JEAN-PIERRE CONZE

Le théorème d'isomorphisme d'Ornstein et la classification des systèmes dynamiques en théorie ergodique

Séminaire N. Bourbaki, 1974, exp. nº 420, p. 50-68

<http://www.numdam.org/item?id=SB_1972-1973__15__50_0>

LE THÉORÈME D'ISOMORPHISME D'ORNSTEIN ET LA CLASSIFICATION DES SYSTÈMES DYNAMIQUES EN THÉORIE ERGODIQUE

par Jean-Pierre CONZE

Introduction, systèmes dynamiques

Un système dynamique est, ici, la donnée d'un espace X , d'une tribu $\mathfrak{a}$ de parties de X , d'une mesure de probabilité m sur $\mathfrak{a}$ et d'une transformation T de X inversible, $\mathfrak{a}$-mesurable ainsi que son inverse, conservant la mesure m .

Pour éviter les difficultés liées à la structure de l'espace mesuré $(X, \mathfrak{a}, m)$, nous supposerons que cet espace est isomorphe à l'espace mesuré formé par l'intervalle $[0,1]$ muni de la tribu des ensembles mesurables au sens de Lebesgue, et de la mesure de Lebesgue. Dans toute la suite de l'exposé, les ensembles considérés seront supposés mesurables, ou construits comme tels. D'autre part, nous sous-entendrons souvent l'expression " à un ensemble de mesure nulle près ".

La notion de système dynamique et les origines de la Théorie Ergodique sont liées, comme on le sait, à l'étude du système formé par n particules en Mécanique. En effet, sur l'espace X des configurations de ce système, la transformation T qui fait passer de la configuration à l'instant 0 à la configuration à l'instant 1 laisse invariante la mesure induite sur X par la mesure de Lebesgue de $\mathbb{R}^{6n}$.

Après des relations étroites avec la Mécanique, la Théorie Ergodique a été influencée, à une époque récente, par la Théorie de l'Information. Des méthodes nouvelles, qui en sont issues, ont permis d'aborder l'étude de la structure des systèmes dynamiques. La définition par Kolmogorov et Sinaï, en 1958, de l'entropie comme invariant des systèmes dynamiques, a été suivie par les travaux de Sinaï, Rokhlin, Pinsker,..., puis par les progrès décisifs accomplis par Ornstein dans les questions d'isomorphisme de systèmes.

Ce sont ces résultats d'Ornstein, principalement l'isomorphisme des schémas de Bernoulli de même entropie, que nous voudrions exposer ici. Outre les articles d'Ornstein cités en références, nous avons utilisé la présentation qu'en a donnée Smorodinsky dans [14].

2. Définitions, Schémas de Bernoulli et K-systèmes

2.1. Isomorphisme, ergodicité, théorème de Birkhoff

Deux systèmes dynamiques $(X, \mathfrak{a}, m, T)$ et $(X', \mathfrak{a}', m', T')$ sont *isomorphes* s'il existe un isomorphisme φ de l'espace mesuré $(X, \mathfrak{a}, m)$ sur l'espace mesuré $(X', \mathfrak{a}', m')$ tel que $\varphi \circ T = T' \circ \varphi$.

Dans le cas où φ est seulement un homomorphisme de $(X, \mathfrak{a}, m)$ sur $(X', \mathfrak{a}', m')$ tel que $\varphi \circ T = T' \circ \varphi$, nous dirons que $(X', \mathfrak{a}', m', T')$ est un *facteur* de $(X, \mathfrak{a}, m, T)$. Remarquons que la donnée d'un facteur de $(X, \mathfrak{a}, m, T)$ est simplement, à un isomorphisme près, la donnée d'une sous-tribu de $\mathfrak{a}$, invariante par T .

Un système dynamique $(X, \mathfrak{a}, m, T)$ est *ergodique* si toute fonction (mesurable) sur X invariante par T est constante (presque partout).

Rappelons encore le *théorème de Birkhoff* :

Soient $(X, \mathfrak{a}, m, T)$ un système dynamique, et f une fonction dans $L^1(m)$. Alors, pour presque tout x , la limite de $\frac{1}{n} \sum_{i=1}^{n-1} f(T^i x)$ existe, et, si le système est ergodique, est égale à $m(f)$.

2.2. Partitions, indépendance et ε-indépendance

Soit $(X,\mathcal{Q},m,T)$ un système dynamique. Une partition P de X est une famille finie ordonnée de sous-ensembles ($\mathcal{Q}$-mesurables) de X, disjoints deux à deux et de réunion X.

Etant données deux partitions P et Q de X, leur borne supérieure est la partition, ordonnée par l'ordre lexicographique,

$$P \vee Q = \{P_i \cap Q_j , P_i \in P , Q_j \in Q\} .$$

Si les éléments de la partition P sont réunion d'éléments de la partition Q, on dit que P est contenue dans Q, et on note $P \subset Q$.

