Free Energy in Spin-flip Processes is Non-increasing

Jean Moulin Ollagnier[1] and Didier Pinchon[2]

[1] Département de Mathématiques, Université Paris Nord, F-93430 Villetaneuse, France
[2] Laboratoire de Probabilités, Université Paris VI, F-75230 Paris Cedex 05, France

Abstract. In this paper, we propose a new and shorter proof of the following fact: in a spin-flip process on $\{-1, +1\}^S$, where S is a countable set, the free energy is non-increasing.

Free energy is a well defined functional only for invariant measures under a convenient group of bijections of S. We formalize this with the notion of B-amenability of S. This frame contains the usual example of Z^d under translations but also many nice lattices that are not groups under groups of isometries.

For invariant measures, except Gibbs ones, the free energy is strictly decreasing. Among invariant measures, the only stationary measures for the spin-flip process are therefore Gibbs measures. From this result we also deduce an ergodic theorem.

The first result on this subject was obtained by Holley [1] for a finite range potential on $\{-1, +1\}^{Z^d}$ and some extension by Higuchi, Shiga [2].

Let S be a countable set, $X = \{-1, +1\}^S$. If x belongs to S, σ_x is the coordinate function at x, τ_x is the homeomorphism of the compact group X defined by: for $\eta \in X$, $\sigma_y \tau_x(\eta) = \sigma_y(\eta)$ if $y \neq x$ and $\sigma_x \tau_x(\eta) = -\sigma_x(\eta)$. G is the group generated by the $(\tau_x, x \in S)$ (finite modifications).

For Λ finite in S and A in Λ, the cylinder (A, Λ) is the subset of X:

$$\{\eta / \sigma_x(\eta) = -1 \text{ if } x \in A \text{ and } \sigma_x(\eta) = +1 \text{ if } x \in \Lambda - A\}.$$

If $x \in \Lambda$, (A_x, Λ) is the cylinder $(A \Delta \{x\}, \Lambda)$.

Let B be a group of bijections of S and in a natural way a group of homeomorphisms of X. In the following, we shall assume the fundamental hypothesis on the action of B:

B is transitive on S.

There is a sequence Λ_n of finite parts of S with the property:

$$\forall x_0 \in S, \forall F \text{ finite } \subset S, \lim_{n \to \infty} |x \in \Lambda_n, \exists b \in B, b(x_0) = x, bF \subset \Lambda_n| / |\Lambda_n| = 1.$$

We say that S is B-amenable and that Λ_n is a B-ameaning sequence.

Examples. $S = Z^d$, $B = Z^d$. In this case a B-ameaning sequence is exactly a sequence A_n that tends to infinity in the sense of Van Hove.

Let S be a lattice (of crystal type) in R^n.

We then consider the group B of its isometries. B acts transitively on S. B is amenable as a finite extension of the abelian subgroup of translations of S.
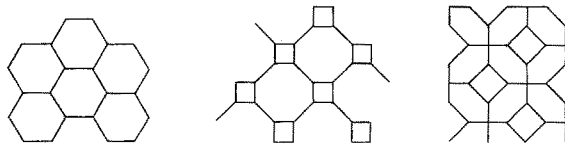
in R^2:

Fig. 1

in R^3: tetraedric lattice, ...

Let $f \in C(X)$; we note $\hat{f}(A)$, A finite $\subset S$, the Fourier coefficient $\hat{f}(A) = \int f \cdot \sigma_A dh$ where h is the Haar measure of the compact group X and σ_A is the character: σ_A

$$= \prod_{x \in A} \sigma_x.$$

The continuous function f is the uniform limit of the partial sums $\sum_{A \subset \Lambda} \hat{f}(A) \sigma_A$ when Λ tends to S (see for instance [4] Proposition II.1.).

If $(V_x, x \in S)$ is a family of continuous functions which generates a G-cocycle on X, i.e. verifies the compatibility conditions:

$$\forall x \in S, \quad V_x + V_x \circ \tau_x = 0,$$

$$\forall x, y \in S, \quad V_x + V_y \circ \tau_x = V_y + V_x \circ \tau_y,$$

then exist numbers $J(A)$ such that: $\hat{V}_x(A) = -2J(A)$ if $x \in A$ and $\hat{V}_x(A) = 0$ if $x \notin A$.

