Z. Wahrscheinlichkeitstheorie verw. Gebiete
55, 11–23 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Mesures de Gibbs invariantes et mesures d'equilibre

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

[1] 14 Villa d'Este, Appt. 2306, F-75648 Paris Cedex 13, France
[2] 62 Rue Edouard Vaillant, F-94140 Alfortville, France

Abstract. We are interested here in the characterization on a symbolic space, of invariant Gibbs measures as equilibrium measures. The first result in this topic was obtained by Lanford and Ruelle (see for example [6]).

This problem involves different objects that can all be defined by using the only amenability of the translation group and the only continuity of the local specification. We therefore tried to state our theorems in this general frame. Among the elements of our proof, there is the use of the information gain introduced by H. Föllmer [1] and some arguments similar to those of C.J. Preston in [5]. But the amenability techniques that we widely develop in [2], [4] and [7] are decisive tools for getting the result.

The corresponding problem for subshifts is not considered in the present paper so symbolic spaces are product spaces.

I. Préliminaires

Nous rappelons quelques définitions et certains résultats de base sur les mesures de Gibbs et les groupes moyennables.

Mesures de Gibbs

Les objets que nous étudions dans cet article ont été introduits comme des modèles simplifiés de quelques problèmes de mécanique statistique sur un réseau. Les point ou sites du réseau sont représentés par les éléments d'un ensemble dénombrable G. L'ensemble I des configurations possibles en un site donné est fini et cet ensemble est le même pour tous les sites. L'espace X des configurations est l'espace produit I^G; c'est un espace compact métrisable.

0044-3719/81/0055/0011/$02.60

Notations. Soit S une partie de G. Posons $X_S = I^S$ et notons ξ_S l'image d'une configuration ξ par la projection naturelle de X sur X_S.

Si S et T sont deux parties d'intersection vide, si α appartient à X_S et β à X_T, $\alpha\beta$ désigne l'unique élément de $X_{S \cup T}$ dont les projections sur X_S et X_T sont α et β.

Soit ξ un élément de X et A une partie finie de G. $A(\xi)$ désigne l'ensemble des configurations η ayant la même projection que ξ sur le complémentaire de A

$$A(\xi) = \{\eta \in X, \eta_{G-A} = \xi_{G-A}\}.$$

Définition. Une spécification locale π est une famille $\{\pi_A\}$ de fonctions continues strictement positives sur X, indexée par l'ensemble des parties finies de G, qui vérifie les relations de cohérence

$$\pi_{A'}(\xi) = \pi_A(\xi) \cdot \sum_{\eta \in A(\xi)} \pi_{A'}(\eta) \qquad \text{lorsque } A' \text{ contient } A. \tag{1}$$

On obtient comme cas particulier, en faisant $A = A'$, les relations de normalisation.

A chaque fonction π_A on associe la contraction positive π_A sur l'espace de Banach $C(X)$ des fonctions réelles continues sur X ainsi définie

$$\pi_A(f)(\xi) = \sum_{\eta \in A(\xi)} \pi_A(\eta) \cdot f(\eta).$$

Des relations de cohérence entre les fonctions π_A on déduit les relations de commutation entre les opérateurs correspondants

$$\pi_A \circ \pi_{A'} = \pi_{A'} \circ \pi_A = \pi_{A'} \qquad \text{lorsque } A' \text{ contient } A. \tag{2}$$

On remarque en particulier que ces opérateurs sont des projecteurs et que $\pi_\emptyset$ est l'identité.

Définition. On dit d'une probabilité de Radon, μ, sur X que c'est une mesure de Gibbs pour la spécification π si, pour toute fonction continue f sur X et pour toute partie finie A de G, on a $\mu(f) = \mu(\pi_A(f))$.

L'ensemble convexe, faiblement compact K_A des mesures de probabilité sur X vérifiant $\mu \circ \pi_A = \mu$ est l'image par le projecteur π_A^* du convexe compact de toutes les probabilités sur X; cet ensemble n'est donc pas vide. D'après les relations de commutations (2), les K_A ont la propriété d'intersection finie non vide. L'ensemble $\mathscr{G}(\pi)$ des mesures de Gibbs pour π n'est donc pas vide. C'est un convexe compact.

Groupes moyennables

La notion d'action moyennable permet de définir des valeurs moyennes lorsque le réseau est homogène sous l'action d'un groupe géométrique; cette hypothèse est raisonnable car elle permet de modéliser la limite thermodynamique. Pour simplifier l'exposition des résultats, nous ne considérons que l'action d'un groupe moyennable dénombrable G par translations sur l'espace $X=I^G$. On peut trouver dans [2] la définition générale d'un réseau moyennable; les résultats que nous montrons ici sont encore valables dans ce cadre.

Les groupes moyennables peuvent être caractérisés par plusieurs propriétés équivalentes. La propriété la plus utile ici est celle de point fixe: toute action de G par homéomorphismes affines sur un convexe compact d'un espace vectoriel topologique localement convexe laisse fixe au moins un point de ce convexe compact.

