1.$$  

This is stronger than the big images property used in [ADU], and for $X$ topologically mixing, is equivalent to the finite primitivity condition used by Mauldin and Urbański in [MU].

40
4.1. GIBBS MEASURES

Theorem 4 Let \( X \) be topologically mixing and let \( \phi : X \to \mathbb{R} \) be some function with summable variations and finite Gurevich pressure. Then \( \phi \) has an invariant Gibbs measure, iff \( X \) has the BIP property and \( V_1(\phi) < \infty \).

Proof of sufficiency. Assume that \( V_1(\phi) < \infty \) and that \( X \) has the BIP property. We show that \( \phi \) is positive recurrent. Fix some \( a \in S \) and set \( \lambda := e^{P_0(\phi)} \). By theorem 3, positive recurrence is equivalent to the existence of \( N_1 \in \mathbb{N} \), such that

\[
\inf \{ \lambda^{-n} Z_n(\phi, a) : n \geq N_1 \} > 0
\]

This is because (4.3) implies recurrence and rules out null recurrence, since for every \( x \in [a] \), \( \lambda^{-n} L^1_{\phi,1}(x) = \lambda^{-n} Z_n(\phi, a) \neq o(1) \).

Let \( W_n := \{ a \in S^n : [a] \neq \emptyset \} \) and set for every \( a \in W_n \),

\[
\phi_n(a_0, \ldots, a_{n-1}) := \inf \{ \phi_n(x) : x \in [a] \}
\]

\[
\phi_n(a_0, \ldots, a_{n}) := \sup \{ \phi_n(x) : x \in [a] \}
\]

These are finite, since \( V_1(\phi) < \infty \). In fact, if \( B_0 := \exp \sum_{k>0} V_k(\phi) \), then for all \( \mu \in W_1 \), \( |\phi_n(\mu) - \phi_n(\mu')| \leq \log B_0 \).

According to the BIP property \( \exists b_1, \ldots, b_N \in S \) such that (4.2) holds. Since \( X \) is topologically mixing there is some \( n_1 \) and admissible words of length \( n_1 \)

\( \omega_{b_i, \omega_{b_i}, a} \in W_{n_i}, \) such that \( (a, \omega_{b_i, b_j}) \) and \( (b_j, \omega_{b_j, a}) \) are admissible for \( i, j = 1, \ldots, N \). Set \( N_1 := 2(n_1 + 2) - 1 = 2n_1 + 3. \) By (4.2) for every \( n > N_1 \) and \( \mu \in W_n \) there are \( 1 \leq i, j \leq N \) such that

\[
(a, \omega_{b_i, b_j}, \omega_{b_j, a}) \in W_{n + N_1 + 1}.
\]

Set

\[
C := \min \{ e^{\phi_n + \phi_{n+1}(b_i, \omega_{b_i})}, e^{\phi_{n+1}(b_j, \omega_{b_j})} : i, j = 1, \ldots, N \}
\]

Then for \( C_1 := C \lambda^{N_1} \),

\[
\lambda^{(n+N_1)} Z_{n+N_1}(\phi, a) \geq C_1 \lambda^n \sum_{\omega \in W_n} e^{\phi_n(\omega)} \geq \frac{C_1}{B_0} \lambda^n \sum_{\omega \in W_n} e^{\phi_n(\omega)}
\]

We claim that the last expression is bounded from below. Indeed, assume by way of contradiction that \( \exists n_0 \) such that \( \lambda^{-n_0} \sum_{\omega \in W_{n_0}} e^{\phi_{n_0}(\omega)} < \frac{1}{2} \). Then for all \( k \)

\[
\lambda^{-k n_0} Z_{k n_0}(\phi, a) \leq \left( \lambda^{-n_0} \sum_{\omega \in W_{n_0}} e^{\phi_{n_0}(\omega)} \right)^k < \frac{1}{2^k}.
\]

This, however, is impossible since by the definition of the Gurevich pressure \( \lambda^{-k n} Z_{k n}(\phi, a) \) cannot decay exponentially fast. This proves (4.3), whence the positive recurrence of \( \phi \).

\[\text{This result appears in a different form in [MU].}\]
Let $\nu$ and $h$ be the eigenmeasure and eigenfunction given by theorem 3. We show that $\sup h < \infty$. Fix some $a \in S$. By positive recurrence,

$$h(x) = \frac{1}{\nu[a]} \lim_{n \to \infty} \lambda^n L_{\phi}^n 1_{[a]}(x) \leq \frac{1}{\nu[a]} \sup_{n \geq 1} \left\{ \lambda^n \sum_{w \in W_n} e^{\phi_n[w]} \right\}$$

By (4.4) and (3.3) this supremum is finite whence $\sup h < \infty$.

We show that $h(x) > 0$. For every $x \in X \exists i \leq N$, such that $(b_i, x)$ is admissible. Then $h(x) = \lambda^{-1} (L_{\phi} h)(x) \geq \lambda^{-1} e^{\phi(b_i, x)} h(b_i x)$. The last quantity is bounded from below, since there are only finitely many $b_i$'s, and since $h$ is uniformly bounded from below on partition sets, by (3.12).

This shows that $0 < h(x) < \sup h < \infty$ whence $\exists H > 1$ such that $\forall x \in X$,

$$\frac{1}{H} < h(x) < H.$$ 

We use this and the BIP property to deduce that $dm = h \nu$ is an invariant Gibbs measure. To see invariance note that for every $f$ continuous and $m$ integrable,

$$\int f \circ T dm = \int \lambda^{-1} L_{\phi} (h \circ T) \nu = \int f \cdot \lambda^{-1} L_{\phi} h \nu = \int f dm.$$ 

We show (4.1). Fix some $a = (a_0, \ldots, a_{n-1})$ and $x \in [a]$. Then

$$m[a_0, \ldots, a_{n-1}] = \int \lambda^n L_{\phi}^n (h 1_{[a]}) \nu = H^{-1} \int \lambda^n L_{\phi}^n 1_{[a]} \nu = (HB_c)^{-1} e^{\phi_n(a)} n_{\log \lambda} \nu(T^n [a]) = (H^2 B_c)^{-1} e^{\phi_n(a)} n_{\log \lambda} m(T^n [a]).$$

It is, therefore, enough to show that $m(T^n [a])$ is bounded away from zero and infinity. Boundedness from above is clear, since by positive recurrence $m(X) < \infty$. Boundness from below follows from the BIP property which implies that

$$m(T^n [a]) \geq \min\{m[b_i] : i = 1, \ldots, N\} > 0.$$ 

A by product of the proof is that the $P$ in (4.1) is equal to the Gurevich pressure of $\phi$.

**Proof of necessity.** Assume now that there is an invariant Gibbs measure $m$. By (4.1) if $x_0 = y_0$, then $e^{\phi(x)}, e^{\phi(y)} = M^{\pm 1} e^{P m[x_0]}$ whence $|\phi(x) - \phi(y)| \leq 2 \log M$. Thus

$$V_1(\phi) < \infty.$$ 

\footnote{This follows from lemma 2.1 in [ADU].}
4.1. GIBBS MEASURES

We establish the BIP property. By (4.1), and since $V_1 (\phi) < \infty$, there is some global constant $C$, such that for all $p, q \in S$, such that $[p, q] \neq \emptyset$,

$$m[p, q] = C^{\pm 1} e^{\phi(p)} \cdot e^{\phi(q)}$$

Summing over all possibilities, first for $p$ (fixing $q$) and then for $q$ (fixing $p$), we have

$$m(T^{-1}[q]) = C^{\pm 1} e^{\phi(q)} \sum_{p: [p, q] \neq \emptyset} e^{\phi(p)}$$

$$m[p] = C^{\pm 1} e^{\phi(p)} \sum_{q: [p, q] \neq \emptyset} e^{\phi(q)}$$

whence, since $m(T^{-1}[q]) = m[q] = M^{\pm 1} e^{k \phi(q)}$ and $m[p] = M^{\pm 1} e^{k \phi(p)}$,

$$\inf \left\{ \sum_{a: [a, q] \neq \emptyset} e^{\phi(a)} : \sum_{a: [p, a] \neq \emptyset} e^{\phi(a)} : p, q \in S \right\} > 0.$$ (4.5)

Inducing an order on $S$, we assume without loss of generality that $S = \mathbb{N}$. Assume by way of contradiction that BIP fails. Then for every $k$, either $\exists q_k \in S$, such that $\{a : [a, q_k] \neq \emptyset\} \subseteq \{a : a > k\}$ or $\exists p_k \in S$, such that $\{a : [p_k, a] \neq \emptyset\} \subseteq \{a : a > k\}$. This means that the infimum in (4.5) is equal to zero, since

$$\sum_{a \in S} e^{\phi(a)} \leq M e^k \sum_{a \in S} m[a] < \infty$$

The BIP property is proved. \[\square\]

It follows from the proof that

**Corollary 2** Let $X$ be a topologically mixing countable Markov shift with the BIP property, and let $\phi$ be some function with summable variations and finite pressure, such that $V_1 (\phi) < \infty$. Then $\phi$ is positive recurrent, the eigenfunction $h$ of $L_\phi$ is bounded away from zero and infinity and $h \, d\mu$ is an invariant Gibbs measure for $\phi$, where $\mu$ is the eigenmeasure of $L^*_\phi$. The $P$ in (4.1) is equal to $P_\phi (\phi)$.

4.1.2 The rate of convergence in the RPF theorem

Recall that $\phi$ is called weakly Hölder continuous with parameter $\theta \in (0, 1)$ if $\exists A > 0$, such that for all $n \geq 2$, $V_n (\phi) \leq A \theta^n$. Let $\beta$ be the partition generated by the image sets $\beta = \sigma \{T[a] : a \in S\}$. $\beta$ may be infinite. Set

$$D_\beta f := \sup_{b \in \beta} \sup_{x, y \in b} \frac{|f(x) - f(y)|}{\theta^{|x, y|}}$$

and $\mathcal{L} := \{ f \in C(X) : \|f\| := \|f\|_{\infty} + D_\beta f < \infty \}$. The following theorem follows from results in [AD].
Theorem 5 Let $X$ be a topologically mixing countable Markov shift with the BIP property. Let $\phi$ be a weakly Hölder continuous function with finite Gurevich pressure, such that $V_1(\phi) < \infty$. Let $\lambda$, $h$ and $\nu$ be as in Theorem 3 and assume $\int h \, d\nu = 1$. Then there are $K > 0$ and $r \in (0, 1)$ such that for every $f \in \mathcal{L}$,

$$\left\| \lambda^{-n} L^n_\phi f - h \int f \, d\nu \right\|_{\mathcal{L}} < Kr^n \| f \|_{\mathcal{L}}$$

Proof. According to corollary 2, $\phi$ is positive recurrent, $h$ and $\nu$ exist and $\| \log h \|_{\infty} < \infty$. Set $dm = h \, d\nu$ and $g^{(n)} = dm / dm \circ T^n$ (see section 3.3.1 for a definition of $m \circ T^n$). Then $g^{(n)} = \lambda^{-n} e^{h \phi} h / h \circ T^n$. Fix an admissible $\underline{p}$ of length $|\underline{p}| = n$. Then for every $x, y \in T^n [\underline{p}]$

$$\left| \log \frac{g^{(n)}(\underline{px})}{g^{(n)}(\underline{py})} \right| = \left| \phi_n(\underline{px}) - \phi_n(\underline{py}) \right| + \left| \log h(\underline{px}) - \log h(\underline{py}) \right| + \left| \log h(y) - \log h(x) \right|$$

Note, by (3.12), that $\log h$ is weakly Hölder continuous and that $V_1(\log h) < \infty$. By the local Hölder continuity of $\phi$, and since $h$ is bounded away from zero and infinity, there exists some $C$ independent of $\underline{p}, n$ such that for every $x, y \in T^n [\underline{p}]$

$$\left| \log \frac{g^{(n)}(\underline{px})}{g^{(n)}(\underline{py})} \right| < C g^{(x,y)}$$

(The fact that $h$ is bounded away from zero and infinity is needed for the case $x, y \in T^n [\underline{p}], x_0 \neq y_0$.) It follows from this and from the BIP property that $m$ is a Gibbs-Markov measure in the terminology of [AD]. One of the properties of a Gibbs-Markov measure is that its transfer operator $\hat{T}$ satisfies that $\exists K > 0 \forall r \in (0, 1)$ such that $\forall \psi \in \mathcal{L}$

$$\left\| \hat{T}^n \psi - f \psi \, dm \right\|_{\mathcal{L}} < Kr^n \| \psi \|_{\mathcal{L}}$$

(see [AD]). For $dm = h \, d\nu$, $\hat{T} \psi = \lambda^{-1} h \, \hat{L}_\phi (h \psi)$, so $\forall \psi \in \mathcal{L}$

$$\| h^{-1} (\lambda^{-n} L^n_\phi (h \psi) - h \nu (\psi \nu)) \|_{\mathcal{L}} < Kr^n \| h^{-1} (\psi \nu) \|_{\mathcal{L}}$$

We finish by proving the existence of $H > 1$ such that

$$\forall f \in \mathcal{L}, \, \| f \|_{\mathcal{L}} = H^{\pm 1} \| fh^{-1} \|_{\mathcal{L}}$$

(this will allow us to set $\psi = fh^{-1}$ in (4.6) and finish the proof). The existence of $H$ follows from the general inequality $\|fg\|_{\mathcal{L}} \leq \|f\|_{\mathcal{L}} \|g\|_{\mathcal{L}}$ valid for all $f, g \in \mathcal{L}$ and the fact that $\|h\|_{\mathcal{L}}, \|h^{-1}\|_{\mathcal{L}} < \infty$, since $\log h$ is weakly Hölder and $V_G(\log h) \leq 2\|\log h\|_{\mathcal{L}} < \infty$. \qed

4.2 Equilibrium measures

Assume $\phi$ is recurrent, and let $dm = h \, d\nu$ be the invariant measure given by theorem 3. When $X$ is a finite Markov shift, it is known that $dm$ is the equilibrium measure, the unique measure which maximizes $h\mu(T) + \int \phi \, d\mu$. The
4.2. EQUILIBRIUM MEASURES

purpose of this section is to provide a similar result for the countable case. We aim at a result which is also valid in the null recurrent case, when \( m(X) = \infty \).

4.2.1 Entropy for infinite measures

We recall the definition given in [Kr], following the approach of [A2]. Let \((X, \mathcal{B}, \mu, T)\) be an ergodic probability preserving transformation. For every measurable set with positive measure \( A \) one can define the **induced transformation** \( T_A : A \to A \) by \( T_A x = T^{\varphi_A (x)} x \) where \( \varphi_A (x) = \inf \{ n \geq 1 : T^n x \in A \} \) (the Poincaré Recurrence theorem guarantees that \( \varphi_A < \infty \) almost everywhere on \( A \)). It is known that the probability measure \( \mu_A (E) = \mu (E \cap A) / \mu (A) \) is \( T_A \)-invariant and ergodic, and that its entropy is given by the **Abramov Formula** (see [Ab]):

\[
h_{\mu} (T) = \mu (A) h_{\mu_A} (T_A)
\]

If \( \mu \) is infinite, ergodic and conservative, the same method of inducing applies (in this case Poincaré’s theorem is replaced by the conservativity assumption). Applying the Abramov formula to \( T_A, T_B \) as induced versions of \( T_{A \cup B} \) one sees that

\[
0 < \mu (A), \mu (B) < \infty \quad \Rightarrow \quad \mu (A) h_{\mu_A} (T_A) = \mu (B) h_{\mu_B} (T_B).
\]

Thus, the number \( \mu (A) h_{\mu_A} (T_A) \) is independent of the choice of \( A \in \mathcal{B} \) (as long as \( 0 < \mu (A) < \infty \)) and may therefore be used as the definition of the entropy of the infinite conservative ergodic measure \( \mu \).

