MAGIC: Ergodic Theory Lecture 9 - Thermodynamic Formalism

Charles Walkden

April 12, 2013

Introduction

Introduction

Historically ergodic theory was a branch of statistical mechanics. There are many concepts in ergodic theory that reflect this heritage (entropy, pressure, Gibbs states, zero-temperature limits etc).

Introduction

Historically ergodic theory was a branch of statistical mechanics. There are many concepts in ergodic theory that reflect this heritage (entropy, pressure, Gibbs states, zero-temperature limits etc).

This is the so-called "thermodynamic formalism", which we introduce in this lecture. We will not study the connections between ergodic theory and statistical mechanics; instead we set things up in a way that allows us to study hyperbolic dynamics.

Shifts of finite type

Shifts of finite type

Let A be an aperiodic $k \times k$ $0-1$ matrix. Define the one-sided shift of finite type

$$\Sigma = \{(x_j)_{j=0}^{\infty} \mid A_{x_j,x_{j+1}} = 1 \text{ for all } j \geq 0\}.$$

Shifts of finite type

Let A be an aperiodic $k \times k$ $0-1$ matrix. Define the one-sided shift of finite type

$$\Sigma = \{(x_j)_{j=0}^{\infty} \mid A_{x_j, x_{j+1}} = 1 \text{ for all } j \geq 0\}.$$

Fix $\theta \in (0, 1)$. Define a metric on Σ by

$$d_\theta(x, y) = \theta^{n(x,y)}$$

where $n(x, y)$ is the first place in which the sequences x, y disagree.

Shifts of finite type

Let A be an aperiodic $k \times k$ $0-1$ matrix. Define the one-sided shift of finite type

$$\Sigma = \{(x_j)_{j=0}^{\infty} \mid A_{x_j, x_{j+1}} = 1 \text{ for all } j \geq 0\}.$$

Fix $\theta \in (0,1)$. Define a metric on Σ by

$$d_\theta(x,y) = \theta^{n(x,y)}$$

where $n(x,y)$ is the first place in which the sequences x, y disagree.

Define cylinder sets by

$$[i_0, \ldots, i_{n-1}] = \{(x_j)_{j=0}^{\infty} \mid x_j = i_j,\ 0 \leq j \leq n-1\}.$$

Cylinder sets are both open and closed.

Shifts of finite type

Let A be an aperiodic $k \times k$ $0-1$ matrix. Define the one-sided shift of finite type

$$\Sigma = \{(x_j)_{j=0}^{\infty} \mid A_{x_j, x_{j+1}} = 1 \text{ for all } j \geq 0\}.$$

Fix $\theta \in (0,1)$. Define a metric on Σ by

$$d_\theta(x,y) = \theta^{n(x,y)}$$

where $n(x,y)$ is the first place in which the sequences x, y disagree.

Define cylinder sets by

$$[i_0, \ldots, i_{n-1}] = \{(x_j)_{j=0}^{\infty} \mid x_j = i_j,\ 0 \leq j \leq n-1\}.$$

Cylinder sets are both open and closed.

Define the shift map

$$\sigma(x_0, x_1, x_2, \ldots) = (x_1, x_2, \ldots).$$

Functions defined on Σ

Functions defined on Σ

We need a nice class of functions to work with.

Functions defined on Σ

We need a nice class of functions to work with.

Let $f : \Sigma \to \mathbb{R}$ or $\mathbb{C}$. Then f is Hölder of exponent $\theta \in (0, 1)$ if there exists $C > 0$ s.t.

$$|f(x) - f(y)| \leq C d_\theta(x, y) \tag{1}$$

Functions defined on Σ

We need a nice class of functions to work with.

Let $f : \Sigma \to \mathbb{R}$ or $\mathbb{C}$. Then f is Hölder of exponent $\theta \in (0,1)$ if there exists $C > 0$ s.t.

$$|f(x) - f(y)| \leq C d_\theta(x,y) \tag{1}$$

(This is an abuse of notation: one should say that f is Lipschitz w.r.t. d_θ. But $d_\theta(x,y)^\alpha = d_{\theta^\alpha}(x,y)$ so α-Hölder w.r.t. d_θ is the same as Lipschitz w.r.t. d_{θ^α}.)

Functions defined on Σ

We need a nice class of functions to work with.

Let $f : \Sigma \to \mathbb{R}$ or $\mathbb{C}$. Then f is Hölder of exponent $\theta \in (0,1)$ if there exists $C > 0$ s.t.

$$|f(x) - f(y)| \leq C d_\theta(x, y) \tag{1}$$

(This is an abuse of notation: one should say that f is Lipschitz w.r.t. d_θ. But $d_\theta(x,y)^\alpha = d_{\theta^\alpha}(x,y)$ so α-Hölder w.r.t. d_θ is the same as Lipschitz w.r.t. d_{θ^α}.)

Let $|f|_\theta$ denote the least possible constant $C > 0$ in (1). Then $|f|_\theta$ is a semi-norm, but not a norm.

Functions defined on Σ

We need a nice class of functions to work with.

Let $f : \Sigma \to \mathbb{R}$ or $\mathbb{C}$. Then f is Hölder of exponent $\theta \in (0, 1)$ if there exists $C > 0$ s.t.

$$|f(x) - f(y)| \leq C d_\theta(x, y) \tag{1}$$

(This is an abuse of notation: one should say that f is Lipschitz w.r.t. d_θ. But $d_\theta(x, y)^\alpha = d_{\theta^\alpha}(x, y)$ so α-Hölder w.r.t. d_θ is the same as Lipschitz w.r.t. d_{θ^α}.)

Let $|f|_\theta$ denote the least possible constant $C > 0$ in (1). Then $|f|_\theta$ is a semi-norm, but not a norm.

Define $\|f\|_\theta = |f|_\infty + |f|_\theta$.

Functions defined on Σ

We need a nice class of functions to work with.

Let $f : \Sigma \to \mathbb{R}$ or $\mathbb{C}$. Then f is Hölder of exponent $\theta \in (0, 1)$ if there exists $C > 0$ s.t.

$$|f(x) - f(y)| \leq C d_\theta(x, y) \tag{1}$$

(This is an abuse of notation: one should say that f is Lipschitz w.r.t. d_θ. But $d_\theta(x, y)^\alpha = d_{\theta^\alpha}(x, y)$ so α-Hölder w.r.t. d_θ is the same as Lipschitz w.r.t. d_{θ^α}.)

Let $|f|_\theta$ denote the least possible constant $C > 0$ in (1). Then $|f|_\theta$ is a semi-norm, but not a norm.

Define $\|f\|_\theta = |f|_\infty + |f|_\theta$. Then $\| \cdot \|_\theta$ is a norm on the Banach space

$$F_\theta(\mathbb{R}) = \{f : \Sigma \to \mathbb{R} \mid \|f\|_\theta < \infty\}.$$

Locally constant functions

Locally constant functions

Suppose $f : \Sigma \to \mathbb{R}$ depends only on the first n co-ordinates, i.e.

$$f(x) = f(x_0, \ldots, x_{n-1}).$$

We say that f is locally constant.

Locally constant functions

Suppose $f : \Sigma \to \mathbb{R}$ depends only on the first n co-ordinates, i.e.

$$f(x) = f(x_0, \ldots, x_{n-1}).$$

We say that f is locally constant.

Equivalently, f is locally constant if it is constant on cylinders of length n (for some n).

Locally constant functions

Suppose $f : \Sigma \to \mathbb{R}$ depends only on the first n co-ordinates, i.e.

$$f(x) = f(x_0, \ldots, x_{n-1}).$$

We say that f is locally constant.

Equivalently, f is locally constant if it is constant on cylinders of length n (for some n).

Locally constant functions only take finitely many values.

Locally constant functions

Suppose $f : \Sigma \to \mathbb{R}$ depends only on the first n co-ordinates, i.e.

$$f(x) = f(x_0, \ldots, x_{n-1}).$$

We say that f is locally constant.