On note également :

$$T^k P = \{T^k P_i , P_i \in P\} , \qquad k \in \mathbb{Z} ,$$

$$\bigvee_{-m}^{n} T^i P = T^{-m}P \vee \ldots \vee T^{-1}P \vee P \vee TP \vee \ldots \vee T^n P , \qquad \text{pour } n , m \geq 0 .$$

La distribution d'une partition $P = (P_1,\ldots,P_k)$ est le vecteur de probabilité $\quad \operatorname{dist} P = (m(P_1),\ldots,m(P_k))$.

Deux partitions P et Q sont indépendantes, si l'on a

$$m(P_i \cap Q_j) = m(P_i)m(Q_j) , \quad \text{quels que soient } P_i \in P , Q_j \in Q .$$

Supposons que tous les éléments de Q aient une mesure non nulle. Sur chaque élément Q_j de Q, la trace de la partition P a pour distribution le vecteur $\quad (m(P_i \cap Q_j)/m(Q_j) , i = 1,\ldots,k)$.

Si ces distributions coïncident avec celle de P, les partitions P et Q sont indépendantes. Si elles diffèrent peu de celle de P, nous avons une forme faible de l'indépendance, précisée dans la définition suivante :

La partition P est ε-indépendante de la partition $Q = (Q_1,\ldots,Q_r)$ si l'on a $\sum_{i=1}^{k} |m(P_i \cap Q_j)/m(Q_j) - m(P_i)| < \varepsilon$, pour tous les éléments Q_j de Q

n'appartenant pas à un ensemble de mesure au plus ε .

2.3. Dépendance entre partitions

Soient P et Q deux partitions. Quitte à les compléter par des ensembles négligeables, nous pouvons supposer qu'elles ont le même nombre d'éléments. La distance entre P et Q est le nombre

$$|P - Q| = \sum_i (m(P_i \cap Q_i^c) + m(P_i^c \cap Q_i)) .$$

On démontre facilement que, pour cette distance, les partitions de X forment un espace métrique complet.

A l'opposé de la notion d' ε-indépendance, nous avons une notion d' ε-inclusion. Une partition P est ε-contenue dans une partition Q , s'il existe une partition Q' telle que $Q' \subset Q$ et $|P - Q'| < \varepsilon$. On note cette relation $P \overset{\varepsilon}{\subset} Q$.

2.4. Interprétation, exemple, partition génératrice

Supposons que le système dynamique $(X,\mathcal{A},m,T)$ représente un système physique dont l'évolution est décrite par l'action de la transformation T sur X . Une mesure, au sens physique, de notre système peut être considérée comme une application de X dans un espace que nous supposerons fini pour simplifier, soit encore comme une partition P de X . Faire une série d'observations à des temps entiers, revient à déterminer la suite des éléments de la partition P dans lesquels se trouvent les images de x par les itérées de T . Cette suite réalise un "codage" du système, qui le détermine plus ou moins suivant le choix de la partition P .

Dans le meilleur des cas, la partition P engendre le système dans le sens suivant. Désignons par $(P)_T$ la plus petite tribu complète contenant les éléments des partitions $\bigvee_{-n}^{n} T^i P$, pour tout n . La partition P est génératrice si la tribu $(P)_T$ est égale à $\mathcal{A}$. Ceci équivaut à supposer que, sauf pour les points d'un ensemble négligeable dans X , deux points distincts ont des codages par P différents.

En théorie de l'Information et en Calcul des Probabilités, beaucoup de systèmes sont munis par construction d'un codage naturel. C'est en particulier le cas des Schémas de Bernoulli.

La méthode du codage, ou représentation symbolique, a été également utilisée dans l'étude des systèmes dynamiques classiques, en particulier par Morse dans l'étude du flot géodésique sur les variétés de courbure négative. Donnons encore l'exemple géométrique du billard triangulaire.

Considérons une particule se déplaçant à l'intérieur d'un triangle, avec réflexions sur les parois. Le système dynamique associé a pour espace des états X l'ensemble des couples (x,ξ) , où x est un point sur un côté du triangle différent des sommets, et ξ un vecteur unitaire en x orienté vers l'intérieur du triangle. La transformation T est définie sur X par réflexion sur les parois du triangle (T n'est pas définie sur l'ensemble négligeable des couples (x,ξ) qui aboutiraient à un sommet). Elle laisse invariante une mesure équivalente à la mesure induite sur X par la mesure de Lebesgue.

Un codage est obtenu en observant les incidences sur les côtés du triangle. Si c_1 , c_2 , c_3 sont les côtés du triangle, ce codage correspond à la partition P de X dont les éléments sont $P_i = \{(x,\xi) \in X$ tels que $x \in c_i\}$, $i = 1,2,3$. Il donne des renseignements intéressants sur les propriétés ergodiques du "billard".

2.5. <u>Distribution d'une suite</u>, <u>distribution des</u> P-n <u>noms</u>

Soit $K = (k_1,\dots,k_n)$ une suite d'entiers à valeurs dans $\{1,\dots,k\}$. La <u>distribution de</u> K est le vecteur de probabilité dist K de dimension k , dont les composantes sont $\frac{1}{n}$ Card $\{j : k_j = i\}$, $i = 1,\dots,k$.