**Example 1.** (Krengel [Kr]) Let \( (p_{ij}) \) be a null recurrent irreducible stochastic matrix and \( (p_i) \) its stationary vector. Let \( \mu \) be the corresponding invariant

finite Markovian measure. Then \( h_{\mu} = - \sum_i p_i p_{di} \log p_{di} \).

For examples arising from interval maps, see [T1].

4.2.2 The induced system

In this sub-section we collect for future reference some facts and notation on the inducing procedure in the context of topological Markov shifts.

Fix some state \( a \in S \). Set

\[
\overline{S} := \{ [a] : a_i = a \text{ iff } i = 0 ; [a, a] \neq \emptyset \}
\]

Set \( \overline{X} := \overline{S}^{\mathbb{N} \cup \{0\}} \) and let \( \overline{T} : \overline{X} \to \overline{X} \) be the left shift. Recall that

\[
\varphi_a (x) := 1_{[a]} (x) \inf \{ n \geq 1 : T^n x \in [a] \}
\]

where \( \inf \emptyset := \infty \). For every \( \phi : X \to \mathbb{R} \) set

\[
\overline{\phi} := \left( \sum_{k=0}^{n-1} \phi \circ T^k \right) \circ \pi
\]
where \( \pi : \mathcal{X} \to [a] \) is given by \( \pi((x_1, x_2, \ldots)) = \langle x_n, x_{n+1}, \ldots \rangle \). The pair \((\mathcal{X}, \phi)\) is called the induced system and \( \phi \) is called the induced potential (on \([a]\)).

Induced systems are in many cases easier to handle than the original systems, as shown by the following example. A system \((X, \phi)\) is called a Bernoulli system if \( X = S^{N \cup \{0\}} \) and if \( \phi(x) = \phi(x_0) \). A system is called a Markov system if \( X \) is a topological Markov shift and \( \phi(x) = \phi(x_0, x_1) \). In this case \( \phi \) is called a Markov potential. If \( \phi \) is a Markov potential, then \((\mathcal{X}, \phi)\) is a Bernoulli system. The simplicity of the induced system is also demonstrated by the following lemma.

**Lemma 3** Let \( X \) be topologically mixing, \( a \in S \) some fixed state and \( \phi \) some function with summable variations. The induced system \((\mathcal{X}, \phi)\) on \([a]\) has the BIP property and \( V_1(\phi) < \infty \). Furthermore,

1. If \( \sum_{n \geq 2} n V_n(\phi) < \infty \) then \( \phi \) has summable variations.

2. If \( \phi \) is weakly Hölder continuous, then \( \phi \) is weakly Hölder continuous.

**Proof.** The BIP property is clear since \( \mathcal{X} = S^{N \cup \{0\}} \). Standard estimates yield \( V_n(\phi) \leq \sum_{j=n+1}^{\infty} V_j(\phi) \), whence the rest of the lemma.

The existence of pressure is always guaranteed, even if \( \phi \) does not have summable variations:

**Lemma 4** Let \( X \) be topologically mixing, \( a \in S \) some fixed state and assume \( \phi \) has summable variations. Let \((\mathcal{X}, \phi)\) be the induced system on \([a]\). Then the following limit exists for all \([a] \in S\) (although it may be infinite) and is independent of the choice of \([a] \):

\[
P_G(\phi) = \lim_{n \to \infty} \frac{1}{n} \log Z_n(\phi, [a]).
\]

**Proof.** Follows from the proof of proposition 1 and the standard estimate \( V_n(\phi + \phi \circ T + \ldots + \phi \circ T^{n-1}) \leq \sum_{k=0}^{\infty} V_k(\phi) < \infty \).

The following lemma is a version of Kac’s formula. It follows in a standard way from theorem 3 and general results for Markov operators (see theorem VLC in [Fo]).

**Lemma 5** Let \( X \) be topologically mixing, and let \( \phi \) be some function with summable variations. Let \((\mathcal{X}, \phi)\) be the induced system on some fixed state \( a \in S \), and assume \( \phi \) has summable variations. Then \( \phi \) is recurrent with pressure zero iff \( \phi \) is positive recurrent with pressure zero. In this case, if \( L_\phi \nu = \nu \), \( L_\phi h = h \), \( L_\phi \mathcal{P} = \mathcal{P} \) and \( L_\phi \mathcal{H} = \mathcal{H} \), then up to a multiplicative constant

\[
\mathcal{P} = \nu \circ \pi \\
\mathcal{H} = h \circ \pi
\]

\(^3\)This is also true for the larger class of potentials \( \phi \) for which \( \exists a \in S \) such that \( \phi(x) = \phi(x_0, \ldots, x_{v_n}(x)) \) as long as the inducing is done with respect to \([a]\). The state \( a \) can be viewed as a “gap” between non-interacting clusters of interacting particles. Analogous potentials are studied in a different mathematical setting in [FP].
4.2. EQUILIBRIUM MEASURES

and

\[ \nu(A) = \int \left( \sum_{n=0}^{\infty} e^{n \lambda} 1_A \right) \circ \pi \, d\nu \]

\[ h = \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} L_{\phi}^k(h_a 1_{[\nu_a = n]}) \mod \nu \]

where \( \ell \) is the operator \( \ell(f) := L_{\phi}(f \cdot 1_{[a]}^\nu) \) and \( h_a := 1_{[a]}^\pi \circ \pi^{-1} \).

**Proof.** One checks that the above formulas yield eigenfunctions and eigenmeasures for the eigenvalue 1 for \( L_{\phi} \) and \( L_{\phi^{-1}} \), and that \( \int T d\nu < \infty \). \( \square \)

4.2.3 Two variational principles

In this section we restrict ourselves to weakly Hölder continuous potentials.

**Theorem 6** Let \( X \) be topologically mixing and \( \phi \) a recurrent weakly Hölder continuous function with finite Gurevich pressure.

1. For every conservative ergodic invariant measure \( \mu \) which is finite on partition sets, if \( \int (P_G(\phi) - \phi) \, d\mu < \infty \), then \( h_{\mu}(T) \leq \int (P_G(\phi) - \phi) \, d\mu \).

2. Let \( h \) and \( \nu \) be as in theorem 3 and set \( dm = h \, d\nu \). If \( \int (P_G(\phi) - \phi) \, d\nu = 0 \), then \( h_{\mu}(T) = \int (P_G(\phi) - \phi) \, d\mu \).

**Proof of part 1.** Without loss of generality assume that \( P_G(\phi) = 0 \) (we can always pass to the potential \( \phi - P_G(\phi) \)). Fix some invariant measure \( \mu \) which satisfies the assumptions of the theorem and fix some \( a \in S \), such that \( 0 < \mu[a] < \infty \). Let \( \mu_a \) be the probability measure \( \mu_a(E) = \mu([a] \cap E) / \mu[a] \). Let \( T_{\phi} : [a] \to [a] \) be the induced map \( T_{\phi}x = T^{\phi(x)} x \) where \( \phi_a(x) = 1_{[a]}(x) \inf \{ n > 0; T^n x \in [a] \} \). Then \( \mu_a \) is \( T_{\phi} \) invariant. Also, since \( \mu \) is conservative, \( \phi_a < \infty \mu \)-almost everywhere.

Let \( (\overline{X}, \overline{\phi}) \) be the induced system on \( [a] \) and let \( \pi : \overline{X} \to [a] \) be the natural injection. For every \( \mu \) as in the above set \( \overline{\mu} = \mu_a \circ \pi \). It is easy to check that \( \pi \) is a measure theoretic isomorphism between the systems \( ([a], B \cap [a], \mu_a, T_a) \) and \( (\overline{X}, B(\overline{X}), \overline{\mu}, \overline{T}) \) where \( \overline{T} : \overline{X} \to \overline{X} \) is the left shift.

By lemma 3 \( \overline{\phi} \) has summable variations (actually it is weakly Hölder continuous). By lemma 5 and the recurrence of \( \phi \), \( P_G(\overline{\phi}) = 0 \). In particular, \( \overline{\phi} \) has finite pressure. Since \( \overline{X} \) is a full shift and since \( V_1(\overline{\phi}) < \infty \), there is some constant \( C \), such that \( Z_n(\overline{\phi}, \overline{\mu}) = C^{\pm n} \left( \sum_{x \in [n]} \exp \sup \{ \overline{\phi}(x) : x \in [n] \} \right) \). This cannot be finite unless \( \sup \overline{\phi} < \infty \). Thus, by the variational principle for the Gurevich pressure (theorem 2),

\[ \sup \left\{ h_{\mu}(T) + \int \overline{\phi} \, d\mu \right\} = \mu \circ T^{-1} = \mu, \quad \mu(\overline{X}) = 1, \quad - \int \overline{\phi} \, d\mu < \infty \]
For every conservative invariant (possibly infinite) ergodic measure $\mu$ such that $\mu[a] < \infty$ and $-\int \phi \, d\mu < \infty$ the measure $\bar{\mu} = \mu \circ \pi$ is a $T$ invariant ergodic probability measure such that

$$-\int \bar{\phi} \, d\bar{\mu} = -\frac{1}{\mu[a]} \int \sum_{k=0}^{\infty} \phi \circ T^k \, d\mu = -\frac{1}{\mu[a]} \int \phi \, d\mu < \infty$$

Thus, $h_\mu(T) + \int \phi \, d\mu = \mu[a] (h_{\bar{\mu}}(T) + \int \bar{\phi} \, d\bar{\mu}) \leq 0$.

**Proof of part 2.** Again, assume without loss of generality that $P_G(\phi) = 0$. Let $dm = dv$ and assume that $-\int \phi \, dm < \infty$. We show that $h_m(T) + \int \phi \, dm = 0$. By lemma 3 $\bar{\mu}$ satisfies the BIP property and $V(\mu) < \infty$ whence by corollary 2 $\bar{\mu}$ is a Gibbs measure for $\bar{\phi}$: $\exists M > 1$ such that for every $a_0, \ldots, a_n \in \mathcal{S}$ and $x \in [a_0, \ldots, a_n] \subseteq \mathcal{X}$

$$M[a] = M^{\pm 1} \exp \sum_{k=0}^{n} \frac{1}{\phi} (T^k x)$$

Set $\bar{\pi} := \{ [a] : a \in \mathcal{S} \}$. By (4.7),

$$H_{\bar{\mu}}(\bar{\pi}) = -\sum_{[a] \in \bar{\pi}} \bar{\mu} [a] \log \bar{\mu} [a]$$

$$\leq -\sum_{[a] \in \bar{\pi}} \bar{\mu} [a] \left( \frac{1}{\bar{\mu} [a]} \int_{[a]} \bar{\phi} \, d\bar{\mu} + \log M \right)$$

$$= -\int_{\mathcal{X}} \bar{\phi} \, d\bar{\mu} + \log M$$

$$= -\frac{1}{m[a]} \int_{[a]} \sum_{k=0}^{\infty} \phi \circ T^k \, dm + \log M$$

$$= -\frac{1}{m[a]} \int \phi \, dm + \log M$$

whence $H_{\bar{\mu}}(\bar{\pi}) < \infty$. Since $\bar{\pi}$ is a generator with finite entropy, we have by the Rokhlin formula $|\text{Ro}|$ and by lemma 1 that

$$h_{\bar{\mu}}(T) = -\int \log \frac{dm}{d\bar{\mu}} \, T \, dm = -\int (\bar{\phi} + \log \bar{\mu} - \log h \circ T) \, d\bar{\mu} = -\frac{1}{m[a]} \int \phi \, dm$$

where $\bar{h}$ is the eigenfunction of $L_{\phi}^{-1}$. Multiplying both sides by $m[a]$ we have that $h_m(T) = -\int \phi \, dm$ as required. \qed

**Remark.** It follows from the proof that $m$ is the unique up to a constant

\footnote{log $\bar{h}$ is integrable, since $\| \log \bar{h} \|_{\infty} < \infty$.}
4.2. EQUILIBRIUM MEASURES

A conservative ergodic invariant measure, such that \( H_\pi (\pi) < \infty \) and \( h_m (T) = \int (P_G (\phi) - \phi) \, dm \), since by a trivial generalization of an argument of Bowen’s if there exists a probability measure which is Gibbs in the sense of Bowen, with a generator which has finite entropy, then this measure is the unique solution of the variational problem (see [B1]).

The problem with the last theorem is that frequently both \( \int (P_G (\phi) - \phi) \, dm \) and \( h_m (T) \) are infinite. In this situation, the sum \( h_m (T) + \int (\phi - P_G (\phi)) \, dm \) is meaningless. The following theorem treats this case as well.

Set

\[
I_\mu := - \sum_{a \in S} \frac{1}{|a|} \log \mu (|a|) \, T^{-1} B
\]

This is well-defined for every \( \mu \), which is finite on partition sets.

**Theorem 7** Let \( X \) be topologically mixing and \( \phi \) weakly Hölder continuous with finite Gurevich pressure. Assume that \( \phi \) is recurrent, let \( h \) and \( \nu \) be as in theorem 3 and set \( \phi' = \phi + \log h - \log h \circ T \). Then for every conservative invariant measure \( \mu \), which is finite on partition sets, \( I_\mu + \phi' - P_G (\phi') \) is one-sided integrable\(^5\) with respect to \( \mu \) and

\[
-\infty \leq \int (I_\mu + \phi' - P_G (\phi')) \, d\mu \leq 0.
\]

For \( \mu \sim \mu \circ T \), the integral in (4.8) is equal to zero iff \( \mu \) is proportional to \( h \, d\nu \).

**Proof.** Fix a conservative invariant measure \( \mu \) which is finite on partition sets and set \( g_\mu = d\mu / d\mu \circ T \) where \( \mu \circ T \) is given by (3.7). Recall from lemma 1 that the transfer operator \( T \) of \( \mu \) is given by \( L_{\log g_\mu} \) and that

\[
I_\mu = - \log g_\mu
\]

Set \( g = \lambda \, e^{\phi} h / h \circ T \) where \( \lambda = \exp P_G (\phi) \). One checks that \( \sum_{T_\mu = x} g (y) = 1 \) and that \( \sum_{T_\mu = x} g_\mu (y) = 1 \) for \( \mu \) almost all \( x \in X \). The first equality follows from the equation \( L_{\phi} h = \lambda h \); the second follows from the identity

\[
\int f \sum_{T_\mu = x} g_\mu (y) \, d\mu = \int L_{\log g_\mu} (f \circ T) \, d\mu = \int f \, d\mu
\]

which is satisfied for every \( f \in L^1 (\mu) \).

