

Manfred Denker

# Lectures on Probability and Dynamics.

25. Dezember 2009

Penn State University  
Mathematics Department  
Department of Statistics



# Content

<b>1</b>	<b>Thermodynamic Formalism</b> .....	1
1.1	Introduction .....	1
1.2	R-expanding Systems .....	4
1.3	Topological Pressure .....	7
1.4	Gibbs-Measures .....	14
<b>2</b>	<b>Stochastic Laws in Dynamics</b> .....	19
2.1	Introduction .....	19
2.2	Martingale Central Limit Theorem .....	21
2.3	Martingale Approximation for Dynamical Systems ...	25
2.4	The Central Limit Theorem for R-Expanding Systems	28
2.5	Local Limit Theorem .....	32
2.6	Characteristic Function Operators .....	33
2.7	Spectral Theorem .....	36
2.8	Local Limit Theorems for R-Expanding Systems .....	39
	<b>Index</b> .....	41



# 1 Thermodynamic Formalism

This is a minicourse given during the workshop "Dynamical Systems and Related Topics" at Göttingen University, July 2009.

## 1.1 Introduction

The work of Willard Gibbs (1839–1903) and Ludwig Boltzmann (1844–1906) on Thermodynamics have found its trace in the theory of dynamical systems. D. Ruelle succeeded in 1970 introducing the notion of pressure for  $\mathbb{Z}^d$ -shifts. For  $\mathbb{Z}$ -operations, so to speak for general continuous transformations, the pressure function was defined first by P. Walters in 1974. The version we shall be giving below is based on Rufus Bowen's approach to thermodynamic formalism via separating and spanning sets. The notion is a rather straight forward generalization of the notion of topological entropy. Walters also proved a variational principle generalizing the case of topological entropy which had been introduced already as early as 1966 by Adler et al..

The theory of Frobenius-Perron of Bowen and Ruelle will be formulated for open and expanding systems. Existence, uniqueness and invariance of Gibbs measures will be shown. For topological Markov chains the theory was (partly) used by Parry and Sinai around 1965.

Methods of thermodynamic formalism are successfully applicable for investigations of Fuchsian (Kleinian, hyperbolic) groups, rational maps, (partially) hyperbolic diffeomorphisms and flows and maps of the interval. In particular, it has found applications in various disciplines in mathematics.

Let us first consider a finite particle system.

A particle system consists of finitely many points at locations in some finite dimensional Euclidean space. It can be described microscopically, meaning that for each particle its position, momentum energy are given, and all these quantities follow rules known in classical mechanics. An ideal gas is of this type. Naturally, such systems are hard to analyze and to describe. It turns out that macroscopic descriptions are much better suited, introducing concepts for the total energy, temperature, pressure and entropy. Here the fundamental axioms of thermodynamics come into play: 1. The energy of a

closed system is constant. 2. Entropy is maximized by its equilibrium (the state of the system into which it evolves in time and then stays in that state).

We begin providing a simple model<sup>1</sup>.

Let  $\Omega$  be a finite set of  $n$  points, called the configuration space. We define a state (of the system) to be a probability measure on  $\Omega$ , that is given by a probability vector  $\mu(\omega)$ ,  $\omega \in \Omega$ . Denote by  $\mathcal{M} \subset \mathbb{R}^n$  the set of all states. The entropy of a state is

$$H(\mu) = - \sum_{\omega \in \Omega} \mu(\omega) \log \mu(\omega),$$

and is a continuous function on  $\mu$ .

Entropy is best understood as a measure of uncertainty. For this define the information function

$$I(\omega) = - \log \mu(\omega)$$

as a measure of information which  $\mu$  provides on the chances that the system is in the microscopic state  $\omega$ .  $I(\omega) = 0$  means that for sure the system is in state  $\omega$ . Then entropy is the expected information of state  $\mu$ .  $H(\mu) = 0$  means that the state  $\mu$  is concentrated on a single microscopic ground state  $\omega \in \Omega$ .

Using concavity of the function  $x \mapsto \phi(x) = -x \log x$

$$\frac{1}{n} \sum_{\omega \in \Omega} \phi(\mu(\omega)) \leq \phi\left(\frac{1}{n} \sum_{\omega \in \Omega} \mu(\omega)\right) = \phi\left(\frac{1}{n}\right) = \frac{1}{n} \log n.$$

Therefore

$$H(\mu) \leq \log n$$

with equality if and only if  $\mu(\omega) = \frac{1}{n}$  for all  $\omega \in \Omega$ .

Let us assume that there is an energy function  $\omega \mapsto u(\omega)$  describing the system. The mean energy with respect to the state  $\mu$  is

$$\mu(u) = \sum_{\omega \in \Omega} \mu(\omega) u(\omega).$$

The partition function is defined by

$$Z(\beta) = \sum_{\omega \in \Omega} \exp[-\beta u(\omega)] \quad \beta \in \mathbb{C}.$$

Finally a Gibbs measure has the form

$$\mu_\beta(\omega) = \frac{1}{Z(\beta)} \exp[-\beta u(\omega)].$$

---

<sup>1</sup> after G. Keller: Equilibrium states in ergodic theory. Lecture Notes London Math. Soc. **42**, Cambridge Univ. Press 1997

First observe that for two microscopic states  $\omega, \omega' \in \Omega$  with  $u(\omega) > u(\omega')$

$$\lim_{\beta \rightarrow \infty} \frac{\mu_\beta(\omega)}{\mu_\beta(\omega')} = 0,$$

and therefore

$$\lim_{\beta \rightarrow \infty} \mu_\beta$$

is the equidistribution on all minimal values  $\{\omega \in \Omega : u(\omega) = \min_{\omega'} u(\omega')\}$ , while taking the limit for  $\beta \rightarrow -\infty$  yields convergence to the equidistribution on the maximal values  $\{\omega \in \Omega : u(\omega) = \max_{\omega'} u(\omega')\}$ .

**Lemma 1.** *The function  $\log Z(\beta)$  is real analytic and*

$$(\log Z)'(\beta) = -\mu_\beta(u) \quad (\log Z)''(\beta) = V_{\mu_\beta}(u),$$

where the latter notion is the variance of  $u$  with respect to the Gibbs measure. Moreover, equality holds if and only if  $u$  is constant. In particular,  $Z$  is a convex function in  $\beta$ .

*Proof.* Note that  $Z$  is analytic in  $\beta \in \mathbb{R}$ , so is  $\log Z$ .

$$(\log Z)'(\beta) = \frac{Z'(\beta)}{Z(\beta)} = -\frac{1}{Z(\beta)} \sum u(\omega) e^{-\beta u(\omega)} = -\mu_\beta(u).$$

$$(\log Z)''(\beta) = \frac{Z''(\beta)}{Z(\beta)} - \left( \frac{Z'(\beta)}{Z(\beta)} \right)^2 = -\frac{1}{Z(\beta)} \sum u(\omega)^2 e^{-\beta u(\omega)} - \mu_\beta(u)^2 = V_{\mu_\beta}(u).$$

**Theorem 1.** *(Variational principle) The Gibbs measure  $\mu_\beta$  satisfies*

$$H(\mu_\beta) + \mu_\beta(-\beta u) = \log Z(\beta) = \max\{H(\nu) - \beta \nu(u) : \nu \in \mathcal{M}\}.$$

*This means that  $\mu_\beta$  is an equilibrium state, and in fact, the unique equilibrium state.*

*Proof.* Apply Jensen's inequality to the concave function  $x \mapsto \log x$ :

$$\begin{aligned} H(\nu) + \nu(-\beta u) &= -\sum_{\omega} \nu(\omega) (\log \nu(\omega) + \beta u(\omega)) \\ &= \sum_{\omega} \nu(\omega) \log \frac{e^{-\beta u(\omega)}}{\nu(\omega)} \\ &\leq \log \sum_{\omega} \nu(\omega) \frac{e^{-\beta u(\omega)}}{\nu(\omega)} \\ &= \log Z(\beta), \end{aligned}$$

with equality if and only if  $\omega \mapsto \frac{e^{-\beta u(\omega)}}{\nu(\omega)}$  is constant.

**Corollary 1.** *Let  $\min u < E^* < \max u$ . Then there exists a unique  $\beta^* \in \mathbb{R}$  such that  $\mu_{\beta^*}$  has energy  $\mu_{\beta^*}(u) = E^*$  and maximal entropy among all states with energy  $E^*$ .*

*Proof.*  $\beta \mapsto \mu_\beta(u)$  is strictly decreasing and continuous as observed before. Therefore there is a unique  $\beta^*$  such that  $\mu_{\beta^*}(u) = E^*$ .

Let  $\nu \in \mathcal{M}$  and  $\nu(u) = E^*$ . Then by the variational principle

$$H(\nu) - \beta E^* = H(\nu) + \nu(-\beta^* u) \leq H(\mu_{\beta^*}) - \beta^* \mu_{\beta^*}(u).$$

Usually,  $T = 1/\beta$  is called the temperature, so  $\mu_{1/T}$  is the equilibrium at temperature  $T$ .

The free energy of a state  $\nu \in \mathcal{M}$  is defined as

$$F(\nu) = \nu(u) - TH(\nu).$$

Therefore

$$F(\nu) \geq F(\mu_{1/T}) = -T \log Z(1/T),$$

with equality if and only if  $\nu = \mu_{1/T}$ .

Note that  $\frac{d}{d\beta} \log Z(\beta) = -\mu_\beta(u)$ , hence

$$\frac{d}{d\beta} H(\mu_\beta) = \frac{d}{d\beta} (\log Z(\beta) + \beta \mu_\beta(u)) = \beta \frac{d}{d\beta} \mu_\beta(u).$$

It follows that

$$\frac{d}{dT} H(\mu_{1/T}) = \frac{1}{T} \frac{d}{dT} \mu_{1/T}(u).$$

This formula must be compared with the law in thermodynamics:

$$dS = \frac{dQ}{T},$$

where  $S$  denotes entropy,  $Q$  the heat content and  $T$  absolute temperature.

## 1.2 R-expanding Systems

Let  $X$  be a compact metric space with metric  $d$  and  $T : X \rightarrow X$  a continuous (non invertible) map.

**Definition 1.** *A continuous dynamical system  $(X, T)$  is called (positively) expansive, if there is a constant  $a > 0$  such that  $\sup_{n \in \mathbb{N}} d(T^n(x), T^n(y)) > a$  for all  $x \neq y \in X$ . The constant  $a$  is called an expansive constant.*

Note that we use  $\mathbb{N}$  to denote all integers  $\geq 0$ . Since we are considering compact spaces, the definition is independent of the choice of the metric, thus expansive is a conjugacy invariant.

**Exercise 1.** Two dynamical systems  $(X, T)$  and  $(Y, S)$  are called conjugate if there exists a homeomorphism  $h : X \rightarrow Y$  such that  $h \circ T = S \circ h$ . Show that expansiveness is a conjugacy invariant.

**Definition 2.** A dynamical system  $(X, T)$  is called expanding, if there exist an equivalent metric  $d$  and constants  $\Lambda > 1$  and  $a > 0$  such that

$$d(T(x), T(y)) \geq \Lambda d(x, y) \quad \forall x, y \in X \text{ with } d(x, y) < a.$$

This notion is not a conjugacy invariant.

**Proposition 1.** An expanding dynamical system is (positively) expansive.

*Proof.* If the distance between two points is bounded by  $a$ , then the distance expands by  $\Lambda$  in each iteration step by  $T$ , as long as the distance between the images is bounded by  $a$ . Since  $X$  is bounded, it follows that any number  $< a$  is an expansion constant.

**Definition 3.** A dynamical system  $(X, T)$  is called R-expanding, if it is expanding, and if the transformation  $T$  is an open mapping.

**Exercise 2.** Show that a subshift of finite type (topological Markov chain)  $(X, T)$  is R-expanding.

**Lemma 2.** A dynamical system  $(X, T)$  is R-expanding, if and only if it has the following property:

There are constants  $a > 0$  and  $\Lambda > 1$ , such that for  $x, y' \in X$ ,  $d(T(x), y') < a$  there is a unique point  $y \in X$  satisfying  $d(x, y) < a$  and  $T(y) = y'$ ; moreover,  $d(T(x), T(y)) \geq \Lambda d(x, y)$ .

*Proof.* Let  $(X, T)$  be R-expanding. Then  $T$  is injective on every open ball  $K(x, a/2)$ , since otherwise for some  $y, z \in K(x, a/2)$  we have that  $0 = d(T(z), T(y)) \geq \Lambda^{-1}d(z, y)$ . Since  $T$  is an open map, there is some  $a/2 > a' > 0$  small enough, such that  $T(K(x, a/2)) \supset K(T(x), a')$ . This implies the above property.

For the converse, it is left to show openness of  $T$ . Suppose  $y \in T(K(x, a/2))$  can be approximated by a sequence  $y_n \notin T(K(x, a/2))$ . Then there are preimages  $z_n \in K(x, a)$  of  $y_n$ , which converge to the preimage  $z$  of  $y$  in  $K(x, a/2)$ , a contradiction.

For  $n \geq 1$ ,  $r > 0$  and  $x \in X$  let

$$K_n(x, r) = \{y \in X : d(T^j(y), T^j(x)) < r \text{ for every } 0 \leq j < n\}$$

denote the ball with radius  $r$  and center  $x$  in the Bowen-metric

$$d_n(x, y) = \max_{0 \leq j < n} d(T^j(y), T^j(x)).$$

**Lemma 3.** *Let  $(X, T)$  be positively expansive with expansion constant  $\theta^*$ . Then there is a constant  $0 < \theta \leq \theta^*$  with*

$$\limsup_{n \rightarrow \infty} \sup_{x \in X} \{\text{diam}(K_n(x, \theta))\} = 0.$$

In the sequel an expansion constant is always one satisfying the additional property of Lemma 3.

*Proof.* Let  $\{K(x_i, \theta^*) : i = 1, \dots, s\}$  be a finite open cover of  $X$  and  $\theta$  a Lebesgue-number.

Assume, the assertion of the Lemma is not true. Then there are  $a > 0$  and for each  $n \geq 1$  points  $x_n, y_n$  with  $d(x_n, y_n) \geq a$  and  $y_n \in K_n(x_n, \theta)$ . Since  $X$  is compact, we may assume that  $x = \lim_{n \rightarrow \infty} x_n$  and  $y = \lim_{n \rightarrow \infty} y_n$  exist. Let  $j \geq 0$  be fixed. Then  $d(T^j(x_n), T^j(y_n)) < \theta$  for every  $n \geq j$ , and there is some  $i = i(j)$ , such that for infinitely many  $n$   $T^j(x_n), T^j(y_n) \in K(x_i, \theta^*)$ . It follows that  $d(T^j(x), T^j(y)) < \theta^*$ . Since  $T$  is positively expansive,  $x = y$ .

**Theorem 2.** [COVEN, REDDY] *An expansive dynamical system  $(X, T)$  has an expanding metric (in general not equivalent).*

*Proof.* Let  $3\theta > 0$  be an expansion constant for  $(X, T)$ . One can apply the Metrization Lemma of Frink (see Kelley, General Topology, v. Nostrand 1955, p.185):

Let  $U_n$  ( $n \geq 0$ ) be a sequence of open neighborhoods of the diagonal  $\Delta \subset X \times X$  with the following properties:

1.  $U_0 = X \times X$ .
2.  $\bigcap_{n=1}^{\infty} U_n = \Delta$ .
3.  $U_n \circ U_n \circ U_n \subset U_{n-1}$  ( $n \geq 1$ ),  
i.e.  $(u, v), (v, w), (w, x) \in U_n \implies (u, x) \in U_{n-1}$ .

