I. Finite systems (Ruelle, pp. 3-4)

Recall defn. of entropy of probability vector: $\overline{p} = (p_1, \ldots, p_n)$:

$$H(\overline{p}) = -\sum_i p_i \log p_i$$

(p_i is probability of "micro-state" i.

Let $u_i = U(i)$ be energy contribution from state i.

Let μ_U be the probability vector $(p_1, \ldots, p_n)$ defined by

$$p_i = \frac{e^{-u_i}}{Z(U)}$$

where Z is the normalization factor:

$$Z(U) = \sum_j e^{-u_j}.$$

Prop (from class): Let $E = E_{\mu_U}(U)$. Then

$$\max_{\overline{p}:\ E_{\overline{p}}U=E} H(\overline{p}) = \log Z(U) + E$$

and achieved uniquely by μ_U.

Interpretation: Given energy function U and expected value of energy, you get a well-defined "most likely state" μ_U

Prop (Ruelle, bottom of p. 3):

$$\max_{\overline{p}} H(\overline{p}) - E_{\overline{p}}(U) = \log Z(U)$$

and achieved uniquely by μ_U.

Proof of Ruelle Prop:

$$H(\overline{p}) - E_{\overline{p}}(U) = \sum_i p_i(-\log(p_i) - u_i)$$

$$= \sum_i p_i \log \frac{e^{-u_i}}{p_i} \leq \log \sum_i p_i \, \frac{e^{-u_i}}{p_i} = \log Z(U)$$

with equality iff $\frac{e^{-u_i}}{p_i}$ is constant $(Z(U)))$. by Jensen.

Note: class prop follows since

$$\max_{\overline{p}: \;\; E_{\overline{p}}U=E} \; H(\overline{p}) - E \leq \max_{\overline{p}} \; H(\overline{p}) - E_{\overline{p}}(U)$$

and LHS includes μ_U.

Observe:

max is a non-probabilistic quantity.

achieved uniquely by an explicit probability "measure", a "equilibrium state"

maximizing meassure has a certain from, a "Gibbs state"

So, Gibbs states = Equilibrium states

II. Variational Principle

Theorem 1 (Ruelle, p. 6):

Let M be a compact metric space.

Let T be a continuous Z^d-action on M and $f : M \to R$ a continuous function.

Let $\mathcal{M}(T)$ be the set of all Borel probability measures invariant under T.

For $\mu \in \mathcal{M}(T)$, let $h_\mu(T)$ denote the measure-theoretic entropy of T w.r.t. μ.

Let $P(T, f)$ denote the pressure (of f, T). Then

$$P(T, f) = \sup_{\mu \in \mathcal{M}(T)} \; h_\mu(T) + \int f d\mu$$

Dictionary of notation:

These notes	Ruelle
M	Ω
T	τ
f	A
μ	σ
$P(T, f)$	$P(A)$
α	$\mathfrak{U}$
$x, y \in M$	$\xi, \eta \in \Omega$
$\mathbf{m} \in Z^d$	$x \in Z^d$

Note: Under certain conditions, sup is *achieved* and under stronger conditions achieved *uniquely.*

Will now define the terms, with examples, in Theorem 1:

Compact metric space: for Ruelle, *"metrizable"* because particular quantities do not depend upon specific choice of metric.

continuous Z^d-action

Defn: group homomorphism: $Z^d :\to \text{Homeo}(M, M)$

Action generated by pairwise commuting homeos $T_1, \ldots, T_d$ of M and for $(m_1, \ldots, m_d) \in Z^d$,

$$T^{(m_1,\ldots,m_d)}(x) = T_1^{m_1} \circ \ldots \circ T_d^{m_d}(x)$$

$d = 1$:
$T^m(x) = T_1^m(x)$

Main Example: *Full Z-shift:*

$M = F^Z$ with product topology (configurations on Z with finite alphabet F).

Metric: $d(x, y) = 2^{-k}$, where x, y agree on $2k+1$ interval centered at origin but not larger k.