We show that \( I_\mu + \phi' - P_G (\phi') \) is one sided integrable. We use the notation \( \psi^+ := \psi 1_{\psi > 0} \) and show that \( I_\mu + \phi' - P_G (\phi')^+ \) is integrable. Fix a sequence of measurable sets \( A_n \uparrow X \), such that \( 0 < \mu (A_n) < \infty \). Fix an arbitrary

\(^5\)A function \( f \) is called one-sided integrable if at least one of \( \left\| f 1_{f > 0} \right\|_1 \), \( \left\| f 1_{f < 0} \right\|_1 \) is finite. In this case \( \int f \, d\mu \) is well defined even if \( f \) is not integrable.
integrable function \( f \geq 0 \). Set \( A_{s,t,n} := A_n \cap \{ s < \frac{g}{\mu} < t \} \). Using the inequality \( \log x \leq x - 1 \) we see that for every \( s, t, n \),

\[
\int_{A_{s,t,n}} (I_\mu + \phi' - P_G (\phi'))^+ f \circ T d\mu = \\
= \int (-\log g_\mu + \log g)^+ 1_{A_{s,t,n}} f \circ T d\mu \\
= \int [\log (g/g_\mu)]^+ 1_{A_{s,t,n}} f \circ T d\mu \\
\leq \int \left( \frac{g}{g_\mu} - 1 \right)^+ 1_{A_{s,t,n}} f \circ T d\mu \\
= \int f \circ T \cdot E_\mu \left( \left( \frac{g}{g_\mu} - 1 \right)^+ 1_{A_{s,t,n}} \right) d\mu \\
= \int f \circ T \sum_{Ty = Tx} g_\mu (y) 1_{A_{s,t,n} (y)} \left( \frac{g (y)}{g_\mu (y)} - 1 \right)^+ d\mu \\
= \int f \circ T \sum_{Ty = Tx} 1_{A_{s,t,n}} (y) \left( g (y) - g_\mu (y) \right)^+ d\mu.
\]

The last integrand is bounded by \( f \circ T \). Since this is true for all \( s, t, n \) the integral \( \int (I_\mu + \phi' - P_G (\phi'))^+ d\mu \) is finite. This implies that \( I_\mu + \phi' - P_G (\phi') \) is one-sided integrable. Applying the same calculation to \( I_\mu + \phi' - P_G (\phi') \) rather than \( (I_\mu + \phi' - P_G (\phi'))^+ \) yields

\[
\int_{A_{s,t,n}} f \circ T (I_\mu + \phi' - P_G (\phi')) d\mu \\
\leq \int f \circ T \sum_{Ty = Tx} 1_{A_{s,t,n}} (y) \left( g (y) - g_\mu (y) \right) d\mu.
\]

The integrand on the left is bounded in absolute value by the integrable function \( f \circ T \). Its pointwise limit when \( s \downarrow 0 \) and \( t, n \uparrow \infty \) is zero, because

\[
\sum_{Ty = Tx} \left[ g (y) - g_\mu (y) \right] = 0
\]

We may therefore apply the dominated convergence theorem and deduce

\[
\int f \circ T [I_\mu + \phi' - P_G (\phi')] d\mu \leq 0.
\]

Since \( f \) was arbitrary, (4.8) follows.

Assume that \( \mu \sim \mu \circ T \). We show that the integral in (4.8) is equal to zero if and only if \( d\mu \) is proportional to \( hdv \). If \( d\mu \) is proportional to \( hdv \) the integrand in (4.8) is identically zero because then \( I_\mu = -\log g \) where \( g = \lambda^{-1} e^{\phi} h / h \circ T \) (this
follows from lemma 1 and the fact that the transfer operator of any measure proportional to $h \nu$ is given by $f \mapsto \lambda \, 1 h L^1_\psi (hf)$.

We show the reverse implication. Assume that $\mu$ is such that $\mu \sim \mu \circ T$ and that there is an equality in (4.8). A close inspection of the proof shows that this is possible only if $\log(g/g_\mu) = (g/g_\mu) - 1$ $\mu$-almost everywhere. This is possible only if $g_\mu = g \mod \mu$. Since $\mu \sim \mu \circ T$, this implies that $g_\mu = g \mod \mu \circ T$. It follows that $L_{\log g}$ is the transfer operator of $\mu$. Consider the function $\psi = \log g = \phi + \log h - \log h \circ T - \log \lambda$. This is a weakly H"older continuous function, because by (3.12) $\log h$, $\log h \circ T$ are both weakly H"older continuous. It is also clear that $L_\psi 1 = 1, L^*_\psi \mu = \mu$ whence $\psi$ is recurrent. Since it is also true that $L^*_\psi (h \nu) = L^*_\psi (h \nu) = h \nu$ we have by (3.1) that $\mu$ and $h \nu$ are proportional. \hfill \Box

4.3 Notes

The BIP property is a strengthening of the **big images property** which reads $\exists b_1, \ldots, b_N \in S$ such that $\forall a \in S, \exists i$ for which $t_{a b_i} = 1$. This property was used by Aaronson and Denker in [AD] as a part of their definition of a Gibbs-Markov measure.\footnote{Our formulation of the big images condition is different than that used in [AD], though equivalent to it. There it is stated that for some given measure $m$ supported on $X$, $\inf \{ m(T^i \phi) : a \in S \} > 0$. This accounts for the name ‘big images’.} It follows from the results of this chapter that if an invariant Gibbs-Markov probability measure exists, then $X$ must have the BIP property, and not only big images.

It was proved in [S2] that $\phi$ has an invariant Gibbs measure iff $X$ admits the big images property, $V_1(\phi) < \infty$, $\phi$ is positive recurrent and the corresponding eigenfunction $h$ satisfies $\| \log h \|_\infty < \infty$. A recent preprint [MU] of Mauldin and Urbanski implies that positive recurrence follows from the BIP property and $V_1(\phi) < \infty$. This, together with the observation that BIP implies that $\| \log h \|_\infty < \infty$, yields the sufficiency part of theorem 4. Our proof that BIP implies positive recurrence is new.

Theorem 5 follows from results in [AD] and the observation that the BIP property implies that $h$ is bounded away from zero.

The idea prove the variational principle for finite Markov shifts by using Rokhlin’s formula $h_\mu(T) = \int \log \sigma \mu \int \mu(1 \eta) \mu$ and the inequality $\log x \leq x - 1$ is due to Ledrappier [Le]. This idea was later used by Walters in [Wal2] to define equilibrium measures for some non-compact systems as those maximizing $\int (\log h + \phi) d\mu$. Yuri considered this idea for some infinite conservative measures in [Yu1].
Chapter 5

Phase Transitions

The purpose of this chapter is to discuss the existence of critical phenomena for one-parameter families of potentials with summable variations. The critical phenomena we consider are non-analyticity of the pressure function and changes in the existence and finiteness of the corresponding equilibrium measure. The results of this chapter appear in [S4].

5.1 The discriminant: statement of main results

We use the definitions and notation of section 4.2.2.

Definition 4 Let $X$ be topologically mixing and let $\phi : X \to \mathbb{R}$ have summable variations and finite Gurevic pressure. Fix $a \in S$ and let $(\overline{X}, \overline{\phi})$ be the induced system on $[a]$. Set $p_a^*[\phi] := \sup \{ p : P_{\overline{\phi}}(\overline{\phi} + p) < \infty \}$. The $a$-discriminant of $\phi$ is $\Delta_a[\phi] := \sup \{ P_{\overline{\phi}}(\overline{\phi} + p) : p < p_a^*[\phi] \} \leq \infty$.

We remind the reader that, according to lemma 4, the pressure of $\overline{\phi}$ is well-defined whenever $\phi$ has summable variations.

As we shall later see (proposition 10, section 5.2.1 below),

$$\Delta_a[\phi] = P_{\overline{\phi}}(\overline{\phi} + p_a^*[\phi])$$

The discriminant is not a mere abstraction, and can often be estimated, or even calculated. Both $\Delta_a[\phi]$ and $p_a^*[\phi]$ are determined by $\sum \xi^n Z_k^*(\phi, a)$ in the following way. Let $R$ be the radius of convergence of this series. Then

$$\Delta_a[\phi] = - \log \sum_{k=1}^{\infty} R^k Z_k^*(\phi, a) \leq \sum_{k=2}^{\infty} V_k(\phi)$$

$$p_a^*[\phi] = - \limsup_{n \to \infty} \frac{1}{n} \log Z_k^*(\phi, a)$$

52
5.1. THE DISCRIMINANT: STATEMENT OF MAIN RESULTS

Both relations follow from the stronger statement (section 5.2.1, proposition 10 below):

\[(5.4) \quad P_G(\phi + \tilde{\phi}) - \log \sum_{k=1}^{\infty} e^{\beta p} Z_k^*(\phi, a) \leq \sum_{k=2}^{\infty} V_k(\phi).\]

Recall that a Markov potential is a function of the form \(\phi(x) = \phi(x_0, x_1)\). Note that when \(\phi\) is a Markov potential, both (5.4) and (5.2) are equalities, as \(\sum_{k \geq 2} V_k(\phi) = 0\) for Markov potentials. Our basic result is:

**Theorem 8 (Discriminant theorem)** Let \(X\) be a topologically mixing countable Markov shift and let \(\phi : X \to \mathbb{R}\) be some function with summable variations and finite Gurevich pressure. Let \(a \in S\) be some arbitrary fixed state.

1. The equation \(P_G(\phi + \tilde{\phi}) = 0\) has a unique solution \(p(\phi)\) if \(\Delta_\phi[\phi] \geq 0\), and no solution if \(\Delta_\phi[\phi] < 0\). The Gurevich pressure of \(\phi\) is given by

\[(5.5) \quad P_G(\phi) = \begin{cases} -p(\phi) \quad \Delta_\phi[\phi] \geq 0 \\ -p_*[\phi] \quad \Delta_\phi[\phi] < 0 \end{cases}\]

2. \(\phi\) is positive recurrent if \(\Delta_\phi[\phi] > 0\) and transient if \(\Delta_\phi[\phi] < 0\). In the case \(\Delta_\phi[\phi] = 0\), \(\phi\) is either positive recurrent or null recurrent.

The discriminant theorem should be understood in the context of one-parameter families of potentials. Given such a family \(\{\phi_\beta\}\), let \(\{\Delta_\beta\}\) be the corresponding one-parameter family of discriminants. When \(\Delta_\beta\) changes sign, \(\{\phi_\beta\}\) changes its recurrence properties and the case in (5.5) changes. By theorem 3, a change in the mode of recurrence implies a change in the qualitative properties of the equilibrium measure (existence and finiteness). A change of case in (5.5) may imply non-smoothness for \(\beta \mapsto P_G(\phi_\beta)\). This suggests that the search for critical phenomena for one-parameter families may be done by studying the sign changes of the discriminant. This can sometimes be done with the aid of (5.2), as we shall see in chapter 6. The proof of theorem 8 is given in section 5.2.

We now discuss the case when the discriminant does not change sign and remains positive. Let \(\phi\) be some function with summable variations and finite pressure. We say that \(\phi\) is **strongly positive recurrent** if for some state \(a \in S\)

\[\Delta_\phi[\phi] > 0\]

(This generalizes the notion of stable positivity for Markov potentials discussed in [GS].) The Discriminant Theorem implies that every strongly positive recurrent function is positive recurrent. The opposite statement is false (example 2 below).

We are interested in differentiability of the pressure functional, i.e. in the existence of directional derivatives \(\frac{d}{dt} t \mapsto P_G(\phi + t\psi)\). We restrict ourselves to the following set of directions:

\[\text{Dir}(\phi) := \left\{ \psi : \sum_{n=2}^{\infty} V_n(\psi) < \infty, \quad \exists \varepsilon > 0 \text{ s.t. } \forall |t| < \varepsilon, \quad P_G(\phi + t\psi) < \infty \right\}\]
The following theorem completes the discriminant theorem by saying that if the discriminant is positive, then there is no critical phenomena of the sort that can be encountered when $\Delta$ changes sign. Its proof is given in section 5.3.

**Theorem 9 (Strong positive recurrence theorem)** Let $X$ be a topologically mixing and $\phi$ be a weakly Hölder continuous function such that $P_G(\phi) < \infty$. If $\phi$ is strongly positive recurrent, then $\forall \psi \in \text{Dir}(\phi)$ weakly Hölder continuous, $\exists \varepsilon > 0$, such that $\phi + t\psi$ is positive recurrent for all $|t| < \varepsilon$, and such that $t \mapsto P_G(\phi + t\psi)$ is real analytic in $(-\varepsilon, \varepsilon)$.

The case $\psi = \phi$ is particularly interesting, as it appears in the study of the one-parameter family $\{\beta\phi\}_{\beta \geq \beta_0}$. If $P_G(\beta\phi) < \infty$, then $P_G(\beta\phi) < \infty$ for all $\beta > \beta_0$, because by corollary 1, $\phi$ is cohomologous to a non-positive function. Therefore, $\forall \beta > \beta_0, \phi \in \text{Dir}(\beta\phi)$. This, however, may not be true for $\beta = \beta_0$.

**Example 1** Let $X = \mathbb{N}^{\nu(x)}$ and $\phi(x) := -\log(x_0[\log 2x_0]^2)$. Then $P_G(\beta\phi) < \infty$ for $\beta \geq 1$, and $P_G(\phi) = \infty$ for $\beta < 1$.

**Proof.** $P_G(\beta\phi) = \log\sum_{k \geq 1} 1 / [k^\beta(\log 2k)^{2\beta}]$. \hfill $\square$

**Corollary 3** Let $X$ be a topologically mixing and let $\phi$ be weakly Hölder continuous function such that $P_G(\phi) < \infty$ and $\phi \in \text{Dir}(\phi)$. The following conditions are equivalent:

1. $\phi$ is strongly positive recurrent.
2. For every weakly Hölder continuous $\psi \in \text{Dir}(\phi)$ there exists $\varepsilon > 0$ such that $\phi + t\psi$ is positive recurrent for every real $t$ such that $|t| < \varepsilon$.
3. For every $a \in S$ $\Delta_a[\phi] > 0$.

**Proof.** The first statement implies the second by theorem 9. The third statement trivially implies the first. It remains to show that the second statement implies the third. Assume that the second statement is true, but that the third statement is false. Then for some $a \in S$, $\Delta_a[\phi]$ is not positive. Since by our assumptions $\phi$ is positive recurrent, $\Delta_a[\phi]$ cannot be negative, so $\Delta_a[\phi] = 0$. Set $\psi := 1_{[a]}$. Since $\psi = 1$ on $X$, $\Delta_a[\phi + t\psi] = t$. This contradicts the second statement because if $t < 0$ then $\phi + t\psi$ is transient. \hfill $\square$

We remark that the assumptions of theorem 9 can be weakened:

**Theorem 10** Let $X$ be a topologically mixing and $\phi$ be a function with summable variations, such that $P_G(\phi) < \infty$, $\Delta_a[\phi] > 0$ and such that the induced potential on $a$, $\overline{\phi}$, is weakly Hölder. Then $\forall \psi \in \text{Dir}(\phi)$, such that $\overline{\psi}$ is weakly Hölder continuous, $\exists \varepsilon > 0$ such that $\phi + t\psi$ is positive recurrent $\forall |t| < \varepsilon$, and such that $t \mapsto P_G(\phi + t\psi)$ is real analytic in $(-\varepsilon, \varepsilon)$.

In section 5.3 we prove this stronger version.

---

1Such families appear in models for systems whose inverse temperature $\beta$ is changed [E].
5.2 Proof of the discriminant theorem

The proof of the discriminant theorem is based on a generalization of certain renewal theoretic ideas. These are presented in the following subsection. The proof of the discriminant theorem is given in the subsection following it.