Then there is a metric  $\rho$ , which is compatible with the topology on  $X$ , such that for any  $n \geq 1$

$$U_n \subset \{(x, y) : \rho(x, y) < 2^{-n}\} \subset U_{n-1}. \quad (1.1)$$

For every  $n \geq 1$  and  $\gamma > 0$  let

$$V_n(\gamma) = \{(x, y) \in X \times X : d(T^j(x), T^j(y)) < \gamma \text{ for all } j = 0, \dots, n\}.$$

Using Lemma 3 one obtains  $M \geq 1$ , such that

$$V_M(3\theta) \subset \{(x, y) : d(x, y) < \theta\}.$$

Setting  $U_0 = X \times X$  and  $U_n = V_{Mn}(\theta)$  ( $n \geq 1$ ), property 1. follows immediately, and 2. follows from expansiveness. Property 3. is shown by induction. The case  $n = 1$  is clear. Hence let 3. be satisfied for  $n$  and  $(x, u), (u, v), (v, y) \in U_{n+1}$ . It follows that  $d(T^j(y), T^j(x)) < 3\theta$  for all  $j = 0, \dots, (n+1)M$ , consequently  $d(T^j(y), T^j(x)) < \theta$  for all  $j = 0, \dots, Mn$ , or  $(x, y) \in V_{Mn}(\theta) = U_n$ .

Frink's Metrization Lemma shows the existence of a metric  $\rho$  on  $X$  with (1.1).

It is sufficient to show that  $T^{3M}$  is expanding with respect to  $\rho$ . Suppose,  $x, y \in X$  satisfy  $0 < \rho(x, y) < \frac{1}{16}$ . Then there is  $n \geq 0$  satisfying  $(x, y) \in U_n \setminus U_{n+1}$ . Necessarily  $n \geq 3$  because of  $0 < \rho(x, y) < \frac{1}{16}$  and (1.1). Moreover, by the choice of  $n$  and the definition of  $U_n$  and  $V_{Mn}(\theta)$  it follows that there is  $Mn < j \leq (n + 1)M$  with  $d(T^j(y), T^j(x)) \geq \theta$ . Since  $3 \leq n$  it follows now that  $d(T^i(T^{3M}(x)), T^i(T^{3M}(y))) \geq \theta$  for some  $0 \leq i \leq (n - 2)M$ , and therefore  $(T^{3M}(x), T^{3M}(y)) \notin U_{n-2}$ . Using again (1.1), the result is deduced from

$$\rho(T^{3M}(x), T^{3M}(y)) \geq 2^{-(n-1)} = 2 \cdot 2^{-n} > 2\rho(x, y).$$

The following lemma provides the basis for the thermodynamic formalism.

**Lemma 4.** *An R-expanding dynamical system  $(X, T)$  has the following properties:*

1. *The number of preimages of  $T$  is uniformly bounded and locally constant.*
2. *The Frobenius-Perron operator (transfer-operator) for  $\phi \in C(X)$ ,*

$$\mathcal{L}_\phi : C(X) \rightarrow C(X), \quad \mathcal{L}_\phi f(x) = \sum_{T(y)=x} f(y) \exp[-\phi(y)]$$

*is well defined, continuous and positive.*

*Proof.* R-expanding systems are characterized in Lemma 2. Let  $d(x, y) < a/2$ . Then every preimage of  $x$  has a preimage of  $y$  at a distance  $< a/2\lambda$ . If a preimage of  $y$  is assigned to two preimages of  $x$ , then these preimages of  $x$  have distance smaller than  $a$ , so have to be equal. Therefore, the assignment of preimages is injective. By symmetry the number of preimages is locally constant. Since  $X$  is compact, 1. follows. The proof of 2. is now easy.

### 1.3 Topological Pressure

Let  $(X, T)$  be a continuous dynamical system with compact metric space  $X$ . The metric on  $X$  will be denoted by  $d$ . However, the notion of topological pressure will be independent of the chosen metric.<sup>2</sup>

For each  $n \in \mathbb{N}$  define a metric, equivalent to  $d$ , by

$$d_n(x, y) = \max_{0 \leq k < n} d(T^k(x), T^k(y)) \quad (x, y \in X),$$

which previously has been called the Bowen-metric. For  $\epsilon > 0$ , a subset  $E \subset X$  is called  $(n, \epsilon)$ -spanning, if the  $d_n$ -balls of radius  $\epsilon$  with center in

---

<sup>2</sup> We follow M. Denker: Einführung in die Analysis dynamischer Systeme. Springer 2005.

$E$  cover the space  $X$  and  $E$  cannot be made smaller without loosing the covering property. Likewise a set  $F$  is called  $(n, \epsilon)$ -separating if the  $\epsilon/2$ -balls with center in  $F$  are pairwise disjoint and  $F$  cannot be enlarged without loosing this property.

Let  $\epsilon \leq \delta$ . Every  $(n, \epsilon)$ -spanning set can be reduced to a  $(n, \delta)$ -spanning set and every  $(n, \delta)$ -separating set can be enlarged to a  $(n, \epsilon)$ -separating set. Therefore the following limits exist for a function  $f \in C(X)$ . Let  $\{E_n(\epsilon) : \epsilon_0; n \geq 1\}$  (resp.  $\{F_n(\epsilon) : \epsilon > 0; n \geq 1\}$ ) be a family of  $(n, \epsilon)$ -separating (resp. -spanning) sets.

$$P(T, f) = \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \sum_{x \in E_n(\epsilon)} e^{S_n f(x)}$$

$$Q(T, f) = \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \sum_{x \in F_n(\epsilon)} e^{S_n f(x)}.$$

The limits are independent of the choice of separating and spanning sets. The proof of this fact is left as an exercise.

**Exercise 3.** Show that the definitions of  $P(T, f)$  and  $Q(T, f)$  are independent of the choices of separating and spanning sets.

**Lemma 5.**

$$P(T, f) = Q(T, f).$$

*Proof.* If  $E$  is a  $(n, \epsilon)$ -separating set, for each  $x \in X$  there is  $y \in E$  such that  $K_{d_n}(x, \epsilon) \cap K_{d_n}(y, \epsilon) \neq \emptyset$ . This implies that balls with radius  $2\epsilon$  and center in  $E$  cover  $X$ , so contains a  $(2\epsilon, n)$ -spanning set, and it follows immediately that  $Q(T, f) \leq P(T, f)$ .

Let  $E$  be a  $(n, \epsilon)$ -separating set and  $F$  a  $(n, \epsilon/2)$ -spanning set. For  $x \in E$  there is  $y(x) \in F$  such that  $x \in K_{d_n}(y(x), \epsilon/2)$ . It follows that for different  $x \neq x'$  the points  $y(x)$  and  $y(x')$  are different as well. Because of the continuity of  $f$  and the definition of  $d_n$  we obtain that

$$v(\epsilon) := \sup_{n \in \mathbb{N}} \sup_{d_n(u, v) < \epsilon/2} |S_n f(u) - S_n f(v)| \rightarrow 0$$

when  $\epsilon \rightarrow 0$ , and, moreover,

$$\sum_{x \in E} \exp[S_n f(x)] \leq \sum_{y \in F} \exp[S_n f(y) + nv(\epsilon)].$$

This means that  $P(T, f) \leq Q(T, f)$ .

**Definition 4.**  $P(T, f)$  is called the pressure of the continuous function  $f$ . The map  $P(T, \cdot) : C(X) \rightarrow \mathbb{R}$  is called the pressure function. The pressure of the function  $f = 0$  is called the topological entropy  $h_{\text{top}}(T) = P(T, 0)$  of  $(X, T)$ .

**Proposition 2.** *The pressure function is positive, Lipschitz-continuous, convex and subadditive.*

*Proof.* If  $f$  is positive, then  $\sum_{x \in E} \exp S_n f(x) \geq 1$ , and  $P(T, f) \geq 0$ . Since for a  $(n, \epsilon)$ -separating set  $E$

$$\frac{1}{n} [\log \sum_{x \in E} \exp S_n f(x) - \log \sum_{x \in E} \exp S_n g(x)] \leq K \|f - g\|_\infty$$

with a constant  $K$  independent of  $n$ , Lipschitz-continuity follows immediately. the remaining two properties are shown similarly.

**Proposition 3.** *For  $n \in \mathbb{N}_0$*

$$P(T^n, S_n f) = nP(T, f).$$

*Proof.* The proof is left as an exercise.

**Exercise 4.** Give a proof of Proposition 3.

The central result in the theory of thermodynamic formalism is the variational principle. In order to prepare it we need some notion on probability preserving dynamical systems.

*Remark 1.* We equip a dynamical system  $(X, T)$  with the Borel- $\sigma$ -field  $\mathcal{F}$ . A probability measure  $m$  on  $\mathcal{F}$  is called  $T$ -invariant if for all  $F \in \mathcal{F}$ ,

$$m(T^{-1}(F)) = m(F).$$

Denote by  $\mathcal{M}(T)$  the space of all  $T$ -invariant measures on  $X$ . A measurable partition  $\alpha$  is a pairwise disjoint finite or countable family of measurable sets  $A \in \mathcal{F}$ . The refinement of two partitions  $\alpha$  and  $\beta$  is the partition

$$\alpha \wedge \beta = \{A \cap B : A \in \alpha, B \in \beta\}.$$

We denote

$$\alpha_0^n = \alpha \wedge T^{-1}\alpha \wedge \dots \wedge T^{-n+1}\alpha \quad n \geq 1.$$

The entropy of a partition and of  $m \in \mathcal{M}(T)$  is

$$H_m(\alpha) = - \sum_{A \in \alpha} m(A) \log m(A)$$

and the mean entropy of  $\alpha$  and  $m$

$$h_m(T, \alpha) = \lim_{n \rightarrow \infty} \frac{1}{n} H_m(\alpha_0^n).$$

Finally we define the entropy of  $m$  to be

$$h_m(T) = \sup\{h_m(T, \alpha) : H_m(\alpha) < \infty\}.$$

A basic fact from entropy theory states that for any sequence  $\alpha_n$  of finite entropy partitions which generates  $\mathcal{F}$  (i.e.  $\alpha_n \uparrow$  and  $\bigcup_n \sigma(\alpha_n) = \mathcal{F}$ ), satisfies

$$h_m(T) = \lim_{n \rightarrow \infty} h_m(T, \alpha_n).$$

A simple calculation yields the formula

$$h_m(T^n) = nh_m(T) \quad n \geq 0. \quad (1.2)$$

Finally we shall need the estimate

$$H_m(\alpha \wedge \beta) \leq H_m(\alpha) + H_m(\beta). \quad (1.3)$$

**Theorem 3.** [Variational principle] *For every continuous function  $f \in C(X)$  the variational principle holds:*

$$P(T, f) = \sup \left\{ h_m(T) + \int f dm : m \in \mathcal{M}(T) \right\}.$$

*Proof.* 1. In a first step we show that  $P(T, f) \geq h_m(T) + \int f dm$  for an arbitrary measure  $m \in \mathcal{M}(T)$ .

Let  $\alpha$  be a measurable partition. The concavity of the logarithm and the invariance of  $m$  give

$$\begin{aligned} \frac{1}{n} H_m(\alpha_0^n) + \int f dm &= \frac{1}{n} \sum_{A \in \alpha_0^n} \int_A S_n f dm - m(A) \log m(A) \\ &= \frac{1}{n} \sum_{A \in \alpha_0^n} m(A) \log \left[ \frac{1}{m(A)} \exp \left( \frac{1}{m(A)} \int_A S_n f dm \right) \right] \\ &\leq \frac{1}{n} \log \sum_{A \in \alpha_0^n} \exp \left( \frac{1}{m(A)} \int_A S_n f dm \right) \\ &\leq \frac{1}{n} \log \sum_{A \in \alpha_0^n} \exp \left( \sup_{x \in A} S_n f(x) \right). \end{aligned}$$

Now we choose:

1.  $x_A \in \overline{A} \in \alpha_0^n$ , such that  $S_n f(x_A) = \sup_{x \in A} S_n f(x)$ ,
2.  $\epsilon, \delta > 0$  with:  $d(u, v) < \delta \Rightarrow |f(u) - f(v)| < \epsilon$ ,
3. a  $(n, \delta)$ -separating set  $E_n$ .

For every  $x_A$  there is  $y_A \in E_n$  with  $x_A \in K_{d_n}(y_A, \delta)$ , and thus

$$\frac{1}{n} H_m(\alpha_0^n) + \int f dm \leq \frac{1}{n} \log \sum_{A \in \alpha_0^n} \exp(S_n f(y_A) + n\epsilon).$$

If  $M(n, \delta, \alpha)$  is an upper bound for the number of preimages of the map  $x_A \mapsto y_A$ , we obtain, also passing to the limit as  $n \rightarrow \infty$ , that

$$h(T, \alpha) + \int f dm \leq P(T, f) + \epsilon + \limsup_{n \rightarrow \infty} \frac{1}{n} \log M(n, \delta, \alpha).$$

Here we also used that  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \sum_{x \in E_n} \exp[S_n f(x)] \leq P(T, f)$ .

It is sufficient to construct a generating sequence of partitions  $\alpha$  keeping the values of  $M(n, \delta, \alpha)$  under control. Let  $\beta$  be an arbitrary partition of  $X$  into sets  $B$ , of which the topological boundary  $\partial B$  are null-sets. For  $B \in \beta$  define  $A = A(B) = \{x \in B : d(x, \partial B) > \delta\}$ . If  $\delta \rightarrow 0$ , any generating sequence of  $\beta$  produces a generating sequence

$$\alpha = \{A(B) : B \in \beta\} \cup (X \setminus \bigcup_{B \in \beta} A(B)),$$

since any measure is regular. By construction  $\delta$ -balls can intersect at most two sets of  $\alpha$ , therefore  $M(n, \delta, \alpha) \leq 2^n$ . Finally, if  $\epsilon \rightarrow 0$ , it follows that

$$h_m(T) + \int f dm \leq P(T, f) + 2.$$

Replacing  $T$  by any of its iterates  $T^M$  and  $f$  by  $S_M f = f + \dots + T^{M-1} f$ , we obtain from (1.2) and Proposition 3 the claim, provided  $M \rightarrow \infty$ .

