ISRAEL JOURNAL OF MATHEMATICS 91 (1995), 419–428

NOMBRES NORMAUX, ENTROPIE, TRANSLATIONS*

BY

BERNARD HOST

Laboratoire de Mathématiques discrètes
163 av. de Luminy, case 930, 13288 Marseille cedex, France
e-mail : host@lmd.univ-mrs.fr

ABSTRACT

Given a measure μ on the circle, we study the relations between the entropy of the multiplication by an integer p and the conservativity for the translations by the p-adic rational numbers. We get a criterium for μ-almost every point to be normal in a basis q prime to p, and generalizations of the result of D. Rudolph about measures which are invariant by multiplication by p and q.

1. Introduction

1.1 Étant donné un entier $p > 1$, on rappelle qu'un réel x est **normal** en base p si la suite $(p^n x \bmod 1\colon n \geq 0)$ est uniformément distribuée sur $[0, 1[$, c'est à dire si, pour tout intervalle $I \subset [0, 1[$,

$$\frac{1}{N}\,\mathrm{Card}\{n\colon 0 \leq n < N,\ p^n x \in I \bmod 1\} \to |I| \quad \text{quand } N \to \infty.$$

Pour la mesure de Lebesgue, presque tout réel est normal en toute base. Cependant, si p et q sont des entiers multiplicativement indépendants, c'est à dire s'ils ne sont pas des puissances d'un même entier, il existe des réels normaux en base q qui ne sont pas normaux en base p ([7]). Précisant ce résultat, différents auteurs ont construit des mesures de probabilité sur $[0, 1[$ pour lesquelles presque tout nombre vérifie cette propriété. Entre autres :

* Ce travail a été terminé pendant le séjour de l'auteur au Japon à l'invitation de la Japan Society for the Promotion of Sciences.
Received December 19, 1993

• Brown, Moran et Pearce ([1]) ont étudié le cas où la mesure est un produit de Riesz.

• Feldman et Smorodinsky ([2]) ont étudié le cas où les digits en base p sont, relativement à la mesure considérée, des variables aléatoires indépendantes de même distribution, ou forment une chaîne de Markov stationnaire.

On donne ici (paragraphe 3) ici une démonstration très simple d'un théorème qui, lorsque p et q sont premiers entre eux, contient ces différents résultats comme cas particuliers ; malheureusement, les méthodes employées ici ne semblent pas se généraliser au cas multiplicativement indépendant.

Désormais $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ est identifié à l'intervalle $[0, 1[$.

THÉORÈME 1: *Soient p, q deux entiers > 1 premiers entre eux ; μ une mesure de probabilité sur $\mathbb{T}$, invariante et ergodique pour la transformation $T\colon x \mapsto px$. Si l'entropie du système $(\mathbb{T}, \mu, T)$ n'est pas nulle, alors μ-presque tout réel est normal en base q. Si la mesure μ est singulière par rapport à la mesure de Lebesgue, μ-presque tout réel est non normal en base p.*

On a bien sûr le même résultat en supposant seulement que presque tout composant ergodique de μ pour T a une entropie positive.

Le théorème 1 permet de retrouver très facilement un résultat de D.J. Rudolph :

THÉORÈME (D. J. Rudolph, [7]): *Soient p, q deux entiers > 1 premiers entre eux, et μ une mesure de probabilité sur $\mathbb{T}$. Supposons que μ est invariante par $T\colon x \mapsto px$ et par $S\colon x \mapsto qx$, ergodique pour l'action du couple (T, S), et que le système $(\mathbb{T}, T, \mu)$ a une entropie positive. Alors μ est la mesure de Lebesgue.*

Il suffit de remarquer que presque tout composant ergodique de μ pour la transformation T a une entropie positive. Aimee Johnson ([5]) a généralisé ce résultat au cas où p et q sont multiplicativement indépendants ; J. Feldman ([3]) en a donné récemment une démonstration plus simple.