E. Følner a démontré un résultat essentiel: l'équivalence entre la propriété de point fixe et l'existence sur l'ensemble des parties finies du groupe d'un filtre particulier, le filtre moyennant. De manière plus précise, il établit que pour toute partie finie D du groupe G et pour tout réel positif ε, l'ensemble $M(D,\varepsilon)$ des parties invariantes à ε près par les éléments de D est non vide.

$$A\in M(D,\varepsilon)\Leftrightarrow|\{x\in A, \forall d\in D\, dx\in A\}|>(1-\varepsilon)\cdot|A|.$$

Les ensembles $M(D,\varepsilon)$ constituent une base du filtre moyennant $\mathfrak{M}$.

Le long du filtre moyennant, les valeurs moyennes de certaines fonctions d'ensembles convergent. On obtient en particulier le théorème suivant dont on peut trouver la démonstration dans [7].

Théorème 0. *Soit f une fonction à valeurs réelles définie sur l'ensemble des parties finies du groupe moyennable G vérifiant les trois conditions suivantes*

a) $f(\emptyset)=0$

b) *f est invariante à droite: pour toute partie finie A et tout élément a de G, $f(Aa)=f(A)$*

c) *f est complètement sous-additive: pour toute décomposition à coefficients positifs, $1_A=\sum\lambda_i\cdot 1_{A_i}$, on a l'inégalité $f(A)\leqq\sum\lambda_i f(A_i)$.*

Alors, la fonction de $\mathscr{F}(G)$ dans $\mathbb{R}$, $A\to|A|^{-1}\cdot f(A)$, admet une limite selon le filtre moyennant, égale à sa borne inférieure.

La condition de forte sous-additivité d) est plus forte que la condition c) de sous-additivité complète.

d) *pour tout couple (A,B) de parties finies de G, $f(A\cup B)+f(A\cap B)\leqq f(A)+f(B)$*

Le résultat du théorème est donc vrai pour les fonctions fortement sous-additives invariantes à droite; nous l'utilisons dans la suite.

Translations sur X

Dans la suite, le réseau G est un groupe moyennable. A un élément g du groupe on fait correspondre un homéomorphisme de l'espace X de la façon suivante

$$T_g(\xi)_a=\xi_{ag}. \tag{3}$$

Les éléments de X sont des applications de G dans I. T_g est alors la translation à droite du graphe

$$\xi_a = i \Leftrightarrow T_g(\xi)_{ag^{-1}} = i$$

ce qui explique la définition précédente.

On obtient alors une représentation de groupe G comme un groupe d'homéomorphismes de X, que l'on appelle le groupe des translations.

Pression d'une fonction continue

Soit f une fonction continue sur X et A une partie finie du groupe G. On pose

$$P(f, A) = \operatorname{Log} \sum_{\eta \in X_A} \sup_{\xi \in [\eta]} \exp \sum_{a \in A} f \circ T_a(\xi) \tag{4}$$

où $[\eta]$ est l'ensemble de toutes les configurations ξ telles que $\xi_A = \eta$. Les ensembles de ce type, qui sont à la fois ouverts et fermés dans X, sont appelés cylindres.

La fonction $A \to P(f, A)$ est invariante à droite et complètement sous-additive, si bien que $|A|^{-1} \cdot P(f, A)$ admet une limite selon le filtre moyennant $\mathfrak{M}$, d'après le théorème 0. Ce nombre $p(f)$ est appelé la pression de la fonction continue f.

Entropie moyenne d'une probabilité invariante

Soit μ une mesure de probabilité sur X, invariante par translations. Les cylindres correspondants aux éléments de X_A constituent une partition de X; on note $H(\mu, A)$ l'entropie de cette partition

$$H(\mu, A) = \sum_{\eta \in X_A} -\mu([\eta]) \cdot \operatorname{Log} \mu([\eta]).$$

La fonction $A \to H(\mu, A)$ est invariante à droite et fortement sous-additive. On appelle alors entropie moyenne de la probabilité invariante μ le nombre $h(\mu)$, qui est la borne inférieure et la limite selon le filtre moyennant de $|A|^{-1} \cdot H(\mu, A)$.

L'entropie moyenne h est une fonction semi-continue supérieurement sur le compact $\mathscr{M}(X, G)$ des probabilités invariantes.

Principe variationnel

Le principe variationnel est le théorème d'après lequel la pression $p(f)$ d'une fonction continue f est la borne supérieure des sommes $h(\mu) + \mu(f)$ où μ décrit l'ensemble des probabilités invariantes sur X.

La nature particulière du système dynamique (X, G) permet les définitions simplifiées que nous donnons de la pression et de l'entropie moyenne. Les cylindres en e constituent tout à la fois un recouvrement ouvert et une partition; ce recouvrement est séparateur et, pour toute probabilité invariante, cette partition est un générateur au sens des systèmes dynamiques abstraits.

Dans le cadre général d'un groupe moyennable d'homéomorphismes d'un espace compact, les définitions sont plus compliquées; il est cependant possible de démontrer le principe variationnel [3].