2. The converse inequality needs a little more work..

Let  $\mu$  be a probability and  $m = \frac{1}{nM} \sum_{k=0}^{nM-1} \mu \circ T^{-k}$ . For every finite partition  $\alpha$  one has using (1.3) and the concavity of the function  $-x \log x$  that

$$\begin{aligned} H_m(\alpha_0^M) &= - \sum_{A \in \alpha_0^M} \frac{1}{nM} \sum_{k=0}^{nM-1} \mu(T^{-k}(A)) \log \frac{1}{nM} \sum_{k=0}^{nM-1} \mu(T^{-k}(A)) \\ &\geq \frac{1}{nM} \sum_{k=0}^{nM-1} H_\mu(T^{-k} \alpha_0^M) = \frac{1}{nM} \sum_{j=0}^{M-1} \sum_{k=0}^{n-1} H_\mu(T^{-j-kM} \alpha_0^M) \\ &\geq \frac{1}{nM} \sum_{j=0}^{M-1} H_\mu(\alpha_j^{j+nM}) \geq \frac{1}{n} H_\mu(\alpha_0^{nM}) - \frac{M}{n} \log |\alpha|. \end{aligned} \quad (1.4)$$

Let  $\eta > 0$ . Choose  $\epsilon > 0$  and  $K \subset \mathbb{N}$  such that for  $k \in K$  there is a  $(k, \epsilon)$ -separating set  $E_k$  with  $\log \sum_{x \in E_k} \exp[S_k f(x)] \geq k(P(T, f) - \eta)$ . Introduce a measure  $\mu_k$  by

$$\mu_k = \frac{\sum_{x \in E_k} \exp[S_k f(x)] \delta_x}{\sum_{x \in E_k} \exp[S_k f(x)]}$$

and put  $\tilde{m}_k = \frac{1}{k} \sum_{j=0}^{k-1} \mu_k \circ T^{-j}$  ( $k \in K$ ). Here  $\delta_x$  denotes the point measure in  $x \in X$ . Let  $m$  be a weak limit of the measures  $\{\tilde{m}_k : k \in K\}$ , so w.l.o.g.

$\lim_{k \in K} \tilde{m}_k = m$ . Let  $\alpha$  be a partition in sets of diameter smaller than  $\epsilon$  and with boundaries of measure 0. It follows for fixed  $n \in \mathbb{N}$ , that every atom of  $\alpha_0^n$  contains at most one point of a  $(n, \epsilon)$ -separating set  $E_n$ , hence the map  $E_n \rightarrow \alpha_0^n, x \mapsto A(x) \in \alpha_0^n$  with  $x \in A(x)$  is injective. Now let  $M \in \mathbb{N}$  be fixed. Choose a subsequence  $n_l \uparrow \infty$  and  $0 \leq i_l < M$ , such that  $k_l := n_l M + i_l \in K$ . Setting  $m_{n_l M} = \frac{1}{M n_l} \sum_{j=0}^{n_l M - 1} \mu_{k_l} \circ T^{-j}$ , it follows that  $\lim_{l \rightarrow \infty} m_{n_l M} = m$ . Together with (1.4) we arrive at

$$\begin{aligned}
H_m(\alpha_0^M) + M \int f dm &= \lim_{l \rightarrow \infty} H_{m_{n_l M}}(\alpha_0^M) + M \int f dm_{n_l M} \\
&\geq \lim_{l \rightarrow \infty} \frac{1}{n_l} H_{\mu_{k_l}}(\alpha_0^{n_l M}) + \frac{M}{M n_l} \int S_{k_l} f d\mu_{k_l} - \frac{M}{n_l} \log |\alpha| \\
&\geq \lim_{l \rightarrow \infty} \frac{1}{n_l} H_{\mu_{k_l}}(\alpha_0^{n_l M + i_l}) + \frac{1}{n_l} \int S_{k_l} f d\mu_{k_l} - \frac{2M}{n_l} \log |\alpha| \\
&= \lim_{l \rightarrow \infty} \frac{1}{n_l} \sum_{A \in \alpha_0^{k_l}} \mu_{k_l}(A) \log \frac{1}{\mu_{k_l}(A)} \exp \left[ \frac{1}{\mu_{k_l}(A)} \int_A S_{k_l} f d\mu_{k_l} \right] \\
&\geq \lim_{l \rightarrow \infty} \frac{1}{n_l} \sum_{A \in \alpha_0^{k_l}} \mu_{k_l}(A) \log \sum_{x \in E_{k_l}} \exp[S_{k_l} f(x)] \\
&\geq M(P(T, f) - \eta).
\end{aligned}$$

Dividing by  $M$  and letting  $M \rightarrow \infty$  yields

$$h_m(T, \alpha) + \int f dm \geq P(T, f) - \eta.$$

The entropy of a measure can be calculated via another variational principle using the pressure:

**Theorem 4.** *Let  $(X, T)$  be a continuous dynamical system with compact  $X$  and finite topological entropy. The entropy function  $\mathcal{M}(T) \rightarrow \mathbb{R}_+$ ,  $m \mapsto h_m(T)$ , is upper semi-continuous in the weak topology, i.e.  $\limsup_{\mu \rightarrow m} h_\mu(T) \leq h_m(T)$ , if and only if for every  $m \in \mathcal{M}(T)$*

$$h_m(T) = \inf \{ P(T, f) - \int f dm : f \in C(X) \}. \quad (1.5)$$

*Proof.* Omitted

**Definition 5.** *An invariant measure  $m \in \mathcal{M}(T)$  is called an equilibrium for the continuous function  $f \in C(X)$ , if*

$$P(T, f) = h_m(T) + \int f dm.$$

*In case  $f = 0$ , such measures are called measures of maximal entropy.*

**Theorem 5.** *If the entropy function is upper semi-continuous, then there is an equilibrium for every continuous function  $f$ .*

*Proof.* Let  $m_n \in \mathcal{M}(T)$  satisfy  $P(T, f) = \lim_{n \rightarrow \infty} h_{m_n}(T) + \int f dm_n$ . Then every weak limit point of the sequence  $\{m_n : n \in \mathbb{N}\}$  is an equilibrium.

Let  $\mathcal{L}$  be a class of functions. Recall that two functions  $f$  and  $g$  in  $\mathcal{L}$  are called cohomologous, if there is a function  $h \in \mathcal{L}$  with  $f - g = h - h \circ T$ .

**Proposition 4.** *1. Two cohomologous functions  $f$  and  $g$  in  $C(X)$  have the same equilibria.  
2. If  $f \in C(X)$  and  $f(x) < 0$  for every  $x \in X$ , then the function  $s \mapsto P(T, sf)$ ,  $s \geq 0$ , is strictly monotone decreasing and continuous.*

*Proof.* 1. for every measure  $m \in \mathcal{M}(T)$  we have  $\int f dm = \int g + h - h \circ T dm = \int g dm$ .

2. Let  $s < t$ . The variational principle implies that

$$\begin{aligned} P(T, t \cdot f) &= \sup_{m \in \mathcal{M}(T)} h_m(T) + t \int f dm \\ &\leq \sup_{m \in \mathcal{M}(T)} h_m(T) + s \int f dm + (t - s) \max_{x \in X} f(x) < P(T, s \cdot f). \end{aligned}$$

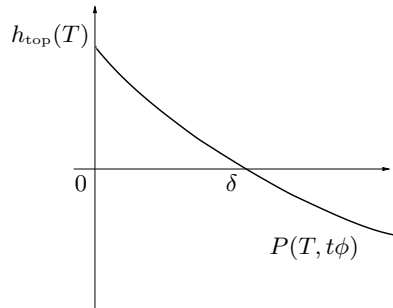
The continuity is a special case of Proposition 2.

**Corollary 2.** *In the situation of Proposition 4 there is a unique  $\delta(f)$  for which the function  $t \mapsto P(T, tf)$  vanishes.*

The formula

$$P(T, \delta(f)f) = 0$$

is called the *Bowen-McClusky-Formula*. The graph of the function  $P(\cdot, f)$  is sketched in figure 1.1.



**Abb. 1.1.** Bowen-McClusky-Formel

### 1.4 Gibbs-Measures

Let  $(X, \mathcal{F}, T, m)$  be a dynamical system on the Lebesgue space  $(X, \mathcal{F})$ , where  $m$  is a non-singular probability measure. Since  $m$  is non-singular,  $m \circ T$  is absolutely continuous with respect to  $m$ , hence has a Radon-Nikodym derivative (density)  $J_m = \frac{dm \circ T}{m}$ . It is called the Jacobi density  $m$ . If  $m$  is forward invariant, the Jacobi density equals 1.

**Definition 6.** Let  $(X, T)$  be a continuous dynamical system. A probability measure  $\mu$  on  $(X, \mathcal{F})$  is called a Gibbs measure for the potential  $\varphi \in C(X)$ , if  $J_m = \exp[-\varphi]$  a.e.

**Theorem 6.** Let  $(X, T)$  be an open and expanding dynamical system. Then, for any  $\varphi \in C(X)$ , there is a Gibbs measure for the potential  $\varphi - P(T, \varphi)$ .

*Proof.* The operator  $\mathcal{L}_{-\varphi}$  is well defined on  $C(X)$  by

$$\mathcal{L}_{-\varphi} f(x) = \sum_{y \in T^{-1}(\{x\})} f(y) \exp[\varphi(y)] \quad x \in X; f \in C(X).$$

By Lemma 2 there is an expansion constant  $a > 0$  and  $\Lambda > 1$ , such that  $T : K(x, a) \rightarrow T(K(x, a))$  is a homeomorphism and  $d(T(y), T(z)) \geq \Lambda d(y, z)$ . Consequently,  $\mathcal{L}_{-\varphi}$  is a well defined, continuous, positive and linear operator (Lemma 4).

The dual operator to  $\mathcal{L}_{-\varphi}$  is denoted by  $\mathcal{L}_{-\varphi}^*$  and operates by Riesz theorem on the space of all finite measures which implies that  $m \mapsto (\int 1 d\mathcal{L}_{-\varphi}^* m)^{-1} \mathcal{L}_{-\varphi}^* m$  is a map which leaves probability measures invariant. Since the space of probabilities is convex and weakly compact, the theorem of Schauder and Tychonoff yields a fixed point  $m$ .  $m$  satisfies  $\mathcal{L}_{-\varphi}^* m = \lambda m$  with eigenvalue  $\lambda = \int 1 d\mathcal{L}_{-\varphi}^* m$ . Let  $f$  be a continuous function vanishing outside of the ball  $K(x, a)$ , and let  $\rho$  denote the inverse function of  $T|_{K(x, a)}$ . It follows that

$$\begin{aligned} \lambda \int f dm &= \int f d\mathcal{L}_{-\varphi}^* m = \int \mathcal{L}_{-\varphi} f dm \\ &= \int f(\rho(z)) \exp[\varphi(\rho(z))] m(dz) = \int f \exp[\varphi] dm \circ T. \end{aligned}$$

Hence the Jacobi density is  $\lambda \exp[-\varphi]$ .

It is left to show that  $\log \lambda = P(T, \varphi)$ . For every  $x \in X$  put  $E_n(x) = T^{-n}(\{x\})$  and choose a  $(n, \epsilon)$ -separating set  $E_n$ . Let  $\omega(\epsilon)$  denote the modulus of continuity of  $\varphi$ . Since

$$\begin{aligned} \lambda^n &= \int \mathcal{L}_{-\varphi}^n 1 dm = \int \sum_{z \in E_n(x)} \exp[S_n \varphi(x)] m(dz) \\ &\leq \sum_{y \in E_n} \exp[S_n \varphi(y) + n\omega(\epsilon)], \end{aligned}$$

it follows that

$$\log \lambda \leq \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \sum_{y \in E_n} \exp[S_n \varphi(y) + n\omega(\epsilon)] = P(T, \varphi).$$

For the proof of the converse inequality one uses the following fact: for every  $\delta > 0$  there is  $m \in \mathbb{N}$ , such that the distance of two points  $y, z \in X$  is bounded by  $\delta$ , provided  $d(T^j(y), T^j(z)) < a$ ,  $j = 0, \dots, m$ . Let  $E_1$  be a  $(1, a)$ -spanning set and  $E_n = T^{-n}E_1$ . Then  $E_{n+m}$  contains a  $(n, \delta)$ -spanning set  $F_n$ , since for  $z \in X$  there is  $x = x(z) \in E_{n+m}$  such that  $d(T^j(x), T^j(z)) < a$  ( $j = 0, \dots, n+m$ ), hence also  $d(T^j(z), T^j(x)) < \delta$  for  $j = 0, 1, \dots, n-1$ . It follows that

$$\sum_{z \in F_n} \exp S_n \varphi(z) \leq \sum_{x \in E_{n+m}} \exp S_n \varphi(x) \leq \sum_{y \in T^{-m}E_1} \mathcal{L}_{-\varphi}^n 1(y).$$

Let  $\alpha$  be a partition of measurable sets of diameter  $< \delta$ , such that each of these sets contains exactly one point of  $T^{-m}E_1$ . It now follows immediately using the modulus of continuity  $\omega_f(\delta) = \sup_{d(x,y) < \delta} |f(x) - f(y)|$  for a function  $f$ , that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{z \in E_{n+m}} e^{S_n \varphi(z)} &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \left[ \sum_{A \in \alpha} \frac{1}{m(A)} \int_A \mathcal{L}_{-\varphi}^n 1 dm \right] + \omega_\varphi(\delta) \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \left[ \int \mathcal{L}_{-\varphi}^n 1 dm \right] + \omega_\varphi(\delta) = \log[\lambda] + \omega(\varphi, \delta). \end{aligned}$$

Letting  $\delta \rightarrow 0$  the proof is complete.

**Proposition 5.** *Let  $(X, T)$  be as in Theorem 6 and topologically transitive. If  $\varphi$  is Hölder continuous, then two Gibbs-measure for the potential  $\varphi - P(T, \varphi)$  are equivalent.*

*Proof.* Consider two Gibbs-measures  $m_1$  and  $m_2$  for the potential  $\varphi - P(T, \varphi)$ . Let  $\alpha$  be a finite Markov partition in sets of diameter  $< a$  having measure zero on their boundaries. Since  $T$  is invertible on each set in  $\alpha$ ,  $T^n$  is as well invertible on each atom  $E \in \alpha_0^n$ , and one has

$$m_i(T^n E) = \int_E \exp[nP(T, \varphi) - S_n \varphi] dm_i \quad i = 1, 2.$$

It is easy to see that  $T^n E$  is an atom of  $\alpha$ . Since  $T$  is topologically transitive, for  $A, B \in \alpha$  there is  $n \in \mathbb{N}$  such that  $C = (A \cap T^{-n}B)^\circ \neq \emptyset$ . It follows that  $C$  is a union of atoms in  $\alpha_0^n$ , and every atom is mapped by  $T^n$  onto  $B$ . If  $B$  has positive measure, so has each of these atoms and thus  $A$ . It follows that  $c := \min\{m_i(A) : i = 1, 2; A \in \alpha\} > 0$ . Choosing points  $x_i \in E$  ( $i = 1, 2$ ) with

$$m_1(E) \exp[nP(T, \varphi) - S_n \varphi(x_1)] \leq \int_E \exp[nP(T, \varphi) - S_n \varphi] dm_1 \leq 1$$

$$c \leq \int_E \exp[nP(T, \varphi) - S_n \varphi] dm_2 \leq m_2(E) \exp[nP(T, \varphi) - S_n \varphi(x_2)],$$

we arrive at

$$\frac{m_1(E)}{m_2(E)} \leq c^{-1} \exp[S_n \varphi(x_1) - S_n \varphi(x_2)].$$

Hölder continuity now yields

$$|S_n \varphi(x_1) - S_n \varphi(x_2)| \leq \sum_{k=0}^{n-1} |\varphi(T^k(x_1)) - \varphi(T^k(x_2))|$$

$$= O\left(\sum_{k=0}^{n-1} d(T^k(x_1), T^k(x_2))^s\right).$$

since the points  $T^k(x_i)$  belong to the same atom of  $\alpha_0^{n-k}$ , and since their diameters decay exponentially fast, the last expression is bounded and independent of  $n$ . It now follows that the Radon-Nikodym density  $\frac{dm_1}{dm_2}$  is bounded from above. Exchanging the role of the two measures shows that both measures are equivalent.