Equivalently, f is locally constant if it is constant on cylinders of length n (for some n).

Locally constant functions only take finitely many values.

If f is locally constant then $f \in F_\theta(\mathbb{R})$ for all $\theta \in (0, 1)$.

Locally constant functions

Suppose $f : \Sigma \to \mathbb{R}$ depends only on the first n co-ordinates, i.e.

$$f(x) = f(x_0, \ldots, x_{n-1}).$$

We say that f is locally constant.

Equivalently, f is locally constant if it is constant on cylinders of length n (for some n).

Locally constant functions only take finitely many values.

If f is locally constant then $f \in F_\theta(\mathbb{R})$ for all $\theta \in (0, 1)$.

In particular, if f is locally constant then f is continuous. (The zero-dimensionality of Σ guarantees the existence of lots of locally constant functions.)

The transfer operator

The transfer operator

Let $f \in F_\theta(\mathbb{R})$.

The transfer operator

Let $f \in F_\theta(\mathbb{R})$. Define the linear operator $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y:\sigma(y)=x} e^{f(y)} w(y) = \sum_{i \text{ s.t. } A_{i,x_0}=1} e^{f(ix)} w(ix).$$

(Here $(ix) = (i, x_0, x_1, \ldots)$.)

The transfer operator

Let $f \in F_\theta(\mathbb{R})$. Define the linear operator $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y:\sigma(y)=x} e^{f(y)} w(y) = \sum_{i \text{ s.t. } A_{i,x_0}=1} e^{f(ix)} w(ix).$$

(Here $(ix) = (i, x_0, x_1, \ldots)$.)

Idea: look at the preimages of x, evaluate w at these preimages, weight them according to the weight function f, then sum.

The transfer operator

Let $f \in F_\theta(\mathbb{R})$. Define the linear operator $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y:\sigma(y)=x} e^{f(y)} w(y) = \sum_{i \text{ s.t. } A_{i,x_0}=1} e^{f(ix)} w(ix).$$

(Here $(ix) = (i, x_0, x_1, \ldots)$.)

Idea: look at the preimages of x, evaluate w at these preimages, weight them according to the weight function f, then sum.

Example: Take Σ =full one-sided 2-shift. Take $f(x) \equiv \log 1/2$. Then

$$L_f w(x) = \frac{1}{2} w(0x) + \frac{1}{2} w(1x).$$

The transfer operator

Let $f \in F_\theta(\mathbb{R})$. Define the linear operator $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y:\sigma(y)=x} e^{f(y)} w(y) = \sum_{i \text{ s.t. } A_{i,x_0}=1} e^{f(ix)} w(ix).$$

(Here $(ix) = (i, x_0, x_1, \ldots)$.)

Idea: look at the preimages of x, evaluate w at these preimages, weight them according to the weight function f, then sum.

Example: Take $\Sigma =$full one-sided 2-shift. Take $f(x) \equiv \log 1/2$. Then

$$L_f w(x) = \frac{1}{2} w(0x) + \frac{1}{2} w(1x).$$

We are interested in the spectral properties of L_f.

An example

An example

Let Σ be the one-sided 2-shift on symbols $0, 1$. Fix $p, q \in (0, 1)$.
Let f be the weight function depending on 2 co-ordinates:

$$f(00) = \log p, \ f(01) = \log(1-p), \ f(10) = \log q, \ f(11) = \log(1-q).$$

An example

Let Σ be the one-sided 2-shift on symbols $0, 1$. Fix $p, q \in (0, 1)$. Let f be the weight function depending on 2 co-ordinates:

$$f(00) = \log p,\ f(01) = \log(1-p),\ f(10) = \log q,\ f(11) = \log(1-q).$$

Consider the action of L_f on functions that only depend on 1 co-ordinate $w(x) = w(x_0)$.

$$\begin{aligned} L_f w(0) &= e^{f(00)} w(0) + e^{f(10)} w(1) = pw(0) + (1-p)w(1) \\ L_f w(1) &= e^{f(01)} w(0) + e^{f(11)} w(1) = qw(0) + (1-q)w(1). \end{aligned}$$

An example

Let Σ be the one-sided 2-shift on symbols $0, 1$. Fix $p, q \in (0,1)$. Let f be the weight function depending on 2 co-ordinates:

$$f(00) = \log p,\ f(01) = \log(1-p),\ f(10) = \log q,\ f(11) = \log(1-q).$$

Consider the action of L_f on functions that only depend on 1 co-ordinate $w(x) = w(x_0)$.

$$\begin{aligned} L_f w(0) &= e^{f(00)} w(0) + e^{f(10)} w(1) = pw(0) + (1-p)w(1) \\ L_f w(1) &= e^{f(01)} w(0) + e^{f(11)} w(1) = qw(0) + (1-q)w(1). \end{aligned}$$

Record this is a matrix

$$\begin{pmatrix} L_f w(0) \\ L_f w(1) \end{pmatrix} = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix} \begin{pmatrix} w(0) \\ w(1) \end{pmatrix} = P \begin{pmatrix} w(0) \\ w(1) \end{pmatrix}$$

An example

Let Σ be the one-sided 2-shift on symbols $0, 1$. Fix $p, q \in (0, 1)$. Let f be the weight function depending on 2 co-ordinates:

$$f(00) = \log p, \ f(01) = \log(1-p), \ f(10) = \log q, \ f(11) = \log(1-q).$$

Consider the action of L_f on functions that only depend on 1 co-ordinate $w(x) = w(x_0)$.

$$\begin{aligned} L_f w(0) &= e^{f(00)} w(0) + e^{f(10)} w(1) = pw(0) + (1-p)w(1) \\ L_f w(1) &= e^{f(01)} w(0) + e^{f(11)} w(1) = qw(0) + (1-q)w(1). \end{aligned}$$

Record this is a matrix

$$\begin{pmatrix} L_f w(0) \\ L_f w(1) \end{pmatrix} = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix} \begin{pmatrix} w(0) \\ w(1) \end{pmatrix} = P \begin{pmatrix} w(0) \\ w(1) \end{pmatrix}$$

What are the eigenvalues of P?

An example

An example

Recall the Perron-Frobenius theorem: if B is an aperiodic non-negative matrix then

An example

Recall the Perron-Frobenius theorem: if B is an aperiodic non-negative matrix then

- there is a maximal eigenvalue $\lambda > 0$ and all other eigenvalues $\lambda_i \in \mathbb{C}$ are s.t. $|\lambda_i| < \lambda$, moreover λ is simple

An example

Recall the Perron-Frobenius theorem: if B is an aperiodic non-negative matrix then

- there is a maximal eigenvalue $\lambda > 0$ and all other eigenvalues $\lambda_i \in \mathbb{C}$ are s.t. $|\lambda_i| < \lambda$, moreover λ is simple
- there are positive left- and right-eigenvectors u, v s.t. $uB = \lambda B$, $Bv = \lambda v$; we can take u, v to be positive.

An example

Recall the Perron-Frobenius theorem: if B is an aperiodic non-negative matrix then

- there is a maximal eigenvalue $\lambda > 0$ and all other eigenvalues $\lambda_i \in \mathbb{C}$ are s.t. $|\lambda_i| < \lambda$, moreover λ is simple
- there are positive left- and right-eigenvectors u, v s.t. $uB = \lambda B$, $Bv = \lambda v$; we can take u, v to be positive.

Note that $P \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ so the maximal eigenvalue is 1.

An example

Recall the Perron-Frobenius theorem: if B is an aperiodic non-negative matrix then

- there is a maximal eigenvalue $\lambda > 0$ and all other eigenvalues $\lambda_i \in \mathbb{C}$ are s.t. $|\lambda_i| < \lambda$, moreover λ is simple
- there are positive left- and right-eigenvectors u, v s.t. $uB = \lambda B$, $Bv = \lambda v$; we can take u, v to be positive.