1.2 La démonstration du théorème 1 utilise les relations liant l'entropie et certaines translations. Soit D un sous-groupe dénombrable de $\mathbb{T}$; on rappelle qu'une mesure de probabilité μ sur $\mathbb{T}$ est **conservative** pour D si

$$\textit{Pour tout borélien } A \textit{ avec } \mu(A) > 0 \textit{ il existe } \alpha \in D,\ \alpha \neq 0,$$
$$\textit{avec } \mu\big(A \cap (A + \alpha)\big) \neq 0.$$

Le lemme suivant est classique ; on en donne une version quantitative dans la section 2 :

LEMME 1: *Soient $p > 1$ un entier et μ une mesure de probabilité sur $\mathbb{T}$, invariante par T: $x \mapsto px$. μ est conservative pour le groupe des rationnels p-adiques si et seulement si presque tout composant ergodique de μ pour T a une entropie positive.*

Dans le résultat de Rudolph, l'hypothèse d'invariance de μ par T semble moins importante que le comportement de cette mesure sous l'action des translations rationnelles p-adiques ; plus précisément, on montre (paragraphe 3) le théorème plus général suivant :

THÉORÈME 2: *Soient p, q deux entiers > 1 premiers entre eux, et μ une mesure de probabilité sur $\mathbb{T}$, conservative pour le groupe des rationnels p-adiques et invariante par la transformation S: $x \mapsto qx$. Alors μ est la mesure de Lebesgue.*

2. Translations et entropie

Désormais, $p > 1$ est un entier et μ une mesure de probabilité sur $\mathbb{T}$; on ne fait pour l'instant aucune hypothèse d'invariance.

2.1 Pour tout $n \geq 1$, on note :

$D_n = \{jp^{-n}: 0 \leq j < p^n\}$;

B_n la (complétée pour μ de la) σ-algèbre formée des boréliens invariants par la translation $x \mapsto x + p^{-n}$;

$$\omega_n = \sum_{\alpha \in D_n} \mu * \delta_\alpha \, ;$$

et $\phi_n(x)$ la dérivée de Radon-Nicodym $\phi_n(x) = \dfrac{d\mu(x)}{d\omega_n(x)}$.

Enfin, $D = \bigcup_n D_n$ est le groupe des rationnels p-adiques.

On remarque immédiatement que la suite (ϕ_n) est décroissante et que $\phi_n(x) = 1$ μ-pp si et seulement si les mesures $(\mu * \delta_\alpha: \alpha \in D_n)$ sont deux à deux mutuellement singulières.

LEMME 2: *La suite $\phi_n(x)$ tend vers 0 μ-presque partout si et seulement si la mesure μ est conservative pour le groupe D des rationnels p-adiques.*

Démonstration (Voir aussi [6], pages 123–125): Pour tout n, $\phi_n(x) > 0$ μ-pp. Soit $K = \{x: \forall n: \phi_n(x) > 0\}$. On a $\mu(K) = 1$ et $\mathbf{1}_K(x) d\omega_n(x) = \phi_n(x)^{-1} d\mu(x)$ pour tout n.

Supposons que la suite $(\phi_n(x))$ ne tend pas vers 0 presque partout ; il existe un borélien $C \subset K$ et $\epsilon > 0$ tels que $\mu(C) > 0$, $\phi_n(x) > \epsilon$ pour tout n et tout

$x \in C$.

Pour tout n,

$$\epsilon^{-1}\mu(C) \geq \int_C \frac{1}{\phi_n(x)}\, d\mu(x) = \omega_n(C) = \int \sum_{\alpha \in D_n} \mathbf{1}_C(x+\alpha) d\mu(x);$$

la suite $\left(\sum_{\alpha\in D_n} \mathbf{1}_C(x+\alpha)\right)$ est donc bornée presque partout, donc est presque partout constante à partir d'un certain rang. Il existe donc n et un borélien $B \subset C$ avec $\mu(B) > 0$ et $\mathbf{1}_C(x+\alpha) = 0$ pour tout $x \in B$ et tout $\alpha \in D \setminus D_n$.