T_1 = left-shift map;

Special Class: *Z-shift of finite type (SFT):*

"Forbid *finitely* many configurations on finite intervals"

Examples:

Golden Mean

$F = \{0, 1\}$: forbid 11

RLL(1,2)

$F = \{0, 1\}$: forbid 11 and 000

Very special class: *Topological Markov Chain (n.n. Z-SFT)*:

Defn: Let C be a square 0-1 (transition) matrix (say $m \times m$).

Let $F = \{0, \ldots, m-1\}$.

Let $M_C = \{x \in F^{\mathbb{Z}} : C_{x_i, x_{i+1}} = 1 \text{ for all } i \in \mathbb{Z}\}$

"allowed" viewpoint

T_C: left shift on M_C

For golden mean:

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

$d = 2$:

Main Example: *Full Z^2-shift:*

$M = F^{Z^2}$ with product topology (configurations on 2-dimensional integer lattice with finite alphabet F).

Metric: $d(x, y) = 2^{-k}$, where x, y agree on a $(2k+1) \times (2k+1)$ square centered at origin but no larger k.

T_1 = left-horizontal shift; T_2 = down-vertical shift, Special Example: Z^2*-shift of finite type (SFT):*

"Forbid finitely many configurations on finite shapes"

Given finite sets $\Delta_1, \ldots, \Delta_n \subset Z^2$ and $u_1 \in F^{\Delta_1}, \ldots, u_n \in F^{\Delta_n}$,

$$M = \{x \in F^{Z^2} : \ \forall \mathbf{m} \in Z^2, \ x_{\Delta_i + \mathbf{m}} \neq u_i, \ i = 1, \ldots, n\}$$

Note: *translation-invariant* condition

Ruelle: defines SFT by "allowed" configs (Ω in mid-page 7):

a finite set $\Delta \subset Z^2$ (think rectangle) and set $G \subset F^\Delta$:

$$M = \{x \in F^{Z^2} : \ \forall \mathbf{m} \in Z^2, \ x_{\Delta + \mathbf{m}} \in G\}$$

Examples:

Hard square

$F = \{0, 1\}$: forbid 11 horizontally and vertically

$RLL(1,2)^{\otimes 2}$

Dominos (Dimers)

$F = \{L, R, T, B\}$: forbid horizontal configs LL, LT, LB, RR, TR, BR, and vertical configs.

Monomer-Dimers

Z^2-TMC (n.n. Z^2-SFT): horizontal and vertical transition matrices

(Topological) Entropy:

Defn: For finite open cover α of M,

$$N(\alpha) = \text{ minimum size of subcover of } \alpha$$

$$H(\alpha) = \log N(\alpha).$$

Defn: for finite open covers α, β of M

$\alpha \vee \beta = \{A_i \cap B_j : \text{ nonempty } \}$

For $\mathbf{m} \in Z^d$, $T^{-\mathbf{m}}(\alpha) = \{T^{-\mathbf{m}}(A_i)\}$

For set $\Lambda \subset Z^d$:

$\alpha_\Lambda = \vee_{\mathbf{m}\in\Lambda} T^{-\mathbf{m}}(\alpha)$.

Consider d-dimensional prisms $\Lambda = \Lambda(a_1, \ldots, a_d)$,

$$h(T, \alpha) = \lim_{a_1,\ldots,a_d \to \infty} (1/|\Lambda|) \log N(\alpha_\Lambda)$$

Defn:

$$h(T) = \sup_\alpha h(T, \alpha)$$

Theorem: If α is a top. generator (i.e., distinguishes points), then $h(T) = h(T, \alpha)$.

Here, "distinguishes points" means: letting $\alpha(x) = \cup_{i:x\in A_i} A_i$, if $x, y \in M$ and $x \neq y$, then for some $u \in Z^d$, $\alpha(T^u(x)) \cap \alpha(T^u(y)) = \emptyset$.