Let  $(X, T)$  be a dynamical system and  $\mu$  a measure. Let  $\mathcal{K}(n, c, \kappa)$  denote the set of all subsets  $C \subset X$  such that:

1.  $T^n : C \rightarrow X$  is injective.
2.  $\text{diam} T^j(C) \leq \kappa^{n-j}$
3.  $\mu(T^n(C)) \geq c$ .

**Theorem 7.** *Let  $(X, T)$  be an open, expanding and topologically transitive dynamical system,  $\varphi \in C(X)$  Hölder continuous and  $\mu$  a Gibbs measure for  $\varphi$ . Then there is a unique invariant probability  $m \sim \mu$  having a Hölder continuous density  $h = \frac{dm}{d\mu}$ .  $m$  is a Gibbs measure for the potential*

$$\varphi - P(T, \varphi) + \log h \circ T - \log h.$$

Moreover, there is a constant  $K$ , such that

$$K^{-1} \leq \frac{m(C)}{\exp[nP(T, \varphi) - S_n \varphi(x)]} \leq K \quad (1.6)$$

for every  $x \in C$  and  $C \in \mathcal{K}(n, c, \kappa)$ .  $m$  is the only invariant measure, which satisfies (1.6) and, in particular, the unique Gibbs measure for a potential, cohomologous to  $\varphi - P(T, \varphi)$  via a bounded coboundary.

The measure  $m_\varphi = m$  is called the invariant Gibbs measure for  $\varphi$ .

*Proof.* Existence and uniqueness will be omitted at this point and deferred to the next chapter. Denote by  $h$  the density  $dm/d\mu$ .

It is easy to show that  $m$  is a Gibbs measure:

$$\begin{aligned} \int \mathcal{L}_{-\varphi+P(T,\varphi)-\log h \circ T + \log h} g dm &= \int \mathcal{L}_{-\varphi+P(T,\varphi)} gh/h \circ T dm \\ &= \int \mathcal{L}_{-\varphi+P(T,\varphi)} gh dm \mu = \int g dm. \end{aligned}$$

The property (1.6) follows from the fact, that for  $x, y \in C \in \mathcal{K}(n, c, \kappa)$

$$|S_n \varphi(x) - S_n \varphi(y)| \leq D_\varphi \sum_{k=0}^{\infty} \kappa^{sk} =: M < \infty,$$

$$c \leq m(T^n(C)) \leq 1$$

and therefore

$$ce^{-M} \leq \mu(C) \exp[S_n \varphi(x) - nP(T, \varphi)] \leq m(T^n(C)) e^M = e^M.$$

The claim follows from this, since  $m$  and  $\mu$  are equivalent. Moreover, by the same inequality, there are no two invariant ergodic measures with the property in (1.6), and therefore invariant Gibbs measures are unique.

We denote by  $\text{Lip}(s)$  the space of Hölder continuous functions with exponent  $s$ .

**Theorem 8.** *The pressure function  $P(\cdot) = P(T, \cdot) : \text{Lip}(s) \rightarrow \mathbb{R}$  for an expanding, topologically transitive and open dynamical system is real analytic. We have  $\varphi, \psi, \psi_1, \psi_2 \in \text{Lip}(s)$ :*

1.  $\frac{d}{dt} P(\varphi + t\psi)|_{t=0} = \int \psi dm_\varphi$ , where  $m = m_\varphi$  is the invariant Gibbs measure for  $\varphi$ .
2.  $\frac{d^2}{dt ds} P(\varphi + t\psi_1 + s\psi_2)|_{t=0} = D_\varphi(\psi_1, \psi_2)$ , where

$$D_\varphi(f, g) = \sum_{k=0}^{\infty} \int (f - \int f dm_\varphi)(g \circ T^k - \int g dm_\varphi) dm_\varphi$$

denotes the asymptotic covariance of the functions  $f$  and  $g$  under the invariant Gibbs measure  $m_\varphi$ .

*Proof.* The proof uses perturbation theory which cannot be developed at this point. We refer to the literature: Parry, Pollicott: Zeta functions and the periodic orbit structure of hyperbolic dynamics. Asterisque **187-8**, 1990.

*Example 1.* Let  $\varphi$  and  $\psi$  be two Hölder continuous functions with  $P(T, \varphi) = 0$  and  $\psi < 0$ . The function

$$s \mapsto P(T, s\psi + q\varphi) =: P(s\psi + p\varphi) \quad q \in \mathbb{R}$$

has, using an appropriate variation of Proposition 4, a unique zero at  $S(q)$ . Since the function  $(s, q) \mapsto s\psi + q\varphi$  is real analytic, so is  $(s, q) \mapsto P(T, s\psi + q\varphi)$  by Theorem 8. Hence the function  $q \mapsto S(q)$  is real analytic by the implicit function theorem, provided the partial derivative for the variable  $s$  does not vanish. According to Theorem 8 we get

$$\frac{d}{dx}P(x\psi + q\varphi)|_{x=s} = \int \psi dm_{s,q}, \quad \frac{d}{dy}P(s\psi + y\varphi)|_{y=q} = \int \varphi dm_{s,q}, \quad (1.7)$$

where  $m_{s,q}$  denotes the invariant Gibbs measure for  $s\psi + q\varphi$ . Since  $\psi < 0$ , the partial derivative for  $s$  does not vanish.

From (1.7) one immediately computes  $S'(q)$ , since

$$0 = \frac{d}{dq}P(S(q)\psi + q\varphi) = \int \psi dm_{S(q),q} S'(q) + \int \varphi dm_{S(q),q}$$

shows that

$$S'(q) = -\frac{\int \varphi dm_{S(q),q}}{\int \psi dm_{S(q),q}} =: -\alpha(q).$$

The second derivative of  $S$  can be obtained in a similar way differentiating twice. Using Theorem 8 one obtains

$$S''(q) = \frac{D_{S(q)\psi + q\varphi}(\varphi - S'(q)\psi, \varphi - S'(q)\psi)}{-\int \psi dm_{S(q),q}}. \quad (1.8)$$

Since  $\psi < 0$ , the second derivative is always positive. Since the numerator vanishes if and only if  $S(q)\psi + q\varphi$  is cohomologous to a constant, it follows that  $S$  is strictly convex, in case it is not cohomologous to a constant.

## 2 Stochastic Laws in Dynamics

This is a minicourse given at the Universidade Federal, Universidade Federal Fluminense and Pontificia Universidade Catolica in Rio de Janeiro in May 2009 and the workshop on "Dynamical Systems and Related Topics" at Göttingen University, July 2009.

### 2.1 Introduction

The central limit theorem (CLT) is one of the oldest theorems in mathematical sciences and goes back to the late 17th, early 18th century when Bernoulli and deMoivre investigated limit laws for random outcomes of win/loss strategies. The theorem states that the deviation from the mean follows a normal law. More precisely: Let  $(X, \mathcal{F}, P)$  be a probability space with sample space  $X$ ,  $\sigma$ -field  $\mathcal{F}$  and probability measure  $P$ . For a sequence of measurable functions  $f_i : X \rightarrow \mathbb{R}$  let

$$S_n = f_1 + \dots + f_n.$$

**Theorem 9.** (CLT for independent, identically distributed functions). Let  $f_i$  be independent and identically distributed measurable functions as above. If  $f_1 \in L_2(P)$  and  $\sigma^2 = \int f_1^2 dP - (\int f_1 dP)^2 > 0$ , then

$$\lim_{n \rightarrow \infty} \sup_{t \in \mathbb{R}} P \left( x \in X : S_n(x) \leq n \int f_1 dP + t\sqrt{n} \right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^t e^{-u^2/(2\sigma^2)} du. \quad (2.1)$$

DeMoivre proved this theorem when the functions  $f_i$  take only two values.

If  $T : X \rightarrow X$  is a measurable and  $P$ -invariant transformation, a special sequence of function is defined by  $f_i = f \circ T^{i-1}$ ,  $i = 1, 2, 3, \dots$ , where  $f : X \rightarrow \mathbb{R}$  is measurable. If (2.1) holds we say that  $f$  satisfies the central limit theorem.

**Corollary 3.** Let  $(X, \mathcal{F}, m, T)$  be a probability preserving dynamical system. Let  $\alpha$  be an independent partition (i.e. the family of  $\sigma$ -algebras generated by  $T^{-1}\alpha$  for  $i \geq 0$  are independent). Then any non-constant function  $f$  which is measurable with respect to  $\alpha$  and is square integrable, satisfies the CLT.

As we know the existence of independent partitions is too restrictive for the CLT to hold. If the entropy of  $T$  is positive one can find independent partitions by Sinai's weak isomorphism theorem, hence functions satisfying the CLT exist. In case the entropy is zero, one can still find a dense set of functions, provided the system is aperiodic. Hence the question to discuss becomes:

When does a function  $f \in L_2(m)$  satisfy the CLT?

The most general approach to this question is given by Gordin's martingale approximation, which we will discuss first.

**Definition 7.** Let  $(\Omega_n, \mathcal{F}_n, P_n)$  ( $n \geq 1$ ) be probability spaces and  $(E, \mathcal{B})$  a measurable space. A family  $\{X_{jn} : j = 1, \dots, k_n, n = 1, 2, 3, \dots\}$  of random elements

$$X_{jn} : \Omega_n \rightarrow E$$

is called an array of  $E$ -valued random elements. It is called independent, if for every  $n \geq 1$  the random elements  $X_{jn}$  ( $j = 1, \dots, k_n$ ) are independent.

**Definition 8.** An array  $\{X_{jn} : 1 \leq j \leq k_n, n \geq 1\}$  of  $\mathbb{R}^d$ -valued random vectors is called asymptotically negligible, if for every  $\epsilon > 0$

$$\lim_{n \rightarrow \infty} \sup_{1 \leq j \leq k_n} P_n(\|X_{jn}\| \geq \epsilon) = 0.$$

This is equivalent to the statement, that the sequences  $(X_{j_n n})_{n \geq 1}$  for arbitrary choices  $1 \leq j_n \leq k_n$  converge to zero uniformly in the sense of stochastic convergence.

**Definition 9.** Let  $\{X_{jn} : 1 \leq j \leq k_n, n \geq 1\}$  be an array of square integrable real valued random variables, i.e.  $X_{jn} \in L_2(P_n)$  ( $1 \leq j \leq k_n, n \geq 1$ ). Denote

$$s_n^2 = \sigma^2(X_{1n} + X_{2n} + \dots + X_{k_n n}) = \sigma^2(X_{1n}) + \dots + \sigma^2(X_{k_n n}).$$

The array is said to satisfy the Lindeberg condition, if for every  $\epsilon > 0$  the quantities

$$\begin{aligned} L_n(\epsilon) &:= \frac{1}{s_n^2} \sum_{j=1}^{k_n} \int_{\{|X_{jn} - E(X_{jn})| \geq \epsilon s_n\}} (X_{jn} - E(X_{jn}))^2 dP_n \\ &= \frac{1}{s_n^2} \sum_{j=1}^{k_n} E \left( 1_{[\epsilon s_n, \infty)}(|X_{jn} - E(X_{jn})|) [X_{jn} - E(X_{jn})]^2 \right) \end{aligned}$$

tend to 0 as  $n \rightarrow \infty$ .

The main theorem in the theory of distributional convergence in probability theory is this:

**Theorem 10.** (*Lindeberg Central Limit Theorem*)

Let  $\{X_{jn} : 1 \leq j \leq k_n, n \geq 1\}$  be an independent array of square integrable random variables with  $\sigma^2(X_{jn}) > 0$  for  $1 \leq j \leq k_n, n \geq 1$ . Then the following two statements are equivalent:

- (1) The array satisfies the Lindeberg condition.
- (2) The array  $\{X_{jn} - E(X_{jn}) : 1 \leq j \leq k_n, n \geq 1\}$  is asymptotically negligible, and the distributions of

$$\frac{X_{1n} - E(X_{1n}) + \dots + X_{k_n n} - E(X_{k_n n})}{s_n}$$

converge weakly to the standard normal distribution  $\mathcal{N}(0, 1)$ .

## 2.2 Martingale Central Limit Theorem

A similar theorem is the Lindeberg central limit theorem for martingale differences. The extension to arrays is known as well, but we restrict to the case of a single sequence. This theorem is the backbone for distributional convergence in dynamical systems.

**Theorem 11.** (*Existence of conditional expectation*)

Let  $\mathcal{A} \subset \mathcal{F}$  be a  $\sigma$ -subalgebra of  $\mathcal{F}$ . Then for every function  $f \in L_1(m)$  there exists a  $\mathcal{A}$ -measurable function  $g \in L_1(m)$ , which satisfies

$$\int_A f \, dm = \int_A g \, dm \tag{BE1}$$

for every  $A \in \mathcal{A}$ . The function  $g$  is a.e. uniquely determined by this equality and by the measurability condition.

**Definition 10.** The function  $g$  is called the conditional expectation of  $f$  given  $\mathcal{A}$  and will be written as

$$g = E(f|\mathcal{A}).$$

In case the  $\sigma$ -algebra  $\mathcal{A}$  is the smallest  $\sigma$ -algebra for which the functions  $h_i$  or random variables  $X_i$  are measurable, or which contains a collection  $\Sigma$  of subsets, these generating quantities appear instead of  $\mathcal{A}$  in the notation: e.g.  $E(X|Y)$ ,  $E(f|h_1, \dots, h_n, \dots)$

*Proof.* Let  $m_0$  be the restriction of the set function  $m$  on  $\mathcal{F}$  to  $\mathcal{A}$ . ( $m_0$  is, of course, again a probability measure.) Define

$$\Lambda(A) = \int_A f \, dm \tag{A \in \mathcal{A}}.$$

$\Lambda$  is a finite signed measure on  $(\Omega, \mathcal{A})$ , and by the theorem of Radon Nikodym there exists a function  $g \in L_1(m_0)$  satisfying

$$\Lambda(A) = \int_A g \, dm_0.$$

$g$  is a.e. uniquely determined, since for another function  $g'$ ,  $\mathcal{A}$ -measurable and having this property, it follows for any  $A \in \mathcal{A}$  that

$$\int_A g \, dm = \int_A g' \, dm.$$

Therefore  $g = g'$  a.e. We also may consider  $g$  as a function in  $L_1(m)$  and hence we have

$$\Lambda(A) = \int_A g \, dm.$$

**Theorem 12.** *The conditional expectation (for  $L_1$ -functions) has the following properties:*

(BE2) *Linearity:*

$$E(\alpha f + \beta g | \mathcal{A}) = \alpha E(f | \mathcal{A}) + \beta E(g | \mathcal{A}) \quad \text{a.e.}$$

for  $f, g \in L_1(m)$  and  $\alpha, \beta \in \mathbb{R}$ .

(BE3) *Positivity:*  $f \geq 0 \implies E(f | \mathcal{A}) \geq 0$ .

(BE4) *Monotonicity:* If  $f \leq g$ , then  $E(f | \mathcal{A}) \leq E(g | \mathcal{A})$  a.e.

(BE5) *Convergence:*  $f_n \rightarrow f$  a.e.,  $|f_n| \leq h \in L_1(m)$  implies that

$$\lim_{n \rightarrow \infty} E(f_n | \mathcal{A}) = E(f | \mathcal{A})$$

a.e. and in mean.