Soient $(X,\mathcal{A},m,T)$ un système dynamique, et P une partition de X .

Nous appellerons P-n nom d'un point $x \in X$ la suite $(k_1, \ldots, k_n)$, telle que

$$x \in T^{-j}P_{k_j} \quad , \qquad j = 0, \ldots, n-1 \ .$$

D'après le théorème de Birkhoff appliqué à la fonction caractéristique de chaque ensemble P_i , pour presque tout x dans X , la distribution du P-n nom de x converge vers la distribution de P , quand n tend vers l'infini.

Soit alors C un élément de la partition $\bigvee_0^{n-1} T^{-j}P$. Son P-n nom (= le P-n nom de ses points) est la suite $(k_1, \ldots, k_n)$ telle que $C = \bigcap_0^{n-1} T^{-j}P_{k_j}$.

D'après ce qui précède, pour tout $\varepsilon > 0$, il existe un entier n tel que la distribution des P-n noms de chaque élément de $\bigvee_0^{n-1} T^jP$ soit égale à ε près à la distribution de P , sauf pour un ensemble d'éléments de mesure totale au plus ε .

Ce résultat montre comment, dans l'étude d'un système physique, la répartition statistique des mesures effectuées au cours d'une série d'observations permet de déterminer la structure du système.

2.6. Schémas de Bernoulli

Soient $\{a_1, \ldots, a_k\}$ un alphabet formé de k lettres, et $\Pi = (\Pi_1, \ldots, \Pi_k)$, $\Pi_i \geq 0$, $i = 1, \ldots, k$, $\sum_{i=1}^{k} \Pi_i = 1$; un vecteur de probabilité. Considérons l'espace des suites bilatères à valeurs dans $\{a_1, \ldots, a_k\}$. On note x_n , la coordonnée d'indice n d'un point x de X . La transformation T est définie sur X par $(Tx)_n = x_{n+1}$, $n \in \mathbb{Z}$, $x \in X$.

L'application de X dans $\{a_1, \ldots, a_k\}$ définie par $x \mapsto x_0$ détermine une partition P de X en les ensembles $P_i = \{x \in X : x_0 = a_i\}$, $i = 1, \ldots, k$.

Rappelons qu'il existe sur la tribu borélienne de X une unique mesure m invariante par T telle que

$$\text{dist } P = \Pi \ ,$$

et, pour tout $n > 0$, T^nP est indépendante de $\bigvee_0^{n-1} T^iP$. Pour obtenir cette mesure, il suffit de poser

$$m\{x \in X : x_{i_1} = a_{i_1}, \dots, x_{i_n} = a_{i_n}\} = m(T^{i_1}P_{i_1} \cap \dots \cap T^{i_n}P_{i_n}) = \Pi_{i_1} \dots \Pi_{i_n} ,$$

pour tous les choix d'indices $i_1, \dots, i_n$, et d'éléments $P_{i_1}, \dots, P_{i_n}$ dans P, et de prolonger m par le théorème de Kolmogorov à la tribu borélienne.

Soit $\mathcal{A}$ la tribu complétée pour la mesure m de la tribu borélienne. A tout vecteur de probabilité Π, on a ainsi associé un système dynamique $(X, \mathcal{A}, m, T)$, appelé schéma de Bernoulli de distribution Π, dont on vérifie qu'il est caractérisé par les propriétés suivantes :

Il existe dans X une partition P telle que

(i) pour tout $n > 0$, T^nP est indépendante de $\bigvee_0^{n-1} T^iP$,

(ii) P est une partition génératrice,

(iii) dist $P = \Pi$.

Dans le cas général, une partition P vérifiant la condition (i) sera dite de Bernoulli.

2.7. K-systèmes et systèmes déterministes

Proches des schémas de Bernoulli, les K-systèmes ont été introduits par Kolmogorov dans l'étude de certains processus stationnaires.

Un système dynamique $(X, \mathcal{A}, m, T)$ est un K-système s'il existe dans X une partition P génératrice telle que, pour tout $k > 0$ et tout $\varepsilon > 0$, il existe n tel que $\bigvee_n^{n+m} T^{-i}P$ soit ε-indépendant de $\bigvee_1^k T^iP$, pour tout $m \geq 0$.

La tribu engendrée par les éléments des partitions T^iP, $i < 0$, (resp. $i > 0$), représente, si l'on interprète l'action de T comme celle du temps, le "passé" (resp. l' "avenir") associé à la partition P dans le système dynamique. La propriété de K-système traduit une condition d'indépendance asymptotique du passé et de l'avenir. Les schémas de Bernoulli, pour lesquels cette indépendance est exactement réalisée, d'après la condition (i) du n° 2.6. sont des K-systèmes. Dans des travaux récents [11,13], il a été prouvé que certains systèmes physiques sont des K-systèmes. Il serait important de savoir si ces systèmes sont, ou non, des schémas de Bernoulli.

A l'opposé des K-systèmes, les systèmes déterministes sont ceux pour lesquels le passé détermine entièrement l'avenir. Plus précisément, un système $(X,\mathfrak{a},m,T)$ est <u>déterministe</u>, si, pour toute partition P, la tribu $(P)_T^-$ engendrée par des partitions T^iP, $i < 0$, contient P, donc est égale à $(P)_T$.

