CONTINUOUS PHASE TRANSITIONS FOR DYNAMICAL SYSTEMS

OMRI SARIG

Abstract. We study the asymptotic expansion of the topological pressure of one–parameter families of potentials at a point of non-analyticity. The singularity is related qualitatively and quantitatively to non–Gaussian limit laws and to slow decay of correlations with respect to the equilibrium measure.

1. Introduction

This paper deals with the thermodynamic formalism of countable Markov shifts. It explores the stochastic implications of non-analyticity for the topological pressure functional, by pursuing an analogy with the theory of continuous phase transitions.

Continuous phase transitions. A *continuous phase transition* (sometimes also called *high–order phase transition*) is a situation where a thermodynamic quantity varies continuously but not analytically when some external parameter of the system is changed. The prototypical example is ferromagnetic material at zero external magnetic field: The magnetic moment per unit volume ('magnetization') decreases continuously as the material is heated, until it completely vanishes at a certain critical temperature T_c; The derivative of the magnetization with respect to temperature ('susceptibility') diverges at T_c.

Systems undergoing a continuous phase transition develop local long–range order. This order can be described in terms of large fluctuations of thermodynamic quantities ('abnormal fluctuations'), and slow decay of correlations ('infinite correlation length'). See [**St**] for examples of continuous phase transitions, and [**BDFN**], [**Hi**] for theoretical treatment.

Most thermodynamic quantities can be expressed as partial derivatives of the Helmholtz or Gibbs Free Energy F. Therefore, a continuous phase transition is sometimes defined as a situation where the free energy is C^1 but not real–analytic. Physicists have found empirically that the free energy $F(t)$ satisfies an asymptotic power law close to the critical point: $F(t) \approx Ct^{\alpha}$+analytic terms for $t = (T - T_c)/T_c$ (the 'reduced temperature'). The parameter α is called a *critical exponent*.

It is not clear how to define $\approx$. In this work (as in [**Hi**]), we formalize $\approx$ by stipulating that $F(t) = \pm t^{\alpha} L(1/t) +$ analytic terms where $L : (c_0, \infty) \to \infty$ is a positive (Borel) function s.t.

$$\frac{L(st)}{L(t)} \xrightarrow[t\to\infty]{} 1 \text{ for all } s > 0. \tag{1}$$

In this case $L(t)$ is called *slowly varying* (s.v.) at infinity, and $t^{\alpha} L(1/t)$ is said to be *regularly varying* (r.v.) with index α, see appendix A.

This work was partially supported by NSF grant DMS–0400687. *Date:* February 15, 2006.

This interpretation of $\approx$ is not completely standard, but it is reasonable because it is equivalent to saying that the singular part $F^*(t)$ of $F(t)$ scales asymptotically like a power: $\frac{F^*(st)}{F^*(t)} \xrightarrow[t\to 0^+]{} s^\alpha$ for all $s>0$, compare with [**BDFN**].[1]

Continuous phase transitions in dynamical systems. Let $T : X \to X$ be some continuous map on a complete metric separable space X (in most examples treated below X is not locally compact). The dynamical counterpart to the free energy is the *topological pressure functional* $\phi \mapsto P_{\text{top}}(\phi)$ defined for continuous $\phi : X \to \mathbb{R}$ s.t. $\sup\phi < \infty$ by

$$P_{\text{top}}(\phi) := \sup\big\{h_\mu(T) + \int \phi d\mu : \mu \text{ is a Borel probability measure} \\ \text{s.t. } \mu\circ T^{-1} = \mu \text{ and } \int \phi d\mu \neq -\infty\big\}.$$

Here and throughout $h_\mu(T)$ is the metric entropy of μ.

The analogy [**Ru**] becomes apparent if one thinks of the metric entropy $h_\mu(T)$ as of entropy per particle, and of $\int \phi d\mu$ as $-\beta\times$energy per particle with $\beta =$inverse temperature (we describe an example in appendix B). With this interpretation, maximizing $h_\mu(T) + \int \phi d\mu$ amounts to minimizing the Helmholtz free energy. The maximizing measure μ (if it exists) is called the *equilibrium measure* of ϕ, and (if unique) is denoted by μ_ϕ.

Definition 1. *Let $T : X \to X$ be a continuous map of a complete metric separable space X, and $\phi_t : X \to \mathbb{R}$ a family of continuous functions, $t \geq 0$.*

1. $\{\phi_t\}_{t\geq 0}$ *is called* regular*, if $\exists\epsilon > 0$ s.t. ϕ_t has an equilibrium measure μ_t for $t \in [0, \epsilon)$.*
2. $\{\phi_t\}_{t\geq 0}$ *is said to undergo a* continuous phase transition *at 0^+, if it is regular, $\exists\epsilon$ s.t. $t \mapsto P_{top}(\phi_t)$ is C^1 on $[0, \epsilon)$, but $\nexists\epsilon > 0$ s.t. $t \mapsto P_{top}(\phi_t)$ extends to a real analytic function on $(-\epsilon, +\epsilon)$.*
3. $\{\phi_t\}_{t\geq 0}$ *is said to exhibit a* critical exponent α *as $t \to 0^+$ if $P_{top}(\phi_t) = \pm t^\alpha L(1/t) + h(t)$ with $h(t)$ analytic at zero, $L(x)$ s.v. at infinity and either $\alpha \notin \mathbb{N}$ or $\alpha \in \mathbb{N}$ and $L(x) \underset{x\to\infty}{\not\longrightarrow}$const.*

Some people would also include in the definition of a continuous phase transition cases when $P_{\text{top}}(\phi_t)$ is equal to two different analytic functions on the two sides of zero, but is differentiable at zero. We do not treat such cases here.

We focus on one–parameter families of the form $\phi_t := \phi + t\psi$. In this case $t \mapsto P_{\text{top}}(\phi + t\psi)$ is convex and this imposes restrictions on the sign in front of $t^\alpha L(1/t)$, see below. The dynamical systems we study are assumed to have countable Markov partitions. This is equivalent to the study of topological Markov shifts.

Topological Markov Shifts. A *topological Markov shift* $(\Sigma^+_{\mathbb{A}}, T)$ with a countable set of *states* S and a *transition matrix* $\mathbb{A} = (t_{ij})_{S\times S}$ is the set $\Sigma^+_{\mathbb{A}} := \{(x_0, x_1, \ldots) \in S^{\mathbb{N}\cup\{0\}} : \forall i,\ t_{x_i x_{i+1}} = 1\}$ together with the map $(Tx)_i = x_{i+1}$.

[1]In fact, this interpretation seems to be implicit in many of the manipulations done in the physical theory of critical phenomena. For example, the standard derivation of the critical exponent identities is done by formal differentiation of a (postulated) asymptotic expansion of the free energy (see e.g. [**BDFN**] §1.5.1). However, if $\alpha > 0$, $f(t) \sim t^\alpha L(1/t)$ and $f'(t) \sim \alpha t^{\alpha-1}L(1/t)$, then $L(1/t)$ *must* be slowly varying, because of Karamata's Theorem (appendix A).

A word $(w_1, \ldots, w_n) \in S^n$ is called *admissible* if $t_{w_i w_{i+1}} = 1$ for all i. A topological Markov shift is called *topologically mixing* if for every $a, b \in S$ there are admissible words beginning with a and ending at b of length n for all large n.

A topological Markov shift is endowed with a metric $d(x,y) := 2^{-\min\{k : x_k \neq y_k\}}$. The resulting topology is generated by the basis of *cylinders*

$$[a_0, \ldots, a_{n-1}] := \{x \in \Sigma_{\mathbb{A}}^+ : x_i = a_i,\ 0 \leq i \leq n-1\}.$$

A function $\phi : \Sigma_{\mathbb{A}}^+ \to \mathbb{R}$ is called *Hölder continuous* if $|\phi(x) - \phi(y)| \leq Ad(x,y)^\kappa$ for some constants $A, \kappa > 0$. This condition is too strong for us, because it implies boundedness. The following notions do not:

1. ϕ is *locally Hölder continuous* if $|\phi(x) - \phi(y)| \leq Ad(x,y)^\kappa$ whenever $x_0 = y_0$;
2. ϕ is *weakly Hölder continuous* if $|\phi(x) - \phi(y)| \leq Ad(x,y)^\kappa$ whenever $x_0 = y_0, x_1 = y_1$;
3. ϕ has *summable variations* if $\sum_{n \geq 2} \mathrm{var}_k \phi < \infty$ where $\mathrm{var}_k \phi := \sup\{|\phi(x) - \phi(y)| : y_i = x_i\ (i = 0, \ldots, k-1)\}$.

Local Hölder continuity is stronger than weak Hölder continuity, and weak Hölder continuity is stronger than summable variations.

We need the *Variational Principle* for countable Markov shifts [**S1**]: Suppose $T : X \to X$ is a topologically mixing topological Markov shift, $\phi : X \to \mathbb{R}$ has summable variations, and $\sup \phi < \infty$; Then for any state a,

$$P_{\mathrm{top}}(\phi) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{\substack{T^n x = x \\ x_0 = a}} \exp \sum_{k=0}^{n-1} \phi(T^k x).$$

The limit on the right hand side is called the *Gurevich pressure* of ϕ in honor of B. Gurevich who proved the variational principle in the case $\phi \equiv 0$ [**Gu**].[2]

Program. In the case $|S| < \infty$, Ruelle [**Ru**] has established the following relation between the analytic properties of $t \mapsto P_{\mathrm{top}}(\phi + t\psi)$ and the statistical properties of the equilibrium measure at $t = 0$:

Theorem 1 (Ruelle)**.** *Suppose $(\Sigma_{\mathbb{A}}^+, T)$ is topologically mixing. If $|S| < \infty$ and $\phi, \psi : \Sigma_{\mathbb{A}}^+ \to \mathbb{R}$ are Hölder continuous, then $t \mapsto P_{top}(\phi + t\psi)$ is real–analytic, and admits the expansion $P_{top}(\phi + t\psi) = P_{top}(\phi) + c_\psi t + \frac{1}{2}\sigma_\psi^2 t^2 + o(t^2)$, where $c_\psi = \mathbb{E}_{\mu_\phi}[\psi]$ and $\frac{1}{\sqrt{n}}\left(\sum_{k=0}^{n-1} \psi \circ T^k - nc_\psi\right) \xrightarrow[n\to\infty]{dist.} N(0, \sigma_\psi^2)$ w.r.t. μ_ϕ.*[3]

Ruelle has also proved exponential decay of correlations in this case [**Ru**]. Thus there can be no phase transitions for short–range interactions when $|S| < \infty$.

Phase transitions *are* possible for short–range interactions when $|S| = \infty$. Indeed, it is well–known that long–range interactions on one–dimensional lattice-gas models may admit phase transitions, and there are cases when such models can be recast as *short*–range interactions on *infinite* state shifts. See appendix B and [**FF**], [**Ho**],[**Lo**],[**PS**],[**Wa1**], [**Wa2**], [**S2**],[**S7**],[**MU2**],[**Y**].

Motivated by the physical analogy described above, we seek a generalization of theorem 1 which relates singular behavior for $P_{\mathrm{top}}(\phi + t\psi)$ ('critical exponents') to

---

[2]The variational principle is stated under a stronger condition on ϕ in [**S1**], but is true with practically the same proof under the assumptions stated above.

[3]Here and throughout, $\mathbb{E}$ denotes expectation, $N(0, \sigma^2)$ is the Gaussian distribution, and $\xrightarrow{\text{dist.}}$ means convergence in distribution, see [**F**], [**GK**].

non-Gaussian distributional limit theorems for $\sum_{k=0}^{n-1}\psi\circ T^k$ and to sub–exponential rates of mixing ('abnormal fluctuations' and 'infinite correlation length').

Such a relation is mentioned in the physics literature, see [**BDFN**] for a renormalization group approach and Hilfer [**Hi**] for a probabilistic point of view very similar to the one we adopt below. Rigorous results are more difficult to find. See section V.8 in [**El**] for a discussion of the Ising model.[4]

2. Statement of Results

Assumptions. Let G_α $(0<\alpha\le 2)$ be the probability distribution with Laplace transform $\int_{\mathbb{R}} e^{s\xi}dG_\alpha(\xi)=\exp[sgn(\alpha-1)s^\alpha]$ when $\alpha\neq 1$ and $\int_{\mathbb{R}} e^{s\xi}dG_\alpha(\xi)=e^{-s}$ when $\alpha=1$. Such distributions exist: When $\alpha\neq 1$, G_α is the standard spectrally negative stable distribution of index α, and when $\alpha=1$ G_α is the degenerate distribution concentrated at $\{-1\}$ (see [**Z**] for details).

Let $(\Sigma_{\mathbb{A}}^+,T)$ be a topologically mixing topological Markov shift with a countable set of states S and a transition matrix $\mathbb{A}=(t_{ij})_{S\times S}$. Our results are simplest to state when $\mathbb{A}$ satisfies the *Big Images and Preimages* property:

$$\exists b_1,\dots,b_N:\forall a\in S,\ \exists i,j \text{ s.t. } t_{b_i a}t_{ab_j}=1. \tag{BIP}$$

This condition appears naturally in the theory of countable Markov shifts, as a necessary and sufficient for the existence of Gibbs measures in the sense of Bowen [**S3**], [**MU1**]. (Equilibrium measures may exist in the absence of (BIP), see [**S4**].)

We can remove the BIP property, at the cost of additional assumptions on ϕ and ψ. Define for a state $a\in S$ the function $r_a(x):=\min\{k:x_k=a\}$, with the convention $\min\varnothing=\infty$. Let μ_ϕ be the equilibrium measure of ϕ (when it exists). We shall impose the following assumption on ϕ:

$$\text{There exists } a\in S \text{ such that } \mathbb{E}_{\mu_\phi}[r_a]<\infty. \tag{Φ}$$

We call a set $E\subseteq\Sigma_{\mathbb{A}}^+$ *bounded*, if $E\subseteq\{x:x_0\in S_0\}$ for some finite set $S_0\subset S$. We shall consider functions ψ for which

$$\psi\in L^1(\mu_\phi), \text{ and } \psi\le\mathbb{E}_{\mu_\phi}[\psi] \text{ outside a bounded set.} \tag{Ψ}$$

Critical exponents and abnormal fluctuations. Throughout this section let $(\Sigma_{\mathbb{A}}^+,T)$ be a topologically mixing topological Markov shift, and suppose $\{\phi+t\psi\}_{t\ge 0}$ is a regular family, where ϕ,ψ are two locally Hölder continuous functions s.t. $\sup\phi<\infty$, $P_{\rm top}(\phi)<\infty$ and $\sup\psi<\infty$.

Theorem 2. *Assume* (Φ) *and* (Ψ). *The following are equivalent for* $1<\alpha<2$*:*

1. **Critical Exponent:** $P_{top}(\phi+t\psi)=P_{top}(\phi)+ct+t^\alpha L(1/t)$ *with* $L(x)$ *slowly varying at infinity.*
2. **Non-Gaussian Fluctuations:** $\frac{1}{B_n}\big(\sum_{k=0}^{n-1}\psi\circ T^k-cn\big)\xrightarrow[n\to\infty]{dist.}G_\alpha$ *w.r.t.* μ_ϕ, *where* $c=\mathbb{E}_{\mu_\phi}[\psi]$, $B_n=n^{\frac{1}{\alpha}}\ell(n)$ *and* $\ell(n)$ *is slowly varying at infinity.*

The following theorems treat the case $\alpha=1,2$.

[4] The case of discontinuous ('first-order') phase transitions is more amenable to rigorous treatment. A discontinuous phase transition is characterized by the lack of differentiability of the free energy. The theory of large deviations can be used to interpret such a singularity as lack of exponential convergence in distribution of an associated macroscopic quantity to a unique thermodynamic value, see Ellis [**El**].

Theorem 3. *Assume* (Φ) *and* (Ψ).

(1) **Taylor Expansion**: $P_{top}(\phi+t\psi) = P_{top}(\phi)+ct+\frac{\sigma^2}{2}t^2+o(t^2)$ *with* $\sigma \neq 0$ *iff* $c = \mathbb{E}_{\mu_\phi}[\psi]$ *and* $\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}(\psi\circ T^k - c) \xrightarrow[n\to\infty]{dist.} N(0,\sigma^2)$. *In this case* $\psi \in L^2(\mu_\phi)$.

(2) **Critical Expansion**: $P_{top}(\phi+t\psi) = P_{top}(\phi)+ct+\frac{1}{2}t^2L(1/t)$ *with* $L(x)$ *s.v. at infinity,* $L(x) \not\to \text{const}$ *iff* $\frac{1}{B_n}\big(\sum_{k=0}^{n-1}\psi\circ T^k - cn\big) \xrightarrow[n\to\infty]{dist.} N(0,1)$, $c = \mathbb{E}_{\mu_\phi}[\psi]$, *and* B_n *is r.v. of index* $\frac{1}{2}$ *s.t.* $\frac{\sqrt{n}}{B_n} \to 0$. *In this case* $L(x) \to \infty$.

Theorem 4.

(1) **Taylor Expansion**: *Assume* (Ψ). *Then* $P_{top}(\phi+t\psi) = P_{top}(\phi)+ct+o(t)$ *with* $c = \mathbb{E}_{\mu_\phi}[\psi]$, *and* $\frac{1}{n}\sum_{k=0}^{n-1}\psi\circ T^k \xrightarrow[n\to\infty]{} \mathbb{E}_{\mu_\phi}[\psi]$ μ_ϕ*-a.s. and in distribution.*

(2) **Critical Expansion**: *Assume* (BIP). *Then* $P_{top}(\phi+t\psi) = P_{top}(\phi)+ct+tL(1/t)$ *with* $|L(x)|$ *s.v. at infinity and* $L(x) \not\to \text{const.}$ *iff* $\psi \notin L^1(\mu_\phi)$ *and* $\exists B_n$ *r.v. of index 1 s.t.* $\frac{n}{B_n} \to 0$ *and* $\frac{1}{B_n}\sum_{k=0}^{n-1}\psi\circ T^k \xrightarrow[n\to\infty]{dist.} G_1$. *In this case* $L(x) \xrightarrow[x\to\infty]{} -\infty$.

To understand the previous results, it is useful to think of $\psi_n := \sum_{k=0}^{n-1}\psi\circ T^k$ as of a 'macroscopic' quantity with average (at equilibrium) $n\mathbb{E}_{\mu_\phi}[\psi]$. In the absence of a phase transition, one expects the fluctuations of ψ_n about its average to be of order $\sqrt{n}$. The previous results say that in the presence of a continuous phase transition with critical exponent $\alpha \leq 2$, the fluctuations are of order B_n with $B_n \gg \sqrt{n}$ (compare with [**Hi**]).

Remark: Theorems 2, 3, 4 remain true if the pair of conditions (Φ) and (Ψ) is replaced by (BIP). Under this new set of assumptions:

(1) Theorem 2 is also valid for $0 < \alpha < 1$, except that in this case $\mathbb{E}_{\mu_\phi}[\psi] = -\infty$, the slow variation of $L(x)$ should be replaced by the slow variation of $-L(x)$, and c can be set to zero (because $ct = o(t^\alpha L(1/t))$, $cn = o(B_n)$);
(2) Case (1) of theorem 3 holds iff $\psi \in L^2(\mu_\phi)$ and ψ is not a measurable coboundary [**AD**], [**Gou2**], and case (2) of theorem 3 holds iff $\psi \notin L^2(\mu_\phi)$ (see theorem 5 below);
(3) Case (1) of theorem 4 holds iff $\psi \in L^1(\mu_\phi)$.

When do different systems exhibit the same asymptotic expansion? In order to answer this question, one needs to clarify what properties of ψ and μ_ϕ are captured by α and $L(x)$. The following is motivated by [**AD**], [**GK**].