5.2.1 A renewal sequence of operators

Let \( a \in S \) be some fixed state. Let \( C_B[a] \) be the Banach space \( C_B[a] := \{ f \in C_B(X) : f(x) = 0 \text{ for } x \not\in [a] \} \) equipped with the supremum norm. Let \( 1, 0 : C_B[a] \to C_B[a] \) be the operators defined by \( 1f = f, 0f = 0 \forall f \in C_B[a] \). Consider the operators \( T_n, R_n : C_B[a] \to C_B[a] \) given by \( T_0 := 1, R_0 := 0 \) and

\[
T_n f := 1_{[a]} L^n \varphi f,
\]

\[
R_n f := 1_{[a]} L^n \varphi(f1_{[\varphi = n]})
\]

(see also [FL] and [PS]). A direct calculation shows that these operators satisfy the following ‘renewal equation’ for \( n \geq 1 \),

\[
T_n = R_1 T_{n-1} + R_2 T_{n-2} + \ldots + R_n T_0
\]

\[
T_n = T_{n-1} R_1 + T_{n-2} R_2 + \ldots + T_0 R_n
\]

Set

\[
T_a[\varphi](z) := 1 + \sum_{n=1}^{\infty} z^n T_n, \quad R_a[\varphi](z) := \sum_{n=1}^{\infty} z^n R_n.
\]

These are well-defined bounded linear operators on \( C_B[a] \) for \( |z| < \lambda \). To see this use the summable bounded linearity property to prove that \( \|T_a[\varphi](z)\| = \|T_a[\varphi](z)1_{[a]}\|_{\infty} \leq B \sum_{n=0}^{\infty} |z|^n Z_n(\varphi, a) \) where \( B := \exp \sum_{n=2}^{\infty} V_n(\varphi) \), and note that by the definition of the Gurevich pressure, the radius of convergence of the series \( \sum z^n Z_n(\varphi, a) \) is \( \lambda \). In terms of these generating functions, we can restate (5.6) in the following form \( \forall |z| < \lambda \),

\[
T_a[\varphi](z) = [1 - R_a[\varphi](z)]^{-1}.
\]

It also follows from (5.6) that for all \( |z| < \lambda \),

\[
T_a[\varphi](z) = 1 + \sum_{n=1}^{\infty} R_a[\varphi](z)^n.
\]

Note that (5.7) is also valid for all \( z \) real, such that \( z \geq \lambda \), as long as both sides are applied to positive functions.

For every bounded linear operator \( S \) on \( C_B[a] \), let \( \rho(S) \) denote the spectral radius of \( S \) (with respect to the supremum norm), with the convention that the ‘operator’ \( Sf = 1_{[a]} \) has an infinite spectral norm. The following two propositions relate the renewal sequence to the discriminant.
Proposition 9 Let \( \overline{\phi} \) denote the induced potential on \([a]\). Then
\[
P_G(\overline{\phi} + p) = \log \rho \left( R_a[\phi](e^p) \right).
\]

**Proof.** Let \( \pi : \overline{X} \to [a] \) be the natural embedding. A calculation shows that for every \( p \) and \( f \in C_B[a] \)
\[
(R_a[\phi](e^p) f) \circ \pi = L_{\overline{\phi} + p}(f \circ \pi).
\]
Fix some \( [a] \in \overline{S} \). Since \( \overline{X} = S^{N \cup \{0\}} \), \( Z_n(\overline{\phi} + p, [a]) \preceq \|L_{\overline{\phi} + p}^n\|_\infty \). The proposition follows from this and (5.8). \( \square \)

Proposition 10 Let \( X \) be topologically mixing countable Markov shift, let \( \phi \) be a function with summable variations and finite Gurevich pressure. Let \( \overline{X} \) and \( \overline{\phi} \) denote the induced pair with respect to \( a \in S \). Then \( P_G(\overline{\phi} + p) \) is convex, strictly increasing and continuous in \( (-\infty, p^* \phi) \). Also, (5.1)-(5.4) hold.

**Proof.** Fix \( a \in S \) and set \( \gamma(p) := P_G(\overline{\phi} + p), p^* := p^*_a[\phi], z := e^p \) and \( R(z) := R_a[\phi](z) \). Proposition 9 and summable variations imply that for all \( x \in [a] \),
\[
\gamma(p) = \lim_{n \to \infty} \frac{1}{n} \log \|R(e^p)^n1_{[a]}\|
= \lim_{n \to \infty} \frac{1}{n} \log R(e^p)^n1_{[a]}(x).
\]
We use this formula to prove that \( \gamma \) is convex, strictly increasing and continuous in \( (-\infty, p^*) \) and that (5.4) and (5.3) hold. The expressions \( R(e^p)^n1_{[a]}(x) \) are convex in \( p \), since they are of the form \( \sum a_i e_k p \), where \( a_i \geq 0 \). Thus, \( \gamma(p) \) is convex in \( p \), being a limit of convex functions. Clearly, \( \forall p < p^*_a \phi \) and all \( 0 \leq f \in C_B[a] \), \( R(e^p)f \geq e^{p} f R(e^p) f \). Iterating this and using the above formula for \( \gamma(p) \) we have that \( \gamma(p) \) is strictly increasing. It follows that \( \gamma \) is finite in \( (-\infty, p^*) \), whence by convexity it is continuous there. Standard estimations show that for every \( 0 \leq f \in C_B[a], R(e^p)f = B^{\geq 1} \sum_{k \geq 1} Z_n[\phi,a] f \) where \( B = \exp \sum_{k \geq 2} V_k(\phi) \). Iterating this, and using the above formula for \( \gamma(p) \), we have (5.4). Clearly, (5.4) implies (5.3), so (5.3) is also proved.

It remains to prove that \( \gamma(p) \) is continuous on the left in \( p^* \). We prove this under the assumption that \( \gamma(p^*) < \infty \) (the proof for the infinite case is essentially the same). By monotonicity, it is enough to prove that for every \( \varepsilon > 0 \) there exists \( p < p^* \) such that \( \gamma(p) > \gamma(p^*) - \varepsilon \). Fix \( x \in [a] \). Setting \( R(e^p) = \sum_{n \geq 1} e^{np} R_n \) in the above formula for \( \gamma \), we have
\[
(5.9) \quad \gamma(p) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{k_1, \ldots, k_n = 1} e^{p(k_1 + \ldots + k_n)} R_{k_1} \cdot \ldots \cdot R_{k_n} 1_{[a]}(x).
\]
By the definition of \( p^* \), there are \( N \) and \( p \) and \( n > \log B / \varepsilon \), such that
\[
a_n := \log \sum_{k_1, \ldots, k_n} e^{p(k_1 + \ldots + k_n)} R_{k_1} \cdot \ldots \cdot R_{k_n} 1_{[a]} \geq n \left( \gamma(p^*) - \varepsilon \right)
\]
By the summable variations property, \( a_{m_1} + a_{m_2} \leq m_{1,m_2} + \log B \) for every \( m_1, m_2 \). Write \( m = kn + r \) where \( 0 \leq r \leq n - 1 \). Then
\[
\frac{a_m}{m} \geq \frac{k a_n + a_r - (k + 1) \log B}{kn + r} \xrightarrow{m \to \infty} \frac{a_n - \log B}{n} \geq \gamma(p) - 2\varepsilon.
\]
whence \( \gamma(p) \geq \gamma(p^*) - 2\varepsilon \). This proves that \( \gamma \) is continuous in \((-\infty, p^*)\]. This also implies that (5.1). This and (5.4) imply (5.2). \( \square \)

5.2. Proof of the discriminant theorem.

Throughout the proof let \( a \) and \( \phi \) be fixed. Let \( T(z) := T_a[\phi](z), R(z) := R_a[\phi](z) \). Let \( B := \exp \sum_{k \geq 1} V_k[\phi] \) and set \( \Delta := \Delta_a[\phi] \), \( \gamma(p) := P^G(\phi + p) \), and \( p^* := p_a[\phi] \).

**Part 1.** Proof of (5.5).

Assume \( \Delta \geq 0 \). According to proposition 10, \( \gamma \) is continuous and strictly increasing in \((-\infty, p^*)\], \( \gamma(p^*) \geq 0 \) and \( \gamma(p) \to -\infty \). Therefore, there exists a unique \( p_a(\phi) \) for which \( \gamma(p_a(\phi)) = 0 \).

We claim that \( p_a(\phi) = -P^G(\phi) \). Fix some \( p < p_a(\phi) \). Since \( \gamma \) is strictly increasing, \( \gamma(p) < 0 \), whence \( \rho(R(e^p)) < 1 \). By (5.7), \( ||T(e^p)1_{[a]}|| \leq 1 + \sum ||R(e^p)|| < \infty \) whence, by summable variations, \( \sum e^{np}Z_n(\phi,a) \) converges. The radius of convergence of this series is \( \exp[-P^G(\phi)] \). Therefore, \( p < -P^G(\phi) \). Taking \( p \uparrow p_a(\phi) \) we have \( p_a(\phi) \leq -P^G(\phi) \). Assume by way of contradiction that \( \exists p_a(\phi) < p^* < -P^G(\phi) \). Then \( ||T(e^{p^*})|| \leq B \sum e^{np^*}Z_n(\phi,a) < \infty \), whence by (5.7), the series \( 1 + \sum R(e^{p^*})^n1_{[a]}(x) \) converges for every \( x \). Summable variations imply that \( \forall x \in [a] \) and \( \forall n \geq 1, ||R(e^{p^*})|| \leq BR(e^{p^*})^n1_{[a]}(x) \). Thus, \( \sum ||R(e^{p^*})|| < \infty \) whence \( \rho(R(e^{p^*})) \leq 1 \), or equivalently, \( \gamma(p^*) < 0 \). This, however, is impossible because \( \gamma \) is strictly increasing, \( \gamma(p_a(\phi)) = 0 \) and \( p^* > p_a(\phi) \).

This proves that \( p_a(\phi) = -P^G(\phi) \) and settles (5.5) for the case \( \Delta > 0 \).

Assume now that \( \Delta < 0 \). In this case there is no solution for the equation \( \gamma(p) = 0 \), because for \( p \leq p^* \), \( \gamma(p) \leq \Delta < 0 \), and for \( p > p^* \), \( \gamma(p) = \infty \). We show that in this case \( P^G(\phi) = -p^* \). By (5.3) and the inequality \( Z_n(\phi,a) < \infty \), \( p^* \geq -P^G(\phi) \). Assume by way of contradiction that \( p^* > -P^G(\phi) \). Then for every \( x \in [a], T(e^{p^*})1_{[a]}(x) \geq B^{-1} \sum e^{np^*}Z_n(\phi,a) = \infty \), whence by (5.7), \( 1 + \sum R(e^{p^*})^n \) diverges everywhere on \( [a] \). Thus \( \rho(R(e^{p^*})) \geq 1 \), whence \( \Delta = \gamma(p^*) \leq -\infty \) in contradiction to our assumptions. This settles the case \( \Delta < 0 \).

**Part 2.** Proof that recurrence is equivalent to \( \Delta \leq 0 \).

Assume that \( \Delta \geq 0 \). By the first part of the theorem \( \gamma(-P^G(\phi)) = 0 \), so the spectral norm of \( R(e^{P^G(\phi)}) \) is equal to 1. Therefore, there exists \( \xi \in [0,2\pi) \), such that \( 1 - e^{\xi}R(e^{P^G(\phi)}) \) does not have a bounded inverse operator. In particular, the series \( 1 + \sum_{k \geq 1} e^{i\xi}R(e^{P^G(\phi)})^k \) does not converge in the strong norm. It follows that there exists \( \varepsilon > 0 \) such that for every \( N \exists n = n(N) > N \).
and \( \exists \eta_n \in C_B[\alpha] \) such that \( \| \eta_n \|_{\infty} = 1 \) and \( \left\| \sum_{k \geq n} e^{ik\xi} R(e^{P\eta(\phi)}^k \eta_n) \right\|_{\infty} > \varepsilon \).

Now, on \([\alpha]\)

\[
\sum_{k=n}^{\infty} R(e^{P\eta(\phi)}^k)1_{[\alpha]} \geq \sum_{k=n}^{\infty} R(e^{P\eta(\phi)}^k)\eta_n \geq \sum_{k=n}^{\infty} e^{ik\xi} R(e^{P\eta(\phi)}^k \eta_n) \geq \varepsilon
\]

whence \( \| \sum_{k \geq n} R(e^{P\eta(\phi)}^k)1_{[\alpha]} \|_{\infty} \geq \varepsilon \) for every \( n \). By the summable variations property this is only possible if \( \sum R(e^{P\eta(\phi)}^k)1_{[\alpha]} \) diverges on \([\alpha]\) whence \( \|T(e^{P\eta(\phi)})1_{[\alpha]}\|_{\infty} = \infty \). This is equivalent to \( \sum e^{k\eta(\phi)} Z_k(\phi,a) = \infty \), so \( \phi \) is recurrent.

Assume that \( \Delta < 0 \). Set \( \rho := \rho[R(e^{P\eta(\phi)})] \). Then \( \rho = \exp \Delta < 1 \). By the definition of the spectral radius, there exists some \( C \) and \( \rho_0 \in (\rho,1) \) such that for every \( n \), \( \|R(e^{P\eta(\phi)})^n\|_{\infty} < C\rho_0^n \). The renewal equation implies that \( \|T(e^{P\eta(\phi)})\| \leq C/(1-\rho_0) \). It follows that \( T(e^{P\eta(\phi)}) \) is bounded, whence \( \sum e^{k\eta(\phi)} Z_k(\phi,a) \) is convergent. By the first part of the theorem, and since \( \Delta < 0 \), \( P^* = -P_G(\phi) \). It follows that \( \phi \) is transient.

**Part 3.** Proof that \( \Delta > 0 \) implies positive recurrence.

Assume that \( \Delta > 0 \). By what we have just proved \( \phi \) is recurrent. Let \( \nu \) and \( h \) be the eigenmeasure and eigenfunction given by theorem 3, and set \( dm = hdm \). Recall that \( m \) an invariant measure, and that \( m(X) < \infty \) if and only if \( \phi \) is positive recurrent. We will prove positive recurrence by showing that \( m(X) < \infty \).

Let \( \nu_a, m_a \) be the measures \( \nu_a(A) := \nu(A \cap [\alpha])/\nu[\alpha] \) and \( m_a(A) := m(A \cap [\alpha])/m[\alpha] \). Let \( T_a := T^{\nu_a} \) be the induced transformation. Since \( m \) is \( T \)-invariant, \( m_a \) is \( T_a \)-invariant. Note that the transfer operator of \( T_a \) with respect to \( \nu_a \) is \( R(\lambda^{-1}) \). To see this note that \( \forall g \in L^\infty(\nu_a), \forall f \in L^1(\nu_a), \nu_a[gR(\lambda^{-1})f] = \sum_{n=1}^{\infty} \nu \left[ \lambda^{-n}L^\phi_1(g \circ T^n \cdot f_{\nu_a=n}) \right] = \nu_a(g \circ T_a \cdot f) \).

Set \( A(x) := [R(\lambda^{-1})\eta_a](x) \). By Kac’s formula, the fact the \( R(\lambda^{-1}) \) acts as the transfer operator of \( \eta_\alpha \) and the boundness of \( h = \frac{dm}{d\eta} \) away from zero and infinity on partition sets, \( \exists C_1 > \) such that

\[
m(X) = \int \varphi_a \, dm_a = C_1^{\pm 1} \int \varphi_a \, d\nu_a = \int_{[\alpha]} A(x) d\nu(x).
\]

Clearly,

\[
A(x) = \sum_{n=1}^{\infty} n \lambda_n \sum_{T^y_\eta=x} e^{\phi_a(y)}1_{[\nu_a=n]}(y) = B^{\pm 1} \sum_{n=1}^{\infty} n \lambda_n^\ast Z_n^*(\phi,a).
\]

Since \( \Delta > 0 \), \( P^* > -P_G(\phi) \) so \( \lambda^{-1} \) is smaller than the radius of convergence of the series \( \sum z^n Z_n^*(\phi,a) \). It follows that \( \sum n \lambda_n^\ast Z_n^*(\phi,a) < \infty \) whence \( \|A(x)\|_{\infty} < \infty \). Since \( \nu[\alpha] < \infty \), \( m(X) < \infty \) as required. \( \square \)
5.3 Proof of the strong positive recurrence theorem

In this section we prove theorem 10, a strengthened version of theorem 9.

5.3.1 Preparations

Let $X$ be topologically mixing and let $a \in S$ be some fixed state. Let $\phi$ be some function with summable variations and finite pressure and let $(\bar{X}, \bar{\phi})$ be the induced system on $[a]$.