On montre facilement que, dans ce cas, la tribu $(P)_T^{-\infty}$, plus grande tribu contenue dans les tribus $T^{-n}(P)_T^-$, pour tout n, est égale à $(P)_T$. Cette tribu $(P)_T^{-\infty}$ s'interprète comme un "passé éloigné" pour la partition P. Ainsi, dans un système déterministe, le passé, et même le passé éloigné, déterminent l'avenir.

Pinsker a montré que tout système dynamique possède un plus grand facteur déterministe. A la suite des travaux de Pinsker, il était tentant de conjecturer que tout système dynamique se factorise en un produit d'un système déterministe et d'un K-système. D'autre part, on pensait alors pouvoir montrer que les K-systèmes sont isomorphes à des schémas de Bernoulli. Nous verrons plus loin comment Ornstein a répondu par la négative à ces deux conjectures.

3. <u>Entropie des systèmes dynamiques, théorème de Mac-Millan</u>

3.1. <u>Entropie d'une partition</u>

L'<u>entropie</u> d'un vecteur de probabilité $\Pi = (\Pi_1,\ldots,\Pi_k)$ est le nombre

$$H(\Pi) = -\sum_{i=1}^{k} \Pi_i \log \Pi_i \ .$$

Traditionnellement, le logarithme est pris en base 2.

L'<u>entropie d'une partition</u> $P = (P_1,\ldots,P_k)$ dans un espace mesuré

$(X,\mathfrak{A},m)$ est l'entropie de sa distribution, soit

$$H(P) = -\sum_{i=1}^{k} m(P_i) \log m(P_i) .$$

Etant données deux partitions P et Q, l'<u>entropie conditionnelle</u> de P par rapport à Q est la moyenne des entropies des distributions de P sur les éléments de Q, soit

$$H(P/Q) = -\sum_{j,i} m(P_i \cap Q_j) \log (m(P_i \cap Q_j)/m(Q_j)) .$$

Il est clair que $H(P/Q) = 0$ si et seulement si, sur chaque élément Q_j de Q, la trace de P est réduite à Q_j, c'est-à-dire si $P \subset Q$. A l'opposé, pour une partition P donnée, $H(P/Q)$ prend sa valeur maximale quand Q est indépendante de P. Dans ce cas, on a $H(P/Q) = H(P)$. On peut donc dire que $H(P) - H(P/Q)$ mesure l'indépendance de P et Q.

<u>Lemme</u>.- Pour tout $\varepsilon > 0$, il existe $\delta = \delta(\varepsilon) > 0$ tel que, si $H(P) - H(P/Q) < \delta$, P est ε-indépendante de Q.

3.2. <u>Entropie</u> $H(P,T)$; <u>entropie d'un système dynamique</u>

Soient $(X,\mathfrak{A},m,T)$ un système dynamique, et P une partition de X. Il résulte facilement des propriétés de l'entropie que la limite

$$\lim_n \frac{1}{n} H\Big(\bigvee_{0}^{n-1} T^{i}P\Big) = \lim_n H\Big(P/\bigvee_{1}^{n-1} T^{-j}P\Big) \text{ existe.}$$

On pose $H(P,T) = \lim_n \frac{1}{n} H\big(\bigvee_{0}^{n-1} T^{-i}P\big)$, et on appelle <u>entropie du système</u> le nombre $h(T) = \sup_P H(P,T)$, où P décrit l'ensemble des partitions de X.

On peut donner des systèmes déterministes et des K-systèmes une caractérisation en termes d'entropie.

L'entropie $H\big(P/\bigvee_{1}^{n-1} T^{-i}P\big)$ tend vers 0 si et seulement si, pour tout $\varepsilon > 0$, il existe un n tel que P soit ε-contenue dans $\bigvee_{1}^{n-1} T^{-i}P$. Il en résulte que $H(P,T) = 0$ si et seulement si P est contenu dans la tribu $(P)_T^-$ du passé de P. Les systèmes d'entropie $h(T)$ nulle sont donc les systèmes déterministes.

D'autre part, d'après un résultat dû à Rokhlin et Sinaï, un système dynamique, tel que $h(T) < \infty$, est un K-système si et seulement si tous ses facteurs non triviaux sont d'entropie > 0.

3.3. L'entropie comme invariant, application aux schémas de Bernoulli

D'après la définition même de l'entropie, deux systèmes dynamiques isomorphes ont la même entropie. L'entropie d'un facteur est inférieure ou égale à l'entropie du système initial. Le résultat suivant, dû à Kolmogorov, permet dans de nombreux cas le calcul explicite de l'entropie.

Théorème.- Soit $(X,\mathfrak{a},m,T)$ un système dynamique. Si P est une partition génératrice, on a $h(T) = H(P,T)$.