Theorem 5. *Let* $(\Sigma_{\mathbb{A}}^+, T)$ *be a topologically mixing topological Markov shift with the BIP property, and suppose* ϕ, ψ *are locally Hölder continuous s.t.* $\sup\phi < \infty$, $P_{top}(\phi) < \infty$, $\sup\psi < \infty$ *and s.t.* ϕ *has an equilibrium measure. The following are equivalent for* $L(x)$ *s.t.* $|L(x)|$ *is s.v. at infinity and* $0 < \alpha \leq 2$:

(1) **Critical Exponent:** $P_{top}(\phi+t\psi) = P_{top}(\phi)+ct+t^\alpha L(1/t)[1+o(1)]$ *as* $t \to 0^+$;
(2) **Domain of Attraction:** *One of the following holds as* $x \to \infty$
  (a) $0 < \alpha < 2$, $\alpha \neq 1$ *and* $\mu_\phi[\psi < -x] \sim -\frac{x^{-\alpha}}{\Gamma(1-\alpha)}L(x)$;

(b) $\alpha = 1$ *and either* $\psi \in L^1$, *and then* $L(x) = \mathbb{E}_{\mu_\phi}[\psi] - c + o(1)$, *or* $\psi \notin L^1$ *and then* $L(x) \sim \mathbb{E}_{\mu_\phi}[\psi \vee (-x)]$;

(c) $\alpha = 2$ *and either* $\psi \in L^2$, *and then* $L(x) = \frac{1}{2}\sigma^2 + o(1)$ *for some* $\sigma \in \mathbb{R}$; *or* $\psi \notin L^2$, *and then* $L(x) \sim \frac{1}{2}\mathbb{E}_{\mu_\phi}\left[\psi^2 1_{[|\psi| \le x]}\right]$.

Here $f(x) \sim g(x)$ means $\frac{f(x)}{g(x)} \xrightarrow[x\to\infty]{} 1$ and $a \vee b := \max\{a, b\}$.

Remark 1: The implication (2) $\Rightarrow$ (1) follows from theorem 2 and the work of Aaronson & Denker [**AD**] who showed that (2) implies a distributional limit theorem. We give an alternative proof below.

Remark 2: Theorem 5 enables one to construct an abundance of ψ's for which $\{\phi + t\psi\}_{t\ge0}$ has a critical exponent. These ψ's are of course unbounded. Indeed, by [**S3**], in the BIP case $P_{\text{top}}(\phi+t\psi)$ is real–analytic whenever $P_{\text{top}}(\phi+t\psi) < \infty$ for all t in some *two*-sided neighbourhood of zero (e.g. bounded ψ's). For shifts without the BIP property, critical exponents are possible for ψ bounded (see below).

Remark 3: The following generalization of theorem 5 to general shifts is a direct consequence of the discussion at the beginning of section 4 and theorem 8 there. Let $(\Sigma_{\mathbb{A}}^+, T)$ be a topologically mixing topological Markov shift, and suppose $\{\phi+t\psi\}_{t\ge0}$ is a regular family, where ϕ, ψ are two locally Hölder continuous functions s.t. $P_{\text{top}}(\phi) < \infty$, $\sup\phi < \infty$, and $\sup\psi < \infty$. We assume (Ψ) (but do *not* assume (Φ) or (BIP)). Let A be the bounded set mentioned in (Ψ), and define

$$\overline{\psi} := 1_A \cdot \sum_{k=0}^{r-1} \psi \circ T^k, \text{ where } r(x) := \inf\{n \ge 1 : T^n x \in A\}.$$

Then $P_{\text{top}}(\phi + t\psi) = P_{\text{top}}(\phi) + ct + t^\alpha L(1/t)[1 + o(1)]$ for $1 < \alpha \le 2$ with L s.v at infinity iff $\overline{\psi}$ satisfies the domain of attraction condition of theorem 5 w.r.t. the normalized restriction of μ_ϕ to A. The random variables $\overline{\psi}$ can be thought of as sums over 'weakly correlated blocks', see [**Hi**], [**FF**] and appendix B.

This explains why in the non–BIP case even bounded ψ's may satisfy non-Gaussian limit laws: $\overline{\psi}$ may have a heavy tail, even if ψ does not, because r may have a heavy tail (of course (Φ) must fail in this case).

Returning to the case treated in theorem 5, we note that the domain of attraction condition is phrased in terms of ψ alone, and is not an asymptotic property of $\sum_{k=0}^{n-1} \psi \circ T^k$ as $n \to \infty$. Of course the thermodynamic limit is still present in the form of the equilibrium measure μ_ϕ. But in the BIP case the equilibrium measure satisfies certain a priori uniform bounds which allow one to deduce the following thermodynamic-limit–free necessary condition for the existence of a critical exponent. Choose some $x_a \in \Sigma_{\mathbb{A}}^+$ s.t. x_a starts at a ($a \in S$).

Corollary 1. *Under the assumptions of theorem* 5, $P_{top}(\phi + t\psi) = P_{top}(\phi) + ct + t^\alpha L(1/t)$ *with* $|L(x)|$ *s.v. at infinity and* $\alpha \in (0, 2) \setminus \{1\}$ *implies:*

$$\sum_{a\in S:\psi(x_a)<-x} e^{\phi(x_a)} \asymp \frac{1}{x^\alpha}|L(x)| \text{ as } x \to \infty.$$

Here and throughout $f(x) \asymp g(x)$ as $x \to \infty$ means: $\exists M$ such that $\frac{1}{M} \le \frac{f(x)}{g(x)} \le M$ for all x large enough.

Critical exponents and slow decay of correlations. The *covariance* of two square integrable functions f, g defined on a probability space $(X, \mathcal{B}, \mu)$ is

$$\mathrm{Cov}_\mu(f,g) := \int fg d\mu - \int f d\mu \int g d\mu.$$

The following result says that under certain assumptions, the existence of a critical exponent implies that the decay of correlations is sub-exponential, as expected from the analogy described in the introduction. We need a strengthening of (Ψ):

$$\psi \in L^1 \text{ and } \exists \epsilon > 0 \text{ s.t. } \psi \leq \mathbb{E}_{\mu_\phi}[\psi] - \epsilon \text{ outside a bounded set.} \tag{Ξ}$$

Theorem 6. *Let $(\Sigma_{\mathbb{A}}^+, T)$ be a topologically mixing topological Markov shift, and suppose $\{\phi + t\psi\}_{t\geq 0}$ is a regular family, where ϕ, ψ are locally Hölder continuous functions s.t. $\sup \phi < \infty$, $P_{top}(\phi) < \infty$, $\|\psi\|_\infty < \infty$, and ψ satisfies (Ξ). If $P_{top}(\phi + t\psi) = P_{top}(\phi) + ct + t^\alpha L(1/t)$ with $1 < \alpha < 2$ and L is s.v. at ∞, then*

$$\mathrm{Cov}_{\mu_\phi}(f, g \circ T^n) \asymp \frac{L(n)}{n^{\alpha-1}} \int f d\mu_\phi \int g d\mu_\phi \text{ as } n \to \infty$$

for all f, g locally Hölder continuous with bounded support and positive expectation.

3. Proofs for shifts satisfying the BIP property

Standing Assumptions. In this section we give the proofs of theorems 2, 3, 4 and 5 in the case of topologically mixing countable Markov shifts with (BIP). Our assumptions on ϕ and ψ are that they are locally Hölder continuous, bounded from above, and that $P_{\mathrm{top}}(\phi) < \infty$. We do not assume (Φ), (Ψ) or that $\{\phi + t\psi\}_{t>0}$ is regular. We do assume that ϕ has an equilibrium measure.[5]

Our results remain unchanged if we add to ϕ a term of the form $h - h \circ T + c$ with h bounded (locally) Hölder continuous and $c \in \mathbb{R}$. It is always possible, by means of such h and c, to change ϕ so that $P_{\mathrm{top}}(\phi) = 0$, $\sup \phi \leq 0$, and

$$\sum_{Ty=x} e^{\phi(y)} = 1 \text{ for all } x.$$

This is Lemma 1 in [**S2**] (the boundedness of h is proved for systems with the BIP property in [**S3**]). Henceforth, we assume that ϕ satisfies these additional assumptions.

Distributional Limit Theorems and Laplace Transforms. We shall study the distributional limit behaviour of $X_n := \frac{1}{B_n}\left(\sum_{k=0}^{n-1} \psi \circ T^k - cn\right)$ by analyzing the behaviour of its Laplace transform $\mathbb{E}_{\mu_\phi}[e^{tX_n}]$:

Proposition 1. *Let X, X_n be random variables such that for some $\omega > 0$, $\mathbb{E}(e^{tX_n})$, $\mathbb{E}(e^{tX})$ are finite for all $0 \leq t \leq \omega$. The following are equivalent:*

1. $\mathbb{E}(e^{tX_n}) \xrightarrow[n\to\infty]{} \mathbb{E}(e^{tX})$ *for all $0 \leq t \leq t_0$ and some $t_0 > 0$;*
2. $X_n \xrightarrow[n\to\infty]{dist.} X$.

Proof. See e.g. Martin–Löf [**ML**]. □

[5] In fact, this assumption can be removed as well: locally Hölder potentials with finite pressure on shifts with (BIP) always have *Gibbs measures* [**S3**], and everything we say below holds with μ_ϕ=Gibbs measure of ϕ.

Nagaev's Method [**N**]**.** This is a method for analyzing the Laplace (or Fourier) transform of the distribution of the sum of *dependent* identically distributed random variables. We need it to analyze the distribution of $\psi_n := \psi + \psi\circ T + \cdots + \psi \circ T^{n-1}$ with respect to μ_ϕ. The idea is to construct a family of operators R_t such that

$$\mathbb{E}_{\mu_\phi}[e^{t\psi_n}] = \mathbb{E}_{\mu_\phi}[R_t^n 1].$$

and use operator theory to analyze the right hand side, see Nagaev [**N**] and Aaronson & Denker [**AD**].

In order to construct R_t, we recall some facts on the structure of equilibrium measures for countable Markov shifts. It was proved in [**S4**] and [**BS**] that the equilibrium measure μ_ϕ must be of the form $hd\nu$ where h is a positive continuous function and ν is a positive measure such that $R_0^*\nu = e^{P_{\text{top}}(\phi)}\nu$ and $R_0 h = e^{P_{\text{top}}(\phi)}h$, where R_0 is *Ruelle's operator*:

$$(R_0 f)(x) = \sum_{Ty=x} e^{\phi(y)} f(y).$$

It is also known that h is, up to a constant, the unique positive continuous function such that $R_0 h = e^{P_{\text{top}}(\phi)}h$. Since by our assumptions on ϕ $P_{\text{top}}(\phi) = 0$ and $R_0 1 \equiv \sum_{Ty=x} e^{\phi(y)} = 1$, we must have $h = \text{const.}$, whence μ_ϕ is a constant times ν. It follows that $R_0^*\mu_\phi = \mu_\phi$. In particular, $\mathbb{E}_{\mu_\phi}[R_0 F] = \mathbb{E}_{\mu_\phi}[F]$ for every bounded continuous function F. Now define the operators

$$R_t f := R_0[e^{t\psi} f].$$

A calculation shows that $R_t^n 1 = R_0^n[e^{t\psi_n}1]$. Passing to expectations with respect to μ_ϕ, we see that $\mathbb{E}_{\mu_\phi}[e^{t\psi_n}] = \mathbb{E}_{\mu_\phi}[R_0^n(e^{t\psi_n})] = \mathbb{E}_{\mu_\phi}[R_t^n 1]$ as required.

Next we seek a Banach space $\mathcal{L}$ such that $R_t : \mathcal{L} \to \mathcal{L}$ has good spectral properties. Such a space was found by Aaronson and Denker [**AD**]. We review their construction.

Recall the metric $d(x,y)$ on $\Sigma_{\mathbb{A}}^+$, and fix some $\kappa > 0$ such that ϕ, ψ are both Hölder continuous with exponent κ with respect to d. Define $\mathcal{L}$ to be the space of functions $f : \Sigma_{\mathbb{A}}^+ \to \mathbb{R}$ such that

$$\|f\|_{\mathcal{L}} := \|f\|_\infty + Df < \infty, \text{ where } Df := \sup\{\frac{|f(x)-f(y)|}{d(x,y)^\kappa} : x \neq y,\ x_0 = y_0\}.$$

This is a Banach space with respect to $\|\cdot\|_{\mathcal{L}}$.

Proposition 2 (Aaronson & Denker)**.** *Suppose $\Sigma_{\mathbb{A}}^+$ has the BIP property, and let ϕ, ψ be two locally Hölder continuous functions such that $\sup\phi \le P_{top}(\phi) = 0$, $R_0 1 = 1$, $\sup\psi < \infty$, and ϕ has an equilibrium measure μ_ϕ. Then:*

1. **Boundedness:** *$R_t(\mathcal{L}) \subseteq \mathcal{L}$ and $R_t : \mathcal{L} \to \mathcal{L}$ are bounded linear operators for all $t \ge 0$.*
2. **Spectral Gap:** *$R_0 = P + N$ where $PR_0 = R_0 P$, $P^2 = P$, $NP = PN = 0$ and the spectral radius of N is less than one. P is given by $Pf := \mathbb{E}_{\mu_\phi}[f]$.*
3. **Continuity:** *$\|R_t - R_s\| = O\left(|t-s| + \mathbb{E}_{\mu_\phi}\left(|1 - e^{|t-s|\psi}|\right)\right)$ for $0 \le t, s \le 1$, where $\|\cdot\|$ is the operator norm.*
4. **Differentiability:** *if $\mathbb{E}_{\mu_\phi}\left[|\psi|\right] < \infty$, then $t \mapsto R_t$ is continuously differentiable on $[0, \delta_0)$ for some $\delta_0 > 0$. The derivative is $R_t' : f \mapsto R_t(\psi f)$ (right derivative is meant at 0).*

Proof. The BIP property implies that any equilibrium measure μ_ϕ of a locally Hölder continuous potential ϕ has the Gibbs property: $\exists G = G(\phi)$ such that

$$G^{-1}\mu_\phi[x_0,\ldots,x_{n-1}] \le e^{\phi_n(x)} \le G\mu_\phi[x_0,\ldots,x_{n-1}], \quad (x \in \Sigma_{\mathbb{A}}^+)$$

where $\phi_n := \phi + \phi\circ T + \cdots + \phi\circ T^{n-1}$ [**S3**]. Thus $e^{\phi(x)} \le G\mu_\phi[x_0]$.

Fix some bounded Lipschitz function $F : (-\infty, \sup\psi] \to \mathbb{R}$ with Lipschitz constant $Lip(F)$, and define the linear operator $R_F : f \mapsto R_0[F\circ\psi\cdot f]$. We need the following estimate: For some constant M independent of F, and $D_a[F\circ\psi] := \sup\{\frac{F(\psi(x))-F(\psi(y))}{d(x,y)^\kappa} : x\neq y,\ x,y\in[a]\}$,

$$\|R_F\| \le M\left(\mathbb{E}_{\mu_\phi}[|F|\circ\psi] + \sum_{a\in S}\mu_\phi[a]D_a(F\circ\psi)\right). \tag{2}$$

To prove this, we must estimate $\|R_F f\|_\infty$, $D[R_F f]$ for $f\in\mathcal{L}$. Fix x,y such that $x_0=y_0$, and let $P(x_0) := \{a\in S : t_{ax_0}=1\}$. Then

$$\begin{aligned}|R_Ff(x)-R_Ff(y)| \le& \sum_{a\in P(x_0)} e^{\phi(ax)}|1-e^{\phi(ay)-\phi(ax)}|\cdot|F(\psi(ax))f(ax)| + \\ &+\sum_{a\in P(x_0)} e^{\phi(ay)}|F(\psi(ax))-F(\psi(ay))||f(ax)| \\ &+\sum_{a\in P(x_0)} e^{\phi(ay)}|F(\psi(ay))||f(ax)-f(ay)|.\end{aligned}$$

If $\phi(ax)\neq\phi(ay)$ then

$$\begin{aligned}|1-e^{\phi(ax)-\phi(ay)}| &\le \frac{|1-e^{\phi(ax)-\phi(ay)}|}{|\phi(ax)-\phi(ay)|}D\phi d(ax,ay)^\kappa \le \\ &\le \sup\{\frac{|1-e^\delta|}{\delta} : |\delta|\le D\phi d(ax,ay)^\kappa\}D\phi d(ax,ay)^\kappa < Kd(x,y)^\kappa\end{aligned}$$

with (for example) $K = D\phi\cdot\sup\{\frac{|1-e^\delta|}{\delta} : |\delta|\le D\phi\}$. Re-define K if necessary to guarantee $K>1$. It is now straightforward to deduce, using the inequality $\sum_{a\in P(x_0)} e^{\phi(ax)} \le G$, that

$$D(R_Ff) \le 2K\|f\|_{\mathcal{L}}\|R_0(|F|\circ\psi)\|_\infty + G\|f\|_{\mathcal{L}}\sum_{a\in S}\mu_\phi[a]D_a(F\circ\psi).$$

It is also clear that $\|R_Ff\|_\infty = \|R_0[F\circ\psi\cdot f]\|_\infty \le \|R_0(|F|\circ\psi)\|_\infty\|f\|_{\mathcal{L}}$. Thus

$$\|R_F\| \le 3K\|R_0(|F|\circ\psi)\|_\infty + G\sum_{a\in S}\mu_\phi[a]D_a(F\circ\psi).$$

We proceed to estimate $\|R_0(|F|\circ\psi)\|_\infty$:

$$\begin{aligned}R_0(|F|\circ\psi)(x) &\le G\sum_{a\in P(x_0)}\mu_\phi[a]\big(\inf_{[a]}|F|\circ\psi + D_a(|F|\circ\psi)\big) \\ &\le G\left(\mathbb{E}_{\mu_\phi}[|F|\circ\psi] + \sum_{a\in S}\mu_\phi[a]D_a(F\circ\psi)\right).\end{aligned}$$

Recalling that $K>1$, we obtain (2) with $M := 3KG$.

Note that for every a, $D_a(|F|\circ\psi) \le Lip(F)D\psi$, so

$$\|R_F\| \le M\left(\mathbb{E}_{\mu_\phi}[|F|\circ\psi] + Lip(F)D\psi\right).$$

The boundedness of R_t is the special case with $F(\xi) = e^{t\xi}$. The spectral gap of R_0 follows from the Ionescu-Tulcea Marinescu theorem and the mixing of μ_ϕ, as in [**AD**]. The modulus of continuity of $t \mapsto R_t$ is obtained by observing that $R_t - R_s = R_F$ with $F(\xi) = e^{t\xi} - e^{s\xi}$.

Differentiability is more difficult. We begin with the continuity of $t \mapsto R_t'$ (defined in part (4)). Write $(R_{t+h}' - R_t')f = R_{F_h}(f)$, where $F_h(\xi) = e^{t\xi}\xi(e^{h\xi} - 1)$. We fix $t > 0$ and show that the norm of this operator tends to zero as $h \to 0$, using (2):

(1) $\mathbb{E}_{\mu_\phi}[|F_h \circ \psi|] \xrightarrow[h\to 0]{} 0$ because $F_h \circ \psi \xrightarrow[h\to 0]{} 0$ pointwise and $|F_h \circ \psi|$ is uniformly bounded for $|h| < \frac{t}{2}$.

(2) $\sum_a \mu_\phi[a] D_a(F_h \circ \psi) \xrightarrow[h\to 0]{} 0$: By the mean value theorem

$$D_a(F_h \circ \psi) \leq D\psi \cdot \sup\{|F_h'(z)| : z \in (\inf \psi[a], \sup \psi[a])\}.$$

The right hand side converges to zero as $h \to 0$, and is uniformly bounded (as a function of a) for all $|h| < \frac{t}{2}$ (direct calculation). The result follows from the bounded convergence theorem.

It follows from (2) that $\|R_{F_h}\| \xrightarrow[h\to 0]{} 0$, whence the continuity of R_t' for $t > 0$. The continuity from the right at $t = 0$ can be proved by repeating the previous argument with $t = 0$ and $h \to 0^+$. The only difference is that now instead of the bounded convergence theorem one has to use the dominated convergence theorem, the integrability of $|\psi|$, and the uniform boundedness of $e^{h\psi}$, $e^{t\psi}t\psi$ for $0 < h, t < 1$.