For every $\bar{x}, \bar{y} \in \bar{X}$ set $\bar{d}(\bar{x}, \bar{y}) := \inf\{n \geq 0 : \bar{x}_n \neq \bar{y}_n\}$. Fix $\theta \in (0, 1)$ and set for every function $f : X \to C$, $Df := \sup\{|f(\bar{x}) - f(\bar{y})|/\theta^d(\bar{x}, \bar{y}) : \bar{x} \neq \bar{y}\}$. Let $\bar{L} = \bar{L}(\theta, a)$ be the space

$$\bar{L}(\theta, a) := \{f \in C([\bar{X}]) : \|f\|_{\bar{L}} := \|f\|_{\infty} + Df < \infty\}$$

A standard argument shows that $\bar{L}$ is a Banach space and that $\|L_{\bar{\phi}}\|_{\infty} < \infty$ then $L_{\bar{\phi}}(\bar{L}) \subset \bar{L}$ and $\|L_{\bar{\phi}}\|_{\mathcal{B}(\bar{L})} < \infty$ where $\mathcal{B}(\bar{L})$ is the space of bounded operators on $\bar{L}$ equipped with the strong operator norm. The following lemma says that the induced system has a spectral gap, and is similar to well-known results in the theory of interval maps with indifferent fixed points ([T1], [T2], [A1], [ADU], [B2], [PS]).

**Lemma 6** Let $X$ be topologically mixing, assume $\phi$ has summable variations and fix some $a \in S$. Let $(\bar{X}, \bar{\phi})$ be the induced system on $[a]$ and assume that $\bar{\phi}$ is weakly Hölder continuous with exponent $\theta \in (0, 1)$ and that $\|L_{\phi}\|_{\infty} < \infty$.

Then $\bar{\phi}$ is positive recurrent and the spectrum of $L_{\bar{\phi}} : \bar{L}(\theta, a) \to \bar{L}(\theta, a)$ consists of a simple eigenvalue $\bar{\lambda}$ and a subset of $\{z : |z| < \tau \bar{\lambda}\}$ where $\tau < 1$. The eigenvalue $\bar{\lambda}$ is equal to $e^{\epsilon \bar{\phi}(\bar{\omega})}$.

**Proof.** The pressure of $\bar{\phi}$ is finite since $\|L_{\phi}\|_{\infty} < \infty$. By lemma 3 in section 4.2.2, $X$ has the BIP property and $V_1(\phi) < \infty$ whence, by theorem 5, $\exists \kappa > 0$, $\exists \tau \in (0, 1)$ such that for every $f \in \bar{L}$

$$\left\|\bar{\lambda}^{-n}(L_{\phi})^n f - \bar{h} \int f d\bar{\mu}\right\|_{\bar{L}} < K\tau^n$$

where $\bar{\lambda} = e^{\epsilon \bar{\phi}(\bar{\omega})}$, $L_{\phi} \bar{h} = \bar{\lambda} \bar{h}$, $L_{\phi} \bar{\mu} = \bar{\lambda} \bar{\mu}$ and $\bar{\mu}(\bar{h}) = 1$. \qed

**Lemma 7** Let $X$ be topologically mixing and let $\phi$ be a function with summable variations such that $P_G(\phi) < \infty$ and $\Delta_\phi[\phi] > 0$. For every $\psi \in \text{Dir}(\phi)$, $\exists \varepsilon > 0$ and $\exists r > 0$ such that $P_G(\phi) - P_G(\psi) > 0$.

**Proof.** Without loss of generality, $P_G(\phi) = 0$ (else pass to $\phi - P_G(\phi)$). Since $\Delta_\phi[\phi] > 0$, $\exists r > \exp[-P_G(\phi)] = 1$ such that $\|R_\phi[\phi](r) I_{[a]}\|_{\infty}$ is finite, or equivalently, $\sum_{n \geq 1} r^n Z^n_\phi(\phi, a) < \infty$. Without loss of generality

$$\limsup_{n \to \infty} -\frac{1}{n} \log Z^n_\phi(\phi, a) < -\log r$$
Set $f_n(t) := (1/n) \log Z_n(\phi + t \psi; a)$ and $f(t) := \limsup_{n \to \infty} f_n(t)$. By Hölder's inequality, $f_n$ are convex, whence so is $f$. Since $\psi \in \text{Dir}(\phi)$, there is some $\varepsilon > 0$ such that $\forall t < 2 \varepsilon, -\infty < f(t) \leq P_G(\phi + t \psi) < \infty$.

By convexity, and since $f < \infty$, either $f(t) = -\infty$ everywhere in $(-2\varepsilon, 2\varepsilon)$, or $\{f(t)\} < \infty$ everywhere in $(-2\varepsilon, 2\varepsilon)$. In the first case the radius of convergence of $\sum_{k \geq 1} x^k Z_k(\phi + t \psi; a)$ is infinite for $t = \pm \varepsilon$, and we are done. In the second case, by convexity and finiteness, $f(t)$ is continuous in $(-2\varepsilon, 2\varepsilon)$. Thus, since $r$ was chosen so that $f(0) < -\log r$, there exists $\varepsilon' < \varepsilon$, such that $\forall t < 2\varepsilon'$, $f(t) < -\log r$. It follows that $r$ is strictly smaller than the radius of convergence of $\sum_{k \geq 1} x^k Z_k(\phi + t \psi; a)$ for $t = \pm \varepsilon'$, and again, we are done.

Recall that function $F : \mathbb{C} \times \mathbb{C} \to B(\mathcal{L})$ is called analytic in a neighborhood of $(z_0, w_0)$ if $\exists F_n \in B(\mathcal{L})$, such that $F(z, w) = \sum_{n, k \geq 0} (w - w_0)^k (z - z_0)^k F_n$ and the series converges in the strong operator norm in a neighborhood of $(z_0, w_0)$.

**Lemma 8** Let $X$ be topologically mixing, let $\phi$ be some function with summable variations such that $P_G(\phi) < \infty$ and let $\psi \in \text{Dir}(\phi)$. Let $a \in S$ be some state such that $\Delta_\phi(\phi) > 0$ and assume that $\phi$ and $\psi$, the induced potentials on $[a]$, are weakly Hölder continuous with parameter $\theta$. Then $F : \mathbb{C} \times \mathbb{C} \to B(\mathcal{L})$ given by $F(z, w) = L_{f + z\psi + \log w}$ is analytic in a neighborhood of $(z, w) = (0, 0, P_G(\phi))$.

**Proof.** Throughout this proof $\| \cdot \|$ denotes the strong operator norm in $B(\mathcal{L})$. We assume, without loss of generality, that $P_G(\phi) = 0$ and prove analyticity in $(0, 1)$. For every function $g : \overline{X} \to \mathbb{C}$ let $M_g$ be the operator $M_g f = gf$. Set $\Lambda_n := \{ x \in X : \varphi_\alpha(\pi(x)) = n \}$. This is a union of partition sets in $X$. Let $\overline{R}_n := L_{\varphi} M_{1_{\Lambda_n}}$. Then,

$$
F(z, w) = \sum_{n = 1}^{\infty} w^n \overline{R}_n^* M_{ev}
$$

$$
= \sum_{n = 1}^{\infty} \sum_{k = 0}^{\infty} \frac{w^n z^k}{k!} \overline{R}_n^* M_{\psi}
$$

We show that this converges in $B(\mathcal{L})$ in some open ball containing $(z, w) = (0, 1)$. Fix some $N$ and set $A_N(x) := e^x - (1 + x + \frac{x^2}{2} + \ldots + \frac{x^N}{N!})$. Then

$$
\left\| \sum_{n, k > N} \frac{w^n z^k}{k!} \overline{R}_n^* M_{\psi} \right\| \leq \sum_{n > N} |w|^n \left\| \overline{R}_n^* M_{A_N(z\psi)} \right\|
$$

We estimate the summands of the last series.

For every $\overline{p} \in \mathcal{S}$ set $Q_{\overline{p}} f(\overline{q}) := f(\overline{p} \overline{q})$. Let $\Lambda'_n := \{ \overline{p} \in \mathcal{S} : |\overline{p}| \subseteq \Lambda_n \}$. By definition, $\overline{R}_n M_{A_N(z\psi)} = \sum_{\overline{p} \in \Lambda'_n} Q_{\overline{p}} M_{A_N(z\psi)} \overline{q}$. Since for every $f, g \in \mathcal{L}$,

$$
\left\| Q_{\overline{p}} M_{A_N(z\psi)} \overline{q} \right\| \leq \left\| Q_{\overline{p}} \right\| \left\| Q_{\overline{p}} A_N(z\psi) \right\| \left\| \overline{q} \right\|
$$

It is standard to check that $\forall x, y \in \mathbb{C}$, $|A_N(x) - A_N(y)| \leq |x - y| (e^{|x|} + e^{|y|})$ and that $\forall x, y \in \mathbb{R}$, $|e^x - e^y| \leq |x - y| (e^x + e^y)$. Using this and the inequality
\[ |A_N(x)| \leq e^{\|x\|}, \text{ it is easy to show that there is some constant } K_1 \text{ (independent of } n \text{ and } N) \text{ such that } \forall |z| < 1, \]
\[ \|Q_n M_{A_N(z \bar{\psi}) \exp \psi}^{\star} \| \leq K \|e^{\|z\|1} \| \infty \|e^{z \bar{\psi}}1 \| \infty \]
Summing over all \( \bar{\psi} \in A'_n \) and using weak Hölder continuity, we have that for some \( K \) independent of \( n \) and \( N \),
\[ \|Q_n M_{A_N(z \bar{\psi})}^{\star} \| \leq K Z_n(\phi + |z| \cdot |\psi|, a) \]
Let \( \varepsilon > 0 \) and \( r > 1 \) be as in lemma 7. Without loss of generality re \( r > 1 \) and \( \varepsilon < 1 \). Then for all \( |z| < \varepsilon \) and \( |w| < r \),
\[ \left\| \sum_{n=N+1}^{\infty} \sum_{k=N+1}^{\infty} \frac{|w|^n |z|^k}{k!} R_n M_{\psi_n}^{\star} \right\| \leq K \sum_{n=N+1}^{\infty} r^n Z_n(\phi + \varepsilon |\psi|, a) \to 0 \] 
whence \( F(z, w) \) is analytic in a neighborhood of \((0, 1)\). \( \square \)

5.3.2 Proof of theorem 10

Let \( \phi \) be a function with summable variations and finite pressure and let \( \psi \in \text{Dir}(\phi) \). Assume that \( 2 \theta \in S \) such that \( \Delta_\theta[\phi] > 0 \) and such that the induced potentials on \([a] \), \( \bar{\phi}, \bar{\psi} \) are weakly Hölder continuous with exponent \( \theta \in (0, 1) \). Without loss of generality, assume that \( P_G(\phi) = 0 \). Set
\[ \Gamma(z, w) := P_G(\phi + z\psi - w) \]
By the discriminant theorem, \( \forall z \in \mathbb{R} \), if \( \exists w \in \mathbb{R} \) such that \( \Gamma(z, w) = 0 \), then \( w = P_G(\phi + z\psi) \). Thus, \( P_G(\phi + z\psi) \) is given implicitly by
\[ \Gamma(z, P_G(\phi + z\psi)) = 0 \]
(5.11)

We will show that \( \Gamma \) has a complex holomorphic extension to a neighborhood of \((z, w) = (0, 0)\) in \( \mathbb{C} \times \mathbb{C} \), and apply the complex implicit function theorem ([Boch], page 39) to deduce that (5.11) defines \( P_G(\phi + z\psi) \) real analytically in a neighborhood of \( z = 0 \). (This theorem applies since \( \forall h > 0, \Gamma(0, h) \leq P_G(\phi - h) = P_G(\phi) - h \) whence \( \Gamma_w(h, 0) \neq 0 \).)

By theorem 8 and lemma 5, since \( \Delta_\theta[\phi] > 0 \) and \( P_G(\phi) = 0 \), \( \phi \) is positive recurrent with pressure zero. By (5.4), \( \sum_{n=1}^{\infty} Z_n^{\star}(\phi, a) < \infty \) whence \( \|L_\phi\| < \infty \). By lemma 6 the spectrum of \( L_\phi^{\star} : \overline{\mathbb{C}} \to \overline{\mathbb{C}} \) consists of the simple isolated eigenvalue 1 and a compact subset of \( \{z : |z| < \tau\} \) for some \( \tau < 1 \). By standard analytic perturbation theory [Ka], there exists \( \delta > 0 \) such that if \( \|L - L_\phi\|_{\mathcal{B}(\overline{\mathbb{C}})} < \delta \) then \( L \) has a (unique) simple eigenvalue \( \lambda(L) \) of maximal magnitude, this eigenvalue is simple, has magnitude larger than \((1 + \tau)/2\), and the rest of the spectrum is contained in \( \{z : |z| < (1 + \tau)/2\} \). Furthermore, the map \( L \to \lambda(L) \) is holomorphic in \( \{L \in \mathcal{B}(\overline{\mathbb{C}}) : \|L - L_\phi\|_{\mathcal{B}(\overline{\mathbb{C}})} < \delta\} \). By lemma 8, \( \exists \varepsilon > 0 \) such that
\((z, w) \mapsto \frac{L_{\phi + z\psi} - w}{w}\) is holomorphic in \(U := \{(z, w) \in \mathbb{C}^2 : |z|, |w| < \varepsilon\}\) and such that \(\|L_{\phi + z\psi} - w - L_{\phi}\|_{B(\mathbb{C})} < \delta\) for all \(|z|, |w| < \varepsilon\). In this neighborhood we define

\[
\hat{\Gamma}(z, w) := \log \lambda \left( \frac{L_{\phi + z\psi} - w}{w} \right)
\]

\(\hat{\Gamma}\) is holomorphic in \(U\). For every \(z, w\) real such that \((z, w) \in U\), the spectrum of \(L_{\phi + z\psi} - w\) consists of a simple eigenvalue \(\lambda(z, w)\) and a compact subset of \(\{\lambda : |\lambda| < |\lambda(z, w)|\}\). By lemma 6, \(\phi + z\psi - w\) is positive recurrent with pressure log \(\lambda(z, w) = \hat{\Gamma}(z, w)\). It follows that \(\hat{\Gamma}\) is a holomorphic extension of \(\Gamma\). This proves that \(t \mapsto P_T(\phi + t\psi)\) is real analytic in \((-\varepsilon, \varepsilon)\).

We show that \(\phi + t\psi\) is positive recurrent for \(|t|\) small. Real analyticity implies continuity, so \(\exists \delta' > 0\) such that \(\forall |t| < \delta', P_T(\phi + t\psi) \in (-\hat{\gamma}, \hat{\gamma})\). Set \(w := -P_T(\phi + t\psi)\). Then \(w - \hat{\gamma} < \varepsilon\) whence \(P_T(\phi + t\psi - w + \hat{\gamma}) = \Gamma(t, w - \hat{\gamma}) < \infty\). Since \(P_T(\phi + t\psi + p)\) is increasing in \(p, P_T(\phi + t\psi + (\hat{\gamma} - w)) > \Gamma(t, w) = 0\) whence \(\Delta_v(\phi + t\psi) > 0\). \(\square\)

5.4 Notes

Although phase transitions are a frequent phenomena for multi-dimensional systems (see [E]), they are relatively rare for one-dimensional systems. For Hölder continuous functions on finite Markov shifts, there are no phase transitions of the form discussed in this chapter (Ruelle [Ru2]). Continuous (but not Hölder continuous) functions on finite Markov shifts with more than one equilibrium measure were constructed by Dyson [D], Bramson and Kalikow [BK] and Hofbauer [Hof].

