A VIEW OF EXTREMALITY IN TERMS OF PROBABILITY KERNELS:
MIXING PROPERTIES OF $\mu^+$ AND $\mu^-$ FOR ISING MODEL ON $\mathbb{Z}^2$

RAIMUNDO BRICEÑO

In these notes we will study the structure of the set of Gibbs measures $\mathcal{G}(\gamma)$ for a given specification $\gamma$. For now we have been studying some particular cases of (Gibbsian) specifications, namely the ones induced by a nearest-neighbour (n.n.) interaction potential $\Phi$. This is an important class, where the Ising model on $\mathbb{Z}^2$ is one of the main examples. By using the idea that the set $\mathcal{G}(\gamma)$ is nothing more than a set of probability measures preserved by a particular family of probability kernels, we will establish properties of the convex set $\mathcal{G}(\gamma)$ and characterize its extreme elements, which will ideally represent an equilibrium state (or phase) of a real system. Using the same formalism, we will also find connections between ergodicity and extremality.

1. Measure kernels

Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be two measurable spaces. A function $\pi : \mathcal{X} \times Y \to [0, \infty]$ is called a (probability) kernel from $\mathcal{Y}$ to $\mathcal{X}$ if:

(i) $\pi(\cdot | y)$ is a (probability) measure on $(X, \mathcal{X})$ for all $y \in Y$, and

(ii) $\pi(A | \cdot)$ is $\mathcal{Y}$-measurable for each $A \in \mathcal{X}$.

Example 1.1. Given $\varphi : Y \to X$ measurable, the function $(A, y) \mapsto 1_A \circ \varphi(y)$ is a kernel from $\mathcal{Y}$ to $\mathcal{X}$.

A kernel $\pi$ from $\mathcal{Y}$ to $\mathcal{X}$ maps measures $\mu$ on $(Y, \mathcal{Y})$ to measures $\mu\pi$ on $(X, \mathcal{X})$, where:

$$
\mu\pi(A) := \int d\mu\pi(A | \cdot), \quad \text{for } A \in \mathcal{X}.
$$

Let $f : X \to \mathbb{R}$ be a measurable function. Then $\pi f : Y \to \mathbb{R}$ is a measurable function given by:

$$
\pi f := \pi(\cdot | y) = \int \pi(dx | \cdot) f(x).
$$

On the other hand, if $f \geq 0$, we let $f\pi$ to be the kernel from $\mathcal{Y}$ to $\mathcal{X}$ defined by:

$$
f\pi(A | \cdot) := \pi(f 1_A) = \int_A \pi(dx | \cdot) f(x), \quad \text{for } A \in \mathcal{X}.
$$

If $(Z, \mathcal{Z})$ is a third measurable space, then the composition $\pi_1 \pi_2$ of a kernel $\pi_1$ from $\mathcal{Z}$ to $\mathcal{Y}$ and a kernel $\pi_2$ from $\mathcal{Y}$ to $\mathcal{X}$ is a kernel from $\mathcal{Z}$ to $\mathcal{X}$ defined by the formula:

$$
\pi_1 \pi_2(A | z) = \int \pi_1(dy | z) \pi_2(A | y), \quad \text{for } A \in \mathcal{X}, \ z \in \mathcal{Z}.
$$

We will focus on a particular family of probability kernels. Let $S$ be a countable infinite set and $(E, \mathcal{E})$ any measurable space. (Think of $S = \mathbb{Z}^2$, $E = \{-1, +1\}$ and $\mathcal{E} = 2^E$.) We consider the product space $(\Omega, \mathcal{F})$, given by $\Omega = E^S = \{\omega = (\omega_i)_{i \in S} : \omega_i \in E\}$ and $\mathcal{F} = \mathcal{E}^S$, the product $\sigma$-algebra. We denote the set of probability measures on $(\Omega, \mathcal{F})$ by $\mathcal{P}(\Omega, \mathcal{F})$ (the set of random fields).