We prove the differentiability of R_t. Set $F_h(\xi) := \frac{e^{h\xi}-1}{h} - \xi$. We have:

$$\left(\frac{R_{t+h} - R_t}{h} - R_t'\right) f \equiv R_{F_h}(e^{t\psi} f).$$

It is easy to check that $e^\psi \in \mathcal{L}$ and that $\|fe^{t\psi}\|_{\mathcal{L}} \leq \|f\|_{\mathcal{L}} \|e^{t\psi}\|_{\mathcal{L}}$. Consequently

$$\left\|\frac{R_{t+h} - R_t}{h} - R_t'\right\| \leq \|e^{t\psi}\|_{\mathcal{L}} \|R_{F_h}\| = O\left(\mathbb{E}_{\mu_\phi}[|F_h(\psi)|] + \sum_{a \in S} \mu_\phi[a] D_a(F_h \circ \psi)\right).$$

By the mean value theorem, the following inequality holds on $[a]$: $D_a(F_h \circ \psi) \leq D\psi |e^{h(\psi + \mathrm{var}_1 \psi)} - 1|$. Consequently,

$$\left\|\frac{R_{t+h} - R_t}{h} - R_t'\right\| = O\left(\mathbb{E}_{\mu_\phi}[|F_h(\psi)|] + \mathbb{E}_{\mu_\phi}[|e^{h(\psi + \mathrm{var}_1 \psi)} - 1|]\right).$$

Now $|F_h| \circ \psi$ is dominated by a constant times $1 + |\psi|$, and $|e^{h(\psi+\mathrm{var}_1\psi)} - 1|$ is uniformly bounded for $0 < h < 1$. Therefore, if $\psi \in L^1$, then $\left\|\frac{R_{t+h}-R_t}{h} - R_t'\right\| \xrightarrow[h\to 0^+]{} 0$.

To see the limit as $h \to 0^-$ (when $t > 0$), write $\tau = t - |h|$, $\tau + |h| = t$ and repeat the previous argument with τ for t, using the continuity of $t \mapsto R_t'$ and the fact that the big Oh in the previous equation is uniform on a neighbourhood of t. $\square$

Spectral gaps are stable under small perturbations [**Ka**]. Therefore, there exists an open neighbourhood U of R_0 in $\mathrm{Hom}(\mathcal{L}, \mathcal{L})$ (the space of bounded linear operators on $\mathcal{L}$ over $\mathbb{C}$) and analytic maps $\lambda : U \to \mathbb{C}$, $P, N : U \to \mathrm{Hom}(\mathcal{L}, \mathcal{L})$ s.t.

$$\left.\begin{array}{c} R = \lambda(R)[P(R) + N(R)] \\ RP(R) = P(R)R = \lambda(R)P(R) \\ P(R)N(R) = N(R)P(R) = 0 \\ P(R)^2 = P(R),\ \dim \mathrm{Im}[P(R)] = 1 \end{array}\right. \quad \text{for all } R \in U,$$

and such that the spectral radius of $N(R)$ is uniformly smaller than one.

Proposition 3. *Under the standing assumptions of this section, there exists* $\epsilon_0 > 0$ *and* $\epsilon(t) \xrightarrow[t\to 0^+]{} 0$ *s.t. for all* $0 \le t \le \epsilon_0$ $\mathbb{E}_{\mu_\phi}[e^{t\psi_n}] = [1 + O(\epsilon(t))]\exp[nP_{top}(\phi + t\psi)]$ *uniformly in* n, *where* $\psi_n := \sum_{k=0}^{n-1} \psi \circ T^k$.

Proof. Fix $\epsilon_0 > 0$ so small that $0 \le t \le \epsilon_0$ implies that $R_t \in U$ and that the spectral radius of N_t is less than $\theta < 1$. This is possible, because $t \mapsto R_t$ is continuous. For such t's $\lambda(t) := \lambda(R_t)$, $P_t := P(R_t)$ and $N_t := N(R_t)$ make sense, and depend continuously on t. In particular, $\mathbb{E}_{\mu_\phi}[P_t 1] \xrightarrow[t\to 0^+]{} \mathbb{E}_{\mu_\phi}[P1] = \mathbb{E}_{\mu_\phi}[1] = 1$. Making ϵ_0 smaller, if necessary, we ensure that $\mathbb{E}_{\mu_\phi}[P_t 1] \neq 0$ for all $0 < t < \epsilon_0$.

Now define $h_t := P_t 1/\mathbb{E}_{\mu_\phi}[P_t 1]$. Recalling that $R_0^* \mu_\phi = \mu_\phi$, we see that

$$\lambda(t)^n = \int \lambda(t)^n h_t d\mu_\phi = \int R_t^n h_t d\mu_\phi = \\ = \int R_0^n[e^{t\psi_n} h_t] d\mu_\phi = \mathbb{E}_{\mu_\phi}[e^{t\psi_n}] + \int e^{t\psi_n}(h_t - 1) d\mu_\phi.$$

Now $|\int e^{t\psi_n}(h_t - 1)d\mu_\phi| \le \mathbb{E}_{\mu_\phi}[e^{t\psi_n}]\|h_t - 1\|_{\mathcal{L}}$, so

$$\lambda(t)^n = [1 + O(\|h_t - 1\|_{\mathcal{L}})]\mathbb{E}_{\mu_\phi}[e^{t\psi_n}].$$

We show that $\|h_t - 1\|_{\mathcal{L}} \xrightarrow[t\to 0^+]{} 0$. Clearly

$$\|h_t - 1\|_{\mathcal{L}} \le \frac{\|P_t 1 - 1\|_{\mathcal{L}} + |1 - \mathbb{E}_{\mu_\phi}(P_t 1)|}{\mathbb{E}_{\mu_\phi}(P_t 1)} \le \frac{2\|1\|_{\mathcal{L}}}{\mathbb{E}_{\mu_\phi}(P_t 1)}\|P(R_t) - P(R_0)\|.$$

The spectral gap of R_0 implies that $\|P(R_t) - P(R_0)\| = O(\|R_t - R_0\|)$ as $t \to 0^+$, so by the previous proposition,

$$\|h_t - 1\|_{\mathcal{L}} = O\left(|t| + \mathbb{E}_{\mu_\phi}(|1 - e^{|t|\psi}|)\right). \tag{3}$$

The bounded convergence theorem now shows that $\|h_t - 1\|_{\mathcal{L}} \xrightarrow[t\to 0^+]{} 0$. We deduce: $\exists \epsilon(t) \xrightarrow[t\to 0^+]{} 0$ such that $\mathbb{E}_{\mu_\phi}[e^{t\psi_n}] = [1 + O(\epsilon(t))]\lambda(t)^n$.

We show that $\lambda(t) = \exp[P_{\text{top}}(\phi + t\psi)]$. Consider the indicator function $1_{[a]}$ of $[a]$ for some $a \in S$ s.t. $\mu_\phi[a] \neq 0$ (in fact every $a \in S$ has this property). Since $1_{[a]} \in \mathcal{L}$, $P_t 1_{[a]} \xrightarrow[t\to 0^+]{\mathcal{L}} P 1_{[a]} = \mathbb{E}_{\mu_\phi}[1_a] \neq 0$. Thus $P_t 1_{[a]} > 0$ for all t small enough.

Fix some $x_a \in [a]$. The commutation relations between R_t, P_t and N_t imply that

$$(R_t^n 1_{[a]})(x_a) = \lambda(t)^n[P_t 1_{[a]} + N_t^n 1_{[a]}](x_a) = \lambda(t)^n[P_t 1_{[a]}(x_a) + o(1)] \sim \lambda(t)^n P_t 1_{[a]}(x_a).$$

We see that for every $x_a \in [a]$,

$$\log \lambda(t) = \lim_{n\to\infty} \frac{1}{n} \log(R_t^n 1_{[a]})(x_a) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{T^n y = x_a} e^{\phi_n(y) + t\psi_n(y)} 1_{[a]}(y) = \\ = \lim_{n\to\infty} \frac{1}{n} \log \sum_{T^n z = z} e^{\phi_n(z) + t\psi_n(z)} 1_{[a]}(z),$$

where the last transition is because the local Hölder continuity of ϕ and ψ allows us to change each $y \in T^{-n}(x_a) \cap [a]$ into $z(y) = (y_0, \ldots, y_{n-1}; y_0, \ldots, y_{n-1}; \ldots)$ without affecting the limit. By the variational principle of [**S1**], $\log \lambda(t) = P_{\text{top}}(\phi + t\psi)$. $\square$

Now that we have related $\mathbb{E}[e^{t\psi_n}]$ to $P_{\rm top}(\phi+t\psi)$ we can proceed as in the case of i.i.d's (see e.g. [**F**]). It is convenient to start with theorem 4, part 1.

Proof of Theorem 4 part 1 for shifts satisfying (BIP). We continue to assume w.l.o.g. that $\phi\le P_{\rm top}(\phi)=0$, $\sum_{Ty=x}e^{\phi(y)}=1$ and $R_0^*\mu_\phi=\mu_\phi$. Subtracting a suitable constant from ψ if necessary, we also assume w.l.o.g. that $\sup\psi<0$.

Recall the notation $R_t,\lambda(t),h_t$ from the proof of proposition 3. We have:

$$\lambda(t)-1=\int R_th_td\mu_\phi-1=\int R_0(e^{t\psi}h_t)d\mu_\phi-\int h_td\mu_\phi=$$
$$=\int(e^{t\psi}-1)h_td\mu_\phi=\mathbb{E}_{\mu_\phi}[e^{t\psi}-1]+\mathbb{E}_{\mu_\phi}[(e^{t\psi}-1)(h_t-1)].$$

Now $|\mathbb{E}_{\mu_\phi}[(e^{t\psi}-1)(h_t-1)]|\le|\mathbb{E}_{\mu_\phi}[e^{t\psi}-1]|\cdot\|h_t-1\|_\infty=o(\mathbb{E}_{\mu_\phi}[e^{t\psi}-1])$, because $e^{t\psi}-1$ doesn't change sign and because $\|h_t-1\|_\infty\le\|h_t-1\|_{\mathcal{L}}\to0$ (see the proof of proposition 3). We conclude that $\lambda(t)-1=[1+o(1)]\mathbb{E}_{\mu_\phi}[e^{t\psi}-1]$.

We have seen in the proof of proposition 3 that $\lambda(t)=\exp P_{\rm top}(\phi+t\psi)$. Since $P_{\rm top}(\phi+t\psi)=o(1)$ as $t\to0^+$,

$$P_{\rm top}(\phi+t\psi)=[1+o(1)]\big(e^{P_{\rm top}(\phi+t\psi)}-1\big)=[1+o(1)]\mathbb{E}_{\mu_\phi}[e^{t\psi}-1],\ \text{ as } t\to0^+. \quad (4)$$

It follows that $P_{\rm top}(\phi+t\psi)=ct+o(t)$ iff $\mathbb{E}_{\mu_\phi}[e^{t\psi}-1]=[ct+o(t)][1+o(1)]$, as $t\to0^+$, which (upon division by t and some rearrangements) is equivalent to

$$\lim_{t\to0^+}\mathbb{E}_{\mu_\phi}\left(\frac{e^{t\psi}-1}{t\psi}\psi\right)=c.$$

It is not difficult to see, using $\psi<0$, that the limit is equal to $\mathbb{E}_{\mu_\phi}[\psi]$. We conclude that $P_{\rm top}(\phi+t\psi)=ct+o(t)$ iff $\psi\in L^1$ and $c=\mathbb{E}_{\mu_\phi}[\psi]$.

In this case $\frac{1}{n}\sum_{k=0}^{n-1}\psi\circ T^k\xrightarrow[n\to\infty]{}\mathbb{E}_{\mu_\phi}[\psi]$ μ_ϕ–almost surely and in distribution, because of the ergodicity of μ_ϕ [**BS**] and the Birkhoff ergodic theorem. This proves the 'Taylor expansion' case of theorem 4 (in the extended form described by the remark after theorem 4).

Proof of Theorem 2 for shifts satisfying (BIP). We keep the standing assumptions of this section. Assume first that $P_{\rm top}(\phi+t\psi)=ct+t^\alpha L(1/t)$ with $|L(x)|$ slowly varying at infinity and $0<\alpha<2$, $\alpha\ne1$ (we are also considering $0<\alpha<1$ because of the remark after theorem 4). Since for every continuous function f and constant C, $P_{\rm top}(f+C)=P_{\rm top}(f)+C$, we can normalize ψ to make $c=0$. The asymptotic relation becomes $P_{\rm top}(\phi+t\psi)=t^\alpha L(1/t)$.

Construct $B_n\to\infty$ such that $\frac{n|L(B_n)|}{B_n^\alpha}\xrightarrow[n\to\infty]{}1$. Here is how to do this: The function $f(x):=x^\alpha/|L(x)|$ is regularly varying at infinity with index $\alpha>0$, and therefore admits a regularly varying asymptotic inverse $g(x)$ (see appendix A). By definition, $(f\circ g)(x)\sim(g\circ f)(x)\sim x$ as $x\to\infty$, so $B_n:=g(n)$ is as required (it tends to infinity, because it is regularly varying with index $1/\alpha>0$).

Assume for the moment that the sign of $L(x)$ converges to $sgn(\alpha-1)$ as $x\to\infty$. Proposition 3 and the expansion of $P_{\rm top}(\phi+t\psi)$ imply that

$$\mathbb{E}_{\mu_\phi}[e^{t\frac{\psi_n}{B_n}}]=[1+O(\epsilon(\tfrac{t}{B_n}))]\exp[nP_{\rm top}(\phi+\tfrac{t}{B_n}\psi)]=$$
$$=[1+O(\epsilon(\tfrac{t}{B_n}))]\exp[t^\alpha\tfrac{nL(B_n)}{B_n^\alpha}\tfrac{L(B_n/t)}{L(B_n)}]\xrightarrow[n\to\infty]{}\exp[sgn(\alpha-1)t^\alpha].$$

The last expression is the Laplace transform of G_α, and so, by proposition 1 (whose conditions hold because $\sup\psi < \infty$), $\frac{1}{B_n}\psi_n \xrightarrow[n\to\infty]{\text{dist.}} G_\alpha$. Finally we observe that when $\alpha > 1$, theorem 4 part 1 applies, and gives $\mathbb{E}_{\mu_\phi}[\psi] = c$.

We now explain why $sgn[L(x)] \xrightarrow[x\to\infty]{} sgn(\alpha - 1)$. Recalling the definition of the topological pressure, we observe that $\varphi(t) := t^\alpha L(1/t) = P_{\text{top}}(\phi + t\psi)$ is convex on $[0,\infty)$. If $\alpha > 1$, then $\varphi(0) = 0$ and $\varphi'_+(0) = 0$ (the right-derivative at zero). Convexity forces φ to be non-negative, whence $L(x) \geq 0$ for all $x > 0$. Since $L(x)$ is eventually non-zero (its absolute value is assumed to be slowly varying), it is eventually positive. If on the other hand $0 < \alpha < 1$, then for any c_0

$$\begin{aligned} P_{\text{top}}(\phi + t(\psi - c_0)) &= P_{\text{top}}(\phi + t\psi) - c_0 t = \\ &= t^\alpha L(1/t)[1 - c_0(1/t)^{\alpha-1}/L(1/t)] = t^\alpha L(1/t)[1 + o(1)]. \end{aligned}$$

If $c_0 > \sup\psi$, then $P_{\text{top}}(\phi + t(\psi - c_0)) < P_{\text{top}}(\phi) = 0$. This forces $L(1/t)$ to be eventually negative. This completes the proof of (1) $\Rightarrow$ (2).

We prove the other direction. Assume $\frac{1}{B_n}(\psi_n - cn) \xrightarrow[n\to\infty]{\text{dist.}} G_\alpha$ for B_n regularly varying with index $1/\alpha$ and $c \in \mathbb{R}$. Again, we can subtract a constant from ψ to make $c = 0$. Our objective is then to show that $P_{\text{top}}(\phi) = t^\alpha L(1/t)$ with $|L(x)|$ slowly varying.

Proposition 1 says that $\mathbb{E}_{\mu_\phi}[e^{t\frac{\psi_n}{B_n}}] \to \exp[sgn(\alpha - 1)t^\alpha]$. Combining this with Proposition 3 gives, since $B_n \to \infty$,

$$\lim_{n\to\infty} nP_{\text{top}}(\phi + \tfrac{t}{B_n}\psi) = sgn(\alpha - 1)t^\alpha \tag{5}$$

on some one–sided right neighbourhood of 0. Applying the sufficient condition for regular variation of appendix A with $f(x) := |P_{\text{top}}(\phi + \frac{1}{x}\psi)|$, $a_n = n$ and $b_n = B_n$, we conclude that $P_{\text{top}}(\phi + t\psi) = t^\rho L(1/t)$, with $|L(x)|$ slowly varying at infinity and some $\rho > 0$. By (5), $\rho = \alpha$. $\square$

Proof of Theorem 4 part 2 for shifts satisfying (BIP). We keep the standing assumptions of this section. Suppose $P_{\text{top}}(\phi + t\psi) = ct + tL(1/t)$ with $|L(x)|$ slowly varying at infinity and $L(x) \not\to \text{const}$.

Changing ψ by a constant, we arrange for $\sup\psi < 0$. Equation (4) holds, and leads to $c + L(1/t) = [1 + o(1)]\mathbb{E}_{\mu_\phi}\left(\frac{e^{t\psi}-1}{t\psi}\psi\right) \xrightarrow[t\to 0^+]{} \mathbb{E}_{\mu_\phi}[\psi]$. Since $L(x) \not\to \text{const.}$, we must have $\mathbb{E}_{\mu_\phi}[\psi] = -\infty$, whence $L(x) \to -\infty$.

As in the proof of theorem 2, we construct B_n regularly varying of index 1 such that $\frac{n|L(B_n)|}{B_n} \to 1$, and observe using proposition 3 that $\mathbb{E}_{\mu_\phi}[e^{\frac{t}{B_n}\psi_n}] \xrightarrow[n\to\infty]{} e^{-t}$. The limit is the Laplace transform of G_1. It follows that $\frac{1}{B_n}\psi_n \xrightarrow[n\to\infty]{\text{dist.}} G_1$. Note that $n/B_n \to 0$, because $|L(x)| \to \infty$. This proves ($\Rightarrow$).

To see ($\Leftarrow$) assume that $\frac{\psi_n}{B_n} \xrightarrow[n\to\infty]{\text{dist.}} G_1$ with B_n r.v. of index one such that $n/B_n \to 0$. Arguing as in the proof of theorem 2, we deduce that $P_{\text{top}}(\phi + t\psi) = tL(1/t)$ with $|L(x)|$ slowly varying at infinity such that $\frac{n|L(B_n)|}{B_n} \xrightarrow[n\to\infty]{} 1$. Since $n/B_n \to 0$, $L(x) \not\to \text{const}$. $\square$

Proof of Theorem 3 for shifts satisfying (BIP). We keep the standing assumptions of this section. Assume first that $\frac{1}{B_n}(\psi_n - cn) \xrightarrow[n\to\infty]{\text{dist.}} N(0,1)$ for some B_n regularly varying of index $\frac{1}{2}$ (this includes the case $B_n = \sigma\sqrt{n}$). Subtracting a suitable constant from ψ we may assume w.l.o.g that $c = 0$ (of course we can no longer assume that $\sup\psi < 0$).

The Laplace transform of $N(0,1)$ is $e^{\frac{1}{2}t^2}$. Arguing as in the proof of theorem 2, we obtain (since $P_{\text{top}}(\phi) = 0$)

$$P_{\text{top}}(\phi + t\psi) = \frac{1}{2}t^2 L(1/t) \tag{6}$$

with $L(x)$ s.v. at infinity such that $\frac{nL(B_n)}{B_n^2} \xrightarrow[n\to\infty]{} 1$.

If $B_n \sim \sigma\sqrt{n}$, then $L(B_n) \xrightarrow[n\to\infty]{} \sigma^2$, and if $\sqrt{n}/B_n \to 0$, then $L(B_n) \xrightarrow[n\to\infty]{} \infty$. The same limits must hold for $L(x)$ as $x \to \infty$, because of the regular variation of B_n and $L(x)$ (use the uniform convergence theorem for slow variation in appendix A). This proves ($\Leftarrow$) in parts (1) and (2).

We prove ($\Rightarrow$). It is enough to treat the case

$$P_{\text{top}}(\phi + t\psi) = \frac{1}{2}t^2 L(1/t)$$

with $L(x) = \sigma^2 + o(1)$ or with $L(x) \not\to$ const., L slowly varying (we can always reduce to this case by subtracting c from ψ). Note that $P_{\text{top}}(\phi + t\psi) = o(t)$, whence by theorem 4 for systems with BIP, $\psi \in L^1$ and $\mathbb{E}_{\mu_\phi}[\psi] = 0$.