A close inspection of Hofbauer’s example shows that the mechanism which produces the phase transition is an embedded countable Markov chain. Wang studied critical phenomena for a specific Markov potential on the renewal shift in [W1, W2]. Other one-dimensional systems with phase transitions which owe their critical phenomena to a ‘hidden’ countable Markov shift were studied by Pomeau and Manneville [PM], Lopes [Lo] and Prellberg and Slawny [PS]. It is interesting to note that although the methods of those papers are different, the countable Markov shift which is responsible for the critical phenomena in all these examples is the same shift, the renewal shift (see section 6.2 in the next chapter).

The possibility that a one-parameter family of matrices \(e^{\phi(i,j)+t\psi(i,j)}\) change its mode of recurrence at \(t = 0\) (in the sense of Vere-Jones [VJ1, VJ2]) was known to Gurevich and Savchenko. A condition preventing this from happening was given by them in [GS]. This condition can be shown to be equivalent to the strong positive recurrence of the potential \(\phi\).
Chapter 6

Examples

The purpose of this chapter is to present some examples which demonstrate the possible range of critical phenomena for countable Markov shifts. We focus on the behavior of one-parameter families of the form \( \{ \beta \phi \}_{\beta > \beta_0} \). These are the one-parameter families one encounters in the study of systems whose temperature is changed (\( \beta \) corresponds to the 'inverse temperature').

The example in section 6.3.5 appears in [84]. The rest of the results, with the exception of those in 6.1, are contained in [85].

6.1 Systems with BIP

Let \( X \) be topologically mixing and let \( \phi \) be some weakly Hölder continuous function with finite pressure and summable variations. In section 4.1 we have seen that when \( X \) satisfies the BIP property (4.2) and \( V_1(\phi) < \infty \), the behavior of \((X, \phi)\) is similar to what happens for finite Markov shifts: \( \phi \) is positive recurrent, has an invariant Gibbs measure, and the convergence in the RPF theorem is exponential. In this section we see that the similarity persists as far as critical phenomena is concerned: under the above assumptions there are no phase transitions, as long as the pressure remains finite.

**Theorem 11** Let \( X \) be topologically mixing and assume \( \phi \) is weakly Hölder continuous with finite Gurevich pressure. If \( X \) satisfies the BIP property (4.2) and \( V_1(\phi) < \infty \), then \( \phi \) is strongly positive recurrent. In particular, \( P_{\beta_0}(\beta \phi) \) is real-analytic in \((1, \infty)\), and \( \beta \phi \) remains positive recurrent in that interval.

**Proof.** Fix some \( a \in S \). We show that \( \Delta_a[\phi] = \infty \). Assume by way of contradiction that \( \Delta_a[\phi] < \infty \). It is easy to verify that if we induce on \( a \), then

\[
\phi + p + t|a| = \phi + p + t,
\]

whence \( P_{\beta_0}(\phi + p + t|a|) = P_{\beta_0}(\phi + p) + t \), whence

\[
\Delta_a[\phi + t|a|] = \Delta_a[\phi] + t.
\]

If \( \Delta_a[\phi] < \infty \), then for some \( t \), \( \Delta_a[\phi + t|a|] < 0 \) whence by the discriminant theorem, \( \phi + t|a| \) is transient. However, \( P_{\beta_0}(\phi + t|a|) \leq P_{\beta_0}(\phi) + |t| < \infty \) and
V_t(\phi + t1_{[n]}) < \infty$, so by the BIP property and corollary 2 \( \phi + t1_{[n]} \) is positive recurrent and not transient.

This shows that every \( \phi \) which satisfies the conditions of the theorem is strongly positive recurrent. By corollary 1, every function with summable variations and finite pressure is cohomologous to a non-positive function plus a constant. Thus, if \( P_G(\phi) < \infty \), then \( P_G(\beta\phi) < \infty \) for all \( \beta > 1 \). Thus, for every \( \beta > 1 \), \( \beta\phi \) satisfies the conditions of the theorem, and by the above discussion, is strongly positive recurrent. The theorem now follows from the strong positive recurrence theorem. \( \square \)

Then next sections contain examples which do admit critical phenomena.

### 6.2 The renewal shift

#### 6.2.1 general results

The examples studied in [PM][Hof],[GW], [W1],[W2], [PS] and [Lo] share the same critical behavior: for some potential \( \phi \), the function \( \beta \mapsto P_G(\beta\phi) \) has one point of non-differentiability \( \beta_c \), and is constant for \( \beta > \beta_c \). A close look at these examples shows that they can be represented as different potentials on the same countable Markov shift, the **renewal shift**. This is the shift with set of states \( S := \mathbb{N} \cup \{0\} \) and transition matrix \( (t_{ij})_{S \times S} \) whose 1 entries are \( t_{oo} \), \( t_{0i} \) and \( t_{ii-1} \) (\( i = 1, 2, 3, \ldots \)). The main result of this section is

**Theorem 12** Let \( X \) be the renewal shift and let \( \phi : X \to \mathbb{R} \) be a function with summable variations such that \( \sup \phi < \infty \), and such that \( \phi \) is weakly Hölder, continuous. Then there exists \( 0 < \beta_c \leq \infty \) such that:

1. \( \beta\phi \) is strongly positive recurrent for \( 0 < \beta < \beta_c \) and transient for \( \beta > \beta_c \).
2. \( P_G(\beta\phi) \) is real analytic in \((0, \beta_c)\) and linear in \((\beta_c, \infty)\). It is continuous but not analytic at \( \beta_c \) (in case \( \beta_c < \infty \)).
3. Set \( A_n := e^{\sup \{ \phi(x) : x \in \mathbb{N} \cup \{0, 1, \ldots, n\} \}} \) and let \( R(\beta) \) be the radius of convergence of \( F_\beta(\xi) := \sum_{n=1}^{\infty} A_n^\beta \xi^n \). If \( F_\beta(R(\beta)) \) is infinite for every \( \beta \) then \( \beta_c = \infty \). If \( \exists \beta > 0 \) such that \( F_\beta(R(\beta)) < 1 \), then \( \beta_c < \infty \).

**Proof.** It is easy to check that \( X \) it topologically mixing. Also, \( \beta\phi \) has finite pressure for all \( \beta \geq 0 \), since \( P_G(\beta\phi) \leq \log \| L_{\beta\phi} \| \leq \log(2e^{\beta \sup \phi}) \). One can easily check that for every function \( f, n \in \mathbb{N} \) and \( \beta > 0 \)

\[
Z_n(\beta f, 0) = Z_n(f, 0)^\beta \quad \text{and} \quad p^n[\beta f] = \beta p^n[f].
\]

Henceforth \( (\mathbb{X}, \phi) \) denotes the induced system on \([0], \) \( P(\beta) := P_G(\beta\phi) \) and \( \Delta[\beta] := \Delta[\beta\phi] \).

If \( \nu[\phi] = \infty \) then \( \Delta[\beta] = \sup \{ P_G(\phi + p) : p < \infty \} = \infty \) because \( P_G(\phi + p) \geq \frac{P_G(\phi)}{\beta} + p \). In this case parts 1 and 2 follow with \( \beta_c = \infty \) from theorem 10 and the discussion following theorem 9. We therefore restrict ourselves to the case \( \nu[\phi] < \infty \).
Without loss of generality, assume that \( p^\delta_0[\phi] = 0 \) (else pass to \( \phi + p^\delta_0[\phi] \) and use (5.3)). By (6.1), \( p^\delta_0[\beta \phi] = 0 \) for all \( \beta > 0 \), whence by (5.1)

\[
\Delta[\beta] = P_G(\beta \phi)
\]

As before, if \( \Delta[\beta] > 0 \) for every \( \beta \), parts 1 and 2 follow with \( \beta_c = \infty \). Assume \( \exists \beta > 0 \) such that \( \Delta[\beta] \leq 0 \) and set \( \beta_c := \inf\{\beta > 0 : \Delta[\beta] \leq 0\} \). Note that \( \beta_c > 0 \), because according to (5.2) and (6.1)

\[
\Delta[\beta] \geq \log \sum_{n=1}^\infty \left( e^{n p^\delta_0[\phi]} Z_n^*(\phi,0) \right)^\beta - \beta \sum_{n=2}^\infty V_n(\phi) \xrightarrow{\beta \to \infty} +\infty.
\]

We claim that \( \Delta[\beta] \to -\infty \) as \( \beta \uparrow +\infty \). Fix some \( \beta_0 \) such that \( \Delta[\beta_0] \leq 0 \). By (6.2), \( \phi \) has finite pressure, whence by corollary 1, \( \phi \) is cohomologous to \( \psi + P_G(\beta_0 \phi) \) where \( \psi \) is weakly H"older continuous (in \( \overline{X} \)) such that \( L_\psi^{-1} \leq 1 \). Since \( \phi \) has summable variations, \( V_1(\overline{\phi}) < \infty \). It follows from corollary 1 that \( V_1(\psi) < \infty \) as well. By (6.2), for all \( t > 1 \),

\[
\Delta[t \beta_0] = P_G(t \overline{\phi}) = P_G(t \overline{\psi}) + t P_G(\beta_0 \phi)
\]

Since \( P_G(\beta_0 \phi) = \Delta[\beta_0] \leq 0 \), we have for all \( t > 1 \),

\[
\Delta[t \beta_0] \leq P_G(t \overline{\psi}) \leq \log ||L_\overline{\psi}||_\infty
\]

By construction, \( L_\overline{\psi}^{-1} \leq 1 \). Therefore, since every \( x \in \overline{X} \) has more than one pre-image, \( \overline{\psi} \) is strictly negative. It follows from this and \( V_1(\overline{\psi}) < \infty \) that \( ||L_\overline{\psi}||_\infty \to 0 \) as \( t \uparrow \infty \). This implies that \( \Delta[\beta] \to -\infty \) as \( \beta \uparrow \infty \).

We show that \( \Delta[\beta] < 0 \) in \( (\beta_c, +\infty) \). By the definition of \( \beta_c \), there are \( \beta_n \downarrow \beta_c \) such that \( \Delta[\beta_n] \leq 0 \). By what we just showed there are \( \beta'_n \uparrow \infty \) such that \( \Delta[\beta'_n] < 0 \). By (6.2) \( \Delta[\beta] = P_G(\beta \overline{\phi}) \), so \( \Delta[\beta] \) is convex in \( (\beta_n, \beta'_n) \). By convexity, \( \Delta[\beta] < 0 \) in \( (\beta_n, \beta'_n) \). Since \( \beta_n \downarrow \beta_c \) and \( \beta'_n \uparrow \infty \), \( \Delta[\beta] \) is strictly negative in \( (\beta_c, +\infty) \).

We have shown that \( \Delta[\beta] < 0 \) in \( (\beta_c, \infty) \). It is obvious that \( \Delta[\beta] > 0 \) in \( (0, \beta_c) \). Part 1 now follows from the discriminant theorem.

We prove part 2. The analyticity of \( P(\beta) \) in \( (0, \beta_c) \) follows from theorem 10 and that fact that \( P(\beta) < \infty \). The discriminant theorem and (6.1) imply that \( \forall \beta > \beta_c, P_G(\beta \phi) = p^\delta_0[\beta \phi] = \beta p^\delta_0[\phi] \) and \( \forall \beta \in (0, \beta_c), P_G(\beta \phi) > \beta p^\delta_0[\phi] \). Thus \( P_G(\beta \phi) \) is linear in \( (\beta_c, \infty) \), but not in \( (0, \beta_c) \). This implies that \( \beta_c \) is a point of non-analyticity. The continuity of \( P(\beta) \) in \( \beta_c \) follows from the convexity of this function.

To prove part 3, recall that \( \Delta[\beta] > 0 \) for \( \beta > 0 \) small, and note by (6.2) that

\[
\log F_\beta(R(\beta)) - \beta \sum_{n=2}^\infty V_n(\phi) \leq \Delta[\beta] \leq \log F_\beta(R(\beta)).
\]

**Example 2** \( \beta_c \phi \) can be positive recurrent, null recurrent or transient. There is a positive recurrent potential which is not strongly positive recurrent.

**Proof.** Let \( \{f_n\}_{n \geq 1} \) be a sequence, such that \( f_n > 0 \) and \( \log f_n = o(n) \). Set \( \phi := \sum_{n \geq 1} f_n \). Then, \( Z_n(\phi,0) = f_n, p^\delta_0[\phi] = 0 \) and \( \sum_{n \geq 2} V_n(\phi) = 0 \), whence by (5.1), \( \Delta[\beta \phi] = \log \sum_{n \geq 1} f_n^\beta \). Let \( \zeta(s) := \sum_{n \geq 1} n^{-s} \).
1. **Positive recurrence.** Set $f_n := \frac{1}{\zeta(3)^n}$. Then $\Delta_0[\beta_\phi] = \log[\zeta(3)/\zeta(3)]$ whence $\beta_\phi = 1$. Note that $\Delta_0[\beta_\phi] = 0$ whence $P_G(\beta_\phi) = -p_G[\beta_\phi] = 0$. It also follows that $\beta_\phi$ is recurrent. Positive recurrence follows from $\sum_{n \geq 1} n^{n \log(n)^2/2} = 0$. Similar calculations show that $\Delta_0[\beta_\phi] = \log 2$ for $\beta_\phi > 1$. Thus, $\beta_\phi = 1$ and $\beta_\phi$ is transient. \qed

2. **Null recurrence.** The same calculations with $f_n = 1/(\zeta(3)^n_\phi)$. Theorem 12, consider the Pomeau-Manneville map $T : [0,1] \to [0,1]$ given by $T(x) = x + x^{1+a}(\mod 1)$ where the value of $T$ at its discontinuity is 0, $T(1) = 1$ and $s > 0$ [PM]. The following theorem generalizes results which are known for $f = -\log[T']$ (see [PM] and [Lo]) to other potentials, whose equilibrium measure is not necessarily equivalent to Lebesgue’s measure.

**Theorem 13 (The Pomeau-Manneville Model).** Let $T$ be the Pomeau-Manneville map and let $f : [0,1] \to \mathbb{R}$ be $C[0,1] \cap C^1(0,1)$ such that $f'(x) \sim c_\alpha x^\alpha$ as $x \to 0$, where $c \neq 0$ and $\alpha > 0$. Set

$$P(\beta) := \sup \left\{ b_m(T) + \beta \int f dm : m \in \mathcal{P}(0,1) \right\}$$

1. There exists $0 < \beta_\phi \leq \infty$ such that $P(\beta)$ is real analytic in $(0, \beta_\phi)$ and linear in $(\beta_\phi, \infty)$. It is continuous but not real-analytic at $\beta_\phi$.

2. $\beta_\phi$ is finite if and only if $\alpha \leq s$ and $c < 0$. In particular, it is finite for $f := -\log[T']$.

**Proof.** It is common knowledge that $T$ can be described symbolically as a renewal shift. We check that the symbolic representation of $f$ has summable variations and apply theorem 12. To do this we recall some facts on the natural Markov partition of $T$ (see [I], lemma 4.8.6 in [A1] and [T1]).