En particulier, l'entropie d'un schéma de Bernoulli construit sur un vecteur de probabilité Π est égale à $H(\Pi)$. Il existe donc des schémas de Bernoulli deux à deux non isomorphes. Ce résultat n'était pas connu avant la définition de l'entropie. Dix ans plus tard, Ornstein a montré que l'entropie est un invariant complet dans la classe des schémas de Bernoulli.

3.4. Le théorème de Mac-Millan

Le théorème de Mac-Millan est fondamental en Théorie de l'Information, dans les problèmes de codages. Il permet de classer les suites de longueur n émises par une source stationnaire ergodique, pour n suffisamment grand, en deux classes complémentaires. L'une de grande mesure rassemble des suites de probabilités élevées et voisines, l'autre est de petite mesure.

En terme de partitions, il s'énonce ainsi : soient $(X,\mathfrak{a},m,T)$ un système ergodique, et P une partition de X. Soit $C_n(x)$ l'élément de la partition $\bigvee_0^{n-1} T^i P$ contenant x. Alors, pour presque tout x, $-\frac{1}{n}\log m(C_n(x))$ converge vers $h(P,T)$.

On en déduit que, pour tout $\varepsilon > 0$, il existe un entier n et des éléments $C_1,\dots,C_\ell$ dans la partition $\bigvee_0^{n-1} T^i P$, tels que l'on ait

$$2^{-n(H(P,T)+\varepsilon)} < m(C_i) < 2^{-n(H(P,T)-\varepsilon)} \quad , \ i = 1,\dots,\ell \ ,$$

et $m(\bigcup_{i=1}^{\ell} C_i) > 1 - \varepsilon$.

On remarque que ces conditions impliquent que $\ell < 2^{n(H(P,T)+\varepsilon)}$.

4. Isomorphisme des schémas de Bernoulli

4.1. Enoncé des résultats

Nous en venons maintenant au théorème d'Ornstein : Deux schémas de Bernoulli de même entropie sont isomorphes [4].

En 1962, Sinaï [12] avait montré qu'étant donnés un système dynamique ergodique $(X,\mathfrak{a},m,T)$ d'entropie $h(T)$ et un vecteur de probabilité Π tel que $h(T) \geq H(\Pi)$, il existe un schéma de Bernoulli de distribution Π facteur de $(X,\mathfrak{a},m,T)$. Ce résultat est obtenu, par une méthode différente, au cours de la démonstration du théorème d'Ornstein.

Nous allons donner les idées de la démonstration, en détaillant plus celle du théorème de Sinaï. Dans toute la suite, $(X,\mathfrak{a},m,T)$ sera un système dynamique ergodique fixé, et les partitions considérées seront prises dans X .

4.2. Lemme

Soient Π un vecteur de probabilité, P une partition à k éléments et $\varepsilon > 0$ un nombre donné. Il existe des entiers n et r , tels que, si $K_1 = (k_1^1,\dots,k_n^1)$,..., $K_r = (k_1^r,\dots,k_n^r)$ sont r suites d'entiers à valeurs dans $\{1,\dots,k\}$, deux à deux distinctes, telles que $|\text{dist}\, K_i - \Pi| < \varepsilon$, $i = 1,\dots,r$, alors il existe une partition R vérifiant $|\text{dist}\, R - \Pi| < 2\varepsilon$, et

$H(R,T) > H(P,T) - \varepsilon$.

Démonstration. Nous allons préciser le rôle des suites K_i dans la construction de R . D'après le théorème de Mac-Millan, δ étant donné, il existe un entier n et des éléments $C_1,\dots,C_\ell$ dans la partition $\bigvee_0^{n-1} T^i P$, tels que l'on ait $\ell < 2^{n(H(P,T)+\delta)}$, et $m(A) > 1 - \delta$, où $A = \bigcup_{j=1}^{\ell} C_j$. Le nombre δ sera choisi par la suite, en fonction de ε , et on suppose que l'on a $r \geq \ell$.

Considérons un sous-ensemble F tel que $F, T^{-1}F, \dots, T^{-n+1}F$ soient disjoints, et $m(B) > 1 - \delta$, où $B = \bigcup_0^{n-1} T^{-i}F$. L'existence de l'ensemble F est assurée par un lemme de Rokhlin, classique en théorie ergodique. On peut supposer de plus que l'on a $m(F \cap A^c) < \delta/n$.

Les ensembles $T^i(F \cap C_j)$, $i = 0,\dots,n-1$, $j = 1,\dots,\ell$, recouvrent B à δ près. On définit la partition R sur B en choisissant $R_1,\dots,R_k$ de la façon suivante : R_s est formé des ensembles $T^{-i}(F \cap C_j)$ tels que $k_i^j = s$, $s = 1,\dots,k$. Pour chaque j , les ensembles $T^{-i}(F \cap C_j)$, $i = 1,\dots,n$, ont la même mesure. La distribution de la partition R , là où elle est définie, reproduit donc celles des suites K_i . Elle diffère de la distribution de Π de moins de ε . Nous complétons la définition de R en ajoutant à R_1 l'ensemble de mesure inférieure à 2δ sur lequel elle n'a pas été définie.