As before, the asymptotic expansion above implies the existence of B_n regularly varying of order $\frac{1}{2}$ such that $\frac{1}{B_n}\psi_n \xrightarrow[n\to\infty]{\text{dist.}} N(0,1)$, and B_n is determined up to asymptotic equivalence by the condition $\frac{nL(B_n)}{B_n^2} \xrightarrow[n\to\infty]{} 1$. In the Taylor expansion case $L(x) = \sigma^2 + o(1)$, so $B_n \sim \sigma\sqrt{n}$. In the critical expansion case, $L(x) \not\to$ const. We shall see in the next section that this happens iff $\psi \notin L^2$ and $L(x) \to \infty$. In particular $\frac{\sqrt{n}}{B_n} \to 0$, and $L(x) \to$ const. can only happen if $\psi \in L^2$. □

Proof of Theorem 5. We keep the standing assumptions of this section, and begin with the direction (1) ⇒ (2).

Case 1. $0 < \alpha < 1$.

In this case (1) can be rewritten as $P_{\text{top}}(\phi + t\psi) = t^\alpha L(1/t)[1 + o(1)]$, because $P_{\text{top}}(\phi) = 0$ (standing assumptions) and $ct = o(t^\alpha L(1/t))$. We assume without loss of generality that $\sup\psi < 0$ (otherwise subtract a suitable constant c_0 from ψ and pass from $L(x)$ to $L(x) - c_0 t^{1-\alpha} \sim L(x)$). We saw in the proof of theorem 2 that $L(x)$ is eventually negative.

Since $\sup\psi < 0$, (4) holds, and so

$$1 - \mathbb{E}_{\mu_\phi}[e^{t\psi}] \sim t^\alpha |L(1/t)| \text{ as } t \to 0^+. \tag{7}$$

Write $1-\mathbb{E}_{\mu_\phi}[e^{t\psi}] = \mathbb{E}_{\mu_\phi}[1-e^{-t|\psi|}] = \int_0^\infty (1-e^{-tx})dF(x)$, where $F(x) := \mu_\phi\big[|\psi| \leq x\big]$ is the distribution function of $|\psi|$. Now[6]

$$\begin{aligned}\int_0^\infty (1-e^{-tx})dF(x) &= t\int_0^\infty \int_0^x e^{-ty}dydF(x)\\ &= t\int_0^\infty \int_0^\infty e^{-ty}1_{[y<x]}dF(x)dy = t\int_0^\infty e^{-ty}(1-F(y))dy.\end{aligned}$$

Consequently, $\int_0^\infty e^{-ty}dU(y) \sim t^{\alpha-1}|L(1/t)|$ where $U(y) := \int_0^y [1-F(x)]dx$.

By Karamata's Tauberian theorem this is equivalent to $U(x) \sim \frac{x^{1-\alpha}}{\Gamma(2-\alpha)}|L(x)|$ as $x\to\infty$. The monotone density theorem of appendix A applies; Differentiating, we obtain $1-F(x) \sim \frac{x^{-\alpha}}{\Gamma(1-\alpha)}|L(x)|$ as $x\to\infty$, which is case (2) (a) in theorem 5.

Case 2. $\alpha = 1$

According to Theorem 4 and the remark immediately following it, either $\psi \in L^1$ and then $L(x) = \mathbb{E}_{\mu_\phi}[\psi] - c + o(1)$, or $\psi \notin L^1$ and then $L(x) \xrightarrow[x\to\infty]{} -\infty$. In the first case there is nothing further to prove, so we focus on the second.

In this case the asymptotic expansion of the pressure becomes $P_{\text{top}}(\phi + t\psi) \sim tL(1/t)$, because $P_{\text{top}}(\phi) = 0$ and $ct = o(tL(1/t))$. As before, we may assume w.l.o.g. that $\sup\psi < 0$, and this gives us (4) with $\alpha = 1$.

Again, Karamata's Tauberian Theorem leads to $U(x) = \int_0^x [1-F(y)]dy \sim |L(x)|$ with $F(\cdot)$ the distribution function of $|\psi|$. We now observe that

$$\begin{aligned}&\int_0^x [1-F(t)]dt = \int_0^x \left(\int_t^\infty dF(y)\right)dt =\\ &\qquad = \int_0^\infty \int_0^\infty 1_{[t\leq x]}1_{[t<y]}dt\,dF(y) = \int_0^\infty (x\wedge y)dF(y) = \mathbb{E}_{\mu_\phi}\big[|\psi|\wedge x\big],\end{aligned}$$

where $a\wedge b := \min\{a,b\}$. We obtain $\mathbb{E}_{\mu_\phi}\big[|\psi|\wedge x\big] \sim |L(x)|$. Since $L(x)$ is eventually negative and $\sup\psi < 0$, case (2)(b) follows.

Case 3. $1 < \alpha \leq 2$.

By theorem 4, in this case $\psi \in L^1$ and $c = \mathbb{E}_{\mu_\phi}[\psi]$. Assume w.l.o.g. that $\mathbb{E}_{\mu_\phi}[\psi] = 0$. We are left with the expansion $P_{\text{top}}(\phi + t\psi) = t^\alpha L(1/t)[1+o(1)]$. As in the proof of theorem 2, $L(x)$ must be eventually positive.

Proposition 2 says that $t\mapsto R_t$ is differentiable on $[0,\delta_0)$ for some $\delta_0 > 0$, that its derivative there is $R_t' : f\mapsto R_t(\psi f)$, and that this derivative converges to R_0' as $t\to 0^+$. Make δ_0 smaller, if necessary, to ensure that $P_t1 > 0$ for $0 < t < \delta_0$. This is possible, because $P_t1 \to P_01 \equiv 1$ uniformly.

Since $P_t = P(R_t)$ and $P(\cdot)$ is analytic close to R_0, $t\mapsto P_t1$ is differentiable on $[0,\delta_0)$ and its derivative is continuous from the right at zero. It follows that $t\mapsto h_t \equiv P_t1/\mathbb{E}_{\mu_\phi}[P_t1]$ is differentiable on $[0,\delta_0)$ and that its derivative, which we denote by h_t', satisfies $h_t' \xrightarrow[t\to 0^+]{\mathcal{L}} h_0'$.

Differentiation of $R_th_t = \lambda(t)h_t$ gives: $R_0[e^{t\psi}\psi h_t] + R_0[e^{t\psi}h_t'] = \lambda'(t)h_t + \lambda(t)h_t'$. Taking expectations on both sides, we obtain after some re-organization:

$$\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi h_t] = \lambda'(t) + (\lambda(t)-1)\mathbb{E}_{\mu_\phi}[h_t'] + t\mathbb{E}_{\mu_\phi}\left[\frac{1-e^{t\psi}}{t\psi}\psi h_t'\right].$$

[6]Here and throughout Lebesgue-Stieltjes integrals are used with the convention $\int_a^b = \int_{(a,b]}$.

Add $\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi(1-h_t)]$ to both sides to get:

$$\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi] = \lambda'(t) + (\lambda(t)-1)\mathbb{E}_{\mu_\phi}[h_t'] + t\mathbb{E}_{\mu_\phi}\left[\frac{1-e^{t\psi}}{t\psi}\psi h_t' + e^{t\psi}\psi\frac{1-h_t}{t}\right].$$

Since $\sup\psi<\infty$, $|\frac{1-e^{t\psi}}{t\psi}\psi h_t' + e^{t\psi}\psi\frac{1-h_t}{t}|$ is dominated by some constant times $|\psi|$. Since $\psi\in L^1$, $\mathbb{E}_{\mu_\phi}\left[\frac{1-e^{t\psi}}{t\psi}\psi h_t' + e^{t\psi}\psi\frac{1-h_t}{t}\right] \xrightarrow[t\to 0^+]{} -2\mathbb{E}_{\mu_\phi}[\psi h_0']$. It follows that

$$\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi] = \lambda'(t) + (\lambda(t)-1)\big(\mathbb{E}_{\mu_\phi}[h_0'] + o(1)\big) - 2t\mathbb{E}_{\mu_\phi}[\psi h_0'] + o(t). \tag{8}$$

Recalling that $\lambda(t) = \exp P_{\rm top}(\phi+t\psi) = \exp\big([1+o(1)]t^\alpha L(1/t)\big)$, we see that

$$\lambda(t)-1 \sim t^\alpha L(1/t).$$

Now $\lambda(t)-1$ is convex, because $P_{\rm top}(\phi+t\psi)$ is convex. Therefore, its derivative is monotonic, and the Monotone Density Theorem (appendix A) applies; Differentiating, we get $\lambda'(t)\sim\alpha t^{\alpha-1}L(1/t)$. Plugging these relations into (8) gives

$$\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi] = \begin{cases}\alpha t^{\alpha-1}L(1/t)(1+o(1)) & 1<\alpha<2\\ 2t\big(L(1/t)-\mathbb{E}_{\mu_\phi}[\psi h_0']\big)+o(t)+o(tL(1/t)) & \alpha=2.\end{cases} \tag{9}$$

When $\alpha=2$, this relation implies (since $\mathbb{E}_{\mu_\phi}[\psi]=0$)

$$\begin{aligned}(1+o(1))L(1/t) - \mathbb{E}_{\mu_\phi}[h_0'\psi] &= \frac{1}{2t}\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi] + o(1)\\ &= \frac{1}{2}\mathbb{E}_{\mu_\phi}\big[\frac{e^{t\psi}-1}{t\psi}\cdot\psi^2\big] + o(1) \xrightarrow[t\to 0^+]{} \frac{1}{2}\mathbb{E}_{\mu_\phi}[\psi^2],\end{aligned}$$

because $\frac{e^{t\psi}-1}{t\psi}$ is positive and uniformly bounded on $[\psi\neq 0]$ when $0<t<1$. We see that $L(x)\to$ const. or $L(x)\to\infty$ according to whether $\psi\in L^2$ or not.

Consider first the case $\alpha=2$ and $\psi\in L^2$. In this case $L(x)\to$ const. This constant is non-negative, otherwise $P_{\rm top}(\phi+t\psi)=t^2L(1/t)[1+o(1)]$ is not convex (see the proof of Theorem 2). We denote it by $\frac{1}{2}\sigma^2$, and recognize the the first half of (2)(c) in Theorem 5.

Next assume that $\alpha=2$ and $\psi\notin L^2$, or that $1<\alpha<2$. In these cases, (9) becomes $\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi]\sim\alpha t^{\alpha-1}L(1/t)$ (when $\alpha=2$ this is because $L(x)\to\infty$).

We wish to differentiate this asymptotic relation. In order to do this we first need to check that $\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi]$ has a monotonic derivative on some interval $(0,\delta)$. To see this, we use the dominated convergence theorem to see that for every $t>0$

$$\frac{d}{dt}\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi] = \mathbb{E}_{\mu_\phi}\left[\lim_{h\to 0} e^{t\psi}\psi\frac{e^{h\psi}-1}{h}\right] = \mathbb{E}_{\mu_\phi}[e^{t\psi}\psi^2].$$

This function is convex. Therefore, it is monotonic on $(0,\delta)$ for some $\delta>0$, and the monotone density theorem is applicable. Differentiating, we have

$$\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi^2]\sim\alpha(\alpha-1)t^{\alpha-2}L(1/t),\ \text{as } t\to 0^+.$$

The right-hand-side diverges at zero; It follows that $\mathbb{E}_{\mu_\phi}[\psi^2]=\infty$ for $\alpha\in(1,2)$. Since $\sup\psi<\infty$, $\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi^2]\sim\mathbb{E}_{\mu_\phi}[e^{-t|\psi|}\psi^2]$ as $t\to 0^+$, and we obtain:

$$\mathbb{E}_{\mu_\phi}[e^{-t|\psi|}|\psi|^2]\sim\alpha(\alpha-1)t^{\alpha-2}L(1/t),\ \text{as } t\to 0^+.$$

Setting $F(x) := \mu_\phi[|\psi| \leq x]$, we rewrite this in the form

$$\alpha(\alpha-1)t^{\alpha-2}L(1/t) \sim \int_0^\infty e^{-tx}x^2dF(x) \equiv \int_0^\infty e^{-tx}d\left(\int_0^x y^2dF(y)\right).$$

By Karamata's Tauberian Theorem:

$$\int_0^x y^2dF(y) \sim \frac{\alpha(\alpha-1)}{\Gamma(3-\alpha)}x^{2-\alpha}L(x), \text{ as } x \to \infty.$$

When $\alpha = 2$ (and $\psi \notin L^2$), we obtain

$$L(x) \sim \frac{1}{2}\int_0^x y^2dF(y) = \frac{1}{2}\mathbb{E}_{\mu_\phi}[\psi^2 1_{[|\psi|\leq x]}],$$

and we recognize case (2)(c) of theorem 5. When $1 < \alpha < 2$, Feller's theorem (appendix A) gives $1 - F(x) \sim -\frac{1}{\Gamma(1-\alpha)}x^{-\alpha}L(x)$ as $x \to \infty$. Observing that $1 - F(x) = \mu_\phi[\psi < -x]$ for all $x > \sup\psi$, we recognize case (2)(a) in theorem 5.

We now assume part (2) in theorem 5, and prove part (1). As explained before, this follows from Aaronson & Denker in [**AD**] and theorems 2, 3, and 4, but we include the proof anyway, because it is much simpler than in the more general case they treated (more on this below).

Suppose first that $0 < \alpha < 1$, and assume w.l.o.g that $\sup\psi < 0$. Reversing the steps of the proof of case 1 above, we see that $\mu_\phi[\psi < -x] \sim \frac{x^{-\alpha}}{|\Gamma(1-\alpha)|}|L(x)|$ implies that $\mathbb{E}_{\mu_\phi}[e^{t\psi} - 1] \sim t^\alpha|L(1/t)|$ as $t \to 0^+$. This implies the desired expansion of $P_{\text{top}}(\phi + t\psi)$ because of (4).

Now assume that $\alpha = 1$. If $\psi \in L^1$ and $L(x) = \mathbb{E}_{\mu_\phi}[\psi] - c + o(1)$, then the expansion of $P_{\text{top}}(\phi+t\psi)$ follows from the version of theorem 4 for shifts satisfying (BIP). If $\psi \notin L^1$ and $L(x) \sim \mathbb{E}_{\mu_\phi}[\psi \vee (-x)]$ as $x \to \infty$, then necessarily $|L(x)| \to \infty$. This allows us to assume w.l.o.g. that $\sup\psi < 0$, because a subtraction of a constant from ψ does not affect the statements $P_{\text{top}}(\phi + t\psi) = tL(1/t)[1 + o(1)]$ or $L(x) \sim \mathbb{E}_{\mu_\phi}[\psi \vee (-x)]$. We can now reverse the steps of the proofs of case 2, and then of case 1, to obtain $\int_0^\infty(1 - e^{-tx})dF(x) \sim t|L(1/t)|$. This, by (4), implies the desired expansion of $P_{\text{top}}(\phi + t\psi)$.

Now suppose that $1 < \alpha < 2$ and $\mu_\phi[\psi < -x] \sim -\frac{x^{-\alpha}}{\Gamma(1-\alpha)}L(x)$. Since $\sup\psi < \infty$, this implies that $\psi \in L^1$, $\psi \notin L^2$, and that $\mu_\phi[|\psi| > x] \sim -\frac{x^{-\alpha}}{\Gamma(1-\alpha)}L(x)$. We subtract a constant from ψ to ensure that $\mathbb{E}_{\mu_\phi}[\psi] = 0$ (this does not affect the previous assertions). Reversing the asymptotic analysis in case 3, we see that $\mathbb{E}_{\mu_\phi}[e^{-t|\psi|}|\psi|^2] \sim \alpha(\alpha-1)t^{\alpha-2}L(x)$, whence $\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi^2] \sim \alpha(\alpha-1)t^{\alpha-2}L(x)$ (these quantities diverge because $\psi \notin L^2$, and differ by $O(1)$ because $\sup\psi < \infty$). Integrating this relation (using $\mathbb{E}_{\mu_\phi}[\psi] = 0$) we deduce $\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi] \sim \alpha t^{\alpha-1}L(1/t)$ as $t \to 0^+$. By (8)

$$\lambda'(t) \sim \alpha t^{\alpha-1}L(1/t)$$

(all terms on the right hand side of (8) are $O(t)$ except $\lambda'(t)$.) Integrating once more gives by Karamata's theorem $\lambda(t) - 1 \sim t^\alpha L(1/t)$. Since $\lambda(t) = \exp P_{\text{top}}(\phi + t\psi)$ and $P_{\text{top}}(\phi) = 0$, this implies $P_{\text{top}}(\phi + t\psi) = t^\alpha L(1/t)[1 + o(1)]$.

Suppose $\alpha = 2$, $\psi \notin L^2$, and $L(x) \sim \frac{1}{2}\mathbb{E}_{\mu_\phi}[\psi^2 1_{[|\psi|\leq x]}]$. By Karamata's Tauberian theorem, $\int_0^\infty e^{-tx}x^2dF(x) = 2[1+o(1)]L(1/t)$ as $t \to 0^+$, where $F(x) = \mu_\phi[|\psi| \leq x]$.

Integrating both sides w.r.t. t over (t_0, ∞) gives

$$\int_0^\infty e^{-t_0 x} x dF(x) = 2\int_{t_0}^\infty [1+o(1)]L(1/t)dt \equiv 2\int_0^{1/t_0} \frac{[1+o(1)]L(s)}{s^2}ds.$$

It follows that

$$\mathbb{E}_{\mu_\phi}\left[|\psi|\right] = \lim_{t_0\to 0^+}\int_0^\infty e^{-t_0 x} x dF(x) = 2\int_0^\infty \frac{[1+o(1)]L(s)}{s^2}ds < \infty,$$

where the last integral converges at infinity because of the slow variation of L. Now that we know that $\psi \in L^1$ we can assume w.l.o.g that $\mathbb{E}_{\mu_\phi}[\psi] = 0$ (the reader can check that $\mathbb{E}_{\mu_\phi}[(\psi - c)^2 1_{[\psi \le x+c]}]$ is still asymptotic to $L(x)$).

The reader may verify that $\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi^2] \sim \mathbb{E}_{\mu_\phi}[e^{-t|\psi|}\psi^2]$ as $t \to 0^+$, using the assumptions $\sup\psi < \infty$ and $\mathbb{E}_{\mu_\phi}[\psi^2] = \infty$. We have already seen that $\mathbb{E}_{\mu_\phi}[e^{-t|\psi|}\psi^2] = 2[1+o(1)]L(1/t)$ as $t \to 0^+$, because of Karamata's Tauberian theorem, and so $\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi^2] \sim 2L(1/t)$ as $t \to 0^+$. Integrating this gives (since $\mathbb{E}_{\mu_\phi}[\psi] = 0$), $\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi] \sim 2tL(1/t)$ as $t \to 0^+$. We can now deduce the asymptotic expansion of $P_{\rm top}(\phi + t\psi)$ from (8) as before.

It remains to treat the case $\alpha = 2$ and $\psi \in L^2$. Without loss of generality, $\mathbb{E}_{\mu_\phi}[\psi] = 0$. We must prove that $P_{\rm top}(\phi + t\psi) = \frac{\sigma^2}{2}t^2 + o(t^2)$ for some $\sigma \in \mathbb{R}$. Define $L(x)$ by the relation

$$\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi] = 2t\left(L(1/t) - \mathbb{E}_{\mu_\phi}[\psi h_0']\right).$$

By (8), $\lambda'(t) = 2t\Big(L(1/t) + \frac{1-\lambda(t)}{2t}\left(\mathbb{E}_{\mu_\phi}[h_0'] + o(1)\right) + o(1)\Big)$. Recalling that $\lambda(t) = \exp P_{\rm top}(\phi + t\psi)$ and that $P_{\rm top}(\phi + t\psi) = o(t)$ by theorem 4 and the assumption $\mathbb{E}_{\mu_\phi}[\psi] = 0$, we deduce that

$$\lambda'(t) = 2tL(1/t) + o(t).$$

Next, observe that $L(x) \xrightarrow[x\to\infty]{} \frac{1}{2}\mathbb{E}_{\mu_\phi}[\psi^2] + \mathbb{E}_{\mu_\phi}[\psi h_0'] =: \frac{1}{2}\sigma_0$, because $\frac{1}{t}\mathbb{E}_{\mu_\phi}[e^{t\psi}\psi] = \mathbb{E}_{\mu_\phi}[\frac{e^{t\psi}-1}{t\psi}\cdot\psi^2] \to \mathbb{E}_{\mu_\phi}[\psi^2]$ by the dominated convergence theorem. Consequently,

$$\lambda'(t) = \sigma_0 t + o(t).$$

Integrating over $(0,t]$ gives $\lambda(t) - 1 = \frac{\sigma_0}{2}t^2 + o(t^2)$. Now $\lambda(t) - 1 = e^{P_{\rm top}(\phi+t\psi)} - 1 \sim P_{\rm top}(\phi + t\psi)$, so also $P_{\rm top}(\phi + t\psi) = \frac{1}{2}\sigma_0 t^2 + o(t^2)$. The convexity of the topological pressure forces σ_0 to be non-negative. We may therefore write $\sigma_0 = \sigma^2$ for some $\sigma \in \mathbb{R}$, and (1) is proved. $\square$

Final Remarks. Our analysis is simplified by the assumption that $\sup\psi < \infty$. This assumption allows us to use Laplace transforms rather than Fourier transforms as in [**AD**], and this enables us to use the full force of the theory of regular variation. It is likely that $\sup\psi < \infty$ can be relaxed to the (more cumbersome) assumption that $\exists t > 0$ for which $\mathbb{E}_{\mu_\phi}[e^{t\psi}] < \infty$ (I did not check). It makes no sense to go further and consider ψ without exponential moments, because for such ψ's the BIP property implies $P_{\rm top}(\phi + t\psi) = \infty$ for all $t > 0$, and critical exponents are meaningless.