For full shift and SFT's, the standard cover:

$\alpha = \{\{x \in M : x_0 = a\} : a \in F\}$ is a topological generator.

So, $h(T)$ is a growth rate of counts.

$d = 1$:

$$h(T) = \lim_{n\to\infty} (1/n) \log(\# \text{ allowed n-sequences })$$

Proposition: For a Z-TMC M_C,

$$h(T_C) = \log \lambda_C$$

where λ_C is the spectral radius of C, i.e.,

$$\lambda_C = \max\{|\lambda| : \lambda \text{ eigenvalues of } C\}$$

Proof:

$$h(T_C) = \lim_{n\to\infty} (\frac{1}{n}) \log \mathbf{1}(C)^{n-1}\mathbf{1}$$

Golden Mean shift: $h(T) = \log($ golden mean$)$

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Can deduce formula for top. entropy of any Z-SFT,

$d = 2$:

$$h(T) = \lim_{n\to\infty} (1/n^2) \log(\# n \times n \text{ allowed arrays })$$

Hard square: ???

Dominos (Dimers): $(1/4) \int_0^1 \int_0^1 (4-2\cos(2\pi s)-2\cos(2\pi t))\, dsdt$

— Monomer-Dimer: ???

Pressure:

Defn given in Ruelle, p. 5:

Same as top, entropy except:

For $\Lambda = \Lambda(a_1, \ldots, a_d)$, replace $N(\alpha_\Lambda)$ by:

$$Z_\Lambda(f, \alpha) = \min_{\text{subcover } \beta \text{ of } \alpha_\Lambda} \sum_j \exp(\sup_{x\in B_j} \sum_{u\in\Lambda} f(T^u(x)))$$

$\beta = \{B_1, B_2, \ldots B_{n(\beta)}\}$

So,

$$P(T, f, \alpha) = \lim_{a_1,\ldots,a_d\to\infty} (1/|\Lambda|) \log Z_\Lambda(f, \alpha)$$

and

$$P(T, f) = \sup_\alpha P(T, f, \alpha)$$

Theorem: If α is a topological generator, then $P(T, f) = P(T, f, \alpha)$.

Note:

$$Z_\Lambda(0, \alpha) = N(\alpha_\Lambda).$$

$$P(T, 0) = h(T).$$

Examples:

$d = 1$:

T_C is TMC:

$f(x) = f(x_0 x_1)$

$$P(T_C, f) = \lim_{n\to\infty} (1/n) \log\Big(\sum_{x_0 \ldots x_n} \exp(f(x_0 x_1) + \ldots f(x_{n-1} x_n))\Big)$$

Note: No min and No sup.

Prop:

$$P(T_C, f) = \log \lambda(C_f)$$

where

$$(C_f)_{ij} = C_{ij} e^{f(ij)}.$$

Proof:

$$P(T_C, f) = \lim_{n\to\infty} \Big(\frac{1}{n}\Big) \log \mathbf{1}(C_f)^{n-1}\mathbf{1}$$

$d = 2$:

1. Hard square with activity.

T = Hard square SFT

Let $c \in R$ and define

$f_c(x) = c$ if $x_0 = 1$
$f_c(x) = 0$ if $x_0 = 0$.

$P(T, f_c)$ = growth rate of number of allowed arrays, with 1's weight by e^c and 0's weighted by 1.

$a = e^c$; activity level

No exact known formula for $P(T, f_c)$ known.

2. Ising model
T = full shift on $F = \{\pm 1\}$

f: Ising model

Given constants β, J, H,

$$f(x) = \beta(Bx_{0,0} + J(x_{0,0}x_{1,0} + x_{0,0}x_{0,1}))$$

β : inverse temperature

J: interaction strength

B: external magnetic field strength

$P(T, f)$: growth rate of number of allowed arrays, weighted by e^f, which incorporates interactions on adjacent sites (horizontal and vertical) and magnetic field (on individual sites) .

Onsager: exact solution for $P(T, f)$, when $B = 0$.