Nous démontrons le principe variationnel dans le cas particulier qui nous intéresse ici, au début de la partie III.

On appelle mesures d'équilibre pour f les probabilités qui vérifient

$$h(\mu)+\mu(f)=p(f).$$

Dans le cas des espaces symboliques, il y a toujours des mesures d'équilibre, puisque la fonction s.c.s. $\mu \to h(\mu)+\mu(f)$ atteint son maximum sur $\mathcal{M}(X, G)$.

Dans la partie II, nous forgeons des outils utiles à la démonstration du résultat principal:

Etant donnée une spécification locale invariante π, les mesures de Gibbs pour π, invariantes par translations, sont exactement les mesures d'équilibre pour une certaine fonction continue construite à partir de π.

II. Quelques résultats asymptotiques sur les spécifications locales invariantes

Rappelons que nous considérons un espace $X=I^G$, où I est un ensemble fini et G un groupe moyennable qui agit sur X par les homéomorphismes que nous appelons translations.

Définition. Une spécification locale π est dite invariante si, pour tout élément g de G et toute partie A de G, on a l'égalité des fonctions $\pi_A=\pi_{Ag^{-1}}\circ T_g$, c'est-à-dire que le nombre $\pi_A(\xi)$ est invariant si l'on applique simultanément la translation à droite par g^{-1} à la partie A et l'homéomorphisme T_g à ξ.

Pour de telles spécifications, les mesures de Gibbs ne sont pas nécessairement invariantes; mais l'action affine de G par dualité conserve le convexe des mesures de Gibbs $\mathcal{G}(\pi)$, de sorte qu'il existe des mesures de Gibbs invariantes puisque le groupe est moyennable.

Energie moyenne pour une spécification locale invariante

Soit π une spécification locale invariante. Pour une partie finie A de G, on définit la fonction $U(A, \cdot)$ qui représente l'énergie dans la partie A de la façon suivante (la spécification π est sous-entendue)

$$U(A, \xi)=\operatorname{Log}\pi_A(\xi)-|A(\xi)|^{-1}\sum_{\eta\in A(\xi)}\operatorname{Log}\pi_A(\eta). \tag{5}$$

On appelle habituellement fonctions de partition les nombres $Z(A)$ ainsi définis (où, une fois encore, la spécification π est sous-entendue)

$$Z(A)=\sum_{\eta\in X_A}\sup_{\xi\in[\eta]}\exp U(A,\xi). \tag{6}$$

C'est la normalisation additive dans la définition de $U(A,\cdot)$ à partir de $\operatorname{Log}\pi_A$ qui rend possible les résultats asymptotiques qui font l'objet de cette seconde partie

$$\sum_{\eta\in A(\xi)} U(A,\eta)=0. \tag{7}$$

Lorsque A est une partie finie de G et a un élément de G n'appartenant pas à A, on note $I_a(A,\cdot)$ l'accroissement de l'énergie

$$I_a(A,\xi)=U(A\cup a,\xi)-U(A,\xi). \tag{8}$$

La proposition suivante est un résultat important, qui permet la localisation de l'énergie.

Proposition 1. *La fonction $(A,\xi)\to I_a(A,\xi)$ est uniformément continue, la structure uniforme étant le produit de l'unique structure uniforme sur l'espace compact X par la structure induite sur l'ensemble des parties finies de $G\setminus a$ par celle du compact $\mathscr{P}(G\setminus a)$.*

Démonstration. Par définition des fonctions U, on a

$$U(A,\xi)=-|A(\xi)|^{-1}\sum_{\eta\in X_A}\operatorname{Log}\pi_A(\eta\cdot\xi_{G\setminus A})-\operatorname{Log}\pi_A(\xi)$$

$$U(A\cup a,\xi)=-|I|^{-1}|A(\xi)|^{-1}\sum_{\alpha\in X_a}\sum_{\eta\in X_A}\operatorname{Log}\pi_{A\cup a}(\alpha\cdot\eta\cdot\xi_{G\setminus A\cup a})-\operatorname{Log}\pi_{A\cup a}(\xi)$$

d'où une expression de $I_a(A,\xi)$ par différence.

En utilisant quatre fois les relations de cohérence (1), on obtient

$$\pi_{A\cup a}(\alpha\cdot\eta\cdot\xi_{G\setminus A\cup a})=\pi_a(\alpha\cdot\eta\cdot\xi_{G\setminus A\cup a})\cdot\sum_{\beta\in X_A}\pi_{A\cup a}(\beta\cdot\eta\cdot\xi_{G\setminus A\cup a})$$

$$\pi_{A\cup a}(\xi)=\pi_A(\xi)\cdot\sum_{\omega\in X_A}\pi_{A\cup a}(\omega\cdot\xi_{G\setminus A})$$

$$\pi_{A\cup a}(\eta\cdot\xi_{G\setminus A})=\pi_a(\eta\cdot\xi_{G\setminus A})\cdot\sum_{\beta\in X_a}\pi_{A\cup a}(\beta\cdot\eta\cdot\xi_{G\setminus A\cup a})$$