4. Inducing

Every countable Markov shift induces a topological Markov shift with the BIP property, in a sense that is explained below. The proof of theorems 2, 3, and 4 for systems without the BIP property uses this technique to reduce the general case to the BIP case. In this section we explain how to relate information on distributional convergence and asymptotic expansions for the pressure for the original system to that for the induced system.

Inducing. Let $(X, \mathcal{B}, m)$ be a probability space, and $T : X \to X$ a measurable map. Assume that T is probability preserving and ergodic. Fix some $A \in \mathcal{B}$ with positive measure. By Poincaré's Recurrence Theorem, the following functions are finite almost everywhere:

$$r(x) := \min\{n \geq 1 : T^n(x) \in A\};$$

$$\varphi(x) := 1_A(x) \min\{n \geq 1 : T^n(x) \in A\}.$$

The *induced map* on A is $T_A(x) := T^{\varphi(x)}(x)$, defined on the measure space $(A, \mathcal{B}_A, m_A)$ where $\mathcal{B}_A := \{E \in \mathcal{B} : E \subseteq A\}$ and

$$m_A(E) := m(E|A) \equiv \frac{m(E \cap A)}{m(A)}.$$

The following facts are classical (we are assuming that m is ergodic and invariant):

1. m_A is ergodic and invariant w.r.t. T_A;
2. Kac' Formula: $\int_X f dm = \int_A \left(\sum_{k=0}^{\varphi-1} f \circ T^k\right) dm$. In particular, $\mathbb{E}_{m_A}[\varphi] = 1/m(A)$;
3. Abramov's Formula: $h_m(T) = m(A) h_{m_A}(T_A)$.

Inducing Distributional Limit Theorems. Let T be an ergodic probability preserving transformation on a standard probability space $(X, \mathcal{B}, m)$, fix a set of positive measure $A \in \mathcal{B}$, and define $r(x)$, $\varphi(x)$, $(A, \mathcal{B}_A, m_A, T_A)$ as above. Set:

$$\begin{aligned} \overline{\varphi}_n &:= 1_A(x) \sum_{k=0}^{n-1} \varphi \circ T_A^k \\ \overline{r}_n &:= r + \overline{\varphi}_{n-1} \circ T^r. \end{aligned}$$

Melbourne & Török [**MT**] related the Central Limit Theorem for Birkhoff sums of T_A to that for Birkhoff sums of T (see also Gouëzel [**Gou2**]). The following theorem generalizes their result to other distributional limit theorems:

Theorem 7. *Suppose $\exists B_n$ s.t. $\frac{1}{B_n}[\overline{\varphi}_n - n/m(A)]$ is tight on $(A, \mathcal{B}_A, m_A)$. Set $\overline{\psi} := \sum_{k=0}^{\varphi-1} \psi \circ T^k$. If B_n is regularly varying of index $0 < \rho \neq 1$, and $\psi \vee 0 \in L^1$ or $\psi \wedge 0 \in L^1$, then the following are equivalent:*

1. $\frac{1}{B_n} \sum_{k=0}^{n-1} \overline{\psi} \circ T_A^k$ *converges in distribution on* $(A, \mathcal{B}_A, m_A)$*;*
2. $\frac{1}{B_n} \sum_{k=0}^{n-1} \psi \circ T^k$ *converges in distribution on* $(X, \mathcal{B}, m)$*.*

If $\exists \epsilon_0 > 0$ s.t. $\frac{1}{n^{1-\epsilon_0}}[\overline{\varphi}_n - \frac{n}{m(A)}]$ is tight on $(A, \mathcal{B}_A, m_A)$, then the conclusion holds for $\rho = 1$.

Proof. We assume w.l.o.g. that T is invertible (otherwise, pass to the natural extension of T). Of course, if T is invertible, then T_A is invertible. Invertibility allows us to define:

$$\psi_n := \begin{cases} \sum_{k=0}^{n-1} \psi \circ T^k & n > 0 \\ 0 & n = 0 \\ -\psi_{|n|} \circ T^n & n < 0 \end{cases} \text{ and } \overline{\psi}_n := \begin{cases} \sum_{k=0}^{n-1} \overline{\psi} \circ T_A^k & n > 0 \\ 0 & n = 0 \\ -\overline{\psi}_{|n|} \circ T_A^n & n < 0. \end{cases}$$

With these conventions, $\psi_{n+m} = \psi_n + \psi_m \circ T^n$ on X, and $\overline{\psi}_{n+m} = \overline{\psi}_n + \overline{\psi}_m \circ T_A^n$ on A for all $m, n \in \mathbb{Z}$.

Given $x \in A$, let $n[x,N]$ be the unique integer such that $\overline{\varphi}_{n[x,N]}(x) \leq N < \overline{\varphi}_{n[x,N]+1}(x)$ (this makes sense almost everywhere in A). Note that $\frac{\overline{\varphi}_{n[x,N]}(x)}{n[x,N]} \leq \frac{N}{n[x,N]} < \frac{\overline{\varphi}_{n[x,N]+1}(x)}{n[x,N]}$. By the ergodic theorem, $\frac{\overline{\varphi}_\ell}{\ell} \xrightarrow[\ell \to \infty]{} \mathbb{E}_{m_A}[\overline{\varphi}]$, and by Kac' formula $\mathbb{E}_{m_A}[\overline{\varphi}] = 1/m(A)$. It follows that

$$n[x,N] \sim N_A := [Nm(A)] \text{ almost everywhere, as } N \to \infty.$$

Here is an outline of the proof. We start, as in [**MT**], from the following identity on A:

$$\frac{\psi_N}{B_N} = \frac{B_{N_A}}{B_N}\left[\frac{\overline{\psi}_{N_A}}{B_{N_A}} + \frac{1}{B_{N_A}}\left(\overline{\psi}_{n[x,N]} - \overline{\psi}_{N_A}\right)\right] + \frac{1}{B_N}\psi_{N-\overline{\varphi}_{n[x,N]}(x)} \circ T^{\overline{\varphi}_{n[x,N]}(x)}. \tag{10}$$

We shall prove below that $\frac{B_{N_A}}{B_N} \to m(A)^\rho$ (step 1), $\frac{1}{B_{N_A}}\left(\overline{\psi}_{n[x,N]} - \overline{\psi}_{N_A}\right) \xrightarrow[n\to\infty]{\text{dist.}} 0$ on $(A, \mathcal{B}_A, m_A)$ (step 2), and $\frac{1}{B_N}\psi_{N-\overline{\varphi}_{n[x,N]}(x)}(T^{\overline{\varphi}_{n[x,N]}}x) \xrightarrow[n\to\infty]{\text{dist.}} 0$ on $(A, \mathcal{B}_A, m_A)$ (step 3). This implies that $\frac{1}{B_N}\overline{\psi}_N$ converges in distribution on $(A, \mathcal{B}_A, m_A)$ iff $\frac{1}{B_N}\psi_N$ converges in distribution on $(A, \mathcal{B}_A, m_A)$. Eagleson's theorem on distributional convergence implies that $\frac{1}{B_N}\psi_N$ converges in distribution on $(A, \mathcal{B}_A, m_A)$ iff it converges in distribution on $(X, \mathcal{B}, m)$ (step 4). The theorem follows.

Step 1. $\frac{B_{N_A}}{B_N} \xrightarrow[N\to\infty]{} m(A)^\rho$.

Proof. Use the uniform convergence theorem for slow variation (appendix A).

Step 2. If (1) or (2) in theorem 7 hold, then $W_N := \frac{1}{B_{N_A}}\left(\overline{\psi}_{n[x,N]} - \overline{\psi}_{N_A}\right) \xrightarrow[N\to\infty]{\text{dist.}} 0$ on $(A, \mathcal{B}_A, m_A)$. (This is a generalization of Lemma 3.4 in [**MT**].)

Proof. Set $m_0[x,N] := n[x,N] - N_A$ and $m[x,N] := m_0[T_A^{-N_A}x, N]$. By step 1, it is enough to show that $\frac{1}{B_N}\overline{\psi}_{m_0[x,N]}(T_A^{N_A}x) \xrightarrow[N\to\infty]{\text{dist.}} 0$ on A. This the same as

$$\frac{1}{B_N}\overline{\psi}_{m[x,N]} \xrightarrow[N\to\infty]{\text{dist.}} 0 \text{ on } A, \tag{11}$$

because T_A is measure preserving.

Case 1. $\psi \in L^1$.

Suppose first that $\int \psi dm \neq 0$. By Kac' formula, if $\psi \in L^1(X)$, then $\overline{\psi} \in L^1(A)$ and $\int_X \psi dm = m(A)\int_A \overline{\psi} dm_A$. By the ergodic theorem, $\frac{\psi_N}{N}, \frac{\overline{\psi}_N}{N}$ converge pointwise, whence in distribution, to their means. These means are different (otherwise $m(A) = 1$ and there is nothing to prove). Therefore, if $\limsup_{n\to\infty} \frac{n}{B_n} = 0$ then (1) and

(2) both hold, and if $\limsup_{n\to\infty} \frac{n}{B_n} > 0$, $\int \psi \neq 0, \int \overline{\psi} \neq 0$, then both (1) and (2) fail. We may therefore restrict ourselves to the case $\int \psi = \int \overline{\psi} = 0$, $0 < \rho \leq 1$ (if $\rho > 1$ then $\frac{n}{B_n} \to 0$).

Fix some N_0 and $\epsilon > 0$ to be determined later.

$$\begin{aligned}
m_A\left[\tfrac{1}{B_N}|\overline{\psi}_{m[x,N]}| > t\right] &\leq m_A\left[m[x,N] \leq N_0, \tfrac{1}{B_N}\sum_{k=-N_0}^{N_0} |\psi|\circ T^k > t\right] + \\
&\quad + m_A\left[m[x,N] \geq N_0, |\tfrac{m[x,N]}{B_N}|\cdot|\tfrac{1}{m[x,N]}\overline{\psi}_{m[x,N]}| > t\right] \\
&\leq m_A\left[\sum_{k=-N_0}^{N_0} |\psi|\circ T^k > tB_N\right] \\
&\quad + m_A\left[m[x,N] \geq N_0, |\tfrac{1}{m[x,N]}\overline{\psi}_{m[x,N]}| > \epsilon\right] \\
&\quad + m_A\left[|\tfrac{m[x,N]}{B_N}| > t/\epsilon\right].
\end{aligned}$$

The first summand is $o(1)$ as $N \to \infty$. The second summand can be made less than ϵ by choosing N_0 sufficiently large, because $\frac{\overline{\psi}_\ell}{\ell} \xrightarrow[\ell\to\infty]{} \mathbb{E}_{m_A}[\overline{\psi}] = 0$ almost surely, whence uniformly outside a set of measure ϵ. Since $m[\cdot, N]$, $m_0[\cdot, N]$ are equal in distribution on $(A, \mathcal{B}_A, m_A)$, this leaves us with

$$m_A\left[\tfrac{1}{B_N}|\overline{\psi}_{m[x,N]}| > t\right] \leq o(1) + \epsilon + m_A\left[|\tfrac{m_0[x,N]}{B_N}| > t/\epsilon\right] \text{ as } N \to \infty.$$

Since ϵ is arbitrary, (11) reduces to the tightness of $\frac{m_0[x,N]}{B_N}$.

When $\rho \in (0,1)$ we argue as follows. By the definition of $m_0[x,N]$ and $n[x,N]$, $m_0[x,N] > tB_N \Leftrightarrow n[x,N] > [tB_N] + N_A =: \alpha_N(t) \Rightarrow \overline{\varphi}_{\alpha_N(t)} < N$. Therefore,

$$\begin{aligned}
m_A\left[\frac{m_0[x,N]}{B_N} > t\right] &\leq m_A[\overline{\varphi}_{\alpha_N(t)} < N] = \\
&= m_A\left[\frac{\overline{\varphi}_{\alpha_N(t)} - \frac{\alpha_N(t)}{m(A)}}{B_{\alpha_N(t)}} < \beta_N(t)\right], \text{ where } \beta_N(t) := \frac{N - \frac{\alpha_N(t)}{m(A)}}{B_{\alpha_N(t)}}.
\end{aligned}$$

But B_N is regularly varying of index $\rho \in (0,1)$, so $\beta_N(t) \xrightarrow[N\to\infty]{} -\frac{t}{m(A)^{\rho+1}}$. Using this, and the assumption that $\frac{1}{B_N}[\overline{\varphi}_N - N/m(A)]$ is tight, it is easy to see to see that for every $\epsilon > 0$, $\exists t$ so large that $m_A\left[\frac{m_0[x,N]}{B_N} > t\right] < \epsilon$ for all N. A similar estimate of $m_A\left[\frac{m_0[x,N]}{B_N} < t\right]$ for $t \ll 0$ finishes the proof of tightness when $\rho \in (0,1)$.

Now suppose $\rho = 1$. Let ϵ_0 be as in the statement of the theorem. Repeating the same argument, we see that

$$\begin{aligned}
m_A\left[\frac{m_0[x,N]}{B_N} > t\right] &\leq m_A[\overline{\varphi}_{\alpha_N(t)} < N] = \\
&= m_A\left[\frac{\overline{\varphi}_{\alpha_N(t)} - \frac{\alpha_N(t)}{m(A)}}{\alpha_N(t)^{1-\epsilon_0}} < \gamma_N(t)\right], \text{ where } \gamma_N(t) := \frac{N - \frac{\alpha_N(t)}{m(A)}}{\alpha_N(t)^{1-\epsilon_0}}.
\end{aligned}$$

Calculating, we see that

$$\gamma_N = \frac{Nm(A) - [tB_N] - N_A}{m(A)\left([tB_N] + N_A\right)^{1-\epsilon_0}} \sim \frac{-t\frac{B_N}{N}N^{\epsilon_0}}{m(A)\left(t\frac{B_N}{N} + m(A) + o(1)\right)^{1-\epsilon_0}}.$$

Since B_N is regularly varying of index 1, B_N/N is slowly varying. It follows that γ_n is minus a regularly varying sequence of index $\epsilon_0 > 0$, whence $\gamma_N \to -\infty$. Since $\frac{1}{N^{1-\epsilon_0}}[\overline{\varphi}_N - n/m(A)]$ is tight, by assumption, we get $m_A\left[\frac{m_0[x,N]}{B_N} > t\right] \to 0$ for all $t > 0$. A similar argument shows that $m_A\left[\frac{m_0[x,N]}{B_N} < t\right] \to 0$ for all t negative, and we obtain the tightness of $\frac{m_0[x,N]}{B_N}$ when $\rho = 1$. (11) follows.

Case 2. $\psi \notin L^1$.

We prove (11) when $\psi \notin L^1$. By our assumptions one of $\psi \vee 0$, $\psi \wedge 0$ is integrable. Without loss of generality, $\int \psi \vee 0 < \infty$ and $\int \psi \wedge 0 = -\infty$.

By the ergodic theorem $\frac{\psi_N}{N}, \frac{\overline{\psi}_N}{N} \xrightarrow[N\to\infty]{} -\infty$ almost surely, so either (1) and (2) are both false or $N/B_N \to 0$. We restrict ourselves to this case. By the ergodic theorem, $\frac{\psi_N}{B_N} = \frac{(\psi\wedge 0)_N}{B_N} + o(1)$ and $\frac{\overline{\psi}_N}{B_N} = \frac{\overline{(\psi\wedge 0)}_N}{B_N} + o(1)$. We may therefore also assume without loss of generality that $\psi \leq 0$.

We begin by showing that if (1) or (2) holds, then $\frac{\overline{\psi}_N}{B_N}$ is tight. When (1) holds, this is clear, so suppose (2) holds. In this case $\frac{\psi_N}{B_N}$ is tight, and since ψ doesn't change sign,

$$\begin{aligned} m_A\left[\left|\frac{\overline{\psi}_N}{B_N}\right| > t\right] &\leq m_A\left[\left|\frac{\psi_{\overline{\varphi}_{n[x,N]+1}}}{B_N}\right| > t\right] \\ &\leq m_A\left[\overline{\varphi}_{n[x,N]+1} \leq 2N \text{ and } \left|\frac{\psi_{2N}}{B_N}\right| > t\right] + m_A\left[\overline{\varphi}_{n[x,N]+1} > 2N\right] \\ &\leq \frac{1}{m(A)} m\left[\left|\frac{\psi_{2N}}{B_{2N}}\right| > t\frac{B_N}{B_{2N}}\right] + m_A\left[\frac{\overline{\varphi}_{n[x,N]+1}}{N_A} - \frac{1}{m(A)} > \frac{1}{m(A)}\right]. \end{aligned}$$

B_N is regularly varying, so $\frac{B_{2N}}{B_N} \xrightarrow[N\to\infty]{} 2^\rho$. $\frac{\psi_N}{B_N}$ is tight, so we can make the first summand uniformly small by choosing t large. The second summand tends to zero as $N \to \infty$, because by the ergodic theorem $\frac{\overline{\varphi}_{n[x,N]+1}}{N_A} \sim \frac{\overline{\varphi}_{n[x,N]+1}}{n[x,N]+1} \xrightarrow[N\to\infty]{} \frac{1}{m(A)}$ a.e., whence $\frac{\overline{\varphi}_{n[x,N]+1}}{N_A} - \frac{1}{m(A)} \xrightarrow[N\to\infty]{\text{dist.}} 0$. This proves tightness.

Next we observe that $\frac{m_0[x,N]}{B_N} \xrightarrow[N\to\infty]{\text{dist.}} 0$ on A, because $|m_0[x,N]| \leq n[x,N] + N_A \leq N[1 + m(A)]$ and $N/B_N \to 0$ by assumption. Since $m_0[\cdot,N]$ $m[\cdot,N]$ are equal in distribution w.r.t m_A, $\frac{m[x,N]}{B_N} \xrightarrow[N\to\infty]{\text{dist.}} 0$ on A.