Ainsi, $B \cap (B+\alpha) = \emptyset$ pour tout $\alpha \in D \setminus D_n$. Il existe $\beta \in D_n$ tel que l'ensemble $A = B \cap [\beta, \beta + p^{-n}[$ ait une mesure > 0 ; on a alors $A \cap (A+\alpha) = \emptyset$ pour tout $\alpha \in D$ non nul, ce qui contredit la conservativité.

Réciproquement, supposons qu'il existe un borélien A avec $\mu(A) > 0$ et $\mu(A \cap (A+\alpha)) = 0$ pour tout $\alpha \in D$ non nul. On se ramène immédiatement au cas où $A \subset K$ et $A \cap (A+\alpha) = \emptyset$ pour tout $\alpha \in D$ non nul. Pour tout n,

$$1 \geq \mu\Big(\bigcup_{\alpha\in D_n} (A+\alpha)\Big) = \sum_{\alpha \in D_n} \mu(A+\alpha) = \omega_n(A) = \int_A \frac{1}{\phi_n(x)}\, d\mu(x)$$

et la suite $(\phi_n(x))$ ne peut pas tendre vers 0 presque partout sur A. ∎

2.2 Supposons de plus que μ est invariante par $T \colon x \mapsto px$.

Pour tout $n \geq 0$, $B_n = T^{-n}B_0$. On vérifie immédiatement que :

$$\phi_n(x) = \phi_1(x)\phi_1(Tx)\cdots\phi_1(T^{n-1}x). \tag{1}$$

D'autre part, pour toute $f \in L^1(\mu)$, l'espérance conditionnelle de f par rapport à B_n est :

$$E(f \mid B_n)(x) = \sum_{\alpha\in D_n} f(x+\alpha)p(x+\alpha).$$

Ainsi, si $\alpha \in D_n$ et $J = [\alpha, \alpha + p^{-n}[$,

$$\phi_n(x) = P(J|B_n)(x) \qquad \text{pour } \mu\text{-presque tout } x \in J.$$

(Les probabilités et espérances conditionnelles sont relatives à la mesure μ.)

Ainsi, comme la partition $\left([jp^{-1}, (j+1)p^{-1}[\right)$ est génératrice, l'entropie h du système $(\mathbb{T}, \mu, T)$ est (voir [6], sections 5.2. et 5.3.) :

$$h = -\int \log \phi_1(x)\, d\mu(x).$$

L'entropie est donc nulle si et seulement si $\phi_1(x) = 1$ μ-pp, et dans ce cas $\phi_n(x) = 1$ μ-pp pour tout $n > 0$.

Si le système $(\mathbb{T}, \mu, T)$ est ergodique, le théorème ergodique et (1) entraînent que :

$$\phi_n(x)^{1/n} \to e^{-h} \ \mu\text{-presque partout.} \tag{2}$$

Dans le cas général, on peut appliquer cette remarque à chaque composant ergodique de μ : on trouve ainsi une version quantitative du lemme 1.

3. Démonstration des théorèmes 1 et 2

3.1 NOTATIONS. Dans ce paragraphe, p, q sont deux entiers > 1, premiers entre eux, et $a \neq 0$ un entier. Pour chaque $n > 0$ la suite des résidus $(aq^k \bmod p^n \colon k \geq 0)$ est périodique et on note T_n sa période ; ainsi, les résidus $(aq^k \bmod p^n \colon 0 \leq k < T_n)$ sont deux à deux distincts.

LEMME 3: *Pour tout n assez grand, $T_n = C^{te} p^n$.*

La démonstration est laissée au lecteur.

Pour chaque $N > 0$, on note

$$g_N(x) = \frac{1}{N} \sum_{k=0}^{N-1} e(aq^k x)$$

où $e(x) = e^{2\pi i x}$.