$$\pi_{A\cup a}(\eta\cdot\xi_{G\setminus A})=\pi_A(\eta\cdot\xi_{G\setminus A})\cdot\sum_{\omega\in X_A}\pi_{A\cup a}(\omega\cdot\xi_{G\setminus A})$$

puis, par un quotient approprié

$$\frac{\pi_{A\cup a}(\alpha\cdot\eta\cdot\xi_{G\setminus A\cup a})}{\pi_{A\cup a}(\xi)}=\frac{\pi_A(\eta\cdot\xi_{G\setminus A})}{\pi_A(\xi)}\cdot\frac{\pi_a(\alpha\cdot\eta\cdot\xi_{G\setminus A\cup a})}{\pi_a(\eta\cdot\xi_{G\setminus A})}.$$

On en déduit une expression de la fonction $I_a(A,\cdot)$ à l'aide de la fonction π_a.

$$I_a(A,\xi)=-|I|^{-1}\cdot\sum_{\alpha\in X_a}|A(\xi)|^{-1}\sum_{\eta\in X_A}\operatorname{Log}\frac{\pi_a(\alpha\cdot\eta\cdot\xi_{G\setminus A\cup a})}{\pi_a(\eta\cdot\xi_{G\setminus A})}.$$

Il suffit alors, pour conclure, de montrer que si f est une fonction continue sur X, la fonction de deux variables $(A, \xi) \to f(A, \xi)$ est uniformément continue au sens de l'énoncé

$$f(A, \xi) = |A(\xi)|^{-1} \sum_{\eta \in X_A} f(\eta \cdot \xi_{G \setminus A})$$

et d'appliquer ce résultat à l'ensemble fini des fonctions f_α, où α appartient à X_a

$$f_\alpha(\xi) = \operatorname{Log} \pi_a(\alpha \cdot \xi_{G \setminus a}) - \operatorname{Log} \pi_a(\xi)$$

Ceci provient de l'uniforme continuité de f sur X; il est facile de vérifier que si B est une partie finie de G telle que $\xi_B = \xi'_B \Rightarrow |f(\xi) - f(\xi')| < \varepsilon$ et si les parties finies A et A' vérifient $A \cap B = A' \cap B$ on a $|f(A, \xi) - f(A', \xi)| < 2\varepsilon$. C.Q.F.D.

Il est donc possible et d'une manière unique d'étendre la fonction I_a en une fonction continue, encore notée I_a, sur le produit $\mathscr{P}(G \setminus a) \times X$. L'invariance de la spécification conduit à la relation

$$I_a(A, \xi) = I_{ag^{-1}}(Ag^{-1}, T_g(\xi)). \tag{9}$$

Nous sommes à la recherche d'une fonction continue φ sur X telle que les fonctions $|A|^{-1} \cdot U(A, \cdot)$ et $|A|^{-1} \sum_{a \in A} \varphi \circ T_a$ soient asymptotiquement les mêmes. Il est possible de construire φ à partir de l'accroissement I_e de l'énergie, où e est l'élément neutre du groupe G, en utilisant les ordres totaux sur G.

Rappelons, sans démonstration, quelques résultats sur l'ensemble T des ordres totaux sur un groupe moyennable G. On pourra, par exemple, trouver les détails dans [4].

L'ensemble T des ordres totaux sur G peut être muni d'une topologie d'espace compact métrisable (puisque G est dénombrable) sur lequel G agit par des homéomorphismes, les translations à droite. Comme dans le cas de l'espace symbolique X, on effectue les translations sur les graphes.

On désigne par a_τ^- l'ensemble des éléments du groupe G strictement antérieurs à l'élément a dans l'ordre τ. Si P est une probabilité de Radon sur T, invariante par translations, la formule suivante définit une fonction continue sur T.

$$\varphi_P(\xi) = \int_T I_e(e_\tau^-, \xi)\, dP(\tau). \tag{10}$$

Ceci résulte de la continuité de I_e sur le produit $\mathscr{P}(G \setminus e) \times X$ et de la continuité de l'application de T dans $\mathscr{P}(G \setminus e)$, $\tau \to e_\tau^-$.

L'invariance à droite de P et la relation (9) conduisent à

$$\begin{aligned} \varphi_P \circ \mathrm{T}_a(\xi) &= \int_T I_e(e_\tau^-, T_a(\xi))\, dP(\tau) = \int_T I_a(e_\tau^- \cdot a, \xi)\, dP(\tau) \\ &= \int_T I_a(a_\tau^-, \xi)\, dP(\tau). \end{aligned} \tag{11}$$

Remarque. Il n'est pas possible d'obtenir n'importe quelle fonction continue sur X de cette façon à partir d'une spécification locale invariante.

Le théorème suivant est un premier résultat asymptotique.