Since the sign of ψ is constant, for every $\epsilon > 0$

$$\begin{aligned} & m_A\left[\left|\frac{\overline{\psi}_{m[x,N]}}{B_N}\right| > t, m[x,N] < 0\right] \leq \\ &\quad \leq m_A\left[\frac{m[x,N]}{B_N} \in [-\epsilon, 0], \left|\frac{\overline{\psi}_{-m[x,N]}\circ T_A^{m[x,N]}}{B_N}\right| > t\right] + m_A\left[\left|\frac{m[x,N]}{B_N}\right| \geq \epsilon\right] \\ &\quad \leq m_A\left[\left|\frac{\overline{\psi}_{[\epsilon B_N]}\circ T_A^{-[\epsilon B_N]}}{[\epsilon B_N]}\right| > t/\epsilon\right] + m_A\left[\left|\frac{m[x,N]}{B_N}\right| \geq \epsilon\right] \\ &\quad = m_A\left[\left|\frac{\overline{\psi}_{[\epsilon B_N]}}{[\epsilon B_N]}\right| > t/\epsilon\right] + m_A\left[\left|\frac{m[x,N]}{B_N}\right| \geq \epsilon\right], \text{ because } m_A \circ T_A^{[\epsilon B_N]} = m_A. \end{aligned}$$

$$m_A\left[\left|\frac{\overline{\psi}_{m[x,N]}}{B_N}\right| > t \text{ and } m[x,N]\geq 0\right] \leq$$

$$\begin{aligned}&\leq \quad m_A\left[0\leq \frac{m[x,N]}{B_N}\leq\epsilon \text{ and } \left|\frac{\overline{\psi}_{m[x,N]}}{B_N}\right|>t\right]+m_A\left[\frac{m[x,N]}{B_N}\geq\epsilon\right]\\ &\leq \quad m_A\left[\left|\frac{\overline{\psi}_{[\epsilon B_N]}}{[\epsilon B_N]}\right|>t/\epsilon\right]+m_A\left[\frac{m[x,N]}{B_N}\geq\epsilon\right];\end{aligned}$$

Putting this all together, we get

$$m_A\left[\left|\frac{\overline{\psi}_{m[x,N]}}{B_N}\right|>t\right]\leq 2m_A\left[\left|\frac{\overline{\psi}_{[\epsilon B_N]}}{[\epsilon B_N]}\right|>t/\epsilon\right]+2m_A\left[\left|\frac{m[x,N]}{B_N}\right|\geq\epsilon\right].$$

Fix $\delta>0$. Since $\frac{\overline{\psi}_k}{B_k}$ is tight, there exists ϵ so small that the first summand is less than δ for all N. Since $\frac{m[x,N]}{B_N}\xrightarrow[N\to\infty]{\text{dist.}}0$, there exists N_0 s.t. the second summand is less than δ for all $N>N_0$. We deduce that $m_A\left[\left|\frac{\overline{\psi}_{m[x,N]}}{B_N}\right|>t\right]<2\delta$ for N large enough, proving (11) in case 2.

This completes the proof of step 2.

Step 3. $\frac{1}{B_N}\psi_{N-\overline{\varphi}_{n[x,N]}(x)}(T^{\overline{\varphi}_{n[x,N]}(x)}x)\xrightarrow[N\to\infty]{}0$ in distribution on $(A,\mathcal{B}_A,m_A)$.

Proof. We thank the referee for the following short argument. Recall the definition of r from the beginning of section 4, and set $S(x):=T_A^{-1}(T^{r(x)}(x))$ $(x\in X)$. Then

$$|\psi_{N-\overline{\varphi}_{n[x,N]}(x)}(T^{\overline{\varphi}_{n[x,N]}(x)}x)|\leq\Psi(T^Nx),\ \text{where } \Psi(x):=\sum_{k=0}^{\varphi(Sx)}|\psi(T^kSx)|.$$

Now $\frac{1}{B_N}\Psi\circ T^N\xrightarrow[N\to\infty]{\text{dist.}}0$ on $(X,\mathcal{B},m)$, because $m\circ T^{-1}=m$ and $B_N\to\infty$. It follows that $\frac{1}{B_N}\Psi\circ T^N\xrightarrow[N\to\infty]{\text{dist.}}0$ on $(A,\mathcal{B}_A,m_A)$.

Steps 1–3 and (10) show that $\frac{\overline{\psi}_N}{B_N}$ converges in distribution on $(A,\mathcal{B}_A,m_A)$ iff $\frac{\psi_N}{B_N}$ converges in distribution on $(A,\mathcal{B}_A,m_A)$.

Step 4. $\frac{\psi_N}{B_N}$ converges in distribution on $(A,\mathcal{B}_A,m_A)$ iff $\frac{\psi_N}{B_N}$ converges in distribution on $(X,\mathcal{B},m)$, and the limiting distribution is the same.

Proof. Eagleson proves that if X_i is a stationary ergodic stochastic process and $Y_n:=\frac{1}{B_n}(X_1+\cdots+X_n)$ converges in distribution for some $B_n\uparrow\infty$ on $(\Omega,\mathcal{F},\mu)$, then Y_n converges in distribution to the same limit on $(\Omega,\mathcal{F},\mu')$ for all $\mu'\ll\mu$ ([**Ea**], theorem 4). This proves ($\Leftarrow$).

To see the other direction, assume $\frac{\psi_N}{B_N}$ converges in distribution on $(A,\mathcal{B}_A,m_A)$, and consider the following decomposition on $(X,\mathcal{B},m)$ in the limit $N\to\infty$:

$$\begin{aligned}\frac{\psi_N}{B_N} &= 1_{[r<N]}\left\{\frac{\psi_N}{B_N}\circ T^r+\left(\frac{\psi_r}{B_N}-\frac{\psi_{r\circ T^{-N}}}{B_N}\circ T^N\right)\right\}+1_{[r\geq N]}O\left(\frac{|\psi|_r}{B_N}\right)\\ &= 1_{[r<N]}\frac{\psi_N}{B_N}\circ T^r+O\left(\frac{|\psi|_r}{B_N}\right)+O\left(\frac{|\psi|_{r\circ T^{-N}}\circ T^N}{B_N}\right).\end{aligned}$$

The big-Oh terms converge to zero in distribution, and $1_{[r<N]}$ converges a.s. (whence in distribution) to 1. The m–distribution of $\frac{\psi_N}{B_N}\circ T^r$ is equal to the $m\circ(T^r)^{-1}$–distribution of $\frac{\psi_N}{B_N}$. Since $m\circ(T^r)^{-1}\ll m_A$, $\frac{\psi_N}{B_N}\circ T^r$ converges in distribution on $(X,\mathcal{B},m)$ to its m_A–distributional limit. ($\Rightarrow$) follows. □

Remark: The proof shows that the distributional limit of $\frac{\psi_N}{B_N}$ is a $m(A)^\rho$–scaled version of the distributional limit of $\frac{\overline{\psi}_N}{B_N}$, see Step 1.

Inducing Asymptotic Expansions. Throughout this section, let $(\Sigma_{\mathbb{A}}^+,T)$ be a topologically mixing countable Markov shift with set of states S, and let $A\subset S$ be some *finite* union of states.

Define $\varphi(x)$ and $T_A(x):=T^{\varphi(x)}(x)$ as above. The resulting map can be given the structure of of countable Markov shift as follows:

1. *States*: $\overline{S}:=\{[a,\xi_1,\dots,\xi_{n-1},b]:a,b\in A,\ n\geq 1,\ \xi_i\notin A\text{ for all }i\}\setminus\{\varnothing\}$;
2. *Transition matrix*: $\overline{\mathbb{A}}=(t_{[\underline{a}],[\underline{b}]})_{\overline{S}\times\overline{S}}$ with $t_{[\underline{a}],[\underline{b}]}=1$ iff the last symbol in $\underline{a}$ is the first symbol in $\underline{b}$.

We call this shift the *induced shift* (on A), because it is conjugate to the induced map. The conjugacy is $\pi:\Sigma_{\overline{\mathbb{A}}}^+\hookrightarrow A$ given by

$$\pi([a^{(1)},\underline{\xi}^{(1)},b^{(1)}],[a^{(2)},\underline{\xi}^{(2)},b^{(2)}],\dots)=(a^{(1)},\underline{\xi}^{(1)},a^{(2)},\underline{\xi}^{(2)},a^{(3)},\dots).$$

It is easy to verify that the induced shift satisfies the BIP property.

Every $f:\Sigma_{\mathbb{A}}^+\to\mathbb{R}$ induces a function $\overline{f}:\Sigma_{\overline{\mathbb{A}}}^+\to\mathbb{R}$ by $\overline{f}:=\left(\sum_{k=0}^{\varphi-1}f\circ T^k\right)\circ\pi$. We call this function the *induced function* (by f). Define

$$\mathcal{H}_A:=\{f:\Sigma_{\mathbb{A}}^+\to\mathbb{R}|f\text{ has summable variations, }\sup f<\infty\text{ and }\overline{f}\text{ is locally Hölder continuous}\}.$$

It is easy to see that $\mathcal{H}_A$ contains all weakly Hölder continuous functions which are bounded from above.

Theorem 8. *Suppose $\phi,\psi\in\mathcal{H}_A$, and that ψ satisfies (Ψ) with respect to a finite set of states A. If $\{\phi+t\psi\}_{t\geq 0}$ is regular and $P_{top}(\phi)=\mathbb{E}_{\mu_\phi}[\psi]=0$, then*

$$P_{top}(\overline{\phi+t\psi})=\frac{1+o(1)}{\mu_\phi(A)}P_{top}(\phi+t\psi)\ \text{ as }t\to 0^+,$$

where μ_ϕ is the equilibrium measure of ϕ.

Lemma 1. *If $f:\Sigma_{\mathbb{A}}^+\to\mathbb{R}$ belongs to $\mathcal{H}_A$, then*

1. $\mathrm{var}_n(\overline{f})\leq\sum_{k=n+1}^\infty\mathrm{var}_k(f)$;
2. *If, in addition, $P_{top}(f)<\infty$, then $\sup\overline{f-P_{top}(f)}<\infty$;*
3. *If, in addition, f has an equilibrium measure μ, then $P_{top}(\overline{f-P_{top}(f)})=0$, and $\overline{\mu}(E):=\frac{(\mu\circ\pi)(E)}{\mu(A)}$ is an equilibrium measure for $\overline{f-P_{top}(f)}$.*

Proof. Suppose $\overline{x},\overline{y}\in\Sigma_{\overline{\mathbb{A}}}^+$ agree on the first n symbols, and write $\overline{x}=\pi(x),\overline{y}=\pi(y)$. Since $\varphi\circ\pi$ is constant on partition sets in $\Sigma_{\overline{\mathbb{A}}}^+$, $\varphi(x)=\varphi(y)=n_0$. One checks that $x,y\in\Sigma_{\mathbb{A}}^+$ agree on the first $\varphi(x)+\varphi(T_Ax)+\cdots\varphi(T_A^{n-1}x)+1$ symbols (the one at the end is because of the last symbol of the last cylinder). We see that x, y agree on (at least) the first $n_0 + (n-1) + 1 = n_0 + n$ symbols, and so

$$|\overline{f}(\overline{x}) - \overline{f}(\overline{y})| \leq \sum_{k=0}^{n_0-1} \left|f(T^k x) - f(T^k y)\right| \leq \sum_{k=n+1}^{n_0+n} \mathrm{var}_k(f).$$

Part (1) follows.

To see part (2), construct a finite set of admissible words $\{\underline{w}_{ab} : a, b \in A\}$ of length n_{ab} (as words in the alphabet $\overline{S}$) such that $\underline{w}_{ab}$ starts with a and ends with b. Such words exist because of the topological mixing of $\Sigma_{\mathbb{A}}^+$. Set

$$C := \sup\{|(\overline{f - P_{\mathrm{top}}(f)})_{n_{ab}}(\overline{x})| : \overline{x} \in [\underline{w}_{ab}], a, b \in A\}.$$

By part (1), $C < \infty$.

We show that $\sup \overline{f - P_{\mathrm{top}}(f)} \leq C + \sum_{n=2}^{\infty} \mathrm{var}_n(f) =: C_0$. Otherwise $\exists \overline{x} \in \Sigma_{\overline{\mathbb{A}}^+}$ for which $\overline{f - P_{\mathrm{top}}(f)}(\overline{x}) > C_0$. By part (1),

$$\overline{f - P_{\mathrm{top}}(f)} > C \text{ on the partition set which contains } \overline{x}.$$

Denote this partition set by $[\overline{x}_0]$, write $\overline{x}_0 = [b, \underline{\xi}, a]$, and consider the point $\overline{z} := (\overline{x}_0, w_{ab}, \overline{x}_0, w_{ab}, \overline{x}_0, w_{ab}, \ldots)$. This is a periodic point of order $1 + n_{ab}$, and $(\overline{f - P_{\mathrm{top}}(f)})_{1+n_{ab}}(\overline{z}) > C - C = 0$.

Write $z = \pi(\overline{z})$. Then for some N, $T^N(z) = z$ and $\sum_{k=0}^{N-1}[f(T^k z) - P_{\mathrm{top}}(f)] > 0$. The measure $\mu := \frac{1}{N}\sum_{k=0}^{N-1} \delta_{T^k z}$ is T–invariant, has zero entropy, and satisfies

$$h_\mu(T) + \int_{\Sigma_{\mathbb{A}}^+} [f - P_{\mathrm{top}}(f)] d\mu = \frac{1}{N} \sum_{k=0}^{N-1} [f(T^k z) - P_{\mathrm{top}}(f)] > 0.$$

It follows that $h_\mu(T) + \int f d\mu > P_{\mathrm{top}}(f)$, in contradiction to the definition of $P_{\mathrm{top}}(f)$. Part (2) is proved.

Before proving part (3), we recall from [**BS**] that $\mu[a] \neq 0$ for any state $a \in S$ and every equilibrium measure μ of a potential with summable variations on a topologically mixing shift. Therefore, $\overline{\mu}$ is well defined.

Next we note that $\overline{\mu}$ is shift invariant, because $\mu|_A$ is T_A–invariant. The formulæ of Kac and Abramov and the conjugacy between T_A and the induced shift give

$$\begin{aligned} P_{\mathrm{top}}(\overline{f - P_{\mathrm{top}}(f)}) \geq h_{\overline{\mu} \circ \pi^{-1}}(T_A) + \int \overline{f - P_{\mathrm{top}}(f)} d\overline{\mu} = \\ = \frac{1}{\mu(A)} \left[h_\mu(T) + \int f - P_{\mathrm{top}}(f) d\mu \right] = 0. \end{aligned}$$

The other inequality is more delicate, because it is not true that every T_A–invariant probability measure is induced by a T–invariant probability measure: We can only guarantee this for T_A–invariant measures for which φ is integrable.

To deal with this difficulty, we note that since $\overline{f - P_{\mathrm{top}}(f)}$ has summable variations (part 1) and is bounded from above (part 2), then $P_{\mathrm{top}}(\overline{f - P_{\mathrm{top}}(f)})$ is equal to the Gurevich pressure of $\overline{f - P_{\mathrm{top}}(f)}$. Therefore, by theorem 2 of [**S1**],

$$P_{\mathrm{top}}(\overline{f - P_{\mathrm{top}}(f)}) = \sup\{h_m(T) + \int \overline{f - P_{\mathrm{top}}(f)} dm : m \text{ has compact support }\}.$$

For such measures $\varphi \circ \pi$ is essentially bounded, whence integrable. Therefore $P_{\mathrm{top}}(\overline{f - P_{\mathrm{top}}(f)})$ is achieved as a supremum over invariant measures which are induced by shift invariant measures on $\Sigma_{\mathbb{A}}^+$. Such measures $\overline{\nu}$ satisfy

$$h_{\overline{\nu}\circ\pi^{-1}}(T_A) + \int_{\Sigma_{\mathbb{A}}^+} \overline{f - P_{\mathrm{top}}(f)} d\overline{\nu} = \frac{1}{\nu(A)} \left[h_\nu(T) + \int_{\Sigma_{\mathbb{A}}^+} f - P_{\mathrm{top}}(f) d\nu \right] \leq 0.$$

Passing to the supremum, we get $P_{\mathrm{top}}(\overline{f - P_{\mathrm{top}}(f)}) \leq 0$.

In the first part of the proof we saw that $h_{\overline{\mu}\circ\pi^{-1}}(T_A) + \int \overline{f - P_{\mathrm{top}}(f)} d\overline{\mu} = 0$ for $\overline{\mu}$ induced by the equilibrium measure of f. Consequently, this is an equilibrium measure for $\overline{f - P_{\mathrm{top}}(f)}$ (by [**BS**] the only one), and the pressure is zero. □

Proof of Theorem 8. The convexity of $P_{\mathrm{top}}(\phi + t\psi)$ and the assumption that $P_{\mathrm{top}}(\phi) = 0$ imply that either $P_{\mathrm{top}}(\phi + t\psi) = 0$ on some right neighborhood of 0, or $P_{\mathrm{top}}(\phi + t\psi) \neq 0$ for all $t \neq 0$ small. In the first case the theorem holds trivially by lemma 1, part (3). We may therefore assume without loss of generality that $P_{\mathrm{top}}(\phi + t\psi) \neq 0$ for all $t > 0$ small.

Recall the definitions of φ, $\Sigma_{\mathbb{A}}^+$, and of the functions $\overline{\phi}, \overline{\psi}$ induced by ϕ, ψ. By assumption, A is a finite union of states such that $\psi \leq 0 = \mathbb{E}_{\mu_\phi}[\psi]$ outside A. Thus:

$$\sup \overline{\psi} < \infty.$$

To see this write $\overline{\psi} = \sum_{k=0}^{\varphi-1} \psi \circ T^k$, and observe that the first summand is dominated by $\sup \psi$, while the other summands are non-positive (they correspond to the part of the orbit which lies outside A). Note also that by lemma 1 part (2)

$$\sup \overline{\phi} < \infty.$$

Step 1. $P_{\mathrm{top}}(\phi + t\psi) > 0$ for all $t > 0$.

Proof. Kac' formula and the assumption $\mathbb{E}_{\mu_\phi}[\psi] = 0$ imply that $\mathbb{E}_{\overline{\mu_\phi}}[\overline{\psi}] = 0$, where $\overline{\mu_\phi} = \frac{\mu_\phi \circ \pi}{\mu_\phi(A)}$. By lemma 1, $P_{\mathrm{top}}(\overline{\phi}) = 0$, and $\overline{\mu_\phi}$ is the equilibrium measure of $\overline{\phi}$: $\overline{\mu_\phi} = \mu_{\overline{\phi}}$. Consequently $\mathbb{E}_{\mu_{\overline{\phi}}}[\overline{\psi}] = 0$.

By theorem 4 for BIP systems, $P_{\mathrm{top}}(\overline{\phi} + t\overline{\psi}) = o(t)$ as $t \to 0^+$ (note that the assumptions listed at the beginning of section 3 are satisfied). We see that the right–derivative of $t \mapsto P_{\mathrm{top}}(\overline{\phi} + t\overline{\psi})$ at $t = 0$ vanishes. But $t \mapsto P_{\mathrm{top}}(\overline{\phi} + t\overline{\psi})$ is convex, so $P_{\mathrm{top}}(\overline{\phi} + t\overline{\psi}) \geq 0$ for $t \geq 0$.

Lemma 1 tells us that $P_{\mathrm{top}}(\overline{\phi + t\psi - P_{\mathrm{top}}(\phi + t\psi)}) = 0$. If $P_{\mathrm{top}}(\phi + t\psi)$ were negative, then by the properties of the topological pressure and since $\varphi \geq 1$

$$0 = P_{\mathrm{top}}(\overline{\phi + t\psi} - P_{\mathrm{top}}(\phi + t\psi)\varphi) \geq P_{\mathrm{top}}(\overline{\phi + t\psi}) + |P_{\mathrm{top}}(\phi + t\psi)| > 0.$$

Therefore $P_{\mathrm{top}}(\phi + t\psi) \geq 0$ for $t > 0$. The inequality is strict, otherwise by convexity $P_{\mathrm{top}}(\phi + t\psi)$ vanishes on some right-neighbourhood of 0, in contrary to our assumptions.

Step 2. Set $f_t := \phi + t\psi$ and $\phi_t := f_t - P_{\mathrm{top}}(f_t)$. The induced potentials $\overline{\phi_t}$, $\overline{f_t}$ have Gibbs measures $\mu_{\overline{f_t}}$, $\mu_{\overline{\phi_t}}$, and $\mathbb{E}_{\mu_{\overline{\phi_t}}}[\varphi] \leq \frac{P_{\mathrm{top}}(\overline{f_t})}{P_{\mathrm{top}}(f_t)} \leq \mathbb{E}_{\mu_{\overline{f_t}}}[\varphi]$ for all $0 \leq t \leq \epsilon_0$.

Proof. Any locally Hölder continuous potential with finite pressure on a shift with the BIP property has an invariant Gibbs measure [**S3**]. Therefore, since $\Sigma_{\mathbb{A}}^+$ has the BIP property, it is enough to check that $\overline{f_t}, \overline{\phi_t}$ have finite pressure. They do, because $\sup \overline{\psi} < \infty$, $P_{\mathrm{top}}(f_t) \geq 0$ (step 1), and $P_{\mathrm{top}}(\overline{\phi}) = 0 < \infty$ (lemma 1 part 3).