Measure-theoretic entropy

Let T be an MPT Z^d-action on probability space $(X, \mathcal{A}, \mu)$.

Defn: For finite, measurable *partition* α,

$$H_\mu(\alpha) = -\sum_i \mu(A_i) \log \mu(A_i)$$

where $\alpha = \{A_i\}$.
For finite set $\Lambda \subset Z^d$:

$\alpha_\Lambda = \vee_{\mathbf{m}\in\Lambda} T^{-\mathbf{m}}(\alpha)$.

Consider d-dimensional prisms $\Lambda = \Lambda(a_1, \ldots, a_d)$,

$$h_\mu(T, \alpha) = \lim_{a_1,\ldots,a_d \to \infty} (1/|\Lambda|) H_\mu(\alpha_\Lambda)$$

Defn:

$$h_\mu(T) = \sup_\alpha h_\mu(T, \alpha)$$

Theorem: If α is a meas.-theo. generator (i.e., $\alpha_{Z^d} = \mathcal{A}$ a.e.), then $h_\mu(T) = h_\mu(T, \alpha)$.

$d = 1$:

X is a *stationary* process with law μ and and $T =$ left-shift, then

$$h_\mu(T) = h(X) = \lim_{n\to\infty} (1/n) H(X_1, \ldots, X_n)$$

. where $H(X_1, \ldots, X_n)$ is the entropy of $(X_1, \ldots, X_n)$ as a random vector.

Examples:

$\mu = \text{iid}(\overline{p})$:

$$h_\mu(T) = H(\overline{p})$$

μ : stationary (first-order) Markov with probability transition matrix P with stationary vector π:

$$h_\mu(T) = -\sum_{ij} \pi_i P_{ij} \log P_{ij}$$

$d = 2$:

X is a stationary Z^2-process with law μ and $T^{(m,n)}$: shift by translation (m, n). Then

$$h_\mu(T) = h(X) = \lim_{n\to\infty} (1/n^2) H(X_{i,j} : \ 1 \le i, j \le n).$$

where H is the entropy of the random vector (array):

$$X_{i,j} : \ 1 \leq i, j \leq n$$

Examples:

1. $\mu = \text{iid}(p)$:

$$h_\mu(T) = H(p)$$

2. Markov chains replaced by Gibbs measures/Markov random fields.

Few explicit results.

Back to Variational Principle:

$$P(T, f) = \sup_{\mathcal{M}(T)} \ h_\mu(T) + \int f d\mu$$

Defn: An *equilibrium state* for T, f is a measure $\mu \in \mathcal{M}(T)$ which achieves $P(T, f)$.

Let $I_{T,f}$ denote the set of equilibrium states (which an be empty).

Defn: T is *expansive* if there exists $\delta > 0$ s.t. $\forall x \neq y \in M, \exists \mathbf{m} \in Z^d$ s.t. $dist(T^{\mathbf{m}}x, T^{\mathbf{m}}y) > \delta$.

Fact: Any Z^d-SFT is expansive.

Theorem: If T is expansive, then for every continuous f, $I_{T,f} \neq \emptyset$.

Proof uses upper semi-continuity of $\mu \mapsto h_\mu(T)$

Non-uniqueness corresponds to *phase transition.*

$d = 1$:

Special case:

Theorem (Variational Principle for irreducible TMC) Let T be TMC and $f(x) = f(x_0 x_1)$. Let

$$(C_f)_{ij} = C_{ij} e^{f(ij)}$$

Then

$$P(T_C, f) = \log \lambda_{C_f} = \sup_{\mu \in \mathcal{M}} h_\mu(T_C) + \int f d\mu$$

and the sup is achieved uniquely by an explicitly describable Markov chain:

$$P_{ij} = C_{ij} e^{f(ij)} \frac{v_j}{\lambda_{C_f} v_i}$$

where v is a right eigenvector for matrix C_f and eigenvalue λ_{C_f}.