Ainsi, pour δ assez petit, on a construit une partition R vérifiant $|\mathrm{dist}\, R - \Pi| < 2\varepsilon$. Les suites K_i étant deux à deux distinctes, la partition $\bigvee_0^{n-1} T^i(R \vee \{F,F^c\})$ contient les ensembles $F \cap C_j$, $j = 1,\dots,\ell$. Il en résulte que la partition P est ε'-contenue dans $\bigvee_{-n+1}^{n-1} T^i(R \vee \{F,F^c\})$, où ε' dépend de δ . On a donc $H(R \vee \{F,F^c\},T) = H(\bigvee_{-n+1}^{n-1} T^i(R \vee \{F,F^c\},T) > H(P,T) - \varepsilon/2$, si δ

a été choisi assez petit. Enfin, l'entropie $H(R \vee \{F,F^c\},T)$ est inférieure à $H(R,T) + H(\{F,F^c\})$, et comme $H(\{F,F^c\}) \sim -[\frac{1}{n} \log \frac{1}{n} + (1 - \frac{1}{n}) \log(1 - \frac{1}{n})]$ peut être rendu inférieur à $\varepsilon/2$ pour n assez grand, on a bien $H(R,T) > H(P,T) - \varepsilon$, pour un choix convenable de δ et de n .

4.3. Lemme d'approximation

Soient Π un vecteur de probabilité et P une partition tels que

(i) $H(\Pi) = h(T)$,

(ii) $|\text{dist } P - \Pi| < \frac{1}{2}\,\delta(\varepsilon^2/3)$, (δ étant la fonction du lemme 3.1),

(iii) $H(P) - H(P,T) < \frac{1}{2}\,\delta(\varepsilon^2/3)$.

Alors, pour tout nombre $\delta > 0$, il existe une partition P' telle que

(1) $|\text{dist } P' - \Pi| < \delta$

(2) $H(P') - H(P',T) < \delta$,

(3) $|P' - P| < \varepsilon$.

Schéma de la démonstration. Oublions provisoirement la partition P , et cherchons à construire une partition P' vérifiant les conditions (1) et (2) du lemme. Pour δ petit, une telle partition est presque une partition de Bernoulli, puisque d'après l'inégalité $H(P') - H(P',T) > H(P') - H(P'/\bigvee_1^n T^{-i}P')$, pour $n > 0$, P' est ε-indépendante de $\bigvee_1^n T^{-i}P'$, pour tout $n > 0$, si $\delta = \delta(\varepsilon)$ comme dans le lemme 3.1.

Le lemme 4.2 donne un procédé pour construire P' . Considérons une partition Q telle que $H(Q,T) > h(T) - \delta/3$. Appliquons le lemme 4.2 à la partition Q , et à des suites $K_1,\ldots,K_r$ à valeurs dans $1,\ldots,k$, de distributions proches de Π et deux à deux distinctes. (On construit ces suites en appliquant au schéma de Bernoulli de distribution Π' , où Π' est un vecteur de probabilité d'entropie légèrement supérieure à celle de Π , le théorème de Birkhoff, cf. 2.5, et le théorème de Mac-Millan.) On obtient une partition P' de distribution proche de celle de Π , et vérifiant

$$H(P') - H(P',T) = (H(P') - H(\Pi)) + (h(T) - H(Q,T)) + (H(Q,T) - H(P',T)) < \delta \ .$$

La difficulté dans le lemme 4.3 est d'obtenir une partition P' proche de la partition P donnée. Ceci est réalisé par un choix convenable des suites K_i utilisant le "lemme des mariages", et l'application du lemme 4.2.

4.4. Démonstration du théorème de Sinaï

En construisant par le lemme 4.3 une suite convergente de partitions réalisant de mieux en mieux les conditions de Bernoulli, on montre l'existence d'une partition de Bernoulli au voisinage d'une partition qui est presque de Bernoulli :

Si P est une partition et Π un vecteur de probabilité tels que

(i) $H(P) - H(P,T) < \frac{1}{2}\,\delta(\varepsilon^2/3)$,

(ii) $|\text{dist } P - \Pi| < \frac{1}{2}\,\delta(\varepsilon^2/3)$,

(iii) $h(T) = H(\Pi)$,

alors, il existe une partition P' telle que

(1) $\text{dist } P' = \Pi$,

(2) P' est une partition de Bernoulli,

(3) $|P - P'| < 8\varepsilon$.

En particulier, soit ρ un vecteur de probabilité tel que $H(\rho) \leq h(T)$. Considérons une partition Q telle que $H(Q,T) = H(\rho)$. En appliquant ce qui précède au système $(X , (Q)_T , m , T)$, on obtient une partition de Bernoulli R de distribution ρ , ce qui constitue le résultat de Sinaï.

4.5. Démonstration du théorème d'isomorphisme

Soit $(X,\mathfrak{a},m,T)$ un schéma de Bernoulli de distribution Π. Soit ρ un vecteur de probabilité tel que $H(\rho) = H(\Pi)$. Nous savons par le théorème de Sinaï qu'il existe une partition de Bernoulli R dans X de distribution ρ . Le problème est maintenant d'obtenir R génératrice, ce qui prouvera l'isomorphisme entre les schémas de Bernoulli de distributions Π et ρ (le "codage" du premier système associé à la partition R réalisant l'isomorphisme cherché).