The energy function U_Λ of the finite part Λ is defined by: $U_\Lambda = \sum_{B \subset \Lambda} J(B) \sigma_B$.

The invariance of the potential under the action of B is given by the following coherence property: $\forall b \in B$, $\forall x \in S$, $V_{b(x)} = V_x \circ b^{-1}$.

Consider the operator Ω defined on the space of functions depending on only finitely many coordinates by: $\Omega f = \sum_{x \in S} c_x (f \circ \tau_x - f)$ where V_x generates a B-invariant G-cocycle and $c_x = d_x e^{-\beta/2 \cdot V_x}$; $d_x \in C(X)$ are strictly positive and $d_x = d_x \circ \tau_x$, $d_{b(x)} = d_x \circ b^{-1}$, $\forall x \in S$, $\forall b \in B$; β is the inverse temperature.

We suppose that the closure of the operator Ω is the infinitesimal generator of a positive contractions semi-group $S(t)$ on $C(X)$ with domain $D(\Omega)$.

For a probability μ on X, we use μ_t to denote $S(t)\mu$; μ_t is a weakly continuous function of $t \in \mathbb{R}^+$.

Technical Lemma. *Let $(f_x, x \in S)$ be a B-coherent family of continuous functions on X i.e.: $\forall b \in B$, $\forall x \in S$, $f_{b(x)} = f_x \circ b^{-1}$.*

If S is B-amenable and (A_n) is a B-ameaning sequence then:

$$\lim_{n \to \infty} |A_n|^{-1} \cdot \sum_{x \in A_n} \left[\sup_{A \subset A_n} \left(\sup_{[A, A_n]} f_x - \inf_{[A, A_n]} f_x \right) \right] = 0.$$

Proof. Let x_0 be a fixed element in S; f_{x_0} is uniformly continuous on X, i.e.: $\forall \varepsilon > 0$, $\exists F(\varepsilon)$ finite with:

$$\forall F \text{ finite } \supset F(\varepsilon), \; \sup_{A \subset F \setminus [A,F]} \left(\sup_{[A,F]} f_{x_0} - \inf_{[A,F]} f_{x_0} \right) < \varepsilon.$$

Since the sequence (Λ_n) is B-ameaning:

$$\forall \varepsilon' > 0, \; \exists n_0, \; n \geq n_0 \Rightarrow |x \in \Lambda_n, \exists b \in B, b(x_0) = x, bF(\varepsilon) \subset \Lambda_n|/|\Lambda_n| > 1 - \varepsilon'.$$

For such x, since $f_x = f_{x_0} \circ b^{-1}$ and $\Lambda_n \supset bF(\varepsilon)$ then:

$$\sup_{A \subset \Lambda_n \setminus [A, \Lambda_n]} \left(\sup_{[A, \Lambda_n]} f_x - \inf_{[A, \Lambda_n]} f_x \right) < \varepsilon.$$

Hence:

$$n \geq n_0 \Rightarrow |\Lambda_n|^{-1} \sum_{x \in \Lambda_n} \left[\sup_{A \subset \Lambda_n \setminus [A, \Lambda_n]} \left(\sup_{[A, \Lambda_n]} f_x - \inf_{[A, \Lambda_n]} f_x \right) \right] < \varepsilon + \varepsilon' M$$

with $M = \sup f_{x_0} - \inf f_{x_0}$.   Q.E.D.

Let Λ finite $\subset S$. We denote the entropy $- \sum_{A \subset \Lambda} \mu[A, \Lambda] \mathrm{Log} \mu[A, \Lambda]$ by $H(\Lambda, \mu)$ and let $F(\Lambda, \mu) = |\Lambda|^{-1} [\mu(U_\Lambda) - \beta^{-1} H(\Lambda, \mu)]$.