Example: Golden mean with $f = f_c$, $(a = e^c)$.

$$C_f = \begin{bmatrix} a & 1 \\ a & 0 \end{bmatrix}$$

$$\lambda = \frac{a + \sqrt{a^2 + 4a}}{2}$$

$$v = \begin{bmatrix} \lambda \\ a \end{bmatrix}$$

No phase transition!

See lecture notes from Entropy class for proof in case $c = 0$.

$d = 2$:

1. Hard core with activity $a = e^c$: unique equilibrium state up to some critical threshold.

2. Ising model: unique equilibrium state up to some critical threshold in β, when $B = 0$.

Gibbs measures

Let $T : M \to M$ be a nearest neighbour Z^2-SFT.

Let $C_1 = F$, the alphabet (a.k.a. configurations on single nodes)

Let C_2 be all *allowed* configurations on domino shapes (i.e., configurations on 1×2 and 2×1 rectangles).

Let $\Phi : C_1 \cup C_2 \to R$, (*nearest-neighbour interaction*).

A *translation invariant (stationary) nearest-neighbour Gibbs measure* on M is a T-invariant measure μ on M such that for all finite subsets $\Lambda \subset Z^d$ and a.e. $x \in M$,

$$\mu(x|_\Lambda \mid x|_{\Lambda^c}) \sim$$

$$\left(\prod_{v\in\Lambda} \exp(\Phi(x_v))\right) \left(\prod_{i=1}^{d} \prod_{\{v\in\Lambda,\ v+e_i\in\Lambda\cup\partial\Lambda\}} \exp(\Phi((x_v, x_{v+e_i})))\right). \tag{1}$$

In particular, $\mu(x|_\Lambda \mid x|_{\Lambda^c}) = \mu(x \mid x|_{\partial\Lambda})$.

Let

$$f_\Phi(x) = \Phi(x_{(0,0)}) + \Phi(x_{(0,0)}, x_{(1,0)}) + \Phi(x_{(0,0)}, x_{(0,1)})$$

Theorem:

$\{$ Equilibirum states for $f_\Phi\} \subseteq \{$ translation invariant Gibbs states for $\Phi\}$

Assuming Condition D (a mixing condition) on the SFT M (Ruelle, p. 57), in particular, for the full shift,

$\{$ Equilibirum states for $f_\Phi\} = \{$ translation invariant Gibbs states for $\Phi\}$

There is a much more general version of this (see Ruelle, Theorem 3, p.8 and Theorem 4.2, p. 58):

1. Begin with an Interaction: function Φ on allowed configurations on finite sets (see Chapter 1)

2. Form the Energy function: f_Φ, a sum of interaction values
3. A Gibbs measure is a measure μ that satisfies: whenever $x, y \in M$ disagree at only finitely many sites, then

$$\mu(x|_\Lambda \mid x|_{\Lambda^c}) = \left[\sum_{y:\ y_{Z^d \setminus \Lambda} = x_{Z^d \setminus \Lambda}} \prod_{u \in Z^d} \exp(f_\Phi(T^u(y)) - f_\Phi(T^u(x))) \right]^{-1}$$

(in nearest neighbour special case above, this is equivalent to (1)

In Ruelle (pp, 7-8), there is no mention of interaction Φ. Gibbs measure is defined for any function f on M that has exponentially decreasing dependence (equivalently Holder continuous). In Ruelle (chapters 3 and 4), $f = f_\Phi$ where Φ has satisfies a summability condition.

Equilibrium states and derivative of pressure

Let $C^\alpha(M, R)$ denote the set of Holder continuous functions, with exponent α from M to R.

For a topologically mixing Z-SFT $T : M \to M$, and $f, g \in C^\alpha(M, R)$, if μ_f is the unique equilibrium state for T, f, then

$$\frac{d}{dt} P(f + tg) = \int g d\mu_f$$

Thus, a unique equilibrium state can be viewed as a derivative of the pressure map

$$P : C^\alpha \to R$$

Phase transitions correspond to discontinuities in derivative of pressure (as well as non-uniqueness of equilibrium states).