Define by induction $c_0 := 1$ and $c_n = c_{n+1} + c_n + s$. Rewriting this as $c_n = c_n^{*} + c_n^{*} + c_{n+1}^{*} + c_{n+1}^{*}$ we see that $c_n^{*} - c_n^{*} \sim s$; whence $c_n \sim (sn)^{1/s}$. It follows from the recursive relation which defines $\{c_n\}$ that

$$c_n - c_{n+1} \sim \frac{1}{(sn)^{1+1/s}}$$

Set $I[a_n] := (c_{n+1}, c_n]$ and $I[a_n, a_{n+1}] := \bigcap_{k=0}^{n-1} T^k I[a_k]$. One checks that $TI[0] = (0,1]$ and $TI[n+1] = I[n]$, whence $I[a_n, a_{n+1}]$ is not empty iff $(a_n, a_{n+1})$ is an admissible word of the renewal shift.
6.2. THE RENEWAL SHIFT

Claim 1. The diameter of $I[a_0, \ldots, a_{n-1}]$ satisfies for every $\varepsilon > 0$

\begin{equation}
|I[a_0, \ldots, a_{n-1}]| = O \left( \frac{1}{n^{1+1/s} \varepsilon} \right)
\end{equation}

Proof. By the previous discussion,

$$I[a_0, \ldots, a_{n-1}] = I[a_0, \ldots, a_{n-1}; a_{n-1} - 1, \ldots, 0]$$

so we may assume that $a_{n-1} = 0$. Set $M := 1 + \sup \{|a_k| \}$ and $N := |\{k : a_k = 0\}|$. Since $(a_0, \ldots, a_n)$ is admissible with respect to the transition matrix of the renewal shift, $MN \geq n$. Thus, for every $\beta \in (0, 1)$ either $M \geq n^\beta$ or $N \geq n^{1/\beta}$.

Set $m_n := [n^\beta] + 1$. If $M \geq n^\beta$ then for some power $k$, $T^n I[a_0, \ldots, a_{n-1}] \subseteq I[m_n]$ whence since $T^n \geq 1$ $|I[a_0, \ldots, a_{n-1}]| \leq |I[m_n]| = c_{m_n} - c_{m_n+1} = O(n^{-\beta(1+1/s)})$. If $N \geq n^{1/\beta}$ then for every $x \in I[a_0, \ldots, a_{n-1}]$

$$\left( T^n f(x) \right) = \prod_{i=0}^{n-1} T'(T^i x) \geq \left( \inf_{x \in I[a]} T'(x) \right)^N$$

whence for $\theta := 1/\inf_{x \in I[a]} T'(x)$

\begin{equation}
|I[a_0, \ldots, a_{n-1}]| \leq \theta^N
\end{equation}

Since $\theta < 1$ and $N \geq n^{1/\beta}$ we have a gain that $I[a_0, \ldots, a_{n-1}] = O(n^{-\beta(1+1/s)})$. Since $\beta \in (0, 1)$ was arbitrary, the claim is proved.

Let $(X, \sigma)$ be the renewal shift and $\pi_0 : X \rightarrow [0, 1]$ be the map defined by the equation $\{\pi_0(x)\} = \bigcap_{n \geq 0} [x_0, \ldots, x_{n-1}] = \bigcap_{n \geq 0} I[x_0, \ldots, x_{n-1}]$. By (6.3) $\pi_0$ is well defined. It is easy to check that $\pi_0 \circ \sigma = T \circ \pi_0$, that $\pi_0$ is 1-1 and that $\pi_0(X) = [0, 1] \setminus \bigcup_{n \geq 0} T^{-n} [0, 1]$.

Claim 2. Let $f$ be $C[0, 1] \cap C^1(0, 1)$ in $[0, 1]$ such that $f'(x) \sim c x^{-\alpha}$ as $x \downarrow 0$, where $c \neq 0$ and $\alpha > 0$. Then $\phi := f \circ \pi_0$ has summable variations and $\phi$, the induced potential on $[0, 1]$, is weakly Hölder continuous.

Proof. Fix $x, y \in [a_0, \ldots, a_{n-1}]$ where without loss of generality $a_{n-1} = 0$. Fix $\varepsilon > 0$ (to be determined later). Then there exists $\xi \in I[a_0, \ldots, a_{n-1}]$ such that

$$|\phi(x) - \phi(y)| = |f'(\xi)| \cdot |I[a_0, \ldots, a_{n-1}]| = O \left( \frac{\xi^{\alpha}}{n^{1+1/s} \varepsilon} \right)$$

Since $\xi \in I[a_0, \ldots, a_{n-1}] \subseteq (c_{a_0+1}, c_{a_0})$, and since by the structure of the renewal shift $a_0 \leq n - 1$, we have that $\xi^{\alpha} = O(1 + c_{a_0})$ whence

$$V_n(\phi) = O \left( \frac{1 + n^{1+1/s} \varepsilon}{n^{1+1/s} \varepsilon} \right)$$
If $\alpha \geq 1$, the nominator is bounded, and choosing $\varepsilon < 1/(2s)$ we see that $V_n(\phi)$ is summable. If $\alpha < 1$ then the nominator is $O(n^{1-\alpha/s})$ and we have that $V_n(\phi) = O\left(n^{1+\alpha/s} \varepsilon \right)$. Choosing $\varepsilon < \alpha/s$ we see the $\sum V_n(\phi)$ is summable. In any case, $\phi$ has summable variations. The weak Hölder continuity of $\phi$ can be proved in a similar way.

**Claim 3.** $P(\beta) = P_G(\beta \phi)$ where $\phi := f \circ \pi_0$.

**Proof.** Since $\sup f < \infty$ and $\forall x \in T \cdot 1^{|x|} = 2$, $\|L_{\beta \phi} 1\|_{\infty} < \infty$. It follows as in ([S2], theorem 3) that

$$P_G(\beta \phi) = \sup \{ h_\mu(\sigma) + \beta \int \phi d\mu : \mu \in \mathcal{P}_\sigma(X) ; \int \phi d\mu < \infty \}$$

(the argument there works also for functions with summable variations).

The claim follows because $m \leftrightarrow m \circ \pi_0$ is a 1-1 onto correspondence between the sets of measures which define $P(\beta)$ and $P_G(\beta \phi)$.

Claims 2 and 3 show that we can apply theorem 12 to $\phi = f \circ \pi_0$ and deduce the existence of $\beta_\varepsilon$. We check the conditions for the finiteness of $\beta_\varepsilon$. Let $A_n$ be as in theorem 12. Since $\phi$ has summable variations, $A_n \exp f_n(d_n)$ where $d_n \in I[0]$ are defined by $T(d_n) = c_n$. It is easy to check that $d_n \downarrow c_1$, whence $A_n \exp \sum_{k=1}^n f(c_k)$. Without loss of generality, $f(0) = 0$ (addition of constants does not affect the finiteness of $\beta_\varepsilon$). Then by the assumptions on $f$, $f(x) \sim cx^\alpha$. Since $c_k \sim (sk)^{1/s}$, $f(c_k) \sim c(sk)^{\alpha/s}$. Thus $\sum_{k=1}^n f(c_k) \asymp c \int_1^n x^{\alpha/s} dx$. It follows that there exist constants $K_1, K_2, K_3, K_4$, such that

$$K_1 \exp \left( K_2 c \int_1^n \frac{1}{x^{\alpha/s}} dx \right) \leq A_n^{\beta} \leq K_3 \exp \left( K_4 c \int_1^n \frac{1}{x^{\alpha/s}} dx \right)$$

Let $F_{\beta}(\xi)$ and $R(\beta)$ be as in theorem 12. Using the above,

1. If $\alpha > s$, then $A_n^{\beta} \asymp 1$ for every $\beta > 0$. In this case $F_{\beta}(R(\beta)) = \infty$ for every $\beta$, so $\beta_\varepsilon = \infty$.

2. If $\alpha = s$, then $K_1 n^{K_2 c} \leq A_n^{\beta} \leq K_3 n^{K_4 c}$. It follows that $R(\beta) = 1$ and that $F(R(\beta))$ is infinite for every $\beta$ if $c > 0$, and $F(R(\beta)) \rightarrow 0$ if $c < 0$.

   Thus for $\alpha = s$, if $c > 0$ then $\beta_\varepsilon$ is infinite, and if $c < 0$ then $\beta_\varepsilon < \infty$.

3. If $\alpha \in (0, s)$ and $\alpha := 1 - (\alpha/s)$, then for some constants $C_1, C_2, C_3, C_4$, $C_1 e^{C_2 c} \leq A_n^{\beta} \leq C_3 e^{C_4 c}$. Since $\alpha < 1$, $R(\beta) = 1$ for every $\beta$. It follows that if $c > 0$ then $F(R(\beta)) = \infty$ for every $\beta$, and if $c < 0$ then $F(R(\beta)) \rightarrow 0$. Thus for $0 < \alpha < s$, if $c > 0$ then $\beta_\varepsilon$ is infinite and if $c < 0$ then $\beta_\varepsilon$ is finite.

Thus $\beta_\varepsilon < \infty$ if and only if $0 < \alpha \leq s$ and $c < 0$. 

$\square$
6.3 Pathological examples

In this section we construct examples whose critical behavior is different than that of potentials on the renewal shift.

6.3.1 Tool for constructing examples

We say that a one-parameter family of functions $F_\beta(\xi)$ is an exponent power series if it is of the form $F_\beta(\xi) = \sum_{n,k \geq 0} a_{nk} \xi^n$, where $a_{nk} \geq 0$. Clearly, if $F_\beta$ and $G_\beta$ are exponent power series, then so are $F_\beta G_\beta$, $F_\beta \circ G_\beta$ and $c_1 F_\beta + c_2 G_\beta$ where $c_1, c_2$ are positive integers. We say that an exponent power series $F_\beta$ is aperiodic if the power expansion of $F_\beta$ contains two co-prime powers of $\xi$. We say that $F_\beta$ is adequate if it is of the form $c^\beta \xi + \xi^2 G_\beta(\xi)$ where $c \geq 0$ and $G_\beta$ is an exponent power series. Recall that a countable Markov shift is called irreducible, if $\forall a, b \in S$ there is for some $n$ an admissible word $w_n$ such that $|w_n| = n$, $w_0 = a$ and $w_{n-1} = b$.

**Theorem 14** For every adequate exponent power series $F_\beta$ there exists an irreducible countable Markov shift $X$ and a Markov potential $\phi = \phi(x_0, x_1)$ such that for all $\beta$, $P_G(\beta \phi + p) = \log F_\beta(e^p)$. If $F_\beta$ is aperiodic, $X$ is topologically mixing.

**Proof.** Write

$$F_\beta(\xi) = c^\beta \xi + \sum_{n=2}^{\infty} \xi^n \sum_{k=1}^{N_n} a_{nk}^\beta$$

where $0 \leq N_n \leq \infty$. Let $S$ be a countable set indexed in the following way

$$S := \{a\} \cup \bigcup_{n=2}^{\infty} \bigcup_{k=1}^{N_n} \{b_{nk}(1), \ldots, b_{nk}(n-1)\}.$$ 

Let $(t_{ij})_{i,j \in S}$ be the transition matrix whose non zero entries are exactly $t_{a,b_{nk}(1)}$, $t_{b_{nk}(i) b_{nk}(i+1)}$, $t_{b_{nk}(n-1) a}$ for all $n, k \geq 1$ and $i = 1, \ldots, n-1$ with the addition of $t_{aa}$ if and only if $c \neq 0$. Let $X$ be the corresponding topological Markov shift. Define $\phi(x)$ by $\phi(x) := \log a_{nk}$ if $x \in [a, b_{nk}(1)]$, $\phi(x) := \log c$ if $x \in [a, a]$ and $\phi(x) := 0$ otherwise. One checks that

$$F_\beta(\xi) = \sum_{n=1}^{\infty} \xi^n Z_n^*(\beta \phi, a),$$

whence by (5.4) and the fact that $\forall k \geq 2 V_k(\phi) = 0$, $P_G(\beta \phi + p) = \log F_\beta(e^p)$. Note that $X$ is irreducible, because all states connect to $a$ and $a$ connects to all states. It is topologically mixing iff there are two words of co-prime lengths which connect $a$ to $a$. This can be easily seen to be equivalent to the aperiodicity of $\sum \xi^n Z_n^*(\beta \phi, a)$, hence to that of $F_\beta$. \qed
6.3.2 Infinite number of changes in recurrence

The following example shows that \( \{ \beta \phi \}_{\beta > 0} \) can change from recurrent to transient an infinite number of times.

**Example 3** There exists \( X \) topologically mixing and \( \phi = \phi(x_0, x_1) \) such that for some \( \beta_n \downarrow 0 \), \( \beta \phi \) is recurrent in \( (\beta_{i+1}, \beta_i) \) for \( i \) even and transient for \( i \) odd.

**Proof.** Consider the following sequence of numbers

\[
N_n := \frac{2}{n} \left( \frac{2n - 2}{n - 1} \right)
\]

A calculation with Stirling’s formula shows that \( N_n \sim \pi^{-1/2} n^{3/2} 2^{2n-1} \). Another calculation shows that

\[
4^n N_n = \frac{1}{4^{n-1}} \left( \frac{2n - 2}{n - 1} \right) - 1 - \frac{1}{4^n} \left( \frac{2n}{n} \right)
\]

Multiplying both sides of (6.5) by \( 4^n \) we see that \( N_n \) are all natural numbers. Summing both sides of (6.5) over \( n \) we see that \( \sum_{n \geq 2} N_n 4^n = \frac{1}{2} \).

Fix some \( \beta_n \downarrow 0 \) with the property that \( \sum \theta^{1/\beta_n} < \infty \) for all \( \theta \in (0, 1) \) (e.g. \( \beta_n := 1/n \)). Set \( \alpha_n(\beta) := -2(\theta)^{\beta/\beta_n} \) and \( p(\beta) := \prod_{n \geq 1} (1 + \alpha_n(\beta)) \). Then for all \( \beta > 0 \), \( p(\beta) \) is well-defined, non zero for \( \beta \not\in \{ \beta_n \}_n \) and satisfies

\[
p(\beta) = 1 + \alpha_1(\beta) + [1 + \alpha_1(\beta)] \alpha_2(\beta) + \ldots
\]

where the convergence on the right is absolute. Collecting summands with the same sign write \( p(\beta) = A(\beta) - B(\beta) \) where \( A(\beta) = \sum a_n^\beta \) and \( B(\beta) = \sum b_n^\beta \) for some \( a_n, b_n \geq 0 \). If \( \beta \in (\beta_{i+1}, \beta_i) \) then \( \beta > \beta_n \) iff \( n \geq i+1 \) whence

\[
sgn (A(\beta) - B(\beta)) = sgn \left( \prod_{n=1}^i (1 - 2^{1/\beta_n}) \prod_{n=i+1}^\infty (1 - 2^{1/\beta_n}) \right) = (-1)^i
\]

Thus, \( A(\beta) > B(\beta) \) iff \( i \) is even.

Now construct \( X \) and \( \phi = \phi(x_0, x_1) \) such that \( P(\beta \phi + p) = \log F(\phi) \) where \( F(\phi) \) is the exponent power series

\[
F(\phi) = 8 A(\beta) B(\beta) \xi^2 + \sum_{n=2}^\infty N_n B(\beta)^n \xi^n
\]

Since \( N_n \sim \frac{4^n}{n^{1/2}} \) the radius of convergence of \( F(\phi) \) is \( R(\beta) = 1 / 4 B(\beta) \). Thus, by (5.2), and since \( \phi \) is a Markov potential,

\[
\Delta_n[\beta \phi] = \log F(\beta) = \log \left( \frac{8 A(\beta) B(\beta)}{16 B(\beta)^2} + \sum_{n=2}^\infty N_n 4^n \right)
\]

whence \( \Delta_n[\beta \phi] = \log \frac{1}{4} (1 + A(\beta) / B(\beta)) \). This is positive iff \( A(\beta) > B(\beta) \). Thus \( \beta \phi \) is recurrent for \( \beta \in (\beta_{i+1}, \beta_i) \) and \( i \) even, and transient for \( \beta \in (\beta_{i+1}, \beta_i) \) and \( i \) odd. □
6.3. PATHOLOGICAL EXAMPLES

6.3.3 Null recurrence in an interval

We have seen that for potentials \( \phi \) on the renewal shift, \( \beta \phi \) can be null recurrence for at most one value of \( \beta \) (the critical point). Our next example shows that for other topological Markov shifts null recurrence can occur in an entire interval. A trivial example would be a Markov shift for which the potential \( \phi \equiv 0 \) is null recurrent. We therefore restrict ourselves to examples where \( \phi \) is not cohomologous to a constant.