Le résultat essentiel dans cette deuxième partie de la démonstration est le

Lemme.- Soit R une partition de Bernoulli de distribution ρ . Soit P une partition de Bernoulli de distribution Π génératrice dans X . Pour tout $\varepsilon > 0$, il existe une partition de Bernoulli R' de distribution ρ et un entier n tels que $|R' - R| < \varepsilon$ et $P \overset{\varepsilon}{\subset} \bigvee_{-n}^{n} T^i R'$.

A l'aide du lemme, on peut alors construire une partition R génératrice de la façon suivante. Soit R_o une partition de Bernoulli de distribution ρ donnée par le théorème de Sinaï. Si P est la partition de Bernoulli de distribution Π génératrice dans $(X,\mathfrak{a},m,T)$, étant donné $\varepsilon_1 > 0$, il existe, d'après le lemme, un entier n_1 et une partition de Bernoulli R_1 de distribution ρ telle que $|R_o - R_1| < \varepsilon_1$, et $P \overset{\varepsilon_1}{\subset} \bigvee_{k=-n_1}^{n_1} T^k R_1$. Si R_2 est une partition très proche de R , la relation d'inclusion précédente est peu modifiée. En appliquant à nouveau le lemme, on peut donc trouver un nombre $\varepsilon_2 > 0$, un entier n_2 et une partition R_2 de distribution ρ telle que $|R_1 - R_2| < \varepsilon_2$,

$P \overset{\varepsilon_2}{\subset} \bigvee_{k=-n_2}^{n_2} T^k R_2$, et $P \overset{(1+\frac{1}{2})\varepsilon_1}{\subset} \bigvee_{k=-n_1}^{n_1} T^k R_2$. De même, pour un choix convenable

de ε_3 , il existe une partition de Bernoulli R de distribution ρ et un entier n_3 tels que $|R_2 - R_3| < \varepsilon_3$, $P \overset{\varepsilon_3}{\subset} \bigvee_{k=-n_3}^{n_3} T^k R_3$, $P \overset{(1+\frac{1}{2})\varepsilon_2}{\subset} \bigvee_{k=-n_2}^{n_2} T^k R_3$,

$$P \overset{(1+\frac{1}{2}+1/2^2)\varepsilon_1}{\subset} \bigvee_{k=-n_1}^{n_1} T^k R_3 .$$

Par cette méthode, on obtient une suite de partition de Bernoulli R_i de distribution ρ, et des suites de nombres $\varepsilon_i > 0$ et d'entiers n_i tels que $|R_i - R_{i+1}| < \varepsilon_{i+1}$ et $P \overset{2\varepsilon_j}{\subset} \bigvee_{k=-n_j}^{n_j} T^k R_i$, pour $i \geq j$. Comme on a pu choisir la série (ε_i) convergente, et comme l'espace des partitions est complet, la suite (R_i) converge vers une partition de Bernoulli R de distribution ρ telle que $P \overset{2\varepsilon_j}{\subset} \bigvee_{k=-n_j}^{n_j} T^k R$, pour tout j. On a donc $P \subset (R)_T$, d'où $(P)_T = (R)_T$, et P, étant génératrice, R est également génératrice.

5. Nouveaux développements

5.1. Recherche de schémas de Bernoulli

Après la démonstration du théorème d'isomorphisme, Ornstein et Friedman et Ornstein ont introduit les notions suivantes (weak Bernoulli, very weak Bernoulli) [2,7].

Une partition P dans un système dynamique est faiblement de Bernoulli, si pour tout $\varepsilon > 0$, il existe un n tel que $\bigvee_{n}^{n+k} T^i P$ soit ε-indépendante de $\bigvee_{-k}^{0} T^i P$, pour tout $k > 0$.

La propriété de faible Bernoulli est plus faible que celle de Bernoulli, et plus forte que celle de K-système définie en 2.7.

Si $\{P^i\}_1^n = \{P^1,\dots,P^n\}$ et $\{\bar{P}^i\}_1^n = \{\bar{P}^1,\dots,\bar{P}^n\}$ sont deux suites de partitions (éventuellement sur des espaces différents), on définit leur distance $\bar{d}$ par

$$\bar{d}(\{P^i\}_1^n, \{\bar{P}^i\}_1^n) = \inf \frac{1}{n} \sum_{i=1}^{n} |Q^i - \bar{Q}^i|$$

où la borne inférieure est prise sur toutes les suites de partitions $\{Q^i\}_1^n$ et $\{\bar{Q}^i\}_1^n$ telles que $\mathrm{dist}(\bigvee_{i=1}^{n} Q^i) = \mathrm{dist}(\bigvee_{i=1}^{n} P^i)$ et $\mathrm{dist}(\bigvee_{i=1}^{n} \bar{Q}^i) = \mathrm{dist}(\bigvee_{i=1}^{n} \bar{P}^i)$.