Note that $P \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ so the maximal eigenvalue is 1.

The corresponding left-eigenvector is $(p_0, p_1) = \left(\frac{q}{q+1-p}, \frac{1-p}{q+1-p} \right)$ so that $pP = p$.

An example

Recall the Perron-Frobenius theorem: if B is an aperiodic non-negative matrix then

- there is a maximal eigenvalue $\lambda > 0$ and all other eigenvalues $\lambda_i \in \mathbb{C}$ are s.t. $|\lambda_i| < \lambda$, moreover λ is simple
- there are positive left- and right-eigenvectors u, v s.t. $uB = \lambda B$, $Bv = \lambda v$; we can take u, v to be positive.

Note that $P \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ so the maximal eigenvalue is 1.

The corresponding left-eigenvector is $(p_0, p_1) = \left(\frac{q}{q+1-p}, \frac{1-p}{q+1-p}\right)$ so that $pP = p$.

In terms of transfer operators: $L_f 1 = 1$ ($1 =$ the function constantly equal to 1). We say that f is *normalised*.

An example

An example

L_f acts on functions of 1 variable as a stochastic matrix P with right-eigenvector p.

An example

L_f acts on functions of 1 variable as a stochastic matrix P with right-eigenvector p. This defines a Markov measure, defined on cylinders by

$$\mu_f[i_0, i_1, \ldots, i_{n-1}] = p_{i_0} P_{i_0,i_1} P_{i_1,i_2} \cdots P_{i_{n-2},i_{n-1}}.$$

An example

L_f acts on functions of 1 variable as a stochastic matrix P with right-eigenvector p. This defines a Markov measure, defined on cylinders by

$$\mu_f[i_0, i_1, \ldots, i_{n-1}] = p_{i_0} P_{i_0,i_1} P_{i_1,i_2} \cdots P_{i_{n-2},i_{n-1}}.$$

Suppose $w(x) = w(x_0)$ is a function of one co-ordinate. Then

$$\int L_f w \, d\mu_f = \mu_f([0]) L_f w(0) + \mu_f([1]) L_f w(1) = \int w \, d\mu_f.$$

An example

L_f acts on functions of 1 variable as a stochastic matrix P with right-eigenvector p. This defines a Markov measure, defined on cylinders by

$$\mu_f[i_0, i_1, \ldots, i_{n-1}] = p_{i_0} P_{i_0,i_1} P_{i_1,i_2} \cdots P_{i_{n-2},i_{n-1}}.$$

Suppose $w(x) = w(x_0)$ is a function of one co-ordinate. Then

$$\int L_f w \, d\mu_f = \mu_f([0]) L_f w(0) + \mu_f([1]) L_f w(1) = \int w \, d\mu_f.$$

Note that $P_{i,j} = e^{f(i,j)}$. Hence if $x = (x_0, x_1, \ldots) \in \Sigma$ then

$$\mu_f[x_0, x_1, \ldots, x_{n-1}] = p_{x_0} e^{f(i_0,i_1)} e^{f(i_1,i_2)} \ldots e^{f(i_{n-2},i_{n-1})} = p_{x_0} e^{f^n(x)}$$

where $f^n(x) = \sum_{j=0}^{n-1} f(\sigma^j x)$.

An example

An example

Hence there exists $C > 0$ s.t. for all $x = (x_0, x_1, \ldots) \in \Sigma$

$$\frac{1}{C} \leq \frac{\mu_f[x_0, x_1, \ldots, x_{n-1}]}{e^{f^n(x)}} \leq C.$$

Such a measure is called a *Gibbs measure*.

An example

Hence there exists $C > 0$ s.t. for all $x = (x_0, x_1, \ldots) \in \Sigma$

$$\frac{1}{C} \leq \frac{\mu_f[x_0, x_1, \ldots, x_{n-1}]}{e^{f^n(x)}} \leq C.$$

Such a measure is called a *Gibbs measure*.

In summary: for our choice of f

An example

Hence there exists $C > 0$ s.t. for all $x = (x_0, x_1, \ldots) \in \Sigma$

$$\frac{1}{C} \leq \frac{\mu_f[x_0, x_1, \ldots, x_{n-1}]}{e^{f^n(x)}} \leq C.$$

Such a measure is called a *Gibbs measure*.

In summary: for our choice of f

- L_f, acting on a suitable space of functions, is normalised ($L_f 1 = 1$) and has a maximal simple eigenvalue at 1 (all other eigenvalues have modulus < 1)

An example

Hence there exists $C > 0$ s.t. for all $x = (x_0, x_1, \ldots) \in \Sigma$

$$\frac{1}{C} \leq \frac{\mu_f[x_0, x_1, \ldots, x_{n-1}]}{e^{f^n(x)}} \leq C.$$

Such a measure is called a *Gibbs measure*.

In summary: for our choice of f

- L_f, acting on a suitable space of functions, is normalised ($L_f 1 = 1$) and has a maximal simple eigenvalue at 1 (all other eigenvalues have modulus < 1)
- there is a unique invariant measure μ_f such that $L_f^* \mu_f = \mu_f$ (i.e. $\int L_f w \, d\mu_f = \int w \, d\mu_f$)

An example

Hence there exists $C > 0$ s.t. for all $x = (x_0, x_1, \ldots) \in \Sigma$

$$\frac{1}{C} \leq \frac{\mu_f[x_0, x_1, \ldots, x_{n-1}]}{e^{f^n(x)}} \leq C.$$

Such a measure is called a *Gibbs measure*.

In summary: for our choice of f

- L_f, acting on a suitable space of functions, is normalised ($L_f 1 = 1$) and has a maximal simple eigenvalue at 1 (all other eigenvalues have modulus < 1)
- there is a unique invariant measure μ_f such that $L_f^* \mu_f = \mu_f$ (i.e. $\int L_f w \, d\mu_f = \int w \, d\mu_f$)
- the invariant measure μ_f is ergodic, and satisfies the Gibbs property.

An example

Hence there exists $C > 0$ s.t. for all $x = (x_0, x_1, \ldots) \in \Sigma$

$$\frac{1}{C} \leq \frac{\mu_f[x_0, x_1, \ldots, x_{n-1}]}{e^{f^n(x)}} \leq C.$$

Such a measure is called a *Gibbs measure*.

In summary: for our choice of f

- L_f, acting on a suitable space of functions, is normalised ($L_f 1 = 1$) and has a maximal simple eigenvalue at 1 (all other eigenvalues have modulus < 1)
- there is a unique invariant measure μ_f such that $L_f^* \mu_f = \mu_f$ (i.e. $\int L_f w \, d\mu_f = \int w \, d\mu_f$)
- the invariant measure μ_f is ergodic, and satisfies the Gibbs property.

What if we allow non-locally constant functions f?

Ruelle's Perron-Frobenius theorem

Ruelle's Perron-Frobenius theorem

Let Σ be a shift of finite type defined by an aperiodic matrix. Let $f \in F_\theta(\mathbb{R})$ be a weight function. Define $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y).$$

Ruelle's Perron-Frobenius theorem

Let Σ be a shift of finite type defined by an aperiodic matrix. Let $f \in F_\theta(\mathbb{R})$ be a weight function. Define $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y).$$

- $\exists$ a simple maximal positive eigenvalue $\lambda > 0$ of L_f.

Ruelle's Perron-Frobenius theorem

Let Σ be a shift of finite type defined by an aperiodic matrix. Let $f \in F_\theta(\mathbb{R})$ be a weight function. Define $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y).$$

- $\exists$ a simple maximal positive eigenvalue $\lambda > 0$ of L_f.
- $\exists$ a strictly positive eigenfunction $0 < h \in F_\theta(\mathbb{R})$ s.t. $L_f h = \lambda h$.