Théorème 1. *Soit P une probabilité invariante sur T et φ_P la fonction continue sur X définie par la formule* (10) *à partir de P et de la fonction I_e, accroissement de l'énergie. La fonction continue $|A|^{-1}\cdot(U(A,\cdot)-\sum_{a\in A}\varphi_P\circ T_a)$ tend uniformément vers* 0 *selon le filtre moyennant.*

Démonstration. En considérant la trace d'un ordre total τ sur la partie finie A de G, et en convenant que $U(\emptyset,\cdot)$ est la fonction nulle, on obtient

$$U(A\cdot\xi)=\sum_{a\in A} I_a(A\cap a_\tau^-,\xi)$$

et l'on peut intégrer cette égalité pour la probabilité P, ce qui donne

$$U(A,\xi)=\sum_{a\in A}\int_T I_a(A\cap a_\tau^-,\xi)\,dP(\tau).$$

En utilisant l'uniforme continuité de I_e et la relation (9), on peut trouver pour chaque réel positif ε une partie finie F de G telle que

$$|I_a(A\cap a_\tau^-,\xi)-I_a(a_\tau^-,\xi)|<\varepsilon$$

dès que A contient Fa.
D'où la majoration

$$\|A|^{-1}\cdot U(A,\cdot)-\sum_{a\in A}\varphi_P\circ T_a|\leqq\varepsilon+2|A|^{-1}\cdot\|I_e\|\cdot|\{a\in A, Fa\not\subset A\}|.$$

La norme uniforme de la fonction considérée est donc inférieure à 3ε dès que A appartient à $M(F\cup e,\varepsilon/\|I_e\|)$. Ceci achève la démonstration du théorème.

On déduit du théorème 1 que $|A|^{-1}\mu(U(A,\cdot))$ admet une limite selon le filtre moyennant lorsque μ est une probabilité sur X invariante par les translations. Cette limite est égale à $\mu(\varphi_P)$; ce nombre ne dépend donc pas de la probabilité choisie sur l'espace des ordres totaux. On appelle énergie moyenne de la probabilité μ et on note $e(\mu)$ cette grandeur (une fois encore, la spécification est sous-entendue dans cette notation).

Les résultats asymptotiques dont nous avons besoin sont des corollaires du théorème 1 et de la proposition suivante, qui est montrée dans [2] (p. 30, "technical lemma").

Proposition 2. *Soit f une fonction continue sur $X=I^G$, où le groupe G est moyennable; la valeur moyenne $|A|^{-1}\cdot\sum_{a\in A}f\circ T_a$ dépend essentiellement des coordonnées dans A lorsque la partie finie A devient arbitrairement moyennante, c'est-à-dire que l'on a*

$$\lim_{\mathfrak{M}}\sup_{\xi_A=\xi_A}|A|^{-1}\cdot|\sum_{a\in A}f\circ T_a(\xi)-\sum_{a\in A}f\circ T_a(\xi)|=0.$$

Démonstration. Puisque f est uniformément continue, pour tout réel positif ε il existe une partie finie F de G telle que $\xi_F=\zeta_F\Rightarrow|f(\xi)-f(\zeta)|<\varepsilon$. On en déduit $\xi_{Fa}=\zeta_{Fa}\Rightarrow|f\circ T_a(\xi)-f\circ T_a(\zeta)|<\varepsilon$. Ceci permet de majorer l'expression

$$|A|^{-1}\cdot\sup_{\xi_A=\zeta_A}|\sum_{a\in A}(f\circ T_a(\xi)-f\circ T_a(\zeta))|$$

par

$$\varepsilon+2|A|^{-1}\cdot\|f\|\cdot|\{a\in A, Fa\not\subset A\}|$$

ce qui est inférieur à 3ε dès que A appartient à $M(F, \varepsilon/\|f\|)$. C.Q.F.D.

On déduit immédiatement du théorème 1 et de la proposition 2 un corollaire d'après lequel l'énergie moyenne dans une partie finie dépend essentiellement des coordonnées dans cette partie.

Corollaire 1. $\lim_{\mathfrak{M}}(\sup_{\xi_A=\zeta_A}|A|^{-1}\cdot|U(A,\xi)-U(A,\zeta)|)=0.$

On utilise le contrôle des variations de $U(A,\cdot)$ et de $\sum\varphi_P\circ T_a$ pour contrôler celles de logarithmes de sommes d'exponentielles dans les deux corollaires qui suivent.

Corollaire 2. $\lim_{\mathfrak{M}}(|A|^{-1}\cdot\sup_{\xi}(|\operatorname{Log}Z(A)-\operatorname{Log}\sum_{\eta\in A(\xi)}\exp U(A,\eta)|))=0.$

Corollaire 3. $\lim_{\mathfrak{M}}(|A|^{-1}\cdot|\operatorname{Log}Z(A)-\operatorname{Log}\sum_{\eta\in X_A}\sup_{\xi\in[\eta]}\exp(\sum_{a\in A}\varphi_P\circ T_a(\xi)|)=0.$

On déduit simplement des deux corollaires précédents le corollaire qui suit

Corollaire 4. $\lim_{\mathfrak{M}}(|A|^{-1}\cdot\operatorname{Log}Z(A))=p(\varphi_P).$

On remarque au passage que la pression de φ_P ne dépend pas de la probabilité invariante P sur les ordres totaux que l'on a choisi pour construire une fonction continue à partir de la fonction I_e.