Let Λ_n be a given B-ameaning sequence; the free energy of μ is the number

$$F(\mu) = \limsup_{n \to \infty} F(\Lambda_n, \mu).$$

Theorem 1. *If $\mu_t = S(t)\mu_0$, $F(\mu_t)$ is non increasing.*

Proof. The proof uses the following lemmas:

Lemma 1. *Let μ a probability on X and Λ finite $\subset S$.*
If $\forall A \subset \Lambda$, $\mu[A, \Lambda] \neq 0$ then $F(\Lambda, S(a)\mu)$ is derivable at 0.
If not $\limsup\limits_{a \to 0} a^{-1} [F(\Lambda, S(a)\mu) - F(\Lambda, \mu)] = -\infty$.

Proof of Lemma 1. The energy part $\mu(U_\Lambda)$ is always derivable with derivative $\mu(\Omega U_\Lambda)$ since $U_\Lambda \in D(\Omega)$.

For the entropy part

$$\limsup_{a \to 0} a^{-1} [-H(\Lambda, \mu_a) + H(\Lambda, \mu)] = \sum_{\substack{A \subset \Lambda \\ \mu[A, \Lambda] \neq 0}} (\mathrm{Log} \mu[A, \Lambda] - 1) \cdot \mu(\Omega 1_{[A, \Lambda]})$$

$$+ \limsup a^{-1} \sum_{\substack{A \subset \Lambda \\ \mu[A, \Lambda] = 0}} \mu_a[A, \Lambda] \mathrm{Log} \mu_a[A, \Lambda]$$

If for every $A \subset \Lambda$ with $\mu[A, \Lambda] = 0$ and for every x in Λ, $\mu[A_x, \Lambda] = 0$, μ equals zero. Therefore there exists $A_0 \subset \Lambda$ with $\mu[A_0, \Lambda] = 0$ and $\dfrac{d}{dt} \mu_t[A_0, \Lambda] > 0$ for $t = 0$.

There exists ε strictly positive such that $\forall t \in {]}0, \varepsilon{[}$, $\mu_a[A_0, \Lambda] > 0$ and $a^{-1} \cdot \mu_a[A_0, \Lambda] \mathrm{Log} \mu_a[A_0, \Lambda] = [\mathrm{Log} \mu_{\theta a}[A_0, \Lambda] + 1] \cdot \mu_{\theta a}(\Omega 1_{[A_0, \Lambda]})$ with $0 < \theta < 1$. $\mu_{\theta a}(\Omega 1_{[A_0, \Lambda]})$ tends to $\dfrac{d}{dt} \mu_t[A_0, \Lambda]$ for $t = 0$ and $\mathrm{Log} \mu_{\theta a}[A_0, \Lambda]$ tends to $-\infty$.

$$\limsup_{\substack{A \subset \Lambda \\ \mu[A,\Lambda]=0}} \sum \leqq \sum_{\substack{A \subset \Lambda \\ \mu[A,\Lambda]=0}} \limsup,$$ every lim sup is non positive and there is at least one which is infinite. This achieves the proof of the Lemma 1.

Notation. For a B-coherent family $(f_x, x \in S)$ of $C(X)$ we denote by $\|f\|$ the common uniform norm of all f_x.

Lemma 2. *We denote by $G(\Lambda, \mu)$ the right derivative at 0 of $F(\Lambda, S(t)\mu)$ even if it is $-\infty$. For every probability μ and every Λ finite in S, there is an uniform upper bound for $G(\Lambda, \mu)$*

$$G(\Lambda, \mu) \leqq K = 2\|V\| \cdot \|c\| + 2\beta^{-1}\|c\| \cdot \|d\| \cdot \left\|\frac{1}{d}\right\|.$$

Proof of Lemma 2. If $G(\Lambda, \mu) = -\infty$ it's clear.

If not, $\mu[A, \Lambda] > 0$ for every cylinder $[A, \Lambda]$ and

$$G(\Lambda, \mu) = |\Lambda|^{-1}\mu(\Omega U_\Lambda) + |\Lambda|^{-1}\beta^{-1}\sum_{A \subset \Lambda}(\mathrm{Log}\,\mu[A,\Lambda] - 1)\mu(\Omega 1_{[A,\Lambda]})$$

$$= |\Lambda|^{-1}\mu(\Omega U_\Lambda) + |\Lambda|^{-1}\beta^{-1}\sum_{A \subset \Lambda}\mathrm{Log}\,\mu[A,\Lambda]\mu(\Omega 1_{[A,\Lambda]}).$$