Ruelle's Perron-Frobenius theorem

Let Σ be a shift of finite type defined by an aperiodic matrix. Let $f \in F_\theta(\mathbb{R})$ be a weight function. Define $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y).$$

- $\exists$ a simple maximal positive eigenvalue $\lambda > 0$ of L_f.
- $\exists$ a strictly positive eigenfunction $0 < h \in F_\theta(\mathbb{R})$ s.t. $L_f h = \lambda h$.
- The remainder of the spectrum of L_f is contained within a disc of radius $< \lambda$.

Ruelle's Perron-Frobenius theorem

Let Σ be a shift of finite type defined by an aperiodic matrix. Let $f \in F_\theta(\mathbb{R})$ be a weight function. Define $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y).$$

- $\exists$ a simple maximal positive eigenvalue $\lambda > 0$ of L_f.
- $\exists$ a strictly positive eigenfunction $0 < h \in F_\theta(\mathbb{R})$ s.t. $L_f h = \lambda h$.
- The remainder of the spectrum of L_f is contained within a disc of radius $< \lambda$.
- $\exists$ a unique probability measure ν s.t. $L_f^* \nu = \lambda \nu$ (i.e. $\int L_f w \, d\nu = \lambda \int w \, d\mu$.

Ruelle's Perron-Frobenius theorem

Let Σ be a shift of finite type defined by an aperiodic matrix. Let $f \in F_\theta(\mathbb{R})$ be a weight function. Define $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y).$$

- $\exists$ a simple maximal positive eigenvalue $\lambda > 0$ of L_f.
- $\exists$ a strictly positive eigenfunction $0 < h \in F_\theta(\mathbb{R})$ s.t. $L_f h = \lambda h$.
- The remainder of the spectrum of L_f is contained within a disc of radius $< \lambda$.
- $\exists$ a unique probability measure ν s.t. $L_f^* \nu = \lambda \nu$ (i.e. $\int L_f w \, d\nu = \lambda \int w \, d\mu$.
- The probability measure $d\mu_f = \frac{1}{\int h \, d\nu} h \, d\nu$ is a σ-invariant measure.

Ruelle's Perron-Frobenius theorem

Let Σ be a shift of finite type defined by an aperiodic matrix. Let $f \in F_\theta(\mathbb{R})$ be a weight function. Define $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ by

$$L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y).$$

- $\exists$ a simple maximal positive eigenvalue $\lambda > 0$ of L_f.
- $\exists$ a strictly positive eigenfunction $0 < h \in F_\theta(\mathbb{R})$ s.t. $L_f h = \lambda h$.
- The remainder of the spectrum of L_f is contained within a disc of radius $< \lambda$.
- $\exists$ a unique probability measure ν s.t. $L_f^* \nu = \lambda \nu$ (i.e. $\int L_f w \, d\nu = \lambda \int w \, d\mu$.
- The probability measure $d\mu_f = \frac{1}{\int h \, d\nu} h \, d\nu$ is a σ-invariant measure.

We call μ_f the *equilibrium state* of f.

Ruelle's Perron-Frobenius theorem

Ruelle's Perron-Frobenius theorem

We write $\lambda = e^{P(f)}$. $P(f)$ is called the *pressure* of f.

Ruelle's Perron-Frobenius theorem

We write $\lambda = e^{P(f)}$. $P(f)$ is called the *pressure* of f.

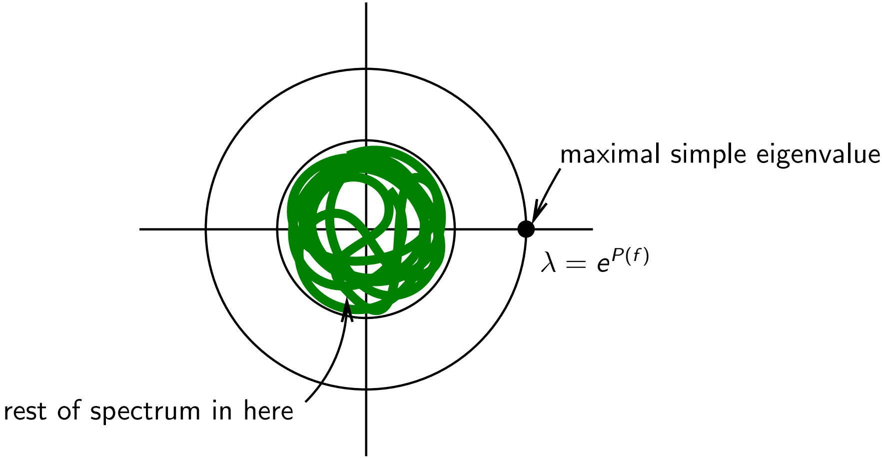
If f is normalised ($L_f 1 = 1$) then $1 = \lambda = e^{P(f)}$, i.e. $P(f) = 0$.

Ruelle's Perron-Frobenius theorem

We write $\lambda = e^{P(f)}$. $P(f)$ is called the *pressure* of f.

If f is normalised ($L_f 1 = 1$) then $1 = \lambda = e^{P(f)}$, i.e. $P(f) = 0$.

The spectrum of L_f acting on F_θ looks like:



Gibbs measures

Gibbs measures

A probability measure (not nec. invariant) m is a *Gibbs measure* if there are constants $A, B > 0$, $C \in \mathbb{R}$ s.t. for all $x \in \Sigma$

$$A \leq \frac{m[x_0, \ldots, n_{n-1}]}{e^{f^n(x) - nC}} \leq B$$

where $f^n(x) = \sum_{j=0}^{n-1} f(\sigma^j(x))$.

Gibbs measures

A probability measure (not nec. invariant) m is a *Gibbs measure* if there are constants $A, B > 0$, $C \in \mathbb{R}$ s.t. for all $x \in \Sigma$

$$A \leq \frac{m[x_0, \dots, n_{n-1}]}{e^{f^n(x) - nC}} \leq B$$

where $f^n(x) = \sum_{j=0}^{n-1} f(\sigma^j(x))$.

Remarks

- Thus if m is a Gibbs measure with potential f if, up to a bounded error, the measure of a cylinder is given by $\exp f^n(x)$ after normalisation by C.

Gibbs measures

A probability measure (not nec. invariant) m is a *Gibbs measure* if there are constants $A, B > 0$, $C \in \mathbb{R}$ s.t. for all $x \in \Sigma$

$$A \leq \frac{m[x_0, \ldots, n_{n-1}]}{e^{f^n(x) - nC}} \leq B$$

where $f^n(x) = \sum_{j=0}^{n-1} f(\sigma^j(x))$.

Remarks

- Thus if m is a Gibbs measure with potential f if, up to a bounded error, the measure of a cylinder is given by $\exp f^n(x)$ after normalisation by C.
- Not all constants are equal. (Orwellian equal, rather than numerically equal!) The values of A, B don't matter. C is the pressure of f.

Gibbs measures

A probability measure (not nec. invariant) m is a *Gibbs measure* if there are constants $A, B > 0$, $C \in \mathbb{R}$ s.t. for all $x \in \Sigma$

$$A \leq \frac{m[x_0, \ldots, n_{n-1}]}{e^{f^n(x) - nC}} \leq B$$

where $f^n(x) = \sum_{j=0}^{n-1} f(\sigma^j(x))$.

Remarks

- Thus if m is a Gibbs measure with potential f if, up to a bounded error, the measure of a cylinder is given by $\exp f^n(x)$ after normalisation by C.
- Not all constants are equal. (Orwellian equal, rather than numerically equal!) The values of A, B don't matter. C is the pressure of f.
- Given a potential f, is there an *invariant* Gibbs measure? Is it unique?

Gibbs measures

Gibbs measures

Let $f \in F_\theta(\mathbb{R})$. By Ruelle's Perron-Frobenius theorem, L_f has a maximal eigenvalue at $e^{P(f)}$. There is a maximal eigenfunction h s.t. $L_f h = e^{P(f)} h$ and a maximal eigenmeasure ν s.t. $L_f^* \nu = e^{P(f)} \nu$. The measure $d\mu_f = \frac{1}{\int h \, d\nu} h \, d\nu$ is a σ-invariant measure.