(BE6) *Connection to integrals:* If  $\mathcal{A} = \{\emptyset, \Omega\}$  mod  $m$ , then

$$E(f | \mathcal{A}) = \int f \, dm \quad (f \in L_1(m)).$$

(BE7) *If  $f$  is  $\mathcal{A}$ -measurable, then  $E(f | \mathcal{A}) = f$  a.e.*

**Definition 11.** *Let  $X_n$  ( $n \geq 1$ ) be a sequence of random variables, adapted to the increasing sequence  $\mathcal{F}_n$  of  $\sigma$  fields in  $\mathcal{F}$ . Then  $(X_n, \mathcal{F}_n)_{n \geq 1}$  is called a martingale difference sequence if for each  $n \in \mathbb{N}$*

$$E(X_{n+1} | \mathcal{F}_n) = 0.$$

Note that the sum  $M_n = X_1 + \dots + X_n$  defines a martingale and that  $E(X_n) = 0$  for every  $n \in \mathbb{N}$ .<sup>1</sup>

<sup>1</sup> A martingale is an integrable sequence  $M_n$  ( $n \geq 1$ ) of random variables such that  $E(M_{n+1} | M_1, \dots, M_n) = M_n$ ,  $n \geq 1$ .

**Theorem 13.** Let  $(X_n, \mathcal{F}_n)$  ( $n \geq 1$ ) be a martingale difference sequence, such that the array  $X_{kn} = X_k$  for  $k = 1, \dots, n$  and  $n \geq 1$  satisfies the Lindeberg condition. Moreover, assume that

$$\sum_{j=1}^n E |E(X_j^2 | \mathcal{F}_{j-1}) - \sigma_j^2| = o(s_n^2),$$

where

$$\sigma_j^2 = E(X_j^2) \quad \text{and} \quad s_n^2 = \sum_{j=1}^n \sigma_j^2.$$

Then

$$\lim_{s_n} \frac{1}{s_n} \sum_{j=1}^n X_j = \mathcal{N}(0, 1)$$

in the weak topology.

*Proof.* In the proof we use the following estimate:

**Lemma 6.** For all  $\delta \in [0, 1]$

$$\left| e^{itx} - \sum_{j=0}^n \frac{(it)^j}{j!} \right| \leq \frac{2^{1-\delta} |t|^{n+\delta}}{(1+\delta)(2+\delta)\dots(n+\delta)}.$$

Now let  $t \in \mathbb{R}$  be fixed. Let  $\epsilon > 0$ . Then, using  $M_n = X_1 + \dots + X_n$  as before,

$$\begin{aligned} & \left| E \left( e^{itM_n/s_n} \right) - e^{t^2/2} \right| \\ &= \left| e^{-t^2/2} \sum_{j=1}^n E \left( e^{itM_j/s_n + s_j^2 t^2 / (2s_n^2)} - e^{itM_{j-1}/s_n + s_{j-1}^2 t^2 / (2s_n^2)} \right) \right|. \end{aligned}$$

Observe that the latter is a telescoping sum! Evaluating the expectations appearing in the sum, we obtain

$$\begin{aligned} & \left| E \left( e^{itM_j/s_n + s_j^2 t^2 / (2s_n^2)} - e^{itM_{j-1}/s_n + s_{j-1}^2 t^2 / (2s_n^2)} \right) \right| \\ &= \left| E \left[ E \left( e^{itM_j/s_n + s_j^2 t^2 / (2s_n^2)} - e^{itM_{j-1}/s_n + s_{j-1}^2 t^2 / (2s_n^2)} \mid \mathcal{F}_{j-1} \right) \right] \right| \\ &= \left| E e^{itM_{j-1}/s_n + s_{j-1}^2 t^2 / (2s_n^2)} \left[ E \left( e^{itX_j/s_n} - e^{-\sigma_j^2 t^2 / (2s_n^2)} \mid \mathcal{F}_{j-1} \right) \right] \right| \\ &\leq e^{t^2/2} E \left| E \left( e^{itX_j/s_n} - e^{-\sigma_j^2 t^2 / (2s_n^2)} \mid \mathcal{F}_{j-1} \right) \right|. \end{aligned}$$

In these estimates we used the facts that  $E(Y) = E(E(Y|\mathcal{A}))$ , and that  $|e^{itM_{j-1}/s_n}| \leq 1$ .

Now we use the lemma with  $\delta = 1$  and  $n = 1, 2$  (three times) to get

$$\begin{aligned}
& \left| E \left( e^{itX_j/s_n} - e^{-\sigma_j^2 t^2 / (2s_n^2)} \middle| \mathcal{F}_{j-1} \right) \right| \\
&= \left| E \left( e^{itX_j/s_n} \middle| \mathcal{F}_{j-1} \right) - e^{-\sigma_j^2 t^2 / (2s_n^2)} \right| \\
&= \left| E \left( e^{itX_j/s_n} - 1 - \frac{itX_j}{s_n} + \frac{t^2 X_j^2}{2s_n^2} \middle| \mathcal{F}_{j-1} \right) + 1 - \frac{\sigma_j^2 t^2}{2s_n^2} - e^{-\sigma_j^2 t^2 / (2s_n^2)} \right. \\
&\quad \left. + \frac{\sigma_j^2 t^2}{2s_n^2} - E \left( \frac{t^2 X_j^2}{2s_n^2} \middle| \mathcal{F}_{j-1} \right) \right| \\
&\leq E \left( \frac{t^2 X_j^2}{s_n^2} \mathbb{I}_{\{|X_j| > \epsilon s_n\}} \middle| \mathcal{F}_{j-1} \right) + E \left( \left| \frac{tX_j}{s_n} \right|^3 \mathbb{I}_{\{|X_j| \leq \epsilon s_n\}} \middle| \mathcal{F}_{j-1} \right) + \frac{\sigma_j^4 t^4}{8s_n^4} + b_j,
\end{aligned}$$

where

$$b_j = \left| E \left( \frac{t^2 X_j^2}{2s_n^2} \middle| \mathcal{F}_{j-1} \right) - \frac{\sigma_j^2 t^2}{2s_n^2} \right|.$$

Next observe that the Lindeberg condition implies that

$$\max_{1 \leq j \leq n} \sigma_j^2 s_n^{-2} \leq \max_{1 \leq j \leq n} s_n^{-2} \left[ \epsilon^2 s_n^2 + \int_{\{|X_j| > \epsilon s_n\}} X_j^2 dP \right] = \epsilon^2 + L_n(\epsilon).$$

Putting everything together we arrive at

$$\begin{aligned}
& \left| E \left( e^{itM_n/s_n} \right) - e^{t^2/2} \right| \\
&\leq e^{-t^2/2} \sum_{j=1}^n e^{t^2/2} E \left| E \left( e^{itX_j/s_n} - e^{-\sigma_j^2 t^2 / (2s_n^2)} \middle| \mathcal{F}_{j-1} \right) \right| \\
&\leq \sum_{j=1}^n E \left( E \left( \frac{t^2 X_j^2}{s_n^2} \mathbb{I}_{\{|X_j| > \epsilon s_n\}} \middle| \mathcal{F}_{j-1} \right) \right) \\
&\quad + E \left( E \left( \left| \frac{tX_j}{s_n} \right|^3 \mathbb{I}_{\{|X_j| \leq \epsilon s_n\}} \middle| \mathcal{F}_{j-1} \right) \right) + \frac{\sigma_j^4 t^4}{8s_n^4} + b_j \\
&\leq \sum_{j=1}^n E \left( \frac{t^2 X_j^2}{s_n^2} \mathbb{I}_{\{|X_j| > \epsilon s_n\}} \right) \\
&\quad + E \left( \left| \frac{tX_j}{s_n} \right|^3 \mathbb{I}_{\{|X_j| \leq \epsilon s_n\}} \middle| \mathcal{F}_{j-1} \right) + \max_{1 \leq l \leq n} \sigma_l^2 s_n^{-2} \frac{\sigma_j^2 t^4}{8s_n^2} + b_j \\
&\leq L_n(\epsilon) + \epsilon |t|^3 + (L_n(\epsilon) + \epsilon^2) t^4 / 8 + \sum_{j=1}^n b_j.
\end{aligned}$$

Observe that  $\sum_{j=1}^n b_j$  tends to zero as  $n$  tends to infinity by assumption. Hence, letting  $n$  tend to infinity and then  $\epsilon$  to zero, shows that

$$\lim_{n \rightarrow \infty} E \left( e^{itM_n/s_n} \right) = e^{t^2/2}$$

for every  $t \in \mathbb{R}$ . By the continuity theorem of Lévy, a probability distribution is determined by its characteristic function and weak convergence is equivalent to the convergence of the corresponding characteristic functions. This implies the convergence to normal distribution.

### 2.3 Martingale Approximation for Dynamical Systems

The CLT for dynamical systems is (at least today) based on Gordin’s approach from 1968, which is the application of the martingale CLT to exact transformations in the non-invertible case (resp.  $\sigma$ -fields with the K-property in the invertible case). Here we shall follow the non-invertible case, because it is a bit easier in notation.

Recall the notion of a probability preserving dynamical system  $(X, \mathcal{F}, P, T)$ :  $T : X \rightarrow X$  is a measurable map and  $(X, \mathcal{F}, P)$  is a probability space such that  $P(T^{-1}(F)) = P(F)$  for all measurable sets  $F \in \mathcal{F}$ . If  $T$  is non-invertible, then

$$T^{-1}\mathcal{F} \subset \mathcal{F}$$

is strictly contained. In fact  $T^{-1}(F)$  contains, together with  $x$ , all points equivalent to  $x$  where two points are equivalent if their images agree. In the remaining part of this chapter we shall assume that  $T$  is non-invertible.

As usual, such systems can be ergodic, weakly mixing, mixing, mixing of all orders or Bernoulli. Here we need another type of mixing property which is described using the notion of tail- $\sigma$ -algebras of transformations. The tail field of  $T$  is

$$\mathcal{F}_\infty = \bigcap_{n=0}^{\infty} T^{-n}\mathcal{F}.$$

**Definition 12.** *The transformation  $T$  is called exact, if its tail field  $\mathcal{F}_\infty$  is trivial.*

**Exercise 5.** Show that the following transformations are exact:

1. Let  $X = \{1, \dots, s\}^{\mathbb{N}}$  and  $T(x_k)_{k \in \mathbb{N}} = (y_l)_{l \in \mathbb{N}}$  with  $y_l = x_{l+1}$ . Let  $P$  be a Markov measure defined a probability vector  $\pi = (\pi_1, \dots, \pi_s)$  and a transition matrix  $\mathbb{P} = (p_{i,j})_{1 \leq i,j \leq s}$ :

$$P(\{x \in X : x_i = a_i, 1 \leq i \leq n\}) = \pi_{a_1} p_{a_1, a_2} \cdot \dots \cdot p_{a_{n-1}, a_n}.$$

2. (difficult) Let  $X$  denote the unit interval  $[0, 1)$  and  $T(x) = \beta x \bmod 1$ . The invariant probability is absolutely continuous with respect to Lebesgue measure.

We denote by  $Uf = f \circ T$  the isometry on each  $L_p(\mathcal{F})$  induced by  $T$  (Koopman operator). An invariant sub- $\sigma$ -algebra  $\mathcal{B}_0 \subset \mathcal{F}$  defines a monotone family of subspaces in  $L_2(P)$  by  $L_2(T^{-k}\mathcal{B}_0) = U^k L_2(\mathcal{B}_0)$ . (We let  $L_2(\mathcal{B})$  denote the space of all square-integrable functions, which are  $\mathcal{B}$ -measurable.) Let  $G(\mathcal{B}_0)$  denote the set of all functions  $g \in U^k L_2(\mathcal{B}_0) \ominus U^l L_2(\mathcal{B}_0)$  for indices  $k \leq l$ . If  $T$  is an endomorphism there is a canonical invariant  $\sigma$ -algebra  $\mathcal{B}_0$ , that is  $\mathcal{F}$ . Let  $\mathbf{P}_k$  denote the projection onto  $U^k L_2(\mathcal{B}_0)$ .

**Theorem 14.** [GORDIN] *Let  $(X, \mathcal{F}, P, T)$  be an ergodic endomorphism. For every function  $f \in L_2(m)$ , satisfying the condition*

$$\inf_{g \in G(\mathcal{B})} \limsup_{n \rightarrow \infty} n^{-1/2} \|S_n(f - g)\|_{L_2(m)} = 0 \quad (2.2)$$

*the sum  $\sigma_f^2 = \lim_{n \rightarrow \infty} n^{-1/2} \|S_n f\|_{L_2(m)}^2$  converges, and  $f$  satisfies the Central Limit Theorem with limiting distribution  $\mathcal{N}(0, \sigma_f^2)$ .*

The latter assertion means the convergence of

$$\lim_{n \rightarrow \infty} P(\{x \in X : \frac{1}{\sqrt{n}} S_n f(x) \leq t\}) = \frac{1}{\sqrt{2\pi\sigma_f^2}} \int_{-\infty}^t \exp[-u^2/2\sigma_f^2] du$$

for every  $t \in \mathbb{R}$ , if  $\sigma_f > 0$ . If  $\sigma_f = 0$ , one has convergence to 0 or 1, depending whether  $t < 0$  or  $t > 0$  (this means convergence to zero in probability). Figure 2.1 shows this convergence for the sums  $\frac{1}{\sqrt{n}} S_n f$ . A histogram is a graph showing the relative frequencies of this function when the starting values is chosen at random.

*Proof.* Let  $\|\cdot\|_2$  denote the norm in  $L_2(P)$  and  $\tilde{\mathbf{P}}_l$  the projection onto the subspace  $U^l L_2(\mathcal{F}) \ominus U^{l+1} L_2(\mathcal{F})$ . Let  $\epsilon > 0$  and  $g \in G(\mathcal{F})$  satisfy  $\limsup_{n \rightarrow \infty} n^{-1/2} \|S_n(f - g)\|_2 < \epsilon$ . Note that the adjoint operator  $U^*$  maps the subspace  $L_2(\mathcal{F}_k) \ominus L_2(\mathcal{F}_{k+1})$  onto  $L_2(\mathcal{F}_{k-1}) \ominus L_2(\mathcal{F}_k)$ . Therefore,  $U^{*k} \tilde{\mathbf{P}}_k g \in L_2(\mathcal{B}) \ominus U L_2(\mathcal{B})$ . Since

$$\begin{aligned} f &= g + f - g = \sum_{l=0}^{\infty} \tilde{\mathbf{P}}_l g + f - g \\ &= \sum_{l=0}^{\infty} U^{*l} \tilde{\mathbf{P}}_l g + \sum_{l=0}^{\infty} \sum_{j=0}^{l-1} U^{*j} \tilde{\mathbf{P}}_l g - U^* \left( \sum_{l=0}^{\infty} \sum_{j=0}^{l-1} U^{*j} \tilde{\mathbf{P}}_l g \right) + f - g \end{aligned}$$

$f$  can be written in the form  $f = h + h_1 - U^* h_1 + f - g$ , hence it follows that  $h \in L_2(\mathcal{F}) \ominus U^* L_2(\mathcal{F})$  and

$$n^{-1/2} \|S_n(f - h)\|_2 \leq n^{-1/2} \|h_1 - U^n h_1\|_2 + n^{-1/2} \|S_n(f - g)\|_2.$$

Consequently  $n^{-1/2} S_n f$  and  $n^{-1/2} S_n h$  have the same limiting distribution, when  $n \rightarrow \infty$  and  $g \rightarrow f$ . For the process  $(U^k h)_{k \geq 0}$  it holds that  $U^k h$  is  $\mathcal{F}_k$ -measurable and for every  $\mathcal{F}_{k+1}$ -measurable function  $u$  one has  $\int u U^k h dP = 0$ .