Gain d'information moyen relativement à une spécification locale invariante

Soit π une spécification locale invariante et ν une mesure de Gibbs pour π. Soit A une partie finie de G et η un élément de X_A. Puisque ν est une mesure de Gibbs pour π, on a $\nu([\eta])=\nu(\pi_A(1_{[\eta]}))$; or, la fonction $\pi_A(1_{[\eta]})$ est strictement positive; donc, la mesure de Gibbs ν donne au cylindre $[\eta]$ une masse strictement positive, ce qui permet la définition qui suit.

Définition. Etant données une probabilité μ sur X et une partie finie A de G, on appelle gain d'information de μ relativement à la mesure de Gibbs ν dans A et on note $I_A(\mu,\nu)$ le nombre

$$I_A(\mu,\nu)=\sum_{\eta\in X_A}\mu([\eta])\cdot\operatorname{Log}\frac{\mu([\eta])}{\nu([\eta])}.$$

De la concavité du logarithme, on déduit que $I_A(\mu,\nu)$ est une fonction croissante de la partie finie A.

De la stricte concavité du logarithme, on déduit que le nombre positif $I_A(\mu,\nu)$ est nul si et seulement si les projections sur X_A des mesures μ et ν coïncident.

D'autre part, il est clair que si les deux probabilités sont invariantes par les translations, la fonction $I_A(\mu, \nu)$ de A est invariante à droite.

Le théorème suivant est essentiel à notre démonstration; c'est une relation asymptotique entre les objets définis à l'aide de la spécification locale invariante et le gain d'information relativement à une mesure de Gibbs pour π.

Théorème 2. *Soit π une spécification locale invariante, μ une probabilité sur X, ν une mesure de Gibbs pour π. Soit $\Delta(A)$ l'expression suivante*

$$\Delta(A) = |A|^{-1} \cdot (\operatorname{Log} Z(A) - (U(A, \cdot)) - H(\mu, A) - I_A(\mu, \nu))$$

$\Delta(A)$ tend vers 0 selon le filtre moyennant.

Démonstration. Calculons $\Delta(A)$ en faisant intervenir les cylindres correspondant aux éléments de X_A

$$\Delta(A) = |A|^{-1} \cdot \sum_{\eta \in X_A} \mu([\eta]) \cdot \left\{ \operatorname{Log} Z(A) - 1/\mu([\eta]) \right.$$

$$\left. \cdot \int_{[\eta]} U(A, \xi)\, d\mu(\xi) + \operatorname{Log} \mu([\eta]) + \operatorname{Log} \frac{\nu([\eta])}{\mu([\eta])} \right\}.$$

Puisque $1/\mu([\eta]) \cdot \int_{[\eta]} U(A, \xi)\, d\mu(\xi)$ est une valeur moyenne sur un cylindre, on peut majorer $\Delta(A)$ en valeur absolue par

$$|A|^{-1} \cdot \sup_{\eta \in X_A} \sup_{\xi \in [\eta]} |\operatorname{Log} Z(A) + \operatorname{Log} \nu([\eta]) - U(A, \xi)|.$$

En utilisant le corollaire 2, il reste à montrer pour conclure que l'expression suivante a une limite nulle selon le filtre moyennant.

$$|A|^{-1} \cdot \sup_{\eta \in X_A} \sup_{\xi \in [\eta]} |\operatorname{Log} \sum_{\zeta \in A(\xi)} \exp U(A, \zeta) - U(A, \xi) + \operatorname{Log} \nu([\eta])|.$$

Puisque ν est une mesure de Gibbs, la masse d'un cylindre peut se calculer ainsi

$$\nu([\eta]) = \int \pi_A(\eta \cdot \xi_{G \setminus A})\, d\nu(\xi).$$

La relation de normalisation, conséquence de (1), permet de retrouver $\pi_A(\xi)$ à partir de l'énergie $U(A, \xi)$.

$$\pi_A(\xi) = \exp U(A, \xi) / \sum_{\zeta \in A(\xi)} \exp U(A, \zeta).$$

On obtient alors

$$\sup_{\xi \in [\eta]} |\operatorname{Log} \nu([\eta]) - \operatorname{Log} \pi_A(\xi)| \leqq \sup_{\xi_A = \zeta_A} |\operatorname{Log} \pi_A(\xi) - \operatorname{Log} \pi_A(\zeta)|$$

ce que l'on peut encore majorer, en regardant séparément numérateurs et dénominateurs, par

$$2 \sup_{\xi_A = \zeta_A} |U(A, \xi) - U(A, \zeta)|.$$

On achève la démonstration du théorème 2 en utilisant le corollaire 1.

De tous ces résultats, on déduit l'énoncé suivant.