Gibbs measures

Let $f \in F_\theta(\mathbb{R})$. By Ruelle's Perron-Frobenius theorem, L_f has a maximal eigenvalue at $e^{P(f)}$. There is a maximal eigenfunction h s.t. $L_f h = e^{P(f)} h$ and a maximal eigenmeasure ν s.t. $L_f^* \nu = e^{P(f)} \nu$. The measure $d\mu_f = \frac{1}{\int h\, d\nu} h\, d\nu$ is a σ-invariant measure.

Then μ_f is the unique σ-invariant Gibbs measure with potential f:

$$A \leq \frac{\mu_f [x_0, \dots, n_{n-1}]}{e^{f^n(x) - nP(f)}} \leq B.$$

Cohomologous functions

Cohomologous functions

Two functions $f, g : \Sigma \to \mathbb{R}$ are *cohomologous* if $\exists u \in F_\theta(\mathbb{R})$ s.t.

$$f(x) = g(x) + u(\sigma(x)) - u(x) \tag{2}$$

Cohomologous functions

Two functions $f, g : \Sigma \to \mathbb{R}$ are *cohomologous* if $\exists u \in F_\theta(\mathbb{R})$ s.t.

$$f(x) = g(x) + u(\sigma(x)) - u(x) \tag{2}$$

Cohomology is the natural equivalence relation between functions in ergodic theory.

Cohomologous functions

Two functions $f, g : \Sigma \to \mathbb{R}$ are *cohomologous* if $\exists u \in F_\theta(\mathbb{R})$ s.t.

$$f(x) = g(x) + u(\sigma(x)) - u(x) \tag{2}$$

Cohomology is the natural equivalence relation between functions in ergodic theory.

Let $f^n(x) = \sum_{j=0}^{n-1} f(\sigma^j x)$. Summing (2) along orbits gives:

$$f^n(x) = g^n(x) + u(\sigma^n(x)) - u(x) = g^n(x) + O(1).$$

Cohomologous functions

Two functions $f, g : \Sigma \to \mathbb{R}$ are *cohomologous* if $\exists u \in F_\theta(\mathbb{R})$ s.t.

$$f(x) = g(x) + u(\sigma(x)) - u(x) \tag{2}$$

Cohomology is the natural equivalence relation between functions in ergodic theory.

Let $f^n(x) = \sum_{j=0}^{n-1} f(\sigma^j x)$. Summing (2) along orbits gives:

$$f^n(x) = g^n(x) + u(\sigma^n(x)) - u(x) = g^n(x) + O(1).$$

Hence, for example,

$$\lim_{n\to\infty} \frac{1}{n} f^n(x) = \lim_{n\to\infty} \frac{1}{n} g^n(x)$$

so the behaviour of cohomologous functions in the ergodic theorem is the same.

Normalising functions

Normalising functions

"Any function is cohomologous to a normalised function plus constant"

Normalising functions

> *"Any function is cohomologous to a normalised function plus constant"*

Let $f \in F_\theta(\mathbb{R})$. Then $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ has a maximal eigenvalue $\lambda > 0$ and positive eigenfunction $h > 0$.

Normalising functions

"Any function is cohomologous to a normalised function plus constant"

Let $f \in F_\theta(\mathbb{R})$. Then $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ has a maximal eigenvalue $\lambda > 0$ and positive eigenfunction $h > 0$.

Let

$$g = f - \log h\sigma + \log h - \log \lambda \in F_\theta(\mathbb{R}).$$

Normalising functions

"Any function is cohomologous to a normalised function plus constant"

Let $f \in F_\theta(\mathbb{R})$. Then $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ has a maximal eigenvalue $\lambda > 0$ and positive eigenfunction $h > 0$.

Let

$$g = f - \log h\sigma + \log h - \log \lambda \in F_\theta(\mathbb{R}).$$

Then

$$L_g 1 \quad =$$

Normalising functions

"Any function is cohomologous to a normalised function plus constant"

Let $f \in F_\theta(\mathbb{R})$. Then $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ has a maximal eigenvalue $\lambda > 0$ and positive eigenfunction $h > 0$.

Let

$$g = f - \log h\sigma + \log h - \log \lambda \in F_\theta(\mathbb{R}).$$

Then

$$L_g 1 = \sum_{y:\sigma y = x} e^{f(y) - \log h(\sigma y) + \log h(y) - \log \lambda}$$

Normalising functions

"Any function is cohomologous to a normalised function plus constant"

Let $f \in F_\theta(\mathbb{R})$. Then $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ has a maximal eigenvalue $\lambda > 0$ and positive eigenfunction $h > 0$.

Let

$$g = f - \log h\sigma + \log h - \log \lambda \in F_\theta(\mathbb{R}).$$

Then

$$\begin{aligned} L_g 1 &= \sum_{y:\sigma y = x} e^{f(y) - \log h(\sigma y) + \log h(y) - \log \lambda} \\ &= \frac{1}{\lambda} \frac{1}{h(x)} \sum_{y:\sigma y = x} e^{f(y)} h(y) \end{aligned}$$

Normalising functions

"Any function is cohomologous to a normalised function plus constant"

Let $f \in F_\theta(\mathbb{R})$. Then $L_f : F_\theta(\mathbb{C}) \to F_\theta(\mathbb{C})$ has a maximal eigenvalue $\lambda > 0$ and positive eigenfunction $h > 0$.

Let

$$g = f - \log h\sigma + \log h - \log \lambda \in F_\theta(\mathbb{R}).$$

Then

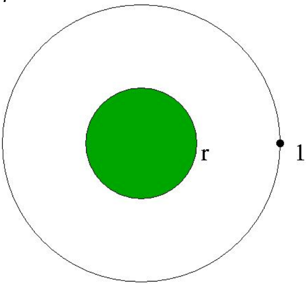
$$\begin{aligned} L_g 1 &= \sum_{y:\sigma y = x} e^{f(y) - \log h(\sigma y) + \log h(y) - \log \lambda} \\ &= \frac{1}{\lambda} \frac{1}{h(x)} \sum_{y:\sigma y = x} e^{f(y)} h(y) \\ &= \frac{1}{\lambda} \frac{1}{h(x)} L_f h(x) = \frac{1}{\lambda} \frac{1}{h(x)} \lambda h(x) = 1. \end{aligned}$$

Hence g is normalised.

How Ruelle's Perron-Frobenius theorem is often used

How Ruelle's Perron-Frobenius theorem is often used

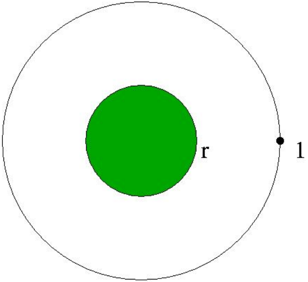
Let $f \in F_\theta(\mathbb{R})$ be normalised. (So $L_f 1 = 1$, $P(f) = 0$.) Let μ be the unique invariant measure such that $L_f^* \mu = \mu$.



How Ruelle's Perron-Frobenius theorem is often used

Let $f \in F_\theta(\mathbb{R})$ be normalised. (So $L_f 1 = 1$, $P(f) = 0$.) Let μ be the unique invariant measure such that $L_f^* \mu = \mu$.

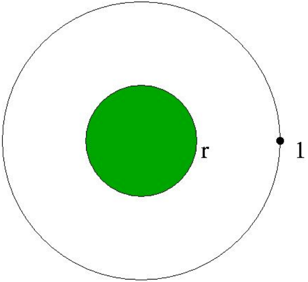
- Let $\mu : F_\theta(\mathbb{R}) \to F_\theta(\mathbb{R})$, $\mu(w) = \int w \, d\mu$ be projection onto the eigenspace corresponding to the eigenvalue 1.