Soient encore μ une mesure de probabilité sur $\mathbb{T}$, et ϕ_n comme au paragraphe précédent.

LEMME 4:
Si $1 \leq N \leq T_n$, alors $$\int \frac{|g_N(x)|^2}{\phi_n(x)}\, d\mu(x) \leq \frac{p^n}{N}.$$

Démonstration: On a $\dfrac{1}{\phi_n(x)} d\mu(x) \leq d\omega_n(x)$ d'où :

$$\begin{aligned} \int \frac{|g_N(x)|^2}{\phi_n(x)}\, d\mu(x) &\leq \int |g_N(x)|^2 d\omega_n(x) = \int \sum_{j=0}^{p^n-1} |g(x + jp^{-n})|^2 d\mu(x) \\ &= \int h(x) d\mu(x) \end{aligned}$$

où $$h(x) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} e\big((q^k - q^l)x\big) \sum_{j=0}^{p^n-1} e\big((q^k - q^l) j p^{-n}\big).$$

Si $0 \leq k, l < N$ avec $k \neq l$, alors par hypothèse $q^k \neq q^l \bmod p^n$ et

$$\sum_{j=0}^{p^n-1} e\big((q^k - q^l)jp^{-n}\big) = 0.$$

On a donc $h(x) = p^n/N$, et $\int h(x)d\mu(x) = p^n/N$, d'où le résultat annoncé. ∎

3.3 Démonstration du théorème 2. On suppose ici de plus que μ vérifie les hypothèses du théorème 2. En utilisant l'inégalité de Cauchy-Schwarz et les lemmes 3 et 4, on obtient que pour tout n assez grand

$$\begin{aligned}\left|\int g_{T_n}(x)\,d\mu(x)\right|^2 &\leq \int \frac{\left|g_{T_n}(x)\right|^2}{\phi_n(x)}\,d\mu(x) \,.\, \int \phi_n(x)\,d\mu(x) \leq \frac{p^n}{T_n}\int \phi_n(x)\,d\mu(x)\\ &\leq C^{te}\int \phi_n(x)d\mu(x)\end{aligned}$$

qui tend vers 0 d'après le lemme 2.

Or, comme μ est invariante par S: $x \mapsto qx$, $\int g_{T_n}(x)\,d\mu(x) = \int e(ax)\,d\mu(x)$ pour tout n ; ainsi $\int e(ax)\,d\mu(x) = 0$. Comme ceci est vrai pour tout entier $a \neq 0$, μ est la mesure de Lebesgue. ∎

3.4 Démonstration du théorème 1. Supposons maintenant que la mesure μ vérifie les hypothèses du théorème 1. Montrons que :

$$g_N(x) \to 0 \quad \mu\text{-pp.} \tag{3}$$

Soient s un entier avec $hs > \log p$, t un réel avec $1 < t < hs/\log p$ et, pour chaque entier $k \geq 1$, n_k l'entier tel que $T_{n_k-1} < k^s \leq T_{n_k}$. Pour tout k, d'après les lemmes 3 et 4,

$$\int \frac{|g_{k^s}(x)|^2}{\phi_{n_k}(x)}d\mu(x) \;\leq\; C^{te} \text{ donc } \sum_{k=1}^{\infty}\int \frac{|g_{k^s}(x)|^2}{k^t\phi_{n_k}(x)}d\mu(x) \;<\; \infty$$

et

$$\frac{|g_{k^s}(x)|^2}{k^t\phi_{n_k}(x)} \;\to\; 0 \quad \mu\text{-pp.}$$

D'autre part, d'après (5), pour presque tout x, $\phi_{n_k}^{1/n_k} \to e^{-h}$ donc, pour tout k assez grand,

$$\phi_{n_k}(x)^{1/n_k} < e^{-(t\log p)/s} = p^{-t/s}$$

et, comme $p^{n_k} \geq C^{te}k^s$,

$$k^t\phi_{n_k}(x) \leq C^{te} \quad \text{et} \quad |g_{k^s}(x)|^2 \leq C^{te}\frac{|g_{k^s}(x)|^2}{k^t\phi_{n_k}(x)}.$$