Fix $\epsilon_0 > 0$ such that $f_t := \phi + t\psi$ has an equilibrium measure for $0 \leq t \leq \epsilon_0$ (regularity). Fix $0 \leq t \leq \epsilon_0$, and consider the function $p(s) := P_{\text{top}}(\overline{f_t} - s\varphi)$ for $s \geq 0$. This is a convex function, and therefore

$$p'_+(0) \leq \frac{p(P_{\text{top}}(f_t)) - p(0)}{P_{\text{top}}(f_t)} \leq p'_+(P_{\text{top}}(f_t)),$$

where p'_+ denotes one-sided derivative from the right (which can be infinite). The term in the middle is $-\frac{P_{\text{top}}(\overline{f_t})}{P_{\text{top}}(f_t)}$ (lemma 1, part (3)). Theorem 4 for BIP systems gives the one-sided derivatives (see the remark after theorem 4):

$$p'_+(0) = \frac{d}{ds}\bigg|_{s=0^+} P_{\text{top}}(\overline{f_t} + s(-\varphi)) = -\mathbb{E}_{\mu_{\overline{f_t}}}[\varphi],$$

$$p'_+(P_{\text{top}}(f_t)) = \frac{d}{ds}\bigg|_{s=0^+} P_{\text{top}}(\overline{\phi_t} + s(-\varphi)) = -\mathbb{E}_{\mu_{\overline{\phi_t}}}[\varphi].$$

This completes the proof.

Step 3. $\mathbb{E}_{\mu_{\overline{f_t}}}[\varphi] \xrightarrow[t\to 0^+]{} \frac{1}{\mu_\phi(A)}$.

Proof. We work on the BIP shift $(\Sigma^{\pm}_{\mathbb{A}}, T)$. Define as in section 3 the space $\mathcal{L}$ and the operators R_0, R_t corresponding to $\overline{\phi}$ and $\overline{\psi}$:

$$R_0(f)(\overline{x}) := \sum_{T\overline{y}=\overline{x}} e^{\overline{\phi}(\overline{y})} f(\overline{y}) \ , \ R_t(f) := R_0[e^{t\overline{\psi}} f].$$

Here and throughout T denotes the shift on $\Sigma^{\pm}_{\mathbb{A}}$.

As in the beginning of section 3, we may assume without loss of generality that $\sum_{Ty=x} e^{\overline{\phi}(y)} = 1$ (otherwise pass to $\overline{\phi} + h - h\circ T$ with some bounded locally Hölder continuous function $h : \Sigma^{\pm}_{A} \to \mathbb{R}$, and note that this does not affect $\mu_{\overline{f_t}}$ or $\mu_{\overline{\phi}}$). This reduction allows us to assume that $R_0 1 = 1$.

By proposition 2 part (3), $\|R_t - R_0\| \xrightarrow[t\to 0^+]{} 0$. It follows that the eigenprojections $P_t := P(R_t)$ are well–defined for t small, and converge in norm to $P_0 := P(R_0)$.

The operator R_t is the Ruelle operator of $\overline{f_t}$. The theory of Ruelle operators for shifts with BIP says that $\lambda(R_t) = \exp P_{\text{top}}(\overline{f_t})$ and that $P_t F = h_t \int F d\nu_t$ where ν_t is an eigenmeasure of R_t, h_t is a positive eigenfunction of R_t, and $\int h_t d\nu_t = 1$. The Gibbs measure of $\overline{f_t}$ is $h_t d\nu_t$. Consequently,

$$\int F d\mu_{\overline{f_t}} = \frac{P_t[F P_t 1]}{P_t 1} \text{ for all } F \in \mathcal{L}. \tag{12}$$

(The RHS is a scalar, because $\dim \operatorname{Im}(P_t) = 1$.)

Since $P_t \to P_0$ in norm and $P_0 1 = 1$ (because $R_0 1 = 1$), $\mathbb{E}_{\mu_{\overline{f_t}}}[F] \xrightarrow[t\to 0^+]{} \mathbb{E}_{\mu_{\overline{\phi}}}[F]$ for all $F \in \mathcal{L}$. In particular, $\mathbb{E}_{\mu_{\overline{f_t}}}[\varphi 1_{[\varphi < N]}] \xrightarrow[t\to 0^+]{} \mathbb{E}_{\mu_{\overline{\phi}}}[\varphi 1_{[\varphi<N]}]$ for every $N \in \mathbb{N}$. We claim that for every $\epsilon > 0$ there exists N such that $\mathbb{E}_{\mu_{\overline{f_t}}}[\varphi 1_{[\varphi \geq N]}] < \epsilon$ for all t in some one–sided neighbourhood of zero (uniform integrability). This will imply that $\mathbb{E}_{\mu_{\overline{f_t}}}[\varphi] \xrightarrow[t\to 0^+]{} \mathbb{E}_{\mu_{\overline{\phi}}}[\varphi]$. Step 3 will then follow from Kac' formula $\mathbb{E}_{\mu_{\overline{\phi}}}[\varphi] = 1/\mu_\phi(A)$.

To prove uniform integrability, we need the *transfer operator* of $\mu_{\overline{f_t}}$, given by

$$T_t F := \frac{\lambda(R_t)^{-1}}{P_t 1} R_t[F P_t 1].$$

It is straightforward to check, using (12), that $\mathbb{E}_{\mu_{\overline{f_t}}}[T_tF] = \mathbb{E}_{\mu_{\overline{f_t}}}[F]$ for all $F \in \mathcal{L}$. It follows that for every $a \in \overline{S}$,

$$\mu_{\overline{f_t}}[a] = \mathbb{E}_{\mu_{\overline{f_t}}}[T_t 1_{[a]}] = \lambda(R_t)^{-1} \int_{T[a]} e^{\overline{f_t}(ax)} \frac{(P_t1)(ax)}{(P_t1)(x)} d\mu_{\overline{f_t}}(x) \leq$$
$$\leq \left[e^{-P_{\mathrm{top}}(\overline{f_t})} \frac{\sup P_t 1}{\inf P_t 1} e^{D\overline{\phi} + t\sup \overline{\psi}} \right] e^{\overline{\phi}(z)} \text{ for all } z \in [a].$$

The term in the brackets converges as $t \to 0^+$ (to $e^{D\overline{\phi}}$), and is therefore uniformly bounded. The term $e^{\overline{\phi}(z)}$ is bounded by $G\mu_{\overline{\phi}}[a]$ where G is as in the proof of proposition 2. Consequently, there exists some constant C_0 such that

$$\mu_{\overline{f_t}}[a] \leq C_0 \mu_{\overline{\phi}}[a] \text{ for all } a \in \overline{S}.$$

Since φ is constant on 1–cylinders in $\Sigma_{\mathbb{A}}^+$, we obtain $\mathbb{E}_{\mu_{\overline{f_t}}}[\varphi 1_{[\varphi \geq N]}] \leq C_0 \mathbb{E}_{\mu_{\overline{\phi}}}[\varphi 1_{[\varphi \geq N]}]$ for all N. The RHS tends to zero as $N \to \infty$, by the dominated convergence theorem. We obtained the uniform integrability of φ w.r.t. $\mu_{\overline{f_t}}$.

Step 4. $\mathbb{E}_{\mu_{\overline{\phi_t}}}[\varphi] \xrightarrow[t \to 0^+]{} \frac{1}{\mu_\phi(A)}$.

Proof. The proof is essentially the same as in the previous step, except that here we need to use the perturbation operators

$$\widetilde{R}_t f := R_t[e^{-P_{\mathrm{top}}(f_t)\varphi} f]$$

(the Ruelle operators of $\overline{\phi_t} = \overline{f_t} - P_{\mathrm{top}}(f_t)\varphi$). We first claim that $\|\widetilde{R}_t - R_0\| \xrightarrow[t \to 0^+]{} 0$.

We need the following generalization of eq. (2): Let $\vec{\psi} = (\psi^{(1)}, \ldots, \psi^{(d)})$ be a vector of real valued functions on $\Sigma_{\mathbb{A}}^+$ and $F(t_1, \ldots, t_d)$ some real valued function such that $F(\vec{\psi}(x))$ is well defined for all $x \in \Sigma_{\mathbb{A}}^+$. Define $R_F f := R_0[F(\vec{\psi})f]$. Then for some constant M which only depends on ϕ,

$$\|R_F\| \leq M \left(\mathbb{E}_{\mu_{\overline{\phi}}}\big[|F(\vec{\psi})|\big] + \sum_{a \in \overline{S}} \mu_{\overline{\phi}}[a] D_a[F(\vec{\psi})] \right). \tag{13}$$

The proof is the same as in the one-dimensional case (as is the constant M).

We now observe that $\widetilde{R}_t - R_0 = R_{F_t}$ with $F_t(\overline{\psi}, \varphi) = e^{t\overline{\psi} - P_{\mathrm{top}}(f_t)\varphi} - 1$. Therefore,

$$\begin{aligned} \|\widetilde{R}_t - R_0\| &\leq M\left(\mathbb{E}_{\mu_{\overline{\phi}}}\left[\left|e^{t\overline{\psi} - P_{\mathrm{top}}(f_t)\varphi} - 1\right|\right] + \sum_{a \in \overline{S}} \mu_{\overline{\phi}}[a] D_a[e^{t\overline{\psi} - P_{\mathrm{top}}(f_t)\varphi}] \right) \\ &\leq M\left(\mathbb{E}_{\mu_{\overline{\phi}}}\left[\left|e^{t\overline{\psi} - P_{\mathrm{top}}(f_t)\varphi} - 1\right|\right] + te^{t\sup\overline{\psi}} D\overline{\psi} \right) \xrightarrow[t \to 0^+]{} 0, \end{aligned}$$

because of the bounded convergence theorem (we are using here the facts that $\sup \overline{\psi} < \infty$ and $P_{\mathrm{top}}(f_t) > 0$).

Now that we know that $\|\widetilde{R}_t - R_0\| \xrightarrow[t \to 0^+]{} 0$ we can proceed exactly as in the previous step, but with the eigenprojections $\widetilde{P}_t := P(\widetilde{R}_t)$ replacing P_t, to deduce that $\mathbb{E}_{\mu_{\overline{\phi_t}}}[\varphi] \xrightarrow[t \to 0^+]{} \mathbb{E}_{\mu_{\overline{\phi}}}[\varphi]$. The theorem follows from step 2. □

5. Proofs for shifts not satisfying the BIP property

Reduction of the General Case to the BIP Case. Let ϕ and ψ be two locally Hölder continuous functions bounded from above and assume (Φ), (Ψ), and that $\phi + t\psi$ has an equilibrium measure for $0 \le t \le \epsilon_0$. We also assume without loss of generality that $P_{\rm top}(\phi) = 0$ and $\mathbb{E}_{\mu_\phi}[\psi] = 0$ (otherwise subtract suitable constants).

Let $A \subset S$ be a finite union of states such that $\psi \le \mathbb{E}_{\mu_\phi}[\psi] = 0$ outside A. Let $a \in S$ be some state such that $\mathbb{E}_{\mu_\phi}[r_a] < \infty$ where $r_a(x) := \min\{k : x_k = a\}$. Without loss of generality, $[a] \subseteq A$ (otherwise add a to A).

Set $\varphi(x) := 1_A(x) \min\{k \ge 1 : T^k x \in A\}$, and let $T_A : A \to A$, $T_A(x) := T^{\varphi(x)}(x)$ be the induced map. We have seen that this map can be coded by a topological Markov shift with the BIP property. Let $\overline{\phi}$ and $\overline{\psi}$ be as before. These are locally Hölder continuous functions, and as in the proof of theorem 8,

$$\sup \overline{\phi} = \sup \overline{\phi - P_{\rm top}(\phi)} < \infty;$$

$$\sup \overline{\psi} = \sup \overline{\psi - \mathbb{E}_{\mu_\phi}[\psi]} < \infty.$$

We conclude that $\Sigma^+_{\overline{\mathbb{A}}}$, $\overline{\phi}$, $\overline{\psi}$ satisfy the standing assumptions listed at the beginning of section 3 – the assumptions needed to prove theorems 2, 3, 4 for BIP systems.

In order to pass from the induced system to the original system, we need to apply theorems 7 and 8. The conditions of theorem 8 are satisfied (by $\Sigma^+_{\mathbb{A}}, \phi, \psi$); We check the conditions of theorem 7. The only thing to check is that the tightness assumption holds in all relevant cases.

If $\alpha \in (1, 2)$ one must show that $\frac{1}{B_n}(\overline{\varphi}_n - n/\mu_\phi(A))$ is tight for any sequence B_n regularly varying of index $\frac{1}{\alpha}$; If $\alpha = 2$ one must check tightness for $B_n = \sqrt{n}$ or for B_n s.t. $\sqrt{n} = o(B_n)$. (The case $\alpha = 1$ does not require theorem 7). We show

$$\frac{1}{B_n}[\overline{\varphi}_n - n/\mu_\phi(A)] \text{ is tight for all } \{B_n\} \text{ positive s.t. } \limsup_{n\to\infty} \frac{\sqrt{n}}{B_n} < \infty. \tag{14}$$

This covers all possibilities.

Observe that $\mathbb{E}_{\mu_{\overline{\phi}}}[\varphi^2] < \infty$. To see this recall from lemma 1 that $\mu_{\overline{\phi}} = \overline{\mu_\phi}$, note that $r_a(x) \ge \min\{k \ge 1 : T^k(x) \in A\} =: r_A(x)$, and observe that

$$\begin{aligned}\mathbb{E}_{\mu_\phi}[\varphi^2] &= \sum_{n=1}^{\infty} n^2 \mu_\phi[\varphi = n] \le \sum_{n=1}^{\infty} \left(2\sum_{k=1}^{n} k\right) \mu_\phi[\varphi = n] = \\ &= 2\sum_{n=1}^{\infty} \left(\int_{[\varphi=n]} \sum_{k=0}^{\varphi-1} r_A \circ T^k\right) d\mu_\phi = 2\int r_A d\mu_\phi \le 2\int r_a d\mu_\phi < \infty.\end{aligned}$$

It follows that $-\overline{\varphi}$ satisfies case (2)(c) of Theorem 5. By theorem 3 for BIP systems, $\overline{\varphi}_n$ satisfies the central limit theorem, and (14) follows.

Proof of Theorem 4 for Systems without the BIP Property. It is enough to treat the case $\mathbb{E}_{\mu_\phi}[\psi], P_{\rm top}(\phi) = 0$. By Lemma 1, $P_{\rm top}(\overline{\phi}) = 0$ and $\mu_{\overline{\phi}} = \overline{\mu_\phi}$. By Kac' formula, $\overline{\psi} \in L^1$, and $\mathbb{E}_{\mu_{\overline{\phi}}}[\overline{\psi}] = 0$. We deduce from theorem 4 in the BIP case that $P_{\rm top}(\overline{\phi + t\psi}) = o(t)$. By theorem 8, $P_{\rm top}(\phi + t\psi) = o(t)$. The remaining part of the theorem is because of the ergodicity of μ_ϕ, see [**BS**]. □

Proof of Theorem 2 for Systems without the BIP Property. It is enough to treat the case $\mathbb{E}_{\mu_\phi}[\psi], P_{\mathrm{top}}(\phi) = 0$. Suppose $P_{\mathrm{top}}(\phi + t\psi) = ct + t^\alpha L(1/t)$ with $c \in \mathbb{R}$, $1 < \alpha < 2$ and $L(x)$ slowly varying at infinity. The previous section shows that $c = 0$.

By theorem 8 $P_{\mathrm{top}}(\overline{\phi + t\psi}) = \frac{1}{\mu_\phi(A)} t^\alpha [1 + o(1)] L(1/t)$. $\overline{L}(x) := \frac{1+o(1)}{\mu_\phi(A)} L(x)$ is slowly varying at infinity, therefore by theorem 2 for BIP systems, $\exists B_n$ regularly varying of index α such that $\frac{1}{B_n}\overline{\psi}_n \xrightarrow[n\to\infty]{\text{dist.}} G_\alpha$.

Since $1 < \alpha < 2$, $\frac{\sqrt{n}}{B_n} \to 0$, so $\frac{1}{B_n}[\overline{\varphi}_n - n/\mu_\phi(A)]$ is tight. Theorem 7 now implies that $\frac{1}{B_n}\psi_n \xrightarrow[n\to\infty]{\text{dist.}} G_\alpha^*$ where G_α^* is equal to G_α up to change of scale. Renormalizing B_n, we obtain (2) in theorem 2, and we proved (1)⇒(2). The other direction is handled in the same way. □

Proof of Theorem 3 for Systems without the BIP Property. It is enough to treat the case $P_{\mathrm{top}}(\phi), \mathbb{E}_{\mu_\phi}[\psi] = 0$. We saw above that $c = 0$.

Part 1. Taylor expansion.

By theorem 8, $P_{\mathrm{top}}(\phi+t\psi) = \frac{1}{2}\sigma^2 t^2 + o(t^2)$ iff $P_{\mathrm{top}}(\overline{\phi + t\psi}) = \frac{1}{2}\big(\frac{\sigma}{\sqrt{\mu_\phi(A)}}\big)^2 t^2 + o(t^2)$. Our results for BIP maps say that this is equivalent to $\frac{1}{\sqrt{n}}\overline{\psi}_n \xrightarrow[n\to\infty]{\text{dist.}} N(0, \frac{\sigma^2}{\mu_\phi(A)})$ w.r.t. $\overline{\mu}_\phi$. By theorem 7, this happens iff $\frac{1}{\sqrt{n}}\psi_n \xrightarrow[n\to\infty]{\text{dist.}} N(0,\sigma^2)$ (see the remark at the end of the proof of theorem 7).

We explain why in this case $\psi \in L^2(\mu_\phi)$. By theorem 5, $\overline{\psi} \in L^2(\overline{\mu}_\phi)$. When we proved (14), we saw that $(\Phi) \Rightarrow \varphi \in L^2(\overline{\mu}_\phi)$. Therefore, $\overline{\psi - \sup\psi} \in L^2$. It follows that $\int \sum_{k=0}^{\varphi-1}(\psi - \sup\psi)^2 \circ T^k d\overline{\mu}_\phi +$positive terms $< \infty$, whence $\overline{(\psi - \sup\psi)^2} \in L^2$. By Kac' formula, $\psi - \sup\psi \in L^2(\mu_\phi)$, and so $\psi \in L^2$.

Part 2. Critical expansion.

By theorem 8, $P_{\mathrm{top}}(\phi + t\psi) = t^2 L(1/t)$ with $L(x)$ slowly varying and not asymptotically constant, iff $P_{\mathrm{top}}(\overline{\phi + t\psi}) = \frac{1+o(1)}{\mu_\phi(A)} t^2 L(1/t)$ with such L. By the BIP property, this is equivalent to the existence of B_n r.v. of index $\frac{1}{2}$ such that $\frac{1}{B_n}\overline{\psi}_n \xrightarrow[n\to\infty]{\text{dist.}} N(0,1)$, $\mathbb{E}_{\mu_{\overline{\phi}}}[\overline{\psi}] = 0$, and $\frac{\sqrt{n}}{B_n} \to 0$. By (14) $\frac{1}{B_n}[\overline{\varphi}_n - n/\mu_\phi(A)]$ is tight, so $\frac{1}{B_n}\overline{\psi}_n \xrightarrow[n\to\infty]{\text{dist.}} N(0,1)$ is equivalent to $\frac{1}{B_n^*}\psi_n \xrightarrow[n\to\infty]{\text{dist.}} N(0,1)$ for B_n^* proportional to B_n. This gives the equivalence in theorem 3, part 2.

To finish the proof, it is enough to observe that the BIP property, the expansion $P_{\mathrm{top}}(\overline{\phi + t\psi}) = \frac{1+o(1)}{\mu_\phi(A)} t^2 L(1/t)$, and theorem 5 case (2)(c) show that $L(x) \to \infty$ whenever it is not asymptotic to a constant. □

Proof of theorem 6. Without loss of generality ϕ has zero pressure, and ψ has zero expectation (and then $P_{\mathrm{top}}(\phi+t\psi) = t^\alpha L(1/t)$). Fix an arbitrary finite union of states A so large that $\psi < \mathbb{E}_{\mu_\phi}[\psi] - \epsilon = -\epsilon$ outside A, and let $\varphi(x) := 1_A(x)\min\{n \geq 1 : T^n(x) \in A\}$. Let $\overline{\psi}$ be the induced version of ψ on A. By theorem 8,

$$P_{\mathrm{top}}(\overline{\phi} + t\overline{\psi}) = \frac{1 + o(1)}{\mu_\phi(A)} t^\alpha L(1/t) \text{ as } t \to \infty.$$

Since $1 < \alpha < 2$, $L(t)$ must be asymptotically non-negative (see section 3). By theorem 5, $\mu_\phi(A \cap [|\overline{\psi}| > t]) \sim \frac{t^{-\alpha}}{|\Gamma(1-\alpha)|} L(t)$ as $t \to \infty$.