How Ruelle's Perron-Frobenius theorem is often used

Let $f \in F_\theta(\mathbb{R})$ be normalised. (So $L_f 1 = 1$, $P(f) = 0$.) Let μ be the unique invariant measure such that $L_f^* \mu = \mu$.

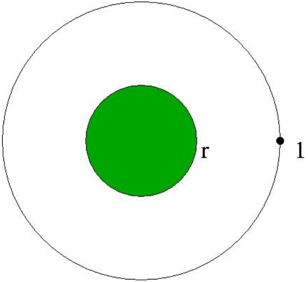
- Let $\mu : F_\theta(\mathbb{R}) \to F_\theta(\mathbb{R})$, $\mu(w) = \int w \, d\mu$ be projection onto the eigenspace corresponding to the eigenvalue 1.
- Let $Q : F_\theta(\mathbb{R}) \to F_\theta(\mathbb{R})$, $Q = L_f - \mu$. Then spec. radius of Q is $< r$, strictly less than 1.



How Ruelle's Perron-Frobenius theorem is often used

Let $f \in F_\theta(\mathbb{R})$ be normalised. (So $L_f 1 = 1$, $P(f) = 0$.) Let μ be the unique invariant measure such that $L_f^* \mu = \mu$.

- Let $\mu : F_\theta(\mathbb{R}) \to F_\theta(\mathbb{R})$, $\mu(w) = \int w \, d\mu$ be projection onto the eigenspace corresponding to the eigenvalue 1.
- Let $Q : F_\theta(\mathbb{R}) \to F_\theta(\mathbb{R})$, $Q = L_f - \mu$. Then spec. radius of Q is $< r$, strictly less than 1.



Then $L_f w = \mu(w) + Q(w)$. As eigenprojections are orthogonal:

$$L_f^n w = \mu(w) + Q^n(w) = \mu(w) + o(r^n).$$

Application: Decay of correlations

Application: Decay of correlations

Let $f \in F_\theta(\mathbb{R})$ be normalised and let μ_f be the equilibrium state corresponding to f (so $\int L_f w \, d\mu = \int w \, d\mu$).

Application: Decay of correlations

Let $f \in F_\theta(\mathbb{R})$ be normalised and let μ_f be the equilibrium state corresponding to f (so $\int L_f w \, d\mu = \int w \, d\mu$).

Recall:

- An invariant measure μ is *strong-mixing* if $\mu(\sigma^{-n} A \cap B) \to \mu(A)\mu(B) \; \forall A, B \in \mathcal{B}$.

Application: Decay of correlations

Let $f \in F_\theta(\mathbb{R})$ be normalised and let μ_f be the equilibrium state corresponding to f (so $\int L_f w \, d\mu = \int w \, d\mu$).

Recall:

- An invariant measure μ is *strong-mixing* if $\mu(\sigma^{-n} A \cap B) \to \mu(A)\mu(B) \; \forall A, B \in \mathcal{B}$. Equivalently:

$$\int v\sigma^n \cdot w \, d\mu \to \int v \, d\mu \int w \, d\mu \; \forall v, w \in L^2(\Sigma, \mathcal{B}, \mu) \quad (3)$$

Application: Decay of correlations

Let $f \in F_\theta(\mathbb{R})$ be normalised and let μ_f be the equilibrium state corresponding to f (so $\int L_f w \, d\mu = \int w \, d\mu$).

Recall:

- An invariant measure μ is *strong-mixing* if $\mu(\sigma^{-n} A \cap B) \to \mu(A)\mu(B)$ $\forall A, B \in \mathcal{B}$. Equivalently:

$$\int v\sigma^n \cdot w \, d\mu \to \int v \, d\mu \int w \, d\mu \ \forall v, w \in L^2(\Sigma, \mathcal{B}, \mu) \tag{3}$$

- If μ is strong-mixing then μ is ergodic.

Application: Decay of correlations

Let $f \in F_\theta(\mathbb{R})$ be normalised and let μ_f be the equilibrium state corresponding to f (so $\int L_f w \, d\mu = \int w \, d\mu$).

Recall:

- An invariant measure μ is *strong-mixing* if $\mu(\sigma^{-n} A \cap B) \to \mu(A)\mu(B) \ \forall A, B \in \mathcal{B}$. Equivalently:

$$\int v\sigma^n \cdot w \, d\mu \to \int v \, d\mu \int w \, d\mu \ \forall v, w \in L^2(\Sigma, \mathcal{B}, \mu) \quad (3)$$

- If μ is strong-mixing then μ is ergodic.

Proposition

The equilibrium state μ_f is strong-mixing (and so is ergodic).

Application: Decay of correlations

Let $f \in F_\theta(\mathbb{R})$ be normalised and let μ_f be the equilibrium state corresponding to f (so $\int L_f w \, d\mu = \int w \, d\mu$).

Recall:

- An invariant measure μ is *strong-mixing* if $\mu(\sigma^{-n} A \cap B) \to \mu(A)\mu(B)$ $\forall A, B \in \mathcal{B}$. Equivalently:

$$\int v\sigma^n \cdot w \, d\mu \to \int v \, d\mu \int w \, d\mu \; \forall v, w \in L^2(\Sigma, \mathcal{B}, \mu) \quad (3)$$

- If μ is strong-mixing then μ is ergodic.

Proposition

The equilibrium state μ_f is strong-mixing (and so is ergodic).

Replacing v, w by $v - \int v \, d\mu$, $w - \int w \, d\mu$, it is easy to see that

$$(3) \Leftrightarrow \int v\sigma^n \cdot w \, d\mu \to 0 \;\; \forall v, w \in L^2 \text{ s.t. } \int v \, d\mu = \int w \, d\mu = 0.$$

Decay of correlations

Decay of correlations

We want to show $\int v\sigma^n \cdot w \, d\mu_f \to 0$ for all $v, w \in L^2$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. It is sufficient to do this for a dense set of functions.

Easy check: $F_\theta(\mathbb{R})$ is dense in L^2.

Decay of correlations

We want to show $\int v\sigma^n \cdot w\, d\mu_f \to 0$ for all $v, w \in L^2$, $\int v\, d\mu_f = \int w\, d\mu_f = 0$. It is sufficient to do this for a dense set of functions.

Easy check: $F_\theta(\mathbb{R})$ is dense in L^2.

Let $v, w \in F_\theta(\mathbb{R})$, $\int v\, d\mu_f = \int w\, d\mu_f = 0$. Then

$$\left| \int v\sigma^n \cdot w\, d\mu_f \right|$$

Decay of correlations

We want to show $\int v\sigma^n \cdot w \, d\mu_f \to 0$ for all $v, w \in L^2$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. It is sufficient to do this for a dense set of functions.

Easy check: $F_\theta(\mathbb{R})$ is dense in L^2.

Let $v, w \in F_\theta(\mathbb{R})$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. Then

$$\left| \int v\sigma^n \cdot w \, d\mu_f \right| = \left| \int L_f^n(v\sigma^n \cdot w) \, d\mu_f \right|$$

Decay of correlations

We want to show $\int v\sigma^n \cdot w \, d\mu_f \to 0$ for all $v, w \in L^2$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. It is sufficient to do this for a dense set of functions.

Easy check: $F_\theta(\mathbb{R})$ is dense in L^2.

Let $v, w \in F_\theta(\mathbb{R})$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. Then

$$\left| \int v\sigma^n \cdot w \, d\mu_f \right| = \left| \int L_f^n (v\sigma^n \cdot w) \, d\mu_f \right| = \left| \int v \cdot L_f^n w \, d\mu_f \right|$$

Decay of correlations

We want to show $\int v\sigma^n \cdot w \, d\mu_f \to 0$ for all $v, w \in L^2$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. It is sufficient to do this for a dense set of functions.

Easy check: $F_\theta(\mathbb{R})$ is dense in L^2.