Ainsi

$$g_{k^s}(x) \;\to\; \mu\text{-pp.}$$

Si $k^s \leq N < (k+1)^s$,

$$|g_N(x) - g_{k^s}(x)| \leq 2\left(\left(\frac{k+1}{k}\right)^s - 1\right).$$

Ainsi, la suite $(g_N(x))$ tend vers 0 en tout x où la suite $(g_{k^s}(x))$ tend vers 0, c'est à dire presque partout, d'où (3).

Comme (3) est vérifié pour tout entier a non nul, le critère de Weyl entraîne que μ-presque tout x est normal en base q. D'autre part, le théorème ergodique entraîne que pour presque tout x la suite $p^n x$ a la distribution μ ; si cette mesure est différente de la mesure de Lebesgue, presque tout x est non normal en base p. Ceci achève la démonstration du théorème 1. ∎

4. Généralisations

4.1 Équidistribution presque-sûre de $\mathbf{f_k x}$.

Dans la démonstration du théorème 1, on a utilisé le fait que la suite (q^k) a une très bonne répartition modulo p^n pour tout n ; cet argument peut être utilisé pour d'autres suites. Soit par exemple (f_k) la suite de Fibonacci. Le résultat suivant est du à T. Kamae :

THÉORÈME 3: *Soient $p > 1$ et μ une mesure de probabilité invariante, ergodique et d'entropie positive pour T: $x \mapsto px$. Pour μ-presque tout x, la suite $(f_k x)$ est équidistribuée modulo* 1.

Résumé de la démonstration: Soit $a \neq 0$ un entier. Pour tout $N > 0$ on note

$$g_N(x) = \frac{1}{N}\sum_{k=0}^{N-1} e(af_k x).$$

On veut montrer que la suite $(g_N(x))$ tend vers 0 μ-pp.

Comme dans la démonstration du lemme 4 , on vérifie que, pour tout $N > 0$ et tout $n > 0$,

$$\int \frac{|g_N(x)|^2}{\phi_n(x)}\, d\mu(x) \leq \frac{p^n}{N^2}\Delta(N,n)$$

où $$\Delta(N,n) = \sum_{j=0}^{p^n-1} \left(\mathrm{Card}\{k: 0 \leq k < N,\ af_k = j \bmod p^n\}\right)^2.$$

D'autre part, pour tout n la suite $(af_k \bmod p^n)$ est périodique, et sa période T_n est égale à $C^{te}p^n$ pour tout n assez grand. On vérifie que $\Delta(T_n, n) \le C^{te}n^hT_n$, où h est le nombre de facteurs premiers distincts de p ; on en déduit que

$$\text{Si } T_n < N \le T_{n+1}, \text{ alors } \frac{p^n}{N^2}\Delta(N,n) \le C^{te}n^h.$$

On termine alors la démonstration comme celle du théorème 1. ∎

4.2 Translations irrationnelles. On donne ici un théorème analogue au théorème 2 pour une translation irrationnelle. Soit $\alpha \in \mathbb{T}$ un irrationnel ; on dit que la mesure de probabilité μ sur $\mathbb{T}$ est **conservative** pour α si

Pour tout borélien A avec $\mu(A) > 0$, il existe un entier $n \neq 0$ avec $\mu\big(A \cap (A + n\alpha)\big) \neq 0$.