By choice of A, $\overline{\psi} = \psi + \sum_{k=1}^{\varphi-1} \psi \circ T^k$, where each summand under the sigma symbol is less than $-\epsilon$. It follows that $\epsilon(\varphi - 1) - \|\psi\|_\infty \leq |\overline{\psi}| \leq \varphi \|\psi\|_\infty$. Since $L(x)$ is slowly varying, $L(\lambda x), L(\lambda + x) \sim L(x)$ as $x \to \infty$ for all $\lambda \in \mathbb{R}^+$, and so

$$\begin{aligned}
\mu_\phi[\varphi > t] &\leq \mu_\phi\big[|\overline{\psi}| > \epsilon(t-1) - \|\psi\|_\infty\big] \sim \frac{\epsilon^{-\alpha} t^{-\alpha} L(t)}{|\Gamma(1-\alpha)|}, \\
\mu_\phi[\varphi > t] &\geq \mu_\phi\big[|\overline{\psi}| > t\|\psi\|_\infty\big] \sim \frac{\|\psi\|_\infty^{-\alpha} t^{-\alpha} L(t)}{|\Gamma(1-\alpha)|}.
\end{aligned}$$

Consequently, $\mu_\phi[\varphi > n] \asymp \frac{L(n)}{n^\alpha}$.

We now appeal to Gouëzel [**Gou1**], Theorem 1.3 (see also [**S5**]), which says that in our context for every f, g locally Hölder continuous supported inside A with non-zero expectation

$$\mathrm{Cov}_{\mu_\phi}(f, g \circ T^n) \sim \sum_{k=n+1}^{\infty} \mu_\phi[\varphi > k] \int f d\mu_\phi \int g d\mu_\phi.$$

The theorem follows from Karamata's Theorem. □

Appendix A. Slow and Regular Variation

Slow and Regular Variation. A positive function $L : (c_0, \infty) \to \mathbb{R}$ is called *slowly varying* (at infinity) if it is Borel measurable and

$$\frac{L(ts)}{L(t)} \xrightarrow[t\to\infty]{} 1 \text{ for all } s > 0.$$

A positive sequence $\{c_n\}_{n\geq 1}$ is called *slowly varying* (at infinity) if $L(t) := c_{[t]}$ is slowly varying (at infinity).

A positive function $f : (c_0, \infty) \to \mathbb{R}$ is called *regularly varying* at infinity with *index* α, if $f(x) = x^\alpha L(x)$ with $L(x)$ slowly varying at infinity. A positive sequence $\{c_n\}_{n\geq 1}$ is called *regularly varying* at infinity with index α, if $f(t) := c_{[t]}$ is *regularly varying* at infinity with *index* α.

For example, $\log x$, $1/\ln\ln x$ are slowly varying at infinity, and $x^\alpha \ln x (\ln\ln x)^2$, $x^\alpha / \ln x$ are regularly varying with index α.

Sufficient Condition for Regular Variation. Let $f(x)$ be a positive continuous function, and $\{a_n\}$, $\{b_n\}$ some positive numbers such that $\limsup_{n\to\infty} b_n = \infty$ and $\limsup_{n\to\infty} \frac{b_{n+1}}{b_n} = 1$. If $\lim_{n\to\infty} a_n f(b_n x)$ exists, is positive, and is continuous on some open interval $(a, b) \subset \mathbb{R}^+$, then $f(x)$ is regularly varying at infinity. ([**BGT**], theorem 1.9.2)

The General Form of Regularly Varying Functions. A Borel function $f(x)$ is regularly varying at infinity with index α iff

$$f(x) = [c + o(1)] x^\alpha \exp \int_1^x \epsilon(u) \frac{du}{u} \qquad \text{as } x \to \infty,$$

where $c > 0$ and $\epsilon(u) \xrightarrow[x\to\infty]{} 0$. ([**BGT**], Theorem 1.3.1)

In particular, any regularly varying function $f(x)$ with index α satisfies $f(x) \to \infty$ when $\alpha > 0$ and $f(x) \to 0$ when $\alpha < 0$. ([**BGT**], Proposition 1.5.1).

Uniform Convergence Theorem. If $L(t)$ is slowly varying at infinity, then $\frac{L(ts)}{L(t)} \xrightarrow[t\to\infty]{} 1$ uniformly on compact subsets of $(0, \infty)$. ([**BGT**], Theorem 1.2.1).

Asymptotic Inversion Theorem. If $f(x)$ is regularly varying at infinity with positive index α, then there exists $g(x)$ regularly varying at infinity with index $1/\alpha$ such that $(f \circ g)(x) \sim (g \circ f)(x) \sim x$ as $x \to \infty$. ([**BGT**], Theorem 1.5.12)

Differentiating Asymptotic Relations: The Monotone Density Theorem. Suppose $U(t) = \int_0^x u(y)dy$, and $L(x)$ is slowly varying at infinity.

(1) If $u(y)$ is monotone at some interval $(0, \delta)$ and $\rho \geq 0$, then

$$U(t) \sim t^\rho L(1/t) \text{ as } t \to 0^+ \text{ implies } u(t) \sim \rho t^{\rho-1} L(1/t) \text{ as } t \to 0^+.$$

(2) If $u(y)$ is monotone at some interval (δ, ∞) and $\rho \in \mathbb{R}$, then

$$U(x) \sim x^\rho L(x) \text{ as } x \to \infty \text{ implies } u(x) \sim \rho x^{\rho-1} L(x) \text{ as } x \to \infty.$$

Here and throughout $f(x) \sim 0 \cdot g(x)$ means $f(x) = o(g(x))$. ([**BGT**], Theorems 1.7.2 and 1.7.2b).

Integrating Asymptotic Relations: Karamata's Theorem. Suppose $L(x)$ is slowly varying at infinity and locally bounded. Then as $x \to \infty$,

$$\int_a^x t^\rho L(t)dt \sim \frac{x^{\rho+1}}{\rho+1} L(x), \quad \text{for all } \rho > -1$$

$$\int_x^\infty t^\rho L(t)dt \sim -\frac{x^{\rho+1}}{\rho+1} L(x), \quad \text{for all } \rho < -1.$$

The converse is also true: Any positive locally bounded $L(x)$ for which one of these relations holds for some $\rho \neq -1$ must be slowly varying. ([**BGT**], theorems 1.5.11 and 1.6.1).

After a change of variables, Karamata's theorem implies that if $L(x)$ is slowly varying at infinity and $\alpha > -1$, then

$$\int_0^t \tau^\alpha L(1/\tau)d\tau \sim \frac{t^{1+\alpha}}{1+\alpha} L(1/t) \text{ as } t \to 0^+.$$

Conversly, if L satisfies the above, then it must be slowly varying at infinity.

Karamata's Tauberian Theorem. Let $U(x)$ be a non-decreasing function on $\mathbb{R}$, which is continuous from the right, and such that $U(0) = 0$. Suppose $L(x)$ is slowly varying at infinity, and $c > 0$, $\rho \geq 0$. The following are equivalent:

$$U(x) \sim \frac{cx^\rho}{\Gamma(1+\rho)} L(x), \quad \text{as } x \to \infty$$

$$\int_0^\infty e^{-tx} dU(x) \sim \frac{c}{t^\rho} L(1/t), \quad \text{as } t \to 0^+.$$

([**BGT**], theorem 1.7.1).

Truncated Variance: Feller's Theorem. Let $F(x)$ be a right continuous probability distribution function such that $F(0) = 0$, and set $U(x) := \int_0^x y^2 dF(y)$. Suppose $L(x)$ is slowly varying at infinity as $x \to \infty$, $c \neq 0$, and $0 < \rho < 2$. The following are equivalent:

$$U(x) \sim cx^\rho L(x), \text{ as } x \to \infty$$
$$1 - F(x) \sim \frac{c\rho}{2-\rho} x^{\rho-2} L(x), \text{ as } x \to \infty.$$

(See Feller [**F**] VIII.9 for generalizations).

Proof: Start with the identity $1 - F(x) = \int_x^\infty y^{-2} d\left(\int_0^y t^2 dF(t)\right) = \int_x^\infty y^{-2} dU(y)$. Integration by parts gives:

$$1 - F(x) = y^{-2}U(y)\Big]_{y=x}^{y=\infty} + 2\int_x^\infty y^{-3}U(y^-)dy.$$

If $U(y) \sim cy^\rho L(y)$, then $U(y^-) \sim cy^\rho L(y)$. By Karamata's theorem:

$$1 - F(x) = -cx^{\rho-2}L(x)[1 + o(1)] + 2cL(x)[1 + o(1)]\int_x^\infty y^{\rho-3}dy =$$
$$= \frac{c\rho}{2-\rho} x^{\rho-2} L(x)[1 + o(1)].$$

To see the other direction, integrate by parts $U(y) = \int_0^x y^2 dF(y)$:

$$U(x) = y^2 F(y)\Big]_{y=0}^{y=x} - 2\int_0^x yF(y^-)dy = x^2F(x) - 2\int_0^x yF(y^-)dy =$$
$$= -x^2[1 - F(x)] + 2\int_0^x y[1 - F(y^-)]dy.$$

Now plug into this expression the asymptotic formula for $1 - F(x)$ and conclude as before, using Karamata's theorem.

Appendix B. The Fisher–Felderhoff droplet model

We describe a crude simplification of a model in [**FF**]. A 'vapor' close to the condensation point consists of microscopic droplets. The interaction between particles in different droplets is negligible, but the interaction between particles in the same droplet is strong, and long–range.[7] When two droplets 'touch', they become one. 'Condensation' is the appearance of macroscopic droplets.

Here is a lattice–gas model of this situation. Space is discretized and described by a one–sided one-dimensional string of sites, each of which can be in one of two states: empty (state '0') or occupied ('1'). The configuration space is $\{0, 1\}^{\mathbb{N}_0}$. A 'droplet' is a maximal string of occupied sites.

We describe the interaction by prescribing the function

$$\phi(x_0, x_1, \ldots) := -\beta U(x_0|x_1, x_2, \ldots)$$

where β is a constant ('inverse temperature') and $U(x_0|x_1, x_2, \ldots)$ is the energy due to the interaction of site zero and the other sites.[8] Note that the energy due to the interaction between the first n sites and the rest is minus the n–th Birkhoff sum of

[7] One example of long–range interactions in liquid droplets is 'surface tension'.

[8] It is useful to think of $U(x_0|x_1, x_2, \ldots)$ as of the energy cost of separating site zero from sites n, $n > 0$, and moving it to infinity – that is, if site zero is occupied.

ϕ. It follows that the Helmholtz free energy(=Energy $-\frac{1}{\beta}\times$Entropy) per site is up to a constant

$$\lim_{n\to\infty}\frac{1}{n}\left[-\int\sum_{k=0}^{n-1}\phi\circ T^k d\mu_\phi - \sum_{n\text{–cylinders}}\mu_\phi[\underline{a}]\log\frac{1}{\mu_\phi[\underline{a}]}\right]=$$

$$=-\left(h_{\mu_\phi}(T)+\int\phi d\mu_\phi\right)=-P_{\text{top}}(\phi),$$

at least when ϕ has an equilibrium measure μ_ϕ.

Since different droplets do not interact, ϕ takes the form

$$\phi(0,*)=0\ ,\ \phi(\underbrace{1,1,\ldots,1}_{n},0,*):=f(n)$$

for some function $f(n)$. If the interaction is 'long range', then this function is not locally Hölder, because the effect of far away sites is not exponentially small.

Consider now the following re-coding of a configuration: $(x_0,x_1,\ldots)\mapsto(y_0,y_1,\ldots)$, where $x_i=0\Rightarrow y_i=0$, and $x_i=1\Rightarrow y_i=1$+number of occupied sites to the right of i until the first unoccupied site, for example:

$$(0,1,1,0,0,1,0,1,1,1,0,\ldots)\mapsto(0,2,1,0,0,1,0,3,2,1,0,\ldots).$$

In this coding, the configuration space becomes the *renewal shift*: the topological Markov shift with state space $\mathbb{N}\cup\{0\}$ and transition matrix

$$\mathbb{A}=(t_{ij})\text{ where }t_{ij}=\begin{cases}1 & i=0;\\ 1 & i>0, j=i-1;\\ 0 & \text{otherwise.}\end{cases}$$

In the new coordinates the interaction becomes locally Hölder ('short range'):

$$\widetilde{\phi}(y_0,y_1,\ldots)=\begin{cases}f(y_0) & y_0\neq 0;\\ 0 & y_0=0.\end{cases}$$

Thus a compact shift with a long range potential is recoded as a non-compact shift with a short range potential.

The critical phenomena for the Fisher–Felderhoff model for various choices of $f(n)$ is described in [**FF**] and [**Wa1**, **Wa2**].

Acknowledgements. The author wishes to thank the referee for his careful reading of the manuscript and for many valuable suggestions, and Michael Fisher and David Ruelle for useful discussions.

References

[**ADU**] J. Aaronson, M. Denker, M. Urbański: *Ergodic theory for Markov fibred systems and parabolic rational maps.* Trans. AMS **337** (1993), 495–548.

[**AD**] J. Aaronson, M. Denker: *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps.* Stoch. Dyn. **1** (2001), no. 2, 193–237.

[**BG**] P. Bálint and S. Gouëzel: *Limit theorems in the stadium billiard.* To appear in Commun. Math. Phys.

[**BGT**] N.H. Bingham, C.M. Goldie, and J.L. Teugels: Regular variation. Encyclopedia of Math. and its Appl. **27**, Cambridge Univ. Press (1987).

[**BDFN**] J.J. Binney. N.J. Dowrick, A.J. Fisher, M.E.J. Newman: The theory of critical phenomena, an introduction to the renormalization group. Oxford Science Publications, Oxford University Press (1992).

[**Bo**] R. Bowen: *Equilibrium states and the ergodic theory of Anosov diffeomorphisms.* Lecture Notes in Mathematics, Vol. **470**. Springer-Verlag, Berlin-New York, 1975.

[**BS**] J. Buzzi and O. Sarig: *Uniqueness of equilibrium measures for countable Markov shifts and multi-dimensional piecewise expanding maps.* Erg. Th. Dynam. Sys. **23**, 1383-1400 (2003).

[**Ea**] G. K. Eagleson: *Some simple conditions for limit theorems to be mixing.* (Russian) Teor. Verojatnost. i Primenen. **21** (1976), no. 3, 653–660. Engl. Transl. in Theor. Probability Appl. **21** (1976), no. 3, 637–642 (1977).

[**El**] R. S. Ellis: Entropy, large deviations, and statistical mechanics. Grund. Math. Wissenschaften **271** Springer Verlag (1985).

[**F**] W. Feller: An introduction to probability theory and its applications. Volume II, second edition, John Wiley & Sons, 1971.

[**FF**] M.E. Fisher and B.U. Felderhof: *Phase transition in one–dimensional cluster–interaction fluids: IA. Thermodynamics, IB. Critical behavior. II. Simple logarithmic model* Ann. Physics **58**, 177–280 (1970).

[**GK**] B.V. Gnedenko and A.N. Kolmogorov: Limit distributions for sums of independent random variables. Translated and annotated by K.L. Chung, with an Appendix by J.L. Doob. Addison–Wesley Publishing Company, (1954).

[**Gou1**] S. Gouëzel: *Sharp polynomial estimates for the decay of correlations.* Israel J. Math. **139** (2004) 29-65.

[**Gou2**] S. Gouëzel: *Central limit theorem and stable laws for intermittent maps.* Probab. Theory Related Fields **128** (2004), no. 1, 82–122.

[**Gu**] B. M. Gurevič: *Topological entropy of a countable Markov chain.* Dokl. Akad. Nauk SSSR **187** 715–718 (1969).

[**Hi**] R. Hilfer: *Classification theory for anequilibrium phase transitions.* Phys. Rev. E **48** (4) 2466–2475 (1993).

[**Ho**] F. Hofbauer: *Examples for the non–uniqueness of the equilibrium states.* Trans. AMS **228** 223–241 (1977).

[**Ka**] T. Kato: Perturbation theory for linear operators. Reprint of the 1980 edition. Classics in Mathematics. Springer-Verlag, Berlin, 1995.

[**Ke**] M. Keane: *Strongly mixing g-measures.* Invent. Math. **16** (1972), 309–324.

[**Lo**] A.O. Lopes: *The Zeta function, non–differentiability of pressure, and the critical exponent of transition.* Adv. Math. **101** no. 2, 133–165 (1993).

[**ML**] A. Martin–Löf: *Mixing properties, differentiability of the free energy and the central limit theorem for a pure phase in the Ising model at low temperature.* Comm. Math. Phys. **32** (1973), 75–92.

[**MU1**] M. Urbański and R.D. Mauldin: *Gibbs states on the symbolic space over an infinite alphabet.* Israel J. Math. **125** (2001), 93–130.

[**MU2**] R.D. Mauldin, M. Urbański: *Graph directed Markov systems; Geometry and dynamics of limit sets.* Cambridge Tracts in Mathematics, **148** Cambridge University Press, Cambridge, 2003.

[**MT**] I. Melbourne and A. Török: *Statistical limit theorems for suspension flows.* Israel Journal of Math. **194** (2004) 191-210.

[**N**] S.V. Nagaev: *Some limit theorems for stationary Markov chains.* (Russian) Teor. Veroyatnost. i Primenen. **2** (1957) 389–416.

[**PS**] T. Prellberg, J. Slawny: *Maps of intervals with indifferent fixed points: thermodynamic formalism and phase transitions.* J. Statist. Phys. **66** (1992), no. 1-2, 503–514.

[**Ru**] D. Ruelle: Thermodynamic Formalism, The mathematical structures of equilibrium statistical mechanics. 2nd Ed. . Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2004.

[**S1**] O.M. Sarig: *Thermodynamic Formalism for Countable Markov shifts.* Erg. Th. Dyn. Sys. **19** (1999) 1565-1593.

[**S2**] O. Sarig: *Phase Transitions for Countable Markov Shifts.* Commun. Math. Phys. **217** 555–577 (2001).

[**S3**] O. Sarig: *Characterization of existence of Gibbs measures for Countable Markov shifts.* Proc. of AMS. **131** (no. 6), 1751-1758 (2003).

[**S4**] O. Sarig: *Thermodynamic formalism for null recurrent potentials.* Israel J. Math. **121**, 285-311 (2001)

[**S5**] O. Sarig: *Subexponential decay of correlations.* Invent. Math. **150**, 629-653 (2002).

[**S6**] O. Sarig: *Thermodynamic formalism for countable Markov shifts.* Tel-Aviv University Thesis (2000).

[**S7**] O. Sarig: *On an example with a non-analytic topological pressure.* C. R. Acad. Sci. Paris Sér. I Math. **330** (2000), no. 4, 311–315.

[**St**] H.E. Stanley: Introduction to phase transitions and critical phenomena. Oxford University Press (1971).

[**W**] P. Walters: *Ruelle's operator theorem and g-measures.* Trans. Amer. Math. Soc. **214** (1975), 375–387.

[**Wa1**] X.-J. Wang: *Abnormal fluctuations and thermodynamic phase transition in dynamical systems.* Phys. Review A **39** no. 6, 3214–3217.

[**Wa2**] X.-J. Wang: *Statistical physics of temporal intermittency*, Phys. Review A **40** no. 11, 6647–6661 (1989).

[**Y**] M. Yuri: *Thermodynamic formalism for countable to one Markov systems.* Trans. Amer. Math. Soc. **355** (2003), no. 7, 2949–2971.

[**Z**] V.M. Zolotarev: *One–dimensional stable distributions.* Transl. Math. Monog. AMS **65** (1986).

Mathematics Department, Pennsylvania State University, University Park, PA 16802, U.S.A.

E-mail address: `sarig@math.psu.edu`