Let us calculate the energy part

$$|\Lambda|^{-1}\mu(\Omega U_\Lambda) = |\Lambda|^{-1}\sum_{x \in \Lambda}\mu\left(c_x\left(\sum_{A \subset \Lambda}J(A)(\sigma_A \circ \tau_x - \sigma_A)\right)\right)$$

$$= |\Lambda|^{-1}\sum_{x \in \Lambda}\mu\left(c_x\sum_{\substack{A \subset \Lambda \\ A \ni x}}-2J(A)\sigma_A\right) = |\Lambda|^{-1}\beta^{-1}\sum_{x \in \Lambda}\mu(c_x\beta V_x^\Lambda).$$

After obvious calculations on the entropy part

$$G(\Lambda, \mu) = \beta^{-1}|\Lambda|^{-1}\sum_{x \in \Lambda}\sum_{A \subset \Lambda}\mu(c_x 1_{[A,\Lambda]})\left(\beta V_x^\Lambda + \mathrm{Log}\frac{\mu[A_x,\Lambda]}{\mu[A,\Lambda]}\right).$$

On the cylinder $[A, \Lambda]$, βV_x^Λ and $-\mathrm{Log}\dfrac{\mu(c_x 1_{[A,\Lambda]})}{\mu(d_x 1_{[A,\Lambda]})} + \mathrm{Log}\dfrac{\mu(c_x 1_{[A,\Lambda]})}{\mu(d_x 1_{[A_x,\Lambda]})}$ are two

constants between the bounds of βV_x on this cylinder and therefore their difference is uniformly bounded by $2\beta\|V\|$.

$$G(\Lambda, \mu) = \beta^{-1}|\Lambda|^{-1}\sum_{x \in \Lambda}\sum_{A \subset \Lambda}\mu(c_x 1_{[A,\Lambda]})\left(\mathrm{Log}\frac{\mu(c_x 1_{[A_x,\Lambda]})}{\mu(c_x 1_{[A,\Lambda]}}\right.$$

$$\left. + \mathrm{Log}\frac{\mu(d_x 1_{[A,\Lambda]})}{\mu(d_x 1_{[A_x,\Lambda]})} + \mathrm{Log}\frac{\mu[A_x,\Lambda]}{\mu[A,\Lambda]}\right) + T_1(\Lambda, \mu)$$

where the corrective term $T_1(\Lambda, \mu)$ is bounded above by $2\|V\| \cdot \|c\|$. Since d_x satisfies $d_x \circ \tau_x = d_x$

$$\left|\mathrm{Log}\frac{\mu(d_x 1_{[A,\Lambda]})}{\mu[A,\Lambda]} - \mathrm{Log}\frac{\mu(d_x 1_{[A_x,\Lambda]})}{\mu[A_x,\Lambda]}\right|$$

$$\leqq \left\|\frac{1}{d}\right\|\left|\frac{\mu(d_x 1_{[A,\Lambda]})}{\mu[A,\Lambda]} - \frac{\mu(d_x 1_{[A_x,\Lambda]})}{\mu[A_x,\Lambda]}\right| \leqq \left\|\frac{1}{d}\right\| \cdot 2 \cdot \|d\|.$$

Thus

$$G(\Lambda, \mu) = \tfrac{1}{2}\beta^{-1}|\Lambda|^{-1} \sum_{x \in \Lambda} \sum_{A \subset \Lambda} [\mu(c_x 1_{[A,\Lambda]}) - \mu(c_x 1_{[A_x, \Lambda]})] \operatorname{Log} \frac{\mu(c_x 1_{[A_x, \Lambda]})}{\mu(c_x 1_{[A,\Lambda]})}$$
$$+ T_1(\Lambda, \mu) + T_2(\Lambda, \mu).$$

$T_2(\Lambda, \mu)$ is bounded above by $2\beta^{-1} \cdot \|c\| \cdot \|d\| \cdot \left\|\dfrac{1}{d}\right\|$.

All the terms in the first sum have the form $(a - b) \operatorname{Log} b/a$, $a, b > 0$ and therefore are non positive.