Let $v, w \in F_\theta(\mathbb{R})$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. Then

$$\begin{aligned} \left| \int v\sigma^n \cdot w \, d\mu_f \right| &= \left| \int L_f^n (v\sigma^n \cdot w) \, d\mu_f \right| = \left| \int v \cdot L_f^n w \, d\mu_f \right| \\ &\leq \|v\|_\infty \|L_f^n w\|_\infty \end{aligned}$$

Decay of correlations

We want to show $\int v\sigma^n \cdot w \, d\mu_f \to 0$ for all $v, w \in L^2$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. It is sufficient to do this for a dense set of functions.

Easy check: $F_\theta(\mathbb{R})$ is dense in L^2.

Let $v, w \in F_\theta(\mathbb{R})$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. Then

$$\begin{aligned}
\left| \int v\sigma^n \cdot w \, d\mu_f \right| &= \left| \int L_f^n (v\sigma^n \cdot w) \, d\mu_f \right| = \left| \int v \cdot L_f^n w \, d\mu_f \right| \\
&\leq \|v\|_\infty \|L_f^n w\|_\infty \leq \|v\|_\infty \|L_f^n w\|_\theta .
\end{aligned}$$

Decay of correlations

We want to show $\int v\sigma^n \cdot w \, d\mu_f \to 0$ for all $v, w \in L^2$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. It is sufficient to do this for a dense set of functions.

Easy check: $F_\theta(\mathbb{R})$ is dense in L^2.

Let $v, w \in F_\theta(\mathbb{R})$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. Then

$$\begin{aligned} \left| \int v\sigma^n \cdot w \, d\mu_f \right| &= \left| \int L_f^n (v\sigma^n \cdot w) \, d\mu_f \right| = \left| \int v \cdot L_f^n w \, d\mu_f \right| \\ &\leq \|v\|_\infty \|L_f^n w\|_\infty \leq \|v\|_\infty \|L_f^n w\|_\theta. \end{aligned}$$

Now $L_f^n w = \mu_f(w) + Q^n w = Q^n w \to 0$ in norm as $n \to \infty$.

Decay of correlations

We want to show $\int v\sigma^n \cdot w \, d\mu_f \to 0$ for all $v, w \in L^2$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. It is sufficient to do this for a dense set of functions.

Easy check: $F_\theta(\mathbb{R})$ is dense in L^2.

Let $v, w \in F_\theta(\mathbb{R})$, $\int v \, d\mu_f = \int w \, d\mu_f = 0$. Then

$$\begin{aligned} \left| \int v\sigma^n \cdot w \, d\mu_f \right| &= \left| \int L_f^n(v\sigma^n \cdot w) \, d\mu_f \right| = \left| \int v \cdot L_f^n w \, d\mu_f \right| \\ &\leq \|v\|_\infty \|L_f^n w\|_\infty \leq \|v\|_\infty \|L_f^n w\|_\theta. \end{aligned}$$

Now $L_f^n w = \mu_f(w) + Q^n w = Q^n w \to 0$ in norm as $n \to \infty$.

Remark

In fact, we've shown that equilibrium states (corresponding to Hölder potentials) have a property known as *exponential decay of correlations* (on Hölder functions).

Another example

Another example

Let $\Sigma = \{(x_j)_{j=0}^{\infty} \mid x_j \in \{0,1\}\}$ be the full one-sided 2-shift. Let $f(x) \equiv 0$.

Another example

Let $\Sigma = \{(x_j)_{j=0}^{\infty} \mid x_j \in \{0,1\}\}$ be the full one-sided 2-shift. Let $f(x) \equiv 0$.

Here $L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y) = w(0x) + w(1x)$.

Another example

Let $\Sigma = \{(x_j)_{j=0}^{\infty} \mid x_j \in \{0,1\}\}$ be the full one-sided 2-shift. Let $f(x) \equiv 0$.

Here $L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y) = w(0x) + w(1x)$.

Let $h \equiv 1$. Then $L_f h = 2h$. Hence the maximal eigenvalue for L_f is $\lambda = 2$, i.e. $P(f) = \log 2$.

Another example

Let $\Sigma = \{(x_j)_{j=0}^{\infty} \mid x_j \in \{0,1\}\}$ be the full one-sided 2-shift. Let $f(x) \equiv 0$.

Here $L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y) = w(0x) + w(1x)$.

Let $h \equiv 1$. Then $L_f h = 2h$. Hence the maximal eigenvalue for L_f is $\lambda = 2$, i.e. $P(f) = \log 2$.

Let ν be the Bernoulli $(1/2, 1/2)$-measure. It is straightforward to check that $\int L_f w \, d\nu = 2 \int w \, d\nu$, so ν is the eigenmeasure for L_f.

Another example

Let $\Sigma = \{(x_j)_{j=0}^{\infty} \mid x_j \in \{0,1\}\}$ be the full one-sided 2-shift. Let $f(x) \equiv 0$.

Here $L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y) = w(0x) + w(1x)$.

Let $h \equiv 1$. Then $L_f h = 2h$. Hence the maximal eigenvalue for L_f is $\lambda = 2$, i.e. $P(f) = \log 2$.

Let ν be the Bernoulli $(1/2, 1/2)$-measure. It is straightforward to check that $\int L_f w \, d\nu = 2 \int w \, d\nu$, so ν is the eigenmeasure for L_f. Hence we get an invariant measure $d\mu_f = h \, d\nu$. So μ_f is also the Bernoulli $(1/2, 1/2)$-measure.

Another example

Let $\Sigma = \{(x_j)_{j=0}^{\infty} \mid x_j \in \{0,1\}\}$ be the full one-sided 2-shift. Let $f(x) \equiv 0$.

Here $L_f w(x) = \sum_{y \text{ s.t. } \sigma(y)=x} e^{f(y)} w(y) = w(0x) + w(1x)$.

Let $h \equiv 1$. Then $L_f h = 2h$. Hence the maximal eigenvalue for L_f is $\lambda = 2$, i.e. $P(f) = \log 2$.

Let ν be the Bernoulli $(1/2, 1/2)$-measure. It is straightforward to check that $\int L_f w \, d\nu = 2 \int w \, d\nu$, so ν is the eigenmeasure for L_f. Hence we get an invariant measure $d\mu_f = h \, d\nu$. So μ_f is also the Bernoulli $(1/2, 1/2)$-measure.

Recall that μ_f is the measure of maximal entropy for σ:

$$P(f) = \log 2 = h_{\mu_f}(\sigma) = \sup\{h_\mu(\sigma) \mid \mu \text{ is } \sigma\text{-invariant}\}$$

(by the variational principle for entropy).

The variational principle for pressure

The variational principle for pressure

Let Σ be an aperiodic shift of finite type. If μ is an invariant measure we write $h(\mu) = h_\mu(\sigma)$ for the entropy.

The variational principle for pressure

Let Σ be an aperiodic shift of finite type. If μ is an invariant measure we write $h(\mu) = h_\mu(\sigma)$ for the entropy.

Theorem (The variational principle)

$$P(f) = \sup\left\{h(\mu) + \int f\, d\mu \mid \mu \textit{ is } \sigma\textit{-invariant}\right\}. \tag{4}$$

The variational principle for pressure

Let Σ be an aperiodic shift of finite type. If μ is an invariant measure we write $h(\mu) = h_\mu(\sigma)$ for the entropy.

Theorem (The variational principle)

$$P(f) = \sup \left\{ h(\mu) + \int f \, d\mu \mid \mu \text{ is } \sigma\text{-invariant} \right\}. \tag{4}$$

If $f \in F_\theta(\mathbb{R})$ then this supremum is achieved by a unique σ-invariant measure, and this measure is μ_f, i.e. $P(f) = h(\mu_f) + \int f \, d\mu_f$.

The variational principle for pressure

Let Σ be an aperiodic shift of finite type. If μ is an invariant measure we write $h(\mu) = h_\mu(\sigma)$ for the entropy.