Consider $(\Omega, \mathcal{F})$ and $\mathcal{B} \subseteq \mathcal{F}$ a sub-$\sigma$-algebra. A probability kernel $\pi : \mathcal{F} \times \Omega \to [0, \infty]$ is a proper probability kernel from $\mathcal{B}$ to $\mathcal{F}$ if:

(i) $\pi(\cdot | \omega)$ is a probability measure on $(\Omega, \mathcal{F})$, for all $\omega \in \Omega$,

(ii) $\pi(A | \cdot)$ is $\mathcal{B}$-measurable, for each $A \in \mathcal{F}$, and

(iii) $\pi(B | \cdot) = 1_B$, for $B \in \mathcal{B}$ (where the word “proper” is coming from).
2. Specifications

Let $S = \{ \Lambda \subset S : 0 < |\Lambda| < \infty \}$ be the set of all non-empty finite subsets of $S$. We denote $\sigma_i : \Omega \to E$, the projection $\omega \mapsto \omega_i$ and $\sigma_\Lambda : \Omega \to E^\Delta$ the natural extension to subsets $\Lambda \subseteq S$. By definition, $F$ is the smallest $\sigma$-algebra on $\Omega$ containing the cylinder events $\{ \sigma_\Lambda \in A \}$, for $\Lambda \in S$ and $A \in E^\Lambda$. For each $\Delta \subseteq S$, we consider the $\sigma$-algebra $F_\Delta$ of all events occurring in $\Delta$ and $T_\Lambda = F_{S\setminus \Lambda}$. The tail $\sigma$-algebra is defined as:

$$T := \bigcap_{\Lambda \in S} T_\Lambda.$$

**Definition 2.1.** A specification with parameter set $S$ and state space $(E, \mathcal{E})$ is a family $\gamma = (\gamma_\Lambda)_{\Lambda \in S}$ of proper probability kernels $\gamma_\Lambda$ from $T_\Lambda$ to $F$ which satisfy the consistency condition $\gamma_\Delta \gamma_\Lambda = \gamma_\Delta$, when $\Lambda \subseteq \Delta$. The random fields in the set:

$$G(\gamma) := \{ \mu \in \mathcal{P}(\Omega, F) : \mu(A|T_\Lambda) = \gamma_\Lambda(A|\cdot) \text{ $\mu$-a.s. for all } A \in F \text{ and } \Lambda \in S \}$$

are then said to be specified or to be admitted by $\gamma$.

**Remark 1.** If $\pi$ is a proper probability kernel from $B$ to $X$ and $\mu \in \mathcal{P}(X, \mathcal{X})$, then:

$$\mu(A|B) = \pi(A|\cdot) \text{ $\mu$-a.s., for all } A \in \mathcal{X} \iff \mu = \pi.$$

**Definition 2.2.** An interaction potential is a family $\Phi = (\Phi_\Lambda)_{\Lambda \in S}$ of functions $\Phi_\Lambda : \Omega \to \mathbb{R}$ such that:

(i) For each $\Lambda \in S$, $\Phi_\Lambda$ is $F_\Lambda$-measurable.

(ii) For all $\Lambda \in S$ and $\omega \in \Omega$, the Hamiltonian $H_\Lambda^{\Phi}(\omega) := \sum_{A \in S, A \neq \emptyset} \Phi_\Lambda(\omega)$ exists.

**Idea (Gibbsian specification):** $\gamma_\Lambda^\Phi(\{ \sigma_\Lambda = \xi_\Lambda \}) \propto \exp[-\beta H_\Lambda^{\Phi}(\xi_\Lambda \omega_\Lambda \cdot)]$, where $\omega$ represents the frozen “surrounding world”. This is equivalent in the n.n. case to ask conditional probabilities of $\mu$ in finite volumes to be proportional to the measures “$\mu_{S, \delta}$” defined in class. The elements of $G(\gamma^\Phi)$ are called Gibbs measures.

3. Extremality

**Definition 3.1.** An element $\mu$ of a convex subset $C$ of any real vector space is said to be extreme in $C$ ($\mu \in \text{ex } C$) if $\mu \neq s\nu + (1-s)\nu'$ for all $0 < s < 1$ and $\nu, \nu' \in C$ with $\nu \neq \nu'$.

**Proposition 3.1.** Let $(\Omega, F)$ be a measurable space, $\pi$ a probability kernel from $F$ to $F$, and $\mu \in \mathcal{P}(\Omega, F)$ with $\mu \pi = \mu$. Then the system:

$$\mathcal{I}_\pi(\mu) = \{ A \in F : \pi(A|\cdot) = 1_A \text{ $\mu$-a.s.} \}$$

of all $\mu$-almost surely $\pi$-invariant sets is a $\sigma$-algebra, and for all measurable $f : \Omega \to [0, \infty]$ we have:

$$(f \mu) \pi = f \mu \iff f \text{ is } \mathcal{I}_\pi(\mu)-measurable.\)$$

**Proof.** Suppose $f$ is $\mathcal{I}_\pi(\mu)$-measurable. We will show that $(f \mu) \pi = f \mu$. $f$ is the limit of an increasing sequence of $\mathcal{I}_\pi(\mu)$-measurable step functions. Therefore it is sufficient to prove that $(1_A \mu) \pi = 1_A \mu$ for all $A \in \mathcal{I}_\pi(\mu)$. We choose any $B \in F$. Then

$$(1_A \mu) \pi(B) = (1_A \mu) \pi(A \cap B) + (1_A \mu) \pi(B \setminus A) \leq \mu \pi(A \cap B) + \mu((1_A \pi(\Omega \setminus A|\cdot))) = \mu(A \cap B) + \mu(1_{A \Omega \setminus A}) = (1_A \mu)(B).$$

Similarly, $(1_A \mu) \pi(\Omega \setminus B)$. This implies $(1_A \mu) \pi(B) = (1_A \mu)(B)$ because

$$(1_A \mu) \pi(B) + (1_A \mu) \pi(\Omega \setminus B) = \mu(A) = (1_A \mu)(B) + (1_A \mu)(\Omega \setminus B).$$

The proof is thus complete. \square

We will say a probability measure $\mu$ is trivial on a $\sigma$-algebra $\mathcal{A}$ if $\mu(A) = 0$ or 1, for all $A \in \mathcal{A}$. 

2
Corollary 1. Let $(\Omega, F)$ be a measurable space, $\Pi$ a non-empty set of probability kernels from $F$ to $F$, and
\[ \mathcal{P}_\Pi = \{ \mu \in \mathcal{P}(\Omega, F) : \mu \pi = \mu \text{ for all } \pi \in \Pi \} \]
the convex set of all $\Pi$-invariant probability measures on $(\Omega, F)$. Let $\mu \in \mathcal{P}_\Pi$ be given and $\mathcal{I}_\Pi(\mu) = \bigcap_{\pi \in \Pi} \mathcal{I}_\pi(\mu)$ be the $\sigma$-algebra of all $\mu$-almost surely $\Pi$-invariant sets. Then $\mu \in \text{ex } \mathcal{P}_\Pi$ if and only if $\mu$ is trivial on $\mathcal{I}_\Pi(\mu)$.

Proof. Suppose there exists a set $A \in \mathcal{I}_\Pi(\mu)$ such that $0 < \mu(A) < 1$. Then we may look at the conditional probabilities
\[ \nu = \mu(\cdot|A) = f \mu, \quad \nu' = \mu(\cdot|A^c) = f' \mu, \]
where $f = 1_A/\mu(A)$ and $f' = 1_{A^c}/\mu(A^c)$. Clearly, $\nu \neq \nu'$ and $\mu = \mu(A)\nu + (1 - \mu(A))\nu'$. But Proposition 3.1 asserts that $\nu, \nu' \in \mathcal{P}_\Pi$ because $f$ and $f'$ are $\mathcal{I}_\sigma(\mu)$-measurable for all $\pi \in \Pi$. Thus $\mu$ is not extreme. \(\square\)

Remark 2. Suppose $\gamma$ is a specification. Then $\mathcal{I}_\gamma = \mathcal{T}$. If $\mu \in \mathcal{G}(\gamma) = \mathcal{P}_\gamma$, then $\mathcal{I}_\gamma(\mu)$ is the $\mu$-completion of $\mathcal{T}$.

Proof. If $A \in \mathcal{T}$ then $\gamma_A(A|\cdot) = 1_A$ for all $\Lambda \in \mathcal{S}$ because all $\gamma_A$'s are proper. Conversely, if $A \in \mathcal{I}_\gamma$ then $A = \{ \gamma_A(A|\cdot) = 1 \} \in \mathcal{T}_\Lambda$ for all $\Lambda \in \mathcal{S}$ and therefore $A \in \mathcal{T}$. \(\square\)

The preceding remark implies in particular that $\mathcal{I}_\gamma(\mu)$ is $\mu$-trivial if and only if $\mathcal{T}$ is $\mu$-trivial.

Theorem 3.2. Let $\gamma$ be a specification. Then the following conclusions hold.