It follows that the process forms a martingale difference sequence and has a normal limiting distribution according to Theorem 13.

**Exercise 6.** Verify that the Martingale Central Limit Theorem applies.

The variance is given by  $\sigma_h^2 = \|h\|_2^2$ . For two approximating functions  $h, h' \in G(\mathcal{F})$  one has

$$\begin{aligned} |\sigma_h - \sigma_{h'}| &\leq \|h - h'\|_2 = \limsup_{n \rightarrow \infty} n^{-1/2} \|S_n(h - h')\|_2 \\ &\leq \limsup_{n \rightarrow \infty} n^{-1/2} (\|S_n(f - h)\|_2 + \|S_n(f - h')\|_2) \rightarrow 0 \end{aligned}$$

whenever  $h, h' \rightarrow f$ . Hence  $\sigma_f^2 = \lim_{n \rightarrow \infty} n^{-1/2} \|S_n f\|_2^2$  exists, and  $n^{-1/2} S_n f$  converges in distribution to the normal probability measure with expectation zero and variance  $\sigma_f^2$  (the degenerate case  $\sigma_f = 0$  is included).

Let  $\mathbf{P}_k$  denote the projection onto  $U^k L_2(\mathcal{F})$ .

**Corollary 4.** *The condition (2.2) of the theorem holds for  $f \in L_{2+\delta}(P)$  ( $\delta \geq 0$ ), if*

$$\sum_{k \geq 0} \|\mathbf{P}_k f\|_{2+\delta} < \infty.$$

Here  $\|h\|_p = \|h\|_{L_p(P)}$  for brevity.

*Proof.* Let  $f_k = \mathbf{P}_k f$ . Then  $f - f_k \in G(\mathcal{F})$ , and using Hölder's inequality in case  $\delta > 0$  (for  $\delta = 0$  use Cauchy-Schwarz) it follows that

$$\begin{aligned} \limsup_{n \rightarrow \infty} n^{-1} \|S_n \mathbf{P}_k f\|_2^2 &= \limsup_{n \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \langle U^i \mathbf{P}_k f, U^j \mathbf{P}_k f \rangle \\ &\leq 2 \limsup_{n \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \sum_{j=0}^i \langle U^{i-j} \mathbf{P}_k f, \mathbf{P}_k f \rangle \\ &\leq 2 \limsup_{n \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \|\mathbf{P}_{k+j} f\|_{2+\delta} \|\mathbf{P}_k f\|_{\frac{2+\delta}{1+\delta}} \\ &\leq 2 \|f\|_{\frac{2+\delta}{1+\delta}} \sum_{i=k}^{\infty} \|P_i f\|_{2+\delta}. \end{aligned}$$

*Example 2.* The approximation of the distribution of the partial sums  $S_n f$  by a normal distribution is shown in figure 2.1 using the  $\beta$ -transformation  $T(x) = 2.3x \bmod 1$ . The function  $\frac{1}{\sqrt{60}} S_{60} \mathbf{I}$  ( $\mathbf{I}$  denotes as usual the indicator function) has been calculated at 300 points, resulting in the shown histogram. It can well be approximated by the density of a normal distribution.

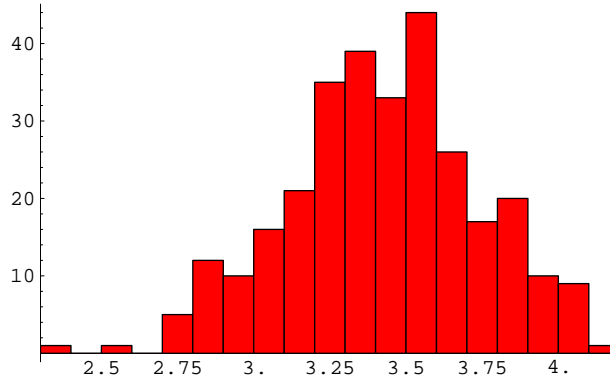


Abb. 2.1. Histogram of the distribution of  $\frac{1}{\sqrt{n}}S_n f$

## 2.4 The Central Limit Theorem for R-Expanding Systems

Let  $(X, T)$  be a topologically mixing R-expanding dynamical system and  $\phi$  a Hölder-continuous function with exponent  $s$  (this means that  $T$  is an open and expanding map). Our goal is to produce an invariant Gibbs measure for  $\phi$  and to show the CLT for every other Hölder continuous function  $f$  with respect to this invariant measure.

The following theorem which we state without proof, permits to construct invariant measures for R-expanding dynamical systems. Let  $E$  be a Banach space with norm  $\|\cdot\|_E$  and  $F \subset E$  closed subspace. Assume that  $F$  is equipped with a further norm  $\|\cdot\|_F$ , such that the identity  $I : (F, \|\cdot\|_F) \rightarrow (F, \|\cdot\|_E)$  is a contraction. A continuous linear operator  $V : F \rightarrow F$  is called *relatively compact*, if  $V$  maps  $\|\cdot\|_F$ -bounded sets onto  $\|\cdot\|_E$ -relatively compact sets. Moreover, an operator  $V$  is called *power bounded*, if  $\sup_{n \in \mathbb{N}} \|V^n\| < \infty$ .

**Theorem 15.** [DOEBLIN, FORTET, IONESCU-TULCEA, MARINESCU] *Let  $V : F \rightarrow F$  continuous in both norms and relatively compact. Moreover, assume that  $V$  is power bounded with respect to the  $\|\cdot\|_E$ . Assume that the Doeblin-Fortet inequality holds:*

$$\|V(x)\|_F \leq r\|x\|_F + R\|x\|_E \quad x \in F \quad (2.3)$$

where  $0 < r < 1$  and  $R \in \mathbb{R}$  are some constants. Then  $V^n$  ( $n \in \mathbb{N}$ ) has a representation

$$V^n = \sum_{i=1}^p \lambda_i^n V_i + W^n$$

with the following properties

1.  $\lambda_1, \dots, \lambda_p \in S^1$  are eigenvalues of  $V$ , and  $V_i : F \rightarrow F(\lambda_i)$  projections onto the eigenspaces  $F(\lambda_i)$  of  $\lambda_i$ , which is finite-dimensional.
2.  $V_i \circ V_j = 0$  ( $1 \leq i \neq j \leq p$ ),  $V \circ V_i = V_i \circ V = \lambda_i V_i$  and  $V_i \circ W = W \circ V_i = 0$  ( $i = 1, \dots, p$ ).
3.  $\|W^n\|_F = O(q^n)$  for some  $q < 1$  and every  $n \in \mathbb{N}$ .

By Lemma 4 the Frobenius-Perron operator for  $\phi \in C(X)$ ,

$$[\mathcal{L}_\phi f](x) = \sum_{T(y)=x} f(y) \exp[-\phi(y)],$$

is well defined on  $C(\Omega)$ . The iterates are given by

$$\mathcal{L}_\phi^n f(x) = \sum_{T^n(y)=x} f(y) \exp[-S_n \phi(y)].$$

The following lemma is not difficult to verify.

**Lemma 7.** 1.  $\mathcal{F}_\phi$  is positive.

2. If  $\sup_{n \in \mathbb{N}} \|\mathcal{L}_\phi^n 1\|_\infty < \infty$ , then  $\mathcal{L}_\phi$  is power bounded.
3. Let  $\mu$  be a non-singular measure and  $m$  absolutely continuous with respect to  $\mu$  and  $T^n$ -invariant. Then  $m' = \frac{1}{n} \sum_{k=0}^{n-1} m \circ T^{-k}$  is  $T$ -invariant and absolutely continuous with respect to  $\mu$ .
4. For  $f, h \in C(\Omega)$  we have  $\mathcal{L}_\phi(f \circ T \cdot h) = f \cdot \mathcal{L}_\phi(h)$ .

Let  $\text{Lip}(s)$  denote the space of all Hölder-continuous functions on  $X$  with exponent  $s$ , equipped with the norm

$$\|f\| = \|f\|_{\text{Lip}(s)} := D_f + \|f\|_\infty; \quad D_f := \sup_{x \neq y \in X} \frac{|f(x) - f(y)|}{d(x, y)^s}.$$

**Lemma 8.** Let  $f, \phi \in \text{Lip}(s)$ . Then for every  $n \geq 1$

$$|\mathcal{L}_\phi^n f(x) - \mathcal{L}_\phi^n f(y)| \leq \left\{ D_f \Lambda^{-ns} + \|f\|_\infty D_\phi \frac{\Lambda^s}{\Lambda^s - 1} \right\} \|\mathcal{L}_\phi^n 1\|_\infty d(x, y)^s.$$

(Here  $\Lambda$  is as in Lemma 2.)

*Proof.* The preimages of  $T^n$  are in one-to-one correspondence on balls of radius  $a/2\Lambda$ : Denote by  $x_1, \dots, x_d$  all preimages of  $x$  and let  $y \in B(x, \frac{a}{2\Lambda})$ . Then there is a unique preimage  $y_i$  of  $y$  with  $d(x_i, y_i) \leq \Lambda^{-n} d(x, y)$ . It follows immediately that

$$\begin{aligned} |\mathcal{L}_\phi^n f(x) - \mathcal{L}_\phi^n f(y)| &\leq \sum_{i=1}^d |f(x_i) - f(y_i)| \exp[-S_n \phi(x_i)] \\ &\quad + |f(y_i)| \exp[-S_n \phi(y_i)] (1 - \exp[S_n \phi(y_i) - S_n \phi(x_i)]) \quad (2.4) \\ &\leq \left\{ D_f \Lambda^{-ns} + \|f\|_\infty D_\phi \frac{\Lambda^s}{\Lambda^s - 1} \right\} \|\mathcal{L}_\phi^n 1\|_\infty d(x, y)^s. \end{aligned}$$

This lemma implies that the Frobenius-Perron operator acts on  $\text{Lip}(s)$  provided  $\phi$  belongs to this space. It is known that every norm-bounded set in  $\text{Lip}(s)$  is relatively compact in  $C(X)$ . Therefore we can apply Theorem 15 to every power  $\mathcal{L}_\phi^n$ , if the DF-inequality (2.3) and  $\sup_n \|\mathcal{L}_\phi^n 1\|_\infty < \infty$  hold.

**Theorem 16.** *Let  $(X, T)$  be a  $R$ -expanding dynamical system. Then there is a probability  $\mu$  with  $\mathcal{L}_\phi^* \mu = \lambda \mu$ . If  $\mu$  is positive on open sets, there is an invariant measure  $m$ , absolutely continuous with respect to  $\mu$  and the Radon-Nikodym derivative  $\frac{dm}{d\mu}$  belongs to  $\text{Lip}(s)$ .*

*Proof.* The map  $\mu \mapsto [\mathcal{L}_\phi^* \mu(\Omega)]^{-1} \mathcal{L}_\phi^* \mu$  is continuous self-map of the weakly compact and convex space of all normed measures in  $C(X)^*$ . Therefore it has a fixed point  $\mu$  (Theorem of Schauder and Tychonov). Passing to  $\phi + \log \lambda$  we may assume that  $\lambda = 1$ . Let  $U_i$  be a partition of  $X$  into sets of positive measure and of diameter  $< a/\Lambda$ . The inequality (2.4), applied to  $f = 1$ , provides a constant  $K$  such that

$$\max_i \sup_{x, y \in U_i} \frac{\mathcal{L}_\phi^n f(x)}{\mathcal{L}_\phi^n f(y)} \leq K,$$

whence

$$1 = \int d\mathcal{L}_\phi^{n*} \mu \geq K^{-1} \sum_i \mu(U_i) \mathcal{L}_\phi^n 1(y_i)$$

for arbitrary  $y_i \in U_i$ . Therefore  $\sup_n \|\mathcal{L}_\phi^n 1\|_\infty < \infty$ . Because of Lemma 8 Theorem 15 is applicable for sufficiently large  $n$ , and there is an eigenfunction  $h \in \text{Lip}(s)$  for the eigenvalue 1 of  $\mathcal{L}_\phi^n$ . For any continuous function  $f$  it thus follows that

$$\begin{aligned} \int f \circ T^n \cdot h \, d\mu &= \int f \circ T^n \cdot h \, d\mathcal{L}_\phi^{n*} \mu = \int \mathcal{L}_\phi^n (f \circ T^n \cdot h) \, d\mu \\ &= \int f \cdot \mathcal{L}_\phi^n (h) \, d\mu = \int f \cdot h \, d\mu, \end{aligned}$$

i.e.  $h d\mu$  is  $T^n$ -invariant, and the theorem follows using Lemma 7.

**Theorem 17.** *Let  $(X, T)$  be an exact  $R$ -expanding dynamical system and  $\phi \in \text{Lip}(s)$ . Then for every function  $f \in \text{Lip}(s)$  the central limit theorem holds with respect to the invariant Gibbs measure  $\mu$  for  $\phi$ .*

*Proof.* The Frobenius-Perron operator for the Gibbs measure  $\mu$  has a representation

$$\mathcal{L}_\phi^n f = \int f \, dm + W^n f.$$

Let  $f \in \text{Lip}(s)$  with  $\int f \, dm = 0$ . The conditional expectation  $E(f|T^{-k}\mathcal{F})$  is given by  $U_T^k \mathcal{L}_\phi^k f$ , since for  $g \in L_\infty(\mu)$  one has

$$\int g \circ T^k U_T^k \mathcal{L}_\phi^k f dP = \int g \mathcal{L}_\phi^k f dP = \int \mathcal{L}_\phi^k [f \cdot g \circ T^k] dP = \int f \cdot g \circ T^k dP.$$

Since the conditional expectation  $E(f|T^{-k}\mathcal{F})$  is as well the orthogonal projection onto the subspaces  $L_2(T^{-k}\mathcal{F})$ , the central limit theorem follows for from corollary 4 and the estimate

$$\mathcal{L}_\phi^k f(T^k(x)) = W^k f(T^k(x)) = O(q^k).$$

## 2.5 Local Limit Theorem

The first local limit theorem has been proven in the 18th century.

**Theorem 18.** (*De Moivre-Laplace 1733*) Let  $X_1, X_2, \dots$  be i.i.d. Bernoulli random variables with parameter  $0 < p < 1$ . Then  $S_n = X_1 + \dots + X_n$  satisfies, as  $\frac{k(n) - E(S_n)}{\sqrt{V(S_n)}} \rightarrow x$ ,

$$\sqrt{V(S_n)}P(S_n = k(n)) \sim \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-x^2}{2}\right].$$

In particular,

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \longrightarrow \mathcal{N}(0, 1) \quad \text{weakly.}$$

Local limit theorems may have two different forms: due to Gnedenko (Kolmogorov) and Stone-Shepp.

**Theorem 19.** (*Gnedenko 1954*) Let  $S_n$  be the sum of i.i.d. random variables with expectation  $\mu$  and variance  $\sigma^2 > 0$  which have a lattice distribution (i.e. take values in  $b + h\mathbb{Z}$  for some  $b \in \mathbb{R}$  and a maximal  $h > 0$ ). Then

$$\sup_N \left| \frac{\sigma\sqrt{n}}{h} P(S_n = nb + Nh) - \frac{1}{\sqrt{2\pi}} e^{-\frac{(nb + Nh - n\mu)^2}{2n\sigma^2}} \right| \rightarrow 0.$$

**Theorem 20.** (*Gnedenko 1954*) Let  $S_n$  be the sum of  $n$  i.i.d. random variables with zero mean and variance 1. Assume that  $S_n/\sqrt{n}$  has density  $f_n$  and let  $\phi$  denote the density of the standard normal distribution. In order that

$$\sup_x |f_n(x) - \phi(x)| \rightarrow 0$$

it is necessary and sufficient that for some  $n$   $f_n$  is bounded.