Proposition 3. *Soit π une spécification locale invariante, ν une mesure de Gibbs invariante pour π, μ une probabilité invariante sur X. La moyenne du gain d'information $|A|^{-1} \cdot I_A(\mu, \nu)$ admet une limite selon le filtre moyennant, qui est égale à la différence $p(\varphi_P) - \mu(\varphi_P) - h(\mu)$; ce nombre est donc indèpendant de la probabilité P sur T et de la mesure de Gibbs invariante ν. On le note $i(\mu, \pi)$ et on l'appelle gain d'information moyen de la mesure invariante μ par rapport à la spécification locale invariante π.*

La démonstration est immédiate.

III. Le résultat principal

Puisque le gain d'information moyen est positif ou nul et qu'il est nul pour les mesures de Gibbs invariantes, on déduit de la proposition 3 le principe variationnel pour la fonction continue φ_P. On en conclut aussi que les mesures de Gibbs invariantes pour π sont des mesures d'équilibre pour φ_P. Il reste à montrer que ce sont les seules; c'est l'objet du théorème suivant.

Théorème 3. *Une mesure d'équilibre μ pour une spécification locale π, c'est-à-dire une probabilité invariante vérifiant $i(\mu, \pi) = 0$, est une mesure de Gibbs pour π.*

Démonstration. Elle se compose de trois arguments essentiels que nous distingons sous la forme de lemmes.

1) On établit d'abord une caractérisation de la propriété de Gibbs à l'aide de l'entropie conditionnelle, le principe variationnel local.

2) On montre ensuite que les nombres intervenant dans le lemme précédent sont les limites des accroissements du gain d'information.

3) Un résultat de pure théorie des groupes permet enfin d'obtenir des propriétés locales à partir de propriétés moyennes.

Lemme 1. *Une mesure de probabilité μ sur X vérifie $\mu \circ \pi_A = \mu$ si et seulement si $-\mu(\operatorname{Log} \pi_A) - H_\mu(A | G \setminus A) = 0$.*

Démonstration. Si l'on désigne par $\mu_A(\eta, \zeta)$ une version mesurable de la probabilité conditionnelle de η dans X_A étant donné ζ dans $X_{G \setminus A}$, on peut écrire l'entropie conditionnelle et l'intégrale de $\operatorname{Log} \pi_A$ à l'aide de la mesure image $\bar{\mu}$ sur $X_{G \setminus A}$.

$$H_\mu(A | G \setminus A) = \int_{X_{G \setminus A}} \sum_{\eta \in X_A} -\mu_A(\eta, \zeta) \cdot \operatorname{Log} \mu_A(\eta, \zeta) \, d\bar{\mu}(\zeta) \tag{12}$$

$$\mu(\operatorname{Log} \pi_A) = \int_{X_{G \setminus A}} \Big[\sum_{\eta \in X_A} \operatorname{Log} \pi_A(\eta \cdot \zeta)\Big] \, d\bar{\mu}(\zeta). \tag{13}$$

L'expression qui nous intéresse est donc égale à l'intégrale d'un gain d'information

$$-\mu(\operatorname{Log} \pi_A) - H_\mu(A | G \setminus A) = \int_{X_{G \setminus A}} \left[\sum_{\eta \in X_A} \mu_A(\eta, \zeta) \cdot \operatorname{Log} \frac{\mu_A(\eta, \zeta)}{\pi_A(\eta \cdot \zeta)}\right] d\bar{\mu}(\zeta).$$

Elle est donc positive ou nulle; elle n'est nulle que si les vecteurs de probabilité $\mu_A(\eta, \zeta)$ et $\pi_A(\eta \cdot \zeta)$ coincident pour $\bar{\mu}$-presque tout ζ. Il est alors clair que si π_A est une version de la probabilité conditionnelle, on a $\mu \circ \pi_A = \mu$.

Lemme 2. *Etant données une mesure de probabilité μ et une mesure de Gibbs ν pour la spécification locale π, l'accroissement du gain d'information $I_{A \cup B}(\mu, \nu) - I_B(\mu, \nu)$ admet une limite lorsque la partie B de $G \setminus A$ tend vers $G \setminus A$, et cette limite est égale à $-\mu(\mathrm{Log}\,\pi_A) - H_\mu(A|G \setminus A)$.*

Démonstration. Un calcul simple de l'accroissement de gain d'information conduit à

$$I_{A \cup B}(\mu, \nu) - I_B(\mu, \nu) = -H_\mu(A|B) - \mu\left(\mathrm{Log}\frac{\nu([\xi_{A \cup B}])}{\nu([\xi_B])}\right).$$

Puisque ν est une mesure de Gibbs pour π, le logarithme à intégrer converge uniformément vers $\mathrm{Log}\,\pi_A$ lorsque B tend vers $G \setminus A$; il est d'autre part bien connu que l'entropie conditionnelle $H_\mu(A|B)$ tend vers $H_\mu(A|G \setminus A)$. D'où le résultat annoncé.

D'après les deux lemmes précédents, une mesure de probabilité μ est de Gibbs pour la spécification locale π si et seulement si, pour toute partie finie A de G, la limite lorsque la partie finie B tend vers $G \setminus A$ de l'accroissement du gain d'information $I_{A \cup B}(\mu, \nu) - I_B(\mu, \nu)$, qui existe toujours, est nulle.