(a) A Gibbs measure $\mu \in \mathcal{G}(\gamma)$ is extreme in $\mathcal{G}(\gamma)$ if and only if $\mu$ is trivial on the tail $\sigma$-field $\mathcal{T}$.
(b) If $\mu \in \mathcal{G}(\gamma)$ and $\nu \in \mathcal{P}(\Omega, F)$ is absolutely continuous with respect to $\mu$ then $\nu \in \mathcal{G}(\gamma)$ if and only if $\nu = f \mu$ for some $\mathcal{T}$-measurable function $f \geq 0$.
(c) Each $\mu \in \mathcal{G}(\gamma)$ is uniquely determined (within $\mathcal{G}(\gamma)$) by its restriction to the tail $\sigma$-field $\mathcal{T}$.
(d) Distinct extreme elements $\mu, \nu$ of $\mathcal{G}(\gamma)$ are mutually singular on $\mathcal{T}$, in that there is some $A \in \mathcal{T}$ with $\mu(A) = 1, \nu(A) = 0$.

Idea: A system’s state is described by a suitable extreme element (phase) of $\mathcal{G}(\gamma)$.

- Microscopic quantities: rapid fluctuations, random, consistent with observed empirical distribution of microscopic variables (Gibbs distribution in finite volumes).
- Macroscopic quantities: constant, non-random, tail measurable functions!

Proposition 3.3. For each $\mu \in \mathcal{P}(\Omega, F)$ the following statements are equivalent.

(i) $\mu$ is trivial on $\mathcal{T}$.
(ii) For all cylinder events $A$ (or, equivalently, for all $A \in \mathcal{F}$),
\[ \lim_{\Lambda \in \mathcal{S}} \sup_{B \in \mathcal{T}_\Lambda} \|\mu(A \cap B) - \mu(A)\mu(B)\| = 0. \]

Proof. Suppose $\mu$ is trivial on $\mathcal{T}$. Let $A \in \mathcal{F}$ be given. The backward martingale convergence theorem asserts that for each cofinal increasing sequence $(\Lambda_n)_{n \geq 1}$ in $\mathcal{S}$
\[ \mu(A|\mathcal{T}_{\Lambda_n}) \to \mu(A|\mathcal{T}) \]
in the $L^1(\mu)$-sense. As $\mu$ is trivial on $\mathcal{T}$, $\mu(A|\mathcal{T}) = \mu(A)$ $\mu$-a.s. Thus for each $\epsilon > 0$ there is some $\Delta \in \mathcal{S}$ such that:
\[ \mu(\|\mu(A|\mathcal{T}_{\Delta}) - \mu(A)\|) < \epsilon. \]

For all $\Delta \subset \Lambda \in \mathcal{S}$ we have:
\[ \sup_{B \in \mathcal{T}_{\Delta}} |\mu(A \cap B) - \mu(A)\mu(B)| \leq \sup_{B \in \mathcal{T}_{\Delta}} \left| \int_B d\mu(A|\mathcal{T}_{\Delta}) - \mu(A) \right| \leq \mu(\|\mu(A|\mathcal{T}_{\Delta}) - \mu(A)\|) < \epsilon. \]

This proves (ii). \(\square\)
4. Ergodic random fields

Fix $S = \mathbb{Z}^d$, the $d$-dimensional integer lattice and consider $\Theta$ the shift group. Then:

$$\mathcal{P}_\Theta(\Omega, \mathcal{F}) = \{\mu \in \mathcal{P}(\Omega, \mathcal{F}) : \theta_i(\mu) = \mu \text{ for all } i \in S\},$$

is the set of shift-invariant random fields on $\mathbb{Z}^d$ (which is always non-empty). We will consider the $\sigma$-algebra

$$\mathcal{I}_\Theta = \{A \in \mathcal{F} : \theta_i A = A \text{ for all } i \in S\}$$

of all shift-invariant events.

**Remark 3.**

1. An $\mathcal{F}$-measurable function $f : \Omega \to \mathbb{R}$ is $\mathcal{I}_\Theta$-measurable if and only if $f$ is invariant, in that $f \circ \theta_i = f$, for all $i \in S$.
2. For each $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$, the $\sigma$-algebra

$$\mathcal{I}_\Theta(\mu) = \{A \in \mathcal{F} : 1_A \circ \theta_i = 1_A \text{ } \mu\text{-a.s. for all } i \in S\}$$

of all $\mu$-almost surely invariant events is the $\mu$-completion of $\mathcal{I}_\Theta$.

We can see the parallel between $\mathcal{G}(\gamma)$ and $\mathcal{T}$, and $\mathcal{P}_\Theta(\Omega, \mathcal{F})$ and $\mathcal{I}_\Theta$ by defining the family of probability kernels $\Pi = \{\hat{\theta}_i : i \in S\}$ from $\mathcal{F}$ to $\mathcal{F}$ given by:

$$\hat{\theta}_i(A|\omega) = 1_A(\theta_i \omega),$$

for $i \in S, A \in \mathcal{F}, \omega \in \Omega$.