**Theorem 21.** (*Stone 1965*) Let  $S_n$  be the sum of  $n$  i.i.d. non-lattice random variables with zero mean and variance  $\sigma^2 > 0$ . Then

$$\sigma\sqrt{n}P(S_n \in [x, x+l]) = l\phi(x/\sigma\sqrt{n}) + o(1)$$

uniformly in  $x \in \mathbb{R}$  and  $l$  in a compact set.

In this lecture we use Stone's form of a local limit theorem. Therefore

**Definition 13.** A  $\mathbb{R}$ -valued stationary stochastic sequence  $X_1, X_2, \dots$  defined on a probability space  $(\Omega, \mathcal{F}, P)$ , is said to satisfy a local limit theorem (of the partial sums  $S_n := X_1 + \dots + X_n$ ), if there exist constants  $A_n, B_n \in \mathbb{R}$ ,  $B_n \rightarrow +\infty$  such that  $\forall \kappa \in \mathbb{R}$  and  $I \subset \mathbb{R}$  (an interval),

$$B_n P(S_n - k_n \in I) \rightarrow |I|g(\kappa) \tag{2.5}$$

as  $\frac{k_n - A_n}{B_n} \rightarrow \kappa$ , where  $g$  is a continuous probability density on  $\mathbb{R}$ .

Local limits are connected to *distributional limits* where

$$\frac{S_n - A_n}{B_n} \rightarrow Y \tag{2.6}$$

in distribution for some limit random variable  $Y$ . This is essentially the idea of De Moivre–Laplace for the proof of the classical central limit theorem.

## 2.6 Characteristic Function Operators

Let  $(X, \mathcal{F}, m, T)$  be an exact, probability preserving  $R$ -expanding map, where  $m$  is an invariant Gibbs measure for the potential  $\phi$ .

For  $\omega : X \rightarrow S^1$  measurable, define

$$\mathcal{L}_\omega f := \mathcal{L}(\omega f) \quad (f \in L^1(m))$$

where  $\mathcal{L}$  denotes the transfer operator (Frobenius-Perron operator), and for  $\phi : X \rightarrow \mathbb{R}$  measurable,  $t \in \mathbb{R}$  set  $\mathcal{L}_t := \mathcal{L}_{\chi_t(\phi)}$  where  $\chi_t(y) := e^{ity}$ .

**Exercise 7.** In the independent case where  $\phi$  is  $\alpha$ -measurable and  $\alpha, T^{-1}\alpha, \dots$  are independent partitions,

$$\mathcal{L}_t 1 = E(e^{it\phi})$$

which is why the  $\mathcal{L}_t$  are sometimes called *characteristic function operators*.

**Proposition 6.** (*D-F inequality*) Suppose that  $\omega : X \rightarrow S^1$  is Hölder continuous with exponent  $s$ , then for  $f \in \text{Lip}(s)$ ,

$$\|\mathcal{L}_\omega^n f\| \leq \left[ D_f \Lambda^{-ns} + \|f\|_\infty (D_\phi + D_\omega) \frac{\Lambda^s}{\Lambda^s - 1} \right] \|\mathcal{L}_\phi^n 1\|_\infty d(x, y)^s,$$

where  $D_f, D_\phi$  and  $D_\omega$  are given by the Hölder norms.

*Proof.* The proof is omitted and we refer to the original article J. Aaronson, M. Denker: Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps. *Stochastics and Dynamics* **1**, 2001. 193–237. The proof follows standard arguments to derive the Doeblin-Fortet inequalities.

**Theorem 22.** (*Continuity*) Suppose that  $f : X \rightarrow \mathbb{R}^d$  is Hölder continuous with exponent  $s$ , then, for some constant  $M$ ,

$$\begin{aligned} & \|\mathcal{L}_s - \mathcal{L}_t\|_{\text{Hom}(L, L)} \\ & \leq M ((2 + 2M + |t|D_f)E|1 - \chi_{t-s}(f)| + (3 + 2M + |t|D_f)|s - t|D_f), \end{aligned}$$

where  $L = \text{Lip}(s)$ .

*Proof.* Let  $\chi_t(x) = \exp[itx]$ . By Proposition 6 and the theorem of Ionescu-Marinescu, for  $g \in L$  and  $t \in \mathbb{R}^d$ , we have that

$$\mathcal{L}_t g = \mathcal{L}(e^{itf} g) = \sum_{i=1}^d \chi_t(f(x_i)) g(x_i) \exp[\phi(x_i)],$$

whence

$$(\mathcal{L}_t - \mathcal{L}_s)g = \sum_{i=1}^d \chi_t(f(x_i)) g(x_i) (1 - \chi_{s-t}(f(x_i))) \exp[\phi(x_i)].$$

We shall use that for  $x, y \in X$ ,

$$|\chi_t(f(x)) - \chi_t(f(y))| \leq |t| D_f d(x, y)^s$$

whence, integrating over a set  $A$  of diameter  $< 1$  containing  $x$ ,

$$|\chi_t(f(x)) - \frac{1}{m(A)} \int_A \chi_t(f) dm| \leq |t| D_f$$

and

$$|1 - \chi_t(f(x))| \leq \frac{1}{m(A)} \int_A |1 - \chi_t(f)| dm + |t| D_f. \quad (2.7)$$

Also  $\forall x, y \in A$ ,

$$\begin{aligned} & |(\chi_t(f(x)) - \chi_s(f(x))) - (\chi_t(f(y)) - \chi_s(f(y)))| \\ & \leq |\chi_{s-t}(f(y)) - \chi_{s-t}(f(x))| + |1 - \chi_{s-t}(f(y))| |\chi_t(f(x)) - \chi_t(f(y))| \\ & \leq d(x, y)^s \left( |s - t| D_f (1 + |t| D_f) + |t| D_f \frac{1}{m(A)} \int_A |1 - \chi_{s-t}(f)| dm \right). \end{aligned} \quad (2.8)$$

Choose a partition  $A_i$  of  $X$  such that  $A_i$  contains the preimage  $x_i$  of  $x$ , is contained in  $K(x, a)$  and  $\min_{1 \leq i \leq d} m(A_i) > 0$ . It follows now that,  $\forall x, y \in A$ , and some constant  $M$

$$\begin{aligned} & |(\mathcal{L}_s - \mathcal{L}_t)g(x)| \\ & \leq \sum_{i=1}^d |1 - \chi_{s-t}(f(x_i)) g(x_i) \exp[\phi(x_i)]| \\ & \leq M \sum_{i=1}^d \left( \int_{A_i} |1 - \chi_{s-t}(f)| dm + m(A_i) |s - t| D_f \right) \|g\|_\infty \\ & \leq M (E(|1 - \chi_{s-t}(f)|) + |s - t| D_f) \|g\|_\infty. \end{aligned}$$

For  $x, y \in X$ ,

$$\begin{aligned}
 & |(\mathcal{L}_s - \mathcal{L}_t)g(x) - (\mathcal{L}_s - \mathcal{L}_t)g(y)| = \\
 & \left| \sum_{i=1}^d (\chi_t(f(x_i)) - \chi_s(f(x_i)))g(x_i) \exp[\phi(x_i)] - \right. \\
 & \left. (\chi_t(f(y_i)) - \chi_s(f(y_i)))g(y_i) \exp[\phi(y_i)] \right| \leq \\
 & \sum_{i=1}^d (\exp[\phi(x_i)]|g(x_i)| |(\chi_t(f(x_i)) - \chi_s(f(x_i))) - (\chi_t(f(y_i)) - \chi_s(f(y_i)))| \\
 & + |1 - \chi_{s-t}(f(y_i))| |\exp[\phi(x_i)]g(x_i) - \exp[\phi(y_i)]g(y_i)|) \\
 & = \sum_{i=1}^d (I_a + II_a).
 \end{aligned}$$

Now

$$\begin{aligned}
 II_a & = |\exp[\phi(x_i)]g(x_i) - \exp[\phi(y_i)]g(y_i)| |1 - \chi_{s-t}(f(x_i))| \\
 & \leq Md(x, y)^s ((M+1)rD_g + M\|g\|_\infty) \left( \int_A |1 - \chi_{s-t}(f)| dm + m(A)|s-t|D_f \right)
 \end{aligned}$$

whence

$$\sum_{i=1}^d II_a \leq Md(x, y)^s ((M+1)rD_g + M\|g\|_\infty) (E|1 - \chi_{s-t}(f)| + |s-t|D_f).$$

Moreover,

$$\begin{aligned}
 I_a & = |\exp[\phi(y_i)]g(y_i)| |(\chi_t(f(x_i)) - \chi_s(f(x_i))) - (\chi_t(f(y_i)) - \chi_s(f(y_i)))| \\
 & \leq M\|g\|_\infty d(x, y)^s \left( m(A)|s-t|D_f(1 + |t|D_f) + |t|D_f \int_A |1 - \chi_{s-t}(f)| dm \right)
 \end{aligned}$$

and

$$\sum_{i=1}^d I_a \leq M\|g\|_\infty d(x, y)^s (|s-t|D_f(1 + |t|D_f) + |t|D_f E(|1 - \chi_{s-t}(f)|)).$$

The conclusion is that

$$\begin{aligned}
 & \|(\mathcal{L}_s - \mathcal{L}_t)g\|_L \\
 & \leq M (E(|1 - \chi_{s-t}(f)|) + |s-t|D_f) \|g\|_\infty \\
 & + M((M+1)rD_g + M\|g\|_\infty) (E|1 - \chi_{s-t}(f)| + |s-t|D_f) \\
 & + M\|g\|_\infty (|s-t|D_f(1 + |t|D_f) + |t|D_f E(|1 - \chi_{s-t}(f)|)) \\
 & \leq M\|g\|_L ((2 + 2M + |t|D_f)E|1 - \chi_{s-t}(f)| + (3 + 2M + |t|D_f)|s-t|D_f).
 \end{aligned}$$

**Exercise 8.** Show that  $t \mapsto \mathcal{L}_t$  is  $C^2$  ( $\mathbb{T} \rightarrow \text{Hom}(L, L)$ ) with

$$\frac{d\mathcal{L}_t}{dt}g = \mathcal{L}(if'e^{itfg})$$

and

$$\frac{d^2\mathcal{L}_t}{dt^2}g = -\mathcal{L}(f'')e^{itfg}.$$

## 2.7 Spectral Theorem

Let  $(X, \mathcal{F}, m, T)$  be an exact R-expanding dynamical system, where  $m$  is an invariant Gibbs measure. We say that a function  $f : X \rightarrow \mathbb{R}$  is aperiodic, if there is no character  $\gamma \in \widehat{\mathbb{R}(\mathbb{Z})}$  so that  $\gamma \circ f$  is  $T$ -cohomologous to a constant, i.e. the equation

$$\gamma \circ f = \frac{\lambda g(x)}{g(T(x))}$$

has no solution  $\lambda \in S^1 := \{z \in \mathbb{C} : |z| = 1\}$ ,  $g : X \rightarrow S^1$  measurable, other than  $\lambda = 1$ , and  $g \equiv 1$ . We say that  $\phi$  is periodic if it is not aperiodic.

**Theorem 23.** 1) *There are constants  $\epsilon > 0$ ,  $K > 0$  and  $\theta \in (0, 1)$ ; and functions  $\lambda : B(0, \epsilon) \rightarrow B_{\mathbb{C}}(0, 1)$ ,  $N : B(0, \epsilon) \rightarrow \text{Hom}(L, L)$  such that*

$$\|\mathcal{L}_t^n h - \lambda(t)^n N(t)h\|_L \leq K\theta^n \|h\|_L \quad \forall |t| < \epsilon, n \geq 1, h \in L$$

where  $\forall |t| < \epsilon$ ,  $N(t)$  is the projection onto the one-dimensional subspace spanned by  $g(t) := N(t)1$ ; and  $g(t)$  satisfies

$$\|g(t) - 1\|_L \leq K(|t| + E(|e^{it\phi} - 1|)).$$

2) *If  $f$  is aperiodic, then  $\forall \widetilde{M} > 0$ ,  $\epsilon > 0$ ,  $\exists K' > 0$  and  $\theta' \in (0, 1)$  such that*

$$\|\mathcal{L}_t^n h\|_L \leq K'\theta'^n \|h\|_L \quad \forall \epsilon \leq |t| \leq \widetilde{M}, h \in L.$$

By Theorem 22,  $t \mapsto \mathcal{L}_t$  is continuous  $\mathbb{R}^d \rightarrow \text{Hom}(L, L)$ , and by Proposition 6  $\mathcal{L}_t$  is a D-F operator  $\forall t \in \mathbb{R}^d$ . The proof of the theorem is established by two lemmas about D-F operators.

The next two lemmas are well known. Similar statements can be found in papers by Nagaev, Parry-Pollicott and Roussaux-Egele. We write  $\|\mathcal{L}\| := \|\mathcal{L}\|_{\text{Hom}(L, L)}$  for  $\mathcal{L} \in \text{Hom}(L, L)$ .

**Lemma 9.** *Suppose that  $\mathcal{L}_0 \in \text{Hom}(L, L)$  satisfies  $\mathcal{L}_0 = \mu_0 + Q_0$  where  $\mu_0^2 = \mu_0$ ,  $\dim \mu_0 L = 1$ ,  $\mu_0 Q_0 = Q_0 \mu_0 = 0$  and such that the spectral radius of  $Q_0$ ,  $r(Q_0) < 1$ , then  $\exists \epsilon > 0$ ,  $\lambda : B(\mathcal{L}_0, \epsilon) \rightarrow \mathbb{C}$ ,  $N_1, Q : B(P_0, \epsilon) \rightarrow \text{Hom}(L, L)$  holomorphic, such that*

$$\mathcal{L}^n = \lambda(\mathcal{L})^n N_1(\mathcal{L}) + Q(\mathcal{L})^n \quad (n \geq 1)$$

and where  $N_1(\mathcal{L})$  is a projection onto a 1-dimensional subspace. Moreover,  $|\lambda(\mathcal{L})| \leq 1$  and  $\exists K \in \mathbb{R}_+$ ,  $\theta \in (0, 1)$  such that  $\|Q(\mathcal{L})^n\| \leq K\theta^n \quad \forall n \geq 1$ ,  $\mathcal{L} \in B(\mathcal{L}_0, \epsilon)$ .

The proof of Lemma 9 is standard using Dunford-Schwartz, chapter VII, section 3.6.

**Lemma 10.** *Suppose that  $\mathcal{K} \subset \text{Hom}(L, L)$  is a compact set of D-F operators, none of which has an L-eigenvalue on  $S^1$  (the unit circle), then  $\exists K \in \mathbb{R}_+$  and  $\theta \in (0, 1)$  such that*

$$\|\mathcal{L}^n\| \leq K\theta^n \quad \forall n \geq 1, \mathcal{L} \in \mathcal{K}.$$

*Proof.* We first show that  $\max_{\mathcal{L} \in \mathcal{K}} \mathbf{r}(\mathcal{L}) < 1$ .