Lemma 3. *The upper bound K for the right derivative $G(\Lambda, \mu)$ is also an upper bound for $t^{-1}(F(\Lambda, S(t)\mu) - F(\Lambda, \mu))$, $t > 0$.*
The proof is straightforwards.

Proof of Theorem 1.

$$F(\mu_t) - F(\mu_0) = \lim \sup F(\Lambda_n, \mu_t) - \lim \sup F(\Lambda_n, \mu_0)$$
$$\leq \lim \sup [F(\Lambda_n, \mu_t) - F(\Lambda_n, \mu_0)]$$

and by continuity

$$F(\Lambda_n, \mu_t) - F(\Lambda_n, \mu_0) = \lim_{a \to 0} a^{-1} \cdot \int_t^{t+a} F(\Lambda_n, \mu_u) du - \int_0^a F(\Lambda_n, \mu_u) du$$
$$= \lim_{a \to 0} \int_0^t a^{-1} \cdot [F(\Lambda_n, \mu_{u+a}) - F(\Lambda_n, \mu_u)] du.$$

Because of the upper bound K (Lemma 3), Fatou's lemma may be applied

$$F(\Lambda_n, \mu_t) - F(\Lambda_n, \mu_0) \leq \int_0^t G(\Lambda_n, \mu_u) du.$$

Applying Fatou's lemma one more time

$$F(\mu_t) - F(\mu_0) \leq \lim \sup \int_0^t G(\Lambda_n, \mu_u) du \leq \int_0^t \lim \sup G(\Lambda_n, \mu_u) du.$$

By applying the technical lemma to the B-coherent families $(\beta V_x, x \in S)$ and $(d_x, x \in S)$, the corrective terms $T_1(\Lambda_n, \mu_u)$ and $T_2(\Lambda_n, \mu_u)$ are shown to tend to zero for an ameaning sequence Λ_n.
Thus

$$F(\mu_t) - F(\mu_0) \leq \int_0^t M(\mu_u) du \leq 0 \quad \text{with} \quad M(\mu) = \lim \sup G_0(\Lambda_n, \mu)$$

and

$$G_0(\Lambda, \mu) = \tfrac{1}{2}\beta^{-1}|\Lambda|^{-1} \sum_{x \in \Lambda} \sum_{A \subset \Lambda} [\mu(c_x 1_{[A,\Lambda]}) - \mu(c_x 1_{[A_x, \Lambda]})] \operatorname{Log} \frac{\mu(c_x 1_{[A_x, \Lambda]})}{\mu(c_x 1_{[A,\Lambda]})}.$$

B-invariant Measures

Ω is such that the semi-group $S(t)$ preserves B-invariant measures.

Proposition 1. *Let μ be a B-invariant probability on X with $\mu[A, \Lambda] > 0$, $\forall A \subset \Lambda$ finite. If μ is not G-quasi-invariant with derivative $\exp(-\beta V_x)$ (Gibbs measure for given energy and temperature) then $M(\mu) < 0$.*

Proof. If μ is not τ_{x_0}-quasi-invariant with derivative $\exp(-\beta V_{x_0})$, there exists a cylinder (A, F) with:

$$[\mu(c_{x_0}1_{[A,F]}) - \mu(c_{x_0}1_{[A_{x_0},F]})] \operatorname{Log} \frac{\mu(c_{x_0}1_{[A_x,F]})}{\mu(c_{x_0}1_{[A,F]})} = -a < 0 \,.$$

The function of two variables $(x, y) \in R^{*+2}$, $(x - y) \operatorname{Log} y/x$ is concave. Thus

$$\sum_{A \subset \Lambda} [\mu(c_{x_0}1_{[A,\Lambda]}) - \mu(c_{x_0}1_{[A_{x_0},\Lambda]})] \operatorname{Log} \frac{\mu(c_{x_0}1_{[A_x,\Lambda]})}{\mu(c_{x_0}1_{[A,\Lambda]})}$$

decreases when Λ increases.