Résumons ce que nous savons sur la fonction $A \to I_A(\mu, \nu)$ lorsque μ est une mesure d'équilibre pour φ_P et ν une mesure de Gibbs invariante pour π.

i) Cette fonction est croissante, invariante à droite, et prend la valeur 0 en $\emptyset$.

ii) La valeur moyenne $|A|^{-1} \cdot I_A(\mu, \nu)$ tend vers 0 selon le filtre moyennant.

iii) Pour toute partie finie A, l'accroissement $I_{A \cup B} - I_A$ admet une limite quand B tend vers $G \setminus A$.

Pour conclure, nous voulons montrer que ces limites sont nulles. Par une transformation de Möbius, la propriété iii) est équivalente à l'existence d'une limite, pour toute couple (a, A) où a appartient à A, lorsque B tend vers $G \setminus A$, pour l'accroissement $I_{a \cup B} - I_B$. Si ces limites sont nulles, on obtient par sommation le résultat.

Il suffit donc, pour aboutir, d'appliquer le lemme suivant.

Lemme 3. *Soit f une fonction croissante, invariante à droite, nulle en $\emptyset$, sur l'ensemble des parties finies d'un groupe moyennable G; on suppose de plus que la valeur moyenne $|A|^{-1} \cdot f(A)$ tend vers 0 selon le filtre moyennant.*

Alors, pour tout couple (a, A) où a appartient à A, la limite inférieure, quand B tend vers $G \setminus A$, de l'accroissement $f(a \cup B) - f(B)$ tend vers 0.

Démonstration. Raisonnons par contraposition et supposons que la limite inférieure de $f(a \cup B) - f(B)$ ne soit pas nulle. Il existe donc un réel strictement positif ε et une partie finie A' de $G \setminus A$ tels que, pour toute partie B de $G \setminus A$ contenent A', on ait

$$f(a \cup B) - f(B) \geqq \varepsilon$$

Désignons par C' l'ensemble des éléments x d'une partie finie C du groupe G tels que $(A \cup A') \cdot a^{-1} \cdot x$ soit contenu dans C. Le rapport des cardinaux $|C'|/|C|$ tend vers 1 selon le filtre moyennant.

Etant donné un ordre total τ sur C, on peut décomposer $f(C)$ de la façon suivante

$$f(C) = \sum_{x \in C} f(x \cup x_\tau^-) - f(x_\tau^-)$$

En faisant la somme de ces expressions pour tous les ordres totaux sur C et en divisant par leur nombre $|C|!$, on obtient la minoration de $f(C)$ suivante.

$$f(C) \geqq \sum_{x \in C'} 1/|C|! \sum_{\tau \in T(x)} f(x \cup x_\tau^-) - f(x_\tau^-)$$

où, lorsque x appartient à C', $T(x)$ désigne l'ensemble des ordres totaux sur C tels que, pour tout élément b de A, $b \cdot a^{-1} \cdot x$ soit supérieur à x, et, pour tout élément c de A', x soit supérieur à $c \cdot a^{-1} \cdot x$. L'inégalité est une conséquence de la croissance de f.

De l'invariance à droite de f on déduit $f(x \cup x_\tau^-) - f(x_\tau^-) = f(a \cup B) - f(B)$, où $B = x_\tau^- \cdot x^{-1} \cdot a$ Cet accroissement est donc minoré par ε quand l'ordre τ appartient à $T(x)$.

Or, le nombre des éléments de $T(x)$ est une proportion constante de $|C|!$.

$$|T(x)| = |C|! \frac{(|A|-1)!\,(|A'|)!}{(|A|+|A'|)!}.$$

La limite inférieure de $|C|^{-1} \cdot f(C)$ n'est donc pas nulle, ce qui contredit l'hypothèse.

Bibliographie

1. Föllmer, H.: On entropy and information gain in random fields. Z. Wahrscheinlichkeitstheorie verw. Gebiete **26**, 207–217 (1973)
2. Moulin Ollagnier, J., Pinchon, D.: Free energy in spin-flip processes is non-increasing. Comm. Math. Phys. **55**, 29–35 (1977)
3. Moulin Ollagnier, J., Pinchon, D.: The variational principle. A paraître dans Studia Mathematica (1980)
4. Moulin Ollagnier, J., Pinchon, D.: Une nouvelle démonstration du théorème de Følner. C.R. Acad. Sci. Paris Sér. A-Bt. **287**, 557–560 (1978)
5. Preston, C.J.: Random fields. Lecture Notes in Math. n° **534**. Berlin-Heidelberg-New York: Springer 1976
6. Ruelle, D.: Thermodynamic formalism. Encyclopedia of Mathematics n° 5. Roading, Mass.: Addison-Wesley 1979
7. Moulin Ollagnier, J., Pinchon, D.: Filtre moyennant et valeur moyenne des capacités invariantes. A paraître. (1979)

Reçu le 30 Novembre 1978; en forme finale le 1er Juillet 1980