For  $\mathcal{L} \in \mathcal{K}$  and  $z \in \rho(\mathcal{L})$

$$R_{\mathcal{L}}(z) = (zI - \mathcal{L})^{-1}.$$

For  $b > \mathbf{r}(\mathcal{L})$ ,

$$M(\mathcal{L}, b) := \sup_{|z| \geq b} \|R_{\mathcal{L}}(z)\| < \infty.$$

If  $\mathcal{L}' \in \text{Hom}(L, L)$  and  $\|\mathcal{L} - \mathcal{L}'\| < M(\mathcal{L}, b)^{-1}$  then  $\forall |z| > b$

$$\sum_{n=1}^{\infty} \|((\mathcal{L}' - \mathcal{L})R_{\mathcal{L}}(z))^n\| < \infty,$$

whence

$$R_{\mathcal{L}}(z) \sum_{n=0}^N ((\mathcal{L}' - \mathcal{L})R_{\mathcal{L}}(z))^n \rightarrow (zI - \mathcal{L}')^{-1}$$

in  $\text{Hom}(L, L)$  as  $N \rightarrow \infty$  and  $B(0, b)^c \subset \rho(\mathcal{L}')$  which implies  $\mathbf{r}(\mathcal{L}') \leq b$ .

For each  $\mathcal{L} \in \mathcal{K}$ , choose  $r_{\mathcal{L}} \in (\mathbf{r}(\mathcal{L}), 1)$ . As above, for each  $\mathcal{L} \in \mathcal{K}$ ,  $\exists \epsilon_{\mathcal{L}} = M(\mathcal{L}, r_{\mathcal{L}})^{-1}$  such that

$$\mathbf{r}(Q) \leq r_{\mathcal{L}} \quad \forall Q \in B(\mathcal{L}, \epsilon_{\mathcal{L}}).$$

By compactness of  $\mathcal{K}$ ,  $\exists \mathcal{L}_1, \dots, \mathcal{L}_N \in \mathcal{K}$  such that

$$\mathcal{K} \subset \bigcup_{k=1}^N B(\mathcal{L}_k, \epsilon_{\mathcal{L}_k})$$

with the consequence that

$$\mathbf{r}(\mathcal{L}) \leq r_0 := \max_{1 \leq k \leq N} r_{\mathcal{L}_k} < 1 \quad \forall \mathcal{L} \in \mathcal{K}.$$

To complete the proof choose

$$\max_{\mathcal{L} \in \mathcal{K}} \mathbf{r}(\mathcal{L}) < b < 1.$$

We have that  $(z, \mathcal{L}) \mapsto (zI - \mathcal{L})^{-1}$  is continuous  $\{|z| = b\} \times \mathcal{K} \rightarrow \text{Hom}(L, L)$ .  
Therefore

$$\sup_{|z|=b, \mathcal{L} \in \mathcal{K}} \|(zI - \mathcal{L})^{-1}\| =: K < \infty.$$

Now, for  $n \geq 1$ ,

$$\mathcal{L}^n = \frac{1}{2\pi i} \oint_{|z|=b} (zI - \mathcal{L})^{-1} z^n dz = \frac{1}{2\pi} \int_0^{2\pi} (be^{it}I - \mathcal{L})^{-1} b^{n+1} e^{i(n+1)t} dt$$

whence

$$\|\mathcal{L}^n\| \leq \frac{1}{2\pi} \int_0^{2\pi} \|(be^{it}I - \mathcal{L})^{-1}\| b^{n+1} dt \leq K\theta^{n+1}.$$

*Proof.* (of theorem 4.1.) The first statement follows from Lemma 9, and Theorem 22.

The aperiodicity of  $\phi$  implies that for  $t \neq 0$ ,  $\mathcal{L}_t$  has no eigenvalue on  $S^1$ .

By Theorem 22,  $\{\mathcal{L}_t : \epsilon \leq |t| \leq \widetilde{M}\}$  is compact in  $\text{Hom}(L, L)$ , and so by Lemma 10,  $\exists \theta \in (0, 1)$  and  $K > 0$  so that  $\|\mathcal{L}_t^n\|_{\text{Hom}(L, L)} \leq K\theta^n$  for  $n \geq 1$ ,  $\epsilon \leq |t| \leq \widetilde{M}$ .

*Remark 2.* It follows from Lemma 9 that if  $f \in \text{Lip}_2$ , then  $t \mapsto \lambda(t) = \lambda\mathcal{L}(t)$  is  $C^2$ . It can be shown that  $\lambda'(0) = iE(f)$  and  $\lambda''(0) = -\lim_{n \rightarrow \infty} \frac{E(f_n^2)}{n}$ . Thus,  $\lambda(t) = 1 + iE(tf) + \frac{t^2}{2} + o(|t|^2)$  as  $t \rightarrow 0$ .

**Theorem 24.** (*Distributional limit theorem*) Under the above conditions,

$$\frac{f_n - n \int f dm}{\sqrt{n}}$$

converges in distribution to a normal distribution  $\mathcal{N}(0, \sigma^2)$  for some  $\sigma^2 \geq 0$ .

*Proof.* Set  $A_n = n \int f dm$  and  $B_n^2 = n\sigma^2$ . We first note that

$$n \left[ \log \lambda\left(\frac{t}{B_n}\right) \right] - it \frac{A_n}{B_n} \rightarrow \log \widehat{N}(t) \text{ as } n \rightarrow \infty.$$

Using Theorem 23 (1),

$$\begin{aligned} \int_X e^{it\left(\frac{f_n}{B_n} - \frac{A_n}{B_n}\right)} dm &= e^{-it \frac{A_n}{B_n}} \int_X \mathcal{L}^n(e^{it \frac{f_n}{B_n}}) dm \\ &= e^{-it \frac{A_n}{B_n}} \int_X \mathcal{L}_{\frac{t}{B_n}}^n 1 dm \\ &= e^{-it \frac{A_n}{B_n}} \lambda\left(\frac{t}{B_n}\right)^n \int_X g\left(\frac{t}{B_n}\right) dm + O(\theta^n). \end{aligned}$$

The theorem follows since  $\widehat{N}(s) \rightarrow \widehat{N}(0) \equiv 1$  as  $s \rightarrow 0$ .

## 2.8 Local Limit Theorems for R-Expanding Systems

**Theorem 25.** *(Conditional lattice local limit theorem) Suppose that  $f : X \rightarrow \mathbb{Z}$  is aperiodic, let  $A_n, B_n$  be as in the proof of Theorem 24, and suppose that  $k_n \in \mathbb{Z}$ ,  $\frac{k_n - A_n}{B_n} \rightarrow \kappa \in \mathbb{R}$  as  $n \rightarrow \infty$ , then*

$$\|B_n \mathcal{L}_\phi^n(1_{[f_n = k_n]}) - \varphi_N(\kappa)\|_\infty \rightarrow 0 \text{ as } n \rightarrow \infty,$$

and, in particular

$$B_n m([f_n = k_n]) \rightarrow \varphi_N(\kappa) \text{ as } n \rightarrow \infty$$

where  $\varphi_N$  denotes the density of the standard normal distribution.

*Proof.* By Theorem 23,  $\exists \delta > 0$ ,  $\theta \in (0, 1)$  such that  $\forall |t| \leq \delta$ ,

$$\|\mathcal{L}_t^n 1 - \lambda(t)^n g(t)\|_L = O(\theta^n) \quad \forall |t| \leq \delta,$$

and that

$$\|\mathcal{L}_y^n 1\|_L = O(\theta^n) \quad \forall \delta \leq |y| \leq \pi.$$

We also may assume that

$$-\operatorname{Re} \log \lambda(t) \geq \frac{|t|^2}{2} \quad \forall |t| \leq \delta.$$

It follows that, uniformly on  $X$ ,

$$\begin{aligned} 2\pi B_n \mathcal{L}_\phi^n(1_{[f_n = k_n]}) &= B_n \mathcal{L}_\phi^n \left( \int_{-\pi}^{\pi} e^{-itk_n} e^{itf_n} dt \right) \\ &= B_n \int_{-\pi}^{\pi} e^{-itk_n} \mathcal{L}_\phi^n(e^{itf_n}) dt \\ &= B_n \int_{-\pi}^{\pi} e^{-itk_n} \mathcal{L}_t^n 1 dt \\ &= B_n \int_{|t| \leq \delta} e^{-itk_n} \lambda(t)^n g(t) dt + O(B_n \theta^n) \\ &= \int_{-\delta B_n}^{\delta B_n} e^{-it \frac{A_n}{B_n}} \lambda\left(\frac{t}{B_n}\right)^n g\left(\frac{t}{B_n}\right) e^{it \left(\frac{A_n - k_n}{B_n}\right)} dt + o(1) \\ &= \int_{-\delta B_n}^{\delta B_n} e^{-it \frac{A_n}{B_n}} \lambda\left(\frac{t}{B_n}\right)^n e^{it \left(\frac{A_n - k_n}{B_n}\right)} dt + o(1) \\ &\rightarrow \int_{\mathbb{R}} \widehat{N}(t) e^{-i\kappa t} dt \\ &= 2\pi \varphi_N(\kappa) \end{aligned}$$

as  $n \rightarrow \infty$  by dominated convergence, since for  $|t| \leq \delta B_n$ ,

$$\left| \lambda\left(\frac{t}{B_n}\right)^n \right| \leq e^{-\operatorname{frac}|t|^{2+\epsilon} 2},$$

which latter function is integrable on  $\mathbb{R}$ .

**Theorem 26.** (Conditional non-lattice local limit theorem) Suppose that  $f : X \rightarrow \mathbb{R}$  is aperiodic, let  $A_n, B_n$  be as in Theorem 24, let  $I \subset \mathbb{R}$  be an interval, and suppose that  $k_n \in \mathbb{Z}$ ,  $\frac{k_n - A_n}{B_n} \rightarrow \kappa \in \mathbb{R}$  as  $n \rightarrow \infty$ , then

$$B_n \mathcal{L}_\phi^n(1_{[\phi_n \in k_n + I]}) \rightarrow |I| \varphi_N(\kappa) \text{ as } n \rightarrow \infty$$

where  $|I|$  is the length of  $I$ , and in particular

$$B_n m([f_n \in k_n + I]) \rightarrow |I| \varphi_N(\kappa) \text{ as } n \rightarrow \infty.$$

*Proof.* We use the method of Breiman (see Breiman: Probability, 1968).

Suppose that  $h \in L^1(\mathbb{R})$ ,  $\hat{h} \in L^1(\mathbb{R})$ , and that  $\hat{h} \equiv 0$  off  $[-M, M]$ .

Arguing as in the proof of Theorem 25, we obtain  $\delta > 0$ , and  $0 < \theta < 1$  such that, uniformly on  $X$ :

$$\begin{aligned} B_n \mathcal{L}_\phi^n(h(f_n - k_n)) &= \frac{B_n}{2\pi} \int_{-M}^M \hat{h}(x) \mathcal{L}_\phi^n(e^{ix(f_n - k_n)}) dx \\ &= \frac{B_n}{2\pi} \int_{-M}^M \hat{h}(x) e^{-ik_n x} \mathcal{L}_x^n 1 dx \\ &= \frac{B_n}{2\pi} \int_{|x| \leq \delta} \hat{h}(x) e^{-ik_n x} \lambda(x)^n g(x) dx + O(B_n \theta^n) \\ &= \frac{1}{2\pi} \int_{|x| \leq \delta B_n} \hat{h}\left(\frac{x}{B_n}\right) e^{-i\frac{k_n}{B_n} x} \lambda\left(\frac{x}{B_n}\right)^n g\left(\frac{x}{B_n}\right) dx + o(1) \\ &= \frac{1}{2\pi} \int_{|x| \leq \delta B_n} \hat{h}\left(\frac{x}{B_n}\right) e^{-i\frac{k_n}{B_n} x} \lambda\left(\frac{x}{B_n}\right)^n dx + o(1) \\ &\rightarrow \frac{1}{2\pi} \int_{\mathbb{R}} \hat{h}(0) \hat{N}(x) e^{-i\kappa x} dx \\ &= \int_{\mathbb{R}} h(x) dx \varphi_N(\kappa) \end{aligned}$$

by dominated convergence as again, for some  $\epsilon > 0$ ,  $\forall |t| \leq \delta B_n$ ,  $|\lambda(\frac{t}{B_n})^n| \leq e^{-\frac{1}{2}|t|^{p+\epsilon}}$ , which latter function is integrable on  $\mathbb{R}$ . Let  $k(x) = \frac{\sin^2 x}{x^2}$ , then  $k > 0$ ,  $k \in L^1(m_{\mathbb{R}})$  and  $\hat{k}$  has compact support.

It follows from Theorem 10.7 in Breiman that if  $U$  is a vague neighbourhood of  $m_{\mathbb{R}}$ , then  $\exists \eta > 0$  and  $t_1, \dots, t_N \in \mathbb{R}$  such that for  $\mu$  a Radon measure on  $\mathbb{R}$ :

$$\left| \int_{\mathbb{R}} e^{it_j x} k(x) d\mu - \int_{\mathbb{R}} e^{it_j x} k(x) dx \right| < \eta \quad (1 \leq j \leq N) \implies \mu \in U.$$

Thus, for  $h : \mathbb{R} \rightarrow \mathbb{R}$  continuous with compact support,

$$B_n \mathcal{L}_\phi^n(h(f_n - k_n)) \rightarrow \int_{\mathbb{R}} h(x) dx \varphi_N(\kappa)$$

uniformly on  $X$ . The theorem follows from monotone approximation of  $1_I$  by non-negative continuous functions with compact support.

# Index

- $C(X)$  continuous functions on  $X$ , 7
- $\beta$ -transformation, 27
  
- aperiodic function, 35
- array of random elements, 20
- asymptotic negligible, 20
  
- Bowen-McClusky-Formula, 13
- Bowen-Metric, 7
- Bowen-metric, 5
  
- central limit theorem, 19, 21, 23, 26, 30, 38
- characteristic function operator, 32
- cohomologous, 13
- conditional expectation, 21
- continuity theorem of Lévy, 24
  
- deMoivre-Laplace, 31
- distributional convergence, 32
- Doebelin-Fortet inequality, 28
  
- entropy, topological, 8
- equilibrium, 12
- exact transformation, 25
- expanding, 5
- expansion constant, 6
- expansive, 4
- expansive constant, 4
  
- Frobenius-Perron operator, 7, 29
  
- Gibbs measure, 14
- Gnedenko, 31
- Gordin's Theorem, 26
  
- histogram, 27
  
- Jacobi density, 14
  
- Koopman operator, 25
  
- Lindeberg condition, 20
- local limit theorem, 31, 38
- local limit theorem, definition, 31
  
- martingale, 22
- martingale approximation, 25
- martingale difference sequence, 22
- measure of maximal entropy, 12
- Metritzation Lemma of Frink, 6
  
- normal distribution, 26
  
- operator, power bounded, 28
- operator, relatively compact, 28
  
- pressure, 8
- pressure function, 8
  
- R-expanding, 5
- Radon Nikodym theorem, 21
  
- spanning set, 7
- Stone, 31
  
- Theorem of Coven and Reddy, 6
- Theorem of Doebelin, Fortet, Ionescu-Tulcea, Marinescu, 28
- theorem of Doebelin, Fortet, Ionescu-Tulcea, Marinescu, 28
- transfer-operator, 7
  
- Variational principle, 10
  
- weak isomorphism theorem, 20