**Example 4** There exist a topologically mixing countable Markov shift \( X \) and a function \( \phi = \phi(x_0, x_1) \) not cohomologous to a constant such that \( \beta \phi \) is null recurrent for every \( \beta \).

**Proof.** Let \( N_n \) be as in example 3 and set \( f_\beta(p) := 2^\beta(e^p + e^{2p}) \). Construct \( X \) topologically mixing and \( \phi = \phi(x_0, x_1) \) such that

\[
P_\beta(\beta \phi + p) = \log \left( 2 \sum_{n=2}^{\infty} N_n f_\beta(p)^n \right)
\]

Since \( N_n \sim 4^n n^{-3/2} \), \( p^*_n[\beta \phi] \) is determined by the equation \( f_\beta(p^*_n[\beta \phi]) = 1/4 \). It follows from this that \( \Delta_n[\beta \phi] = 0 \). By the Discriminant theorem, for all \( \beta \) \( \beta \phi \) is recurrent and \( P_\beta(\beta \phi) = -p^*_n[\beta \phi] \). It also follows that \( \phi \) is not a linear function of \( \beta \), being given by the equation \( f_\beta[-P_\beta(\beta \phi)] = 1/4 \).

We show that \( \beta \phi \) is null recurrent for all \( \beta \). Since \( V_2(\phi) = 0 \), \( P_\beta(\beta \phi + p) = \log \sum_{n \geq 1} e^{n \beta} Z^*_n(\beta \phi, a) \) whence

\[
\sum_{n=1}^{\infty} e^{n \beta} Z^*_n(\beta \phi, a) = \frac{d}{dp} \bigg|_{p=p^*_n} e^{P_\beta(\beta \phi + p)} = 2 f_\beta(p^*_n[\beta \phi]) \sum_{n=2}^{\infty} n N_n 4 \quad (n \geq 1)
\]

and this diverges, because \( N_n \sim 4^n n^{-3/2} \). \( \square \)

6.3.4 Coexistence of all modes of recurrence

The following example shows that all modes of recurrence can co-exist for interval ranges of inverse temperatures.

**Example 5** There exist \( X \) topologically mixing and \( \phi = \phi(x_0, x_1) \) such that for some \( 1 < \beta_1 < \beta_2 < \infty \), \( \beta \phi \) is null recurrent for \( \beta \in (1, \beta_1) \), positive recurrent for \( \beta \in (\beta_1, \beta_2) \) and transient for \( \beta \in (\beta_2, \infty) \).

**Proof.** Fix some positive \( a_n \sim 1/[2n(\log n)^2] \) such that \( a_1 = \frac{2}{\pi} \), \( \sum_{n=1}^{\infty} a_n = 1 \) and set \( A(\beta) = \sum_{n \geq 1} a_n^\beta, F(\beta) = \sum_{n \geq 1} a_n^\beta A(\beta)^n \xi^{n+1} \) and

\[
G(\beta) = F(\beta) F(2\beta)
\]

This is an adequate aperiodic exponent power series. Let \( X \) and \( \phi \) be the corresponding shift and potential.
Let \( R_F(\beta) \) and \( R_G(\beta) \) denote the radii of convergence of \( F_\beta(\xi) \) and \( G_\beta(\xi) \). Note that \( R_F(\beta) = 1/A(\beta) \) and \( F_\beta(R_F(\beta)) = 1 \). Let \( \beta_2 \) be the solution of \( R_F(\beta_2) = 2 \). Clearly, \( R_F(\beta) < 2 \) for \( \beta < \beta_2 \) and \( R_F(\beta) > 2 \) for \( \beta > \beta_2 \). Thus,

1. If \( \beta \in (1, \beta_2) \), then \( 2F_\beta(R_F(\beta)) = 2 > R_F(\beta) \) so \( R_G(\beta) = F_\beta(1/2R_F(\beta)) \).
   In this case \( G_\beta(R_G(\beta)) = F_\beta(2 \cdot R_F(\beta)) = F_\beta(R_F(\beta)) = 1 \).

2. If \( \beta > \beta_2 \), then \( 2F_\beta(R_F(\beta)) = 2 < R_F(\beta) \) so \( R_G(\beta) = R_F(\beta) \). In this case \( G_\beta(2F_\beta(R_F(\beta))) = F_\beta(R_F(\beta)) = 1 \).

Since \( \Delta_\alpha[\phi] = \log G_\beta(R_G(\beta)) \), \( \beta \phi \) is transient for \( \beta > \beta_2 \) and recurrent for \( \beta \in (1, \beta_2) \).

We check positive recurrence and null recurrence for \( \beta \in (1, \beta_2) \). Fix some \( \beta \in (1, \beta_2) \). Since \( G_\beta(R_G(\beta)) = 1 \), \( \Delta_\alpha[\phi] = 0 \) whence \( e^{P_\alpha(\beta \phi)} = e^{\sum_{n \geq 1} n^\alpha \phi} = R_G(\beta) \). Thus \( \sum_{n \geq 1} n^\alpha \phi \) is finite and \( R_G(\beta) = F_\beta \frac{1}{2} R_F(\beta) \).

Since \( R_G(\beta) < R_F(\beta) \), this is finite iff \( F_\beta(R_F(\beta)) < \infty \), which is comparable to

\[
\frac{1}{A(\beta)} \sum_{n=2}^{\infty} \frac{n}{2^\beta n^\beta (\log n)^{2\beta}}
\]

This sum is infinite for \( \beta \in (1, 2) \) and finite for \( \beta \in (2, \beta_2) \). (Note that \( \beta_2 > 2 \) since \( a_1^\beta > 1 \) whereas \( a_2^\beta < A(\beta_2) = 1/e \).)

\[\square\]

6.3.5 Positive measure set of critical points

**Example 6** There exists a topologically mixing countable Markov shift \( X \) and a potential \( \phi : X \to \mathbb{R} \) such that \( \phi(x) = \phi(x_0, x_1) \) and such that the following set has positive Lebesgue measure:

\[ \{ \beta > 0 : R_G(t \phi) \text{ is not real analytic in a neighbourhood of } t = \beta \} \]

**Proof.** We construct an exponent power series \( F_\beta(\xi) \) with radius of convergence \( R(\beta) \) with the following properties, where \( m \) denotes the Lebesgue measure:

\[ m\{ \beta : R(\beta) \text{ is not analytic in a neighbourhood of } \beta \} > 0 \]

\[ F_\beta(R(\beta)) < 1 \]

Let \( \{ I_n \}_{n=1}^{\infty} \) be disjoint open intervals in \([1, 2]\) such that \( \bigcup_{n \geq 1} I_n \) is dense in \([1, 2]\) but \( m(\bigcup_{n \geq 1} \partial I_n) < 1 \). Then \( \bigcup_{n \geq 1} \partial I_n \) has positive Lebesgue measure. Set
6.3. PATHOLOGICAL EXAMPLES

\( I_n = (a_n, b_n), \quad N_n = 2^n \quad \text{and} \quad p_n (\beta) := -2N_n^{3\beta} \left( N_n^{\beta/a_n} - N_n \right) \left( N_n^{\beta/b_n} - N_n \right). \)

Expanding and collecting all terms with the same sign write

\[ p_n (\beta) = A_n (\beta) - B_n (\beta) \]

where \( A_n (\beta) = 2N_n \left( N_n^{\beta (\varepsilon_n^{-1})} + N_n^{\beta (\varepsilon_n^{-1})} \right) \) and \( B_n (\beta) = 2N_n^{\beta (\varepsilon_n^{-1} + \varepsilon_n^{-1})} + 2N_n^{3\beta} \). Set

\[ B (\beta) := \sum_{n=1}^{\infty} B_n (\beta) \quad \text{and} \quad A^{(n)} (\beta) := A_n (\beta) + B (\beta) - B_n (\beta). \]

Both \( \frac{1}{2} B (\beta) \) and \( \frac{1}{2} A^{(n)} (\beta) \) are convergent exponent sums for \( \beta > \frac{2}{3} \). Also,

\[ p_n (\beta) = A_n (\beta) - B_n (\beta) = A^{(n)} (\beta) - B (\beta). \]

Thus \( A^{(n)} (\beta) > B (\beta) \) if and only if \( \beta \in I_n \), whence

\[ M (\beta) := \max \left\{ B (\beta), A^{(1)} (\beta), A^{(2)} (\beta), \ldots \right\} \]

is given by

\[ M (\beta) = \begin{cases} A^{(n)} (\beta) & \text{if } \beta \in I_n \\ B (\beta) & \text{if } \beta \notin \bigcup_{n \geq 1} I_n \end{cases} \]

(6.8)

Now set \( c_n := \lfloor 2^n/n^3 \rfloor \) and

\[ F_\beta (\xi) := \sum_{n=2}^{\infty} c_n \left( \left( \frac{1}{2} B (\beta) \xi \right)^n + \sum_{j=1}^{n} \left( \frac{1}{2} A^{(j)} (\beta) \xi \right)^n \right) \]

By the preceding discussion, this is an exponent power series. Let \( R (\beta) \) denote its radius of convergence. Then \( R (\beta) = 1/M (\beta) \) where \( M (\beta) \) is given by (6.8). \( M (\beta) \) is not analytic in the neighborhood of each of the points in \( \bigcup_{n \geq 1} \partial I_n \), because \( \bigcup_{n \geq 1} \partial I_n \) are points of non differentiability. Thus (6.6) is satisfied. Also,

\[ F_\beta (R (\beta)) \leq \sum_{n=2}^{\infty} \frac{1}{n^3} \frac{1}{M (\beta)^n} \left( B (\beta)^n + \sum_{j=1}^{n} A^{(j)} (\beta)^n \right) \leq \sum_{n=2}^{\infty} \frac{1}{n^2} \]

and (6.7) follows as well. Let \((X, \phi)\) be the corresponding system. Then \( \beta \phi \) is always transient, whence \( P_G (\beta \phi) = -p_0 [\beta \phi] = -\log R (\beta) \) and this function has a positive Lebesgue measure set of critical points.

\[ \square \]

Remark. The set of points for which the pressure function constructed above is not differentiable is countable (it contains each of the points \( a_n \) and \( b_n \)). This cannot be improved, since the pressure function is by definition convex in \( \beta \).
6.4 Notes

The non-existence of phase transitions for systems with BIP is a generalization of Ruelle’s lack of phase transitions result for finite Markov shifts [Ru2], since every finite Markov shift has the BIP property. Theorem 12 generalizes results on the existence of phase transition for specific examples which were considered in [PM], [Lo], [W1, W2] and [PS]. Theorem 13 is a generalization of results in [PM], [Lo] and [LSV] which were proved for the special case \( f = -\log T' \) (when the conformal measure is Lebesgue’s measure) to the case when the conformal measure is singular. Theorem 14 and the examples of section 6.3 are new. The examples in sections 6.3.2 and 6.3.5 answer a question of Gurevic and Savchenko about the possibility of an infinite number of critical points (private communication).
Bibliography


BIBLIOGRAPHY


Index

admissible word, 4
big images property, 51
BIP property, 40
and Gibbs measures, 41
and spectral gap, 44
lack of phase transitions, 63
conditional expectation
and Ruelle’s operator, 32
conformal measures, 7
and Ruelle’s operator, 32
existence, 26
corollary 1 (cohomology to non-positive functions + constant), 25
corollary 2 (Gibbs measures), 43
corollary 3 ($\Delta_a[\theta] > 0$ is independent of $a$), 54
cylinders, 4
definition 1 (top. Markov shift), 4
definition 2 (Gurevic pressure), 18
definition 3 (recurrence), 24
definition 4 (the discriminant), 52
discriminant
definition, 52
discriminant theorem, 53
formulas for, 52
infinite changes of sign, 70
strict negativity of, 73
strict positivity of, 54
entropy
for infinite measures, 45
topological, 22
equilibrium measure, 7
exponent power series, 69
exponential decay of correlations, 12
Gibbs measures
and BIP property, 41
definition, 40
Gibbs property, 7
Gurevic pressure
and finite subshifts, 19
definition, 18
variational principle, 21
inducing, 45
for topological Markov shifts, 45
induced potential, 45
induced system, 46
Kac formula, 47
information function, 32, 49
irreducibility, 69
lemma 1, 31
lemma 2, 33
lemma 3, 46
lemma 4, 46
lemma 5, 46
lemma 6, 59
lemma 7, 59
lemma 8, 60
length, 4
Markov potential, 46
discriminant for, 53
negative function
cohomology to, 25
non-singular, 31
one-sided integrability, 49

80
INDEX

partition sets, 5
phase transition, 8
pointwise dual ergodicity, 32, 34
return sequence, 35
Pomeau-Manneville map, 66
pressure
  analyticity of, 54
  for induced potentials, 46
Gurevic, 11
topological, 7
proposition 1 (existence of the Gurevic pressure), 18
proposition 2 (properties of the Gurevic pressure), 19
proposition 3 (3 conformal measure \(\Rightarrow\) recurrence), 26
proposition 4 (recurrence \(\Rightarrow\) existence of conformal measure), 27
proposition 5 (recurrence \(\Rightarrow\) existence of eigenfunction), 33
proposition 6 (identification of \(a_n\)), 35
proposition 7 (null recurrence is equivalent to \(\int fd\nu = \infty\)), 36
proposition 8 (null recurrence implies \(a_n = o(1)\)), 37
proposition 9 \((P_\mu (\phi + p)\text{ and the spectral radius of }R_{\mu} [\phi (e^p)])\), 56
proposition 10 (properties of \(p \Rightarrow P_\mu (\phi + p)\)), 56
renewal shift
  critical behavior for, 64
definition, 64
Ruelle operator, 5
  and conditional expectation, 32
  and conformal measures, 7
  and phase transitions, 8
  and thermodynamic limits, 7
Ruelle’s Perron-Frobenius theorem
  countable Markov shifts, 25
  finite Markov shifts, 9
  properties of eigenfunction, 25
  rate of convergence, 44
Schweiger property, 32
states, 5
strong positive recurrence, 53
  and positive recurrence, 65
summable variations, 5
theorem 1 (equivalent definition for the pressure), 19
theorem 2 (Variational principle), 20
theorem 3 (generalized Ruelle’s Perron-Frobenius theorem), 25
theorem 4 (Gibbs measures exist iff BIP and \(V_1 < \infty\)), 41
theorem 5 (BIP \(\Rightarrow\) exponential decay of correlations), 44
theorem 6 \((h\alpha\nu \text{ maximizes } \int (P_\mu (\phi) - \phi) d\mu)\), 47
theorem 7 \((h\alpha\nu \text{ maximizes } \int (I_{\mu} (\phi' - P_\mu (\phi')) d\mu)\), 49
theorem 8 (discriminant theorem), 53
theorem 9 (strong positive recurrence theorem), 54
theorem 10 (stronger version of SPR theorem), 54
theorem 11 (BIP \(\Rightarrow\) no phase transitions), 63
theorem 12 (phase transitions for the renewal shift), 64
theorem 13 (phase transitions for the Pomeau-Manneville map), 66
theorem 14 (tool for constructing pathological examples), 69
thermodynamic formalism, 5
  countable Markov shifts, 9
  finite Markov shifts, 8
  thermodynamic limit, 7
tightness, 28
topological Markov shift, 5
countable, 5
finite, 5
topological mixing, 5
transfer operator, 26
transience, 11
transition matrix, 5

variational principle
  countable Markov shifts, 21
  finite Markov shifts, 8
  Gurevich pressure, 10
    infinite equilibrium measures, 47
variations, 5

weak Holder continuity, 5