**Theorem 4.1.**

1. A probability measure $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$ is extreme in $\mathcal{P}_\Theta(\Omega, \mathcal{F})$ if and only if $\mu$ is trivial on the invariant $\sigma$-algebra $\mathcal{I}_\Theta$.
2. Suppose $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$ and $\nu \in \mathcal{P}(\Omega, \mathcal{F})$ is absolutely continuous relative to $\mu$. Then $\nu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$ if and only if $\nu = f \mu$ for some $\mathcal{I}_\Theta$-measurable function $f$.
3. Each $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$ is uniquely determined (within $\mathcal{P}_\Theta(\Omega, \mathcal{F})$) by its restriction to $\mathcal{I}_\Theta$.
4. Distinct probability measures $\mu, \nu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$ are mutually singular on $\mathcal{I}_\Theta$, in that there exists an $A \in \mathcal{I}_\Theta$ such that $\mu(A) = 1$ and $\nu(A) = 0$.

Notice the similarity between Theorem 4.1 with Theorem 3.2.

**Proposition 4.2.** If $\mu \in \text{ex } \mathcal{G}(\gamma) \cap \mathcal{P}_\Theta(\Omega, \mathcal{F})$, then $\mu \in \text{ex } \mathcal{P}_\Theta(\Omega, \mathcal{F})$.

*Proof.* If $\mu \in \text{ex } \mathcal{G}(\gamma)$, then for all cylinder events $A$, $\lim_{n \to \infty} \sup_{B \in \mathcal{T}_n} \mu(A \cap B) = \mu(A \mu(B)) = 0$. Let $A, C$ be two cylinder events. W.l.o.g., consider $A, C \in \mathcal{F}_\Lambda N$, for some $N$. It suffices to prove that:

$$\forall \epsilon > 0, \exists n_0 \in \mathbb{N} : \|i\|_{\infty} \geq n_0 \implies |\mu(A \cap B) - \mu(A)\mu(B)| < \epsilon.$$

Let $m$ be such that $|\mu(A \cap B) - \mu(A)\mu(B)| < \epsilon$, for all $B \in \mathcal{T}_m$. Taking $n_0 = m + (2N + 1)$, we have that if $\|i\|_{\infty} \geq n_0$, then $\theta_i C \in \mathcal{T}_m$, so:

$$\epsilon > |\mu(A \cap \theta_i C) - \mu(A)\mu(\theta_i C)| = |\mu(A \cap \theta_i C) - \mu(A)\mu(C)|,$$

since $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$.

\[ \square \]

**Corollary 2.** $\mu^+$ and $\mu^-$ are strong mixing.

*Proof.* From Ben’s talk we know that:

$$\mu_\Lambda \leq \mu_\Lambda^+ \leq \mu_\Lambda^+$$

Taking the mean $\int \mu(d\eta)$ in the previous equation, we obtain that $\mu_\Lambda^+ \leq \mu \leq \mu_\Lambda^+$ and since stochastic domination is preserved under weak limits, we end up with:

$$\mu^- \leq \mu \leq \mu^+$$
when $\mu$ is any Ising-model Gibbs measure for fixed $\beta$ and $h$. On the one hand, this shows that $\mu^-$ and $\mu^+$ are extremal. If $\mu^+ = s\mu_1 + (1-s)\mu_2$, for $\mu_1 \neq \mu_2$ and $0 < s < 1$, there must exist an increasing event $A$ such that $\mu_1(A) \neq \mu_2(A)$. W.l.o.g., we can assume that $\mu_1(A) < \mu_2(A)$. Then:

$$\mu^+(A) = s\mu_1(A) + (1-s)\mu_2(A) < s\mu_2(A) + (1-s)\mu_2(A) = \mu_2(A),$$

which this is a contradiction with the fact that $\mu_2 \leq \mu^+$. Since $\mu^-$ and $\mu^+$ are also shift-invariant, both must be strong mixing (in particular, ergodic).

**Theorem 4.3** (Aizenman-Higuchi, 1980). For the Ising model on $\mathbb{Z}^2$ with no external field and inverse temperature $\beta > \beta_c$, $\mu^+$ and $\mu^-$ are the only phases, and any other Gibbs measure is a mixture of these two.

**References**