Theorem (The variational principle)

$$P(f) = \sup \left\{ h(\mu) + \int f \, d\mu \mid \mu \textit{ is } \sigma\textit{-invariant} \right\}. \tag{4}$$

If $f \in F_\theta(\mathbb{R})$ then this supremum is achieved by a unique σ-invariant measure, and this measure is μ_f, i.e. $P(f) = h(\mu_f) + \int f \, d\mu_f$.

Remarks

- When $f \equiv 0$ this is the variational principle that we had in the last lecture: $P(0)$ is the topological entropy and the corresponding equilibrium state is the measure of maximal entropy.

The variational principle for pressure

Let Σ be an aperiodic shift of finite type. If μ is an invariant measure we write $h(\mu) = h_\mu(\sigma)$ for the entropy.

Theorem (The variational principle)

$$P(f) = \sup\left\{ h(\mu) + \int f \, d\mu \mid \mu \text{ is } \sigma\text{-invariant} \right\}. \tag{4}$$

If $f \in F_\theta(\mathbb{R})$ then this supremum is achieved by a unique σ-invariant measure, and this measure is μ_f, i.e. $P(f) = h(\mu_f) + \int f \, d\mu_f$.

Remarks

- When $f \equiv 0$ this is the variational principle that we had in the last lecture: $P(0)$ is the topological entropy and the corresponding equilibrium state is the measure of maximal entropy.
- We can use equation (4) to *define* pressure for an arbitrary continuous function f and continuous transformation T.

Definitions of pressure

Definitions of pressure

Pressure is an ubiquitous quantity in ergodic theory. There are many ways of defining it.

Definitions of pressure

Pressure is an ubiquitous quantity in ergodic theory. There are many ways of defining it.

- as the normalisation constant for Gibbs measure

$$\mu[x_0, \ldots, x_{n-1}] \sim \exp(f^n(x) - nP(f))$$

Definitions of pressure

Pressure is an ubiquitous quantity in ergodic theory. There are many ways of defining it.

- as the normalisation constant for Gibbs measure

$$\mu[x_0, \dots, x_{n-1}] \sim \exp(f^n(x) - nP(f))$$

- from the variational principle

$$P(f) = \sup\{h(\mu) + \int f \, d\mu \mid \mu \text{ is } \sigma\text{-invariant}\},$$

Definitions of pressure

Pressure is an ubiquitous quantity in ergodic theory. There are many ways of defining it.

- as the normalisation constant for Gibbs measure

$$\mu[x_0, \dots, x_{n-1}] \sim \exp(f^n(x) - nP(f))$$

- from the variational principle

$$P(f) = \sup\{h(\mu) + \int f \, d\mu \mid \mu \text{ is } \sigma\text{-invariant}\},$$

- as the log of the maximal eigenvalue for L_f,

Definitions of pressure

Pressure is an ubiquitous quantity in ergodic theory. There are many ways of defining it.

- as the normalisation constant for Gibbs measure

$$\mu[x_0, \ldots, x_{n-1}] \sim \exp(f^n(x) - nP(f))$$

- from the variational principle

$$P(f) = \sup\{h(\mu) + \int f \, d\mu \mid \mu \text{ is } \sigma\text{-invariant}\},$$

- as the log of the maximal eigenvalue for L_f,
- many other expressions: eg:

$$P(f) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{x=\sigma^n(x)} e^{f^n(x)}.$$

Definitions of pressure

Pressure is an ubiquitous quantity in ergodic theory. There are many ways of defining it.

- as the normalisation constant for Gibbs measure

$$\mu[x_0, \ldots, x_{n-1}] \sim \exp(f^n(x) - nP(f))$$

- from the variational principle

$$P(f) = \sup\{h(\mu) + \int f \, d\mu \mid \mu \text{ is } \sigma\text{-invariant}\},$$

- as the log of the maximal eigenvalue for L_f,
- many other expressions: eg:

$$P(f) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{x=\sigma^n(x)} e^{f^n(x)}.$$

(Sp $P(0) = h_{top}(\sigma)$ = exponential rate of growth of the number of periodic orbits of period n.)

Properties of pressure

Properties of pressure

- $P(f)$ is monotone: $f \leq g \Rightarrow P(f) \leq P(g)$.

Properties of pressure

- $P(f)$ is monotone: $f \le g \Rightarrow P(f) \le P(g)$.
- P is convex: $P(\alpha f + (1-\alpha)g) \le \alpha P(f) + (1-\alpha)P(g)$, $0 \le \alpha \le 1$.

Properties of pressure

- $P(f)$ is monotone: $f \leq g \Rightarrow P(f) \leq P(g)$.
- P is convex: $P(\alpha f + (1-\alpha)g) \leq \alpha P(f) + (1-\alpha)P(g)$, $0 \leq \alpha \leq 1$.
- if f is cohomologous to $g + c$ (i.e. $f = g + u\sigma - u + c$) then $P(f) = P(g) + c$.

Properties of pressure

- $P(f)$ is monotone: $f \leq g \Rightarrow P(f) \leq P(g)$.
- P is convex: $P(\alpha f + (1-\alpha)g) \leq \alpha P(f) + (1-\alpha)P(g)$, $0 \leq \alpha \leq 1$.
- if f is cohomologous to $g + c$ (i.e. $f = g + u\sigma - u + c$) then $P(f) = P(g) + c$.
- $f \mapsto P(f)$ is analytic. This is because $P(f)$ is the (log of a) simple isolated eigenvalue of a bounded linear operator. Such eigenvalues perturb analytically as the operator is perturbed. The map $f \mapsto L_f$ is analytic. Hence $f \mapsto P(f)$ is analytic.

Properties of pressure

- $P(f)$ is monotone: $f \le g \Rightarrow P(f) \le P(g)$.
- P is convex: $P(\alpha f + (1-\alpha)g) \le \alpha P(f) + (1-\alpha)P(g)$, $0 \le \alpha \le 1$.
- if f is cohomologous to $g + c$ (i.e. $f = g + u\sigma - u + c$) then $P(f) = P(g) + c$.
- $f \mapsto P(f)$ is analytic. This is because $P(f)$ is the (log of a) simple isolated eigenvalue of a bounded linear operator. Such eigenvalues perturb analytically as the operator is perturbed. The map $f \mapsto L_f$ is analytic. Hence $f \mapsto P(f)$ is analytic.

  Example: fix $f \in F_\theta(\mathbb{R})$. Define $s \mapsto P(-sf) : \mathbb{R} \to \mathbb{R}$. Then this is analytic in s.

Properties of pressure

- $P(f)$ is monotone: $f \leq g \Rightarrow P(f) \leq P(g)$.
- P is convex: $P(\alpha f + (1-\alpha)g) \leq \alpha P(f) + (1-\alpha)P(g)$, $0 \leq \alpha \leq 1$.
- if f is cohomologous to $g + c$ (i.e. $f = g + u\sigma - u + c$) then $P(f) = P(g) + c$.
- $f \mapsto P(f)$ is analytic. This is because $P(f)$ is the (log of a) simple isolated eigenvalue of a bounded linear operator. Such eigenvalues perturb analytically as the operator is perturbed. The map $f \mapsto L_f$ is analytic. Hence $f \mapsto P(f)$ is analytic.

  Example: fix $f \in F_\theta(\mathbb{R})$. Define $s \mapsto P(-sf) : \mathbb{R} \to \mathbb{R}$. Then this is analytic in s.

  (Suppose $f > 0$. Note that when $s = 0$, $P(-sf) = h_{top}(\sigma) > 0$. Also $P(-sf) \searrow -\infty$ as $s \to \infty$. Hence there is a unique s_0 such that $P(-s_0 f) = 0$. This value of s_0, for particular f, is often of great importance in applications.)

Next lecture

Next lecture

In this lecture we have only been interested in shifts of finite type.

Next lecture

In this lecture we have only been interested in shifts of finite type.

In the next lecture we discuss the hyperbolic dynamical systems, and see how one can study the ergodic theory of such systems using symbolic dynamics and thermodynamic formalism.