Une suite de partitions $\{P^i\}_1^n$ est ε-indépendante d'une partition Q s'il existe un sous-ensemble A d'éléments de Q tel que $m(A) > 1 - \varepsilon$, et $\bar{d}(\{P^i/Q_j\}_1^n, \{P^i\}_1^n < \varepsilon$, pour tous les éléments Q_j de Q dans A. (On a désigné par P^i/Q_j la trace de la partition P^i sur Q_j.)

Une partition P dans un système dynamique est <u>très faiblement de Bernoulli</u>, si, pour tout $\varepsilon > 0$, il existe n tel que pour tout $k > 0$ et tout $m > 0$, $\{T^i P\}_0^{k+n}$ est ε-indépendante de $\bigvee_{-m}^{-1} T^i P$.

Ornstein a montré [7] que l'existence d'une partition très faiblement de Bernoulli (<u>a fortiori</u> faiblement de Bernoulli) génératrice dans un système dynamique, implique que ce système est isomorphe à un schéma de Bernoulli.

Lui-même ou ses élèves ont appliqué ces critères pour montrer que les systèmes dynamiques suivants sont isomorphes à des schémas de Bernoulli :

un facteur de schéma de Bernoulli [6],
une limite projective de schémas de Bernoulli [5],
une chaîne de Markov mélangeante [2],
un C-système transitif pour sa mesure d'entropie maximale [1],
un automorphisme ergodique de tore [3], etc...

Les systèmes dynamiques ergodiques classiques (c'est-à-dire réalisés par des difféormorphismes de variétés laissant invariante une mesure différentiable) paraissent devoir être, en général, isomorphes à des schémas de Bernoulli, dès qu'ils satisfont à des conditions d'hyperbolicité suffisantes.

Dans cette direction, mentionnons deux problèmes ouverts :

les extensions à spectre continu par des groupes compacts des schémas de Bernoulli,

les automorphismes ergodiques de nilvariétés,

sont-ils isomorphes à des schémas de Bernoulli ?

Nous n'avons pas abordé ici les résultats d'Ornstein concernant les schémas de Bernoulli d'entropie infinie [5], et le plongement des schémas de Bernoulli dans des flots [7].

5.2. Classification des K-systèmes

Peu après, Ornstein a construit une famille non dénombrable de K-systèmes de même entropie deux à deux non isomorphes, donc en particulier non isomorphes à des schémas de Bernoulli, montrant ainsi que l'entropie n'est pas un invariant complet pour la classe des K-systèmes [8] et (avec P. C. Shields) [10].

Comme corollaire, il a construit des systèmes qui ne se décomposent pas en un produit d'un K-système et d'un système déterministe, ce qui répond par la négative à la conjecture de Pinsker [8].

Ces travaux ouvrent la voie à l'étude de la structure des K-systèmes. En même temps, ils montrent quelle est la complexité de ce problème. Une question particulièrement intéressante dans ce domaine est l'étude de la structure des K-systèmes obtenus comme modèles mathématiques de systèmes physiques.

BIBLIOGRAPHIE

[1] R. AZENCOTT - Difféomorphismes d'Anosov et Schémas de Bernoulli, C. R. Acad. Sci. 270, 1970.

[2] N. A. FRIEDMAN, D. S. ORNSTEIN - On isomorphism of weak Bernoulli transformations, Advances in Math. 5, 3 (1970), p. 365-394.

[3] Y. KATZNELSON - Ergodic automorphisms of T^n are Bernoulli shifts, à paraître.

[4] D. S. ORNSTEIN - Bernoulli shifts with the same entropy are isomorphic, Advances in Math. 4, 3 (1970), p. 337-352.

[5] D. S. ORNSTEIN - Bernoulli shifts with infinite entropy are isomorphic, Advances in Math. 5, 3 (1970), p. 339-348.

[6] D. S. ORNSTEIN - Factors of Bernoulli shifts are Bernoulli shifts, Advances in Math. 5, 3 (1970), p. 349-364.

[7] D. S. ORNSTEIN - Imbedding Bernoulli shifts in flows, Midwest Conference, Ergodic Theory and Probability, Lecture Notes in Maths. 160, p. 178-218, Springer-Verlag.

[8] D. S. ORNSTEIN - A K-automorphism with no square root and Pinsker's conjecture, à paraître (1971).

[9] D. S. ORNSTEIN - An application of Ergodic Theory to Probability Theory, à paraître (1972).

[10] D. S. ORNSTEIN, P. C. SHIELDS - An uncountable family of K-automorphisms, à paraître (1972).

[11] D. RUELLE - Statistical Mechanics of a one dimensional lattice gas, Comm. Math. Phys. 9, (1968), p. 267-278.

[12] Ya. G. SINAÏ - A weak isomorphism of transformations with an invariant measure Dokl. Ak. Nauk, 147 (1962), p. 797-800.

[13] Ya. G. SINAÏ - Dynamical systems with elastic reflections, Usp. Mat. Nauk, (= Russian Math. Surveys), 25, 6 (1970).

[14] M. SMORODINSKY - Ergodic theory, Entropy, Lecture Notes in Maths. 214, 1971, Springer-Verlag.