Using $c_{b(x)} = c_x \circ b^{-1}$, with the B-ameaning sequence Λ_n:

$$\forall \varepsilon > 0, \exists n_0, n \geqq n_0 \Rightarrow |\Lambda_n|^{-1} \sum_{x \in \Lambda_n} \leqq |\Lambda_n|^{-1} \sum_{x \in \Lambda'_n} \leqq -(1 - \varepsilon)a$$

where

$$\Lambda'_n = \{x \in \Lambda_n / \exists b, b(x_0) = x, bF \subset \Lambda_n\} \,.$$

So $M(\mu) \leqq -a < 0$.

Theorem 2. *Let μ be a B-invariant probability on X.*

If μ is not a Gibbs measure, there exist strictly positive constants a, T and a neighbourhood U of μ such that:

$$\forall v \in U, \forall t \in]0, T[, \; F(v_t) - F(v) \leqq -at \,.$$

Proof. Since μ is not a Gibbs measure, there exist $x_0 \in S$ and $[A, F]$ with

$$\mu(c_{x_0}1_{[A,F]}) \neq \mu(c_{x_0}1_{[A_{x_0},F]}) \,.$$

The set V of all probabilities v that satisfy

$$|v(c_{x_0}1_{[A,F]}) - v(c_{x_0}1_{[A_{x_0},F]})| > 1/2 \cdot |\mu(c_{x_0}1_{[A,F]}) - \mu(c_{x_0}1_{[A_{x_0},F]})| = a_1$$

and

$$\sup(v(c_{x_0}1_{[A,F]}), v(c_{x_0}1_{[A_{x_0},F]})) > 1/2 \cdot \sup(\mu(c_{x_0}1_{[A,F]}), \mu(c_{x_0}1_{[A_{x_0},F]})) = a_2$$

is a neighbourhood of μ.

If v in V is diffuse then:

$$[v(c_{x_0}1_{[A,F]}) - v(c_{x_0}1_{[A_{x_0},F]})] \operatorname{Log} \frac{v(c_{x_0}1_{[A_x,F]})}{v(c_{x_0}1_{[A,F]})} \leqq -\frac{a_1^2}{a_2} = -a \,. \tag{1}$$

Since the application $(v, t) \to S(t)v$ is continuous there exist a neighbourhood U of μ and $T > 0$ such that $\forall v \in U, \forall t \in [0, T[, S(t)v \in V$. Given a measure v in U, and t in $[0, T[$, for almost every u in $[0, t[, v_u$ is diffuse, B-invariant and verifies (1).

The proof of Proposition 1 then shows that: $F(v_t) - F(v) \leqq -at$.

Corollary. *Among B-invariant probabilities the only stationary ones are Gibbs probabilities.*

Another consequence of Theorem 2 is the following ergodic theorem:

Theorem 3. *Let μ be a B-invariant probability.*

Let t_n an increasing sequence of real numbers tending to infinity such that μ_{t_n} admit the weak limit v. Then v is Gibbs.

Proof. If v is not Gibbs, from Theorem 2 there exists a convenient subsequence of μ_{t_n}, called also μ_{t_n}, verifying: $F(\mu_{t_{n+1}}) - F(\mu_{t_n}) \leqq - aT$.

In this case $F(\mu_{t_n})$ would tend to infinity: there is a contradiction with the fact that F is a bounded function.

Remark (cf. [4]). The use of Kieffer's technics [3] shows that for B-invariant probabilities $|A_n|^{-1} H(A_n, \mu)$ has a limit, independent of the chosen ameaning sequence A_n. On the other hand it is possible to prove the existence of a limit for $|A_n|^{-1} \mu(U_{A_n})$, independant of A_n.

References

1. Holley, R.: Free energy in a markovian model of lattice spin systems. Commun. math. Phys. **23**, 87—99 (1971)
2. Higushi, Y., Shiga, T.: Some results on Markov processes of infinite lattice spin systems. J. Math. Kyoto Univ. **15**, 211—229 (1975)
3. Kieffer, J. C.: A generalized Shannon-McMillan theorem for the action of an amenable group on a probability space. Ann. Prob. **3**, 1031 – 1037 (1975)
4. Moulin Ollagnier, J., Pinchon, D.: Fonctions thermodynamiques locales en mécanique statistique. Séminaire sur les processus markoviens a une infinité de particules. Ecole Polytechnique (1976)