Théorème 4: *Soient α un irrationnel, $q > 1$ un entier, et μ une mesure de probabilité sur $\mathbb{T}$, invariante par S: $x \mapsto qx$ et conservative pour α. Si $\{q^n\alpha\colon n \ge 0\}$ est dense dans $\mathbb{T}$, μ est la mesure de Lebesgue.*

Démonstration: Pour chaque $n > 0$, soient

$$\nu_n = \sum_{k=0}^{n-1} \mu * \delta_{k\alpha} \quad et \quad \psi_n(x) = \frac{d\mu(x)}{d\nu_n(x)}.$$

Comme μ est conservative pour α, une petite modification de la démonstration du lemme 2 donne :

$$\psi_n(x) \to 0 \quad \mu\text{-presque partout.}$$

Soit $a \neq 0$ un entier ; on veut montrer que $\int e(ax)\, d\mu(x) = 0$. Soit $n > 0$. Comme $\{aq^k\alpha\}$ est dense dans $\mathbb{T}$, il existe des entiers $k_j \ge 0$ $(0 \le j < n)$ tels que

$$\big|e(aq^{k_j}\alpha) - e(\frac{j}{n})\big| < \frac{1}{n^2}.$$

Posons

$$g(x) = \frac{1}{n}\sum_{j=0}^{n-1} e(aq^{k_j}x).$$

On obtient :

$$\begin{aligned}\int \frac{|g(x)|^2}{\psi_n(x)}\, d\mu(x) &\le \int |g(x)|^2\, d\nu_n(x) \\ &= \int \frac{1}{n^2}\sum_{i,j=0}^{n-1} e\big(a(q^{k_i} - q^{k_j})x\big) \sum_{m=0}^{n-1} e\big(a(q^{k_i} - q^{k_j})m\alpha\big)\, d\mu(x) \\ &\le 1 + \frac{1}{n^2}\sum_{i,j=0}^{n-1}\sum_{m=0}^{n-1}\Big|e\big(a(q^{k_i} - q^{k_j})m\alpha\big) - e\Big(\frac{i-j}{n}m\Big)\Big| \le 2.\end{aligned}$$

D'où

$$\Big|\int e(ax)\,d\mu(x)\Big|^2 = \Big|\int g(x)\,d\mu(x)\Big|^2 \le 2\int \psi_n(x)d\mu(x)$$

et enfin $\int e(ax)\,d\mu(x) = 0$, ce qu'il fallait démontrer. ∎

4.3 Un peu de dynamique non singulière. On peut donner dest théorèmes 2 et 4 des démonstrations encore plus courtes, quoique moins élémentaires. Soient μ, p, q comme dans l'énoncé du théorème 2 ; D désigne le groupe des rationnels p-adiques. Soient $(c_\alpha\colon \alpha\in D)$ des réels >0 avec

$$\sum_{\alpha\in D} c_\alpha = 1 \text{ et } \nu = \sum_{\alpha\in D} \mu * \delta_\alpha.$$

La mesure ν est conservative pour D, quasi-invariante par les translations de D, et μ est absolument continue par rapport à ν.

Soit $a\neq 0$ un entier, et supposons que $\int e(ax)\,d\mu(x)\neq 0$. Soient Θ l'adhérence de la suite $(aq^k\colon k\ge 0)$ dans $\mathbb{Z}_p$, $\theta\in\Theta$, et (k_j) une suite d'entiers telle que (aq^{k_j}) tende vers θ dans $\mathbb{Z}_p$; en remplaçant cette suite par une sous-suite, on peut supposer que la suite de fonctions $\big(e(aq^{k_j}x)\big)$ converge faiblement dans $L^\infty(\nu)$. La limite f de cette suite n'est pas identiquement nulle car $\int f(x)\,d\mu(x) = \int e(ax)\,d\mu(x)\neq 0$. D'autre part, pour tout $\alpha\in D$, la fonction $e(x+\alpha)$ est la limite faible des fonctions $e\big(aq^{k_j}(x+\alpha)\big) = e\big(aq^{k_j}x\big)e\big(aq^{k_j}\alpha\big)$ donc $f(x+\alpha) = f(x)\,(\theta|\alpha)$ μ-pp, où $(.|.)$ désigne la dualité entre $\mathbb{Z}_p$ et D. Ainsi chaque $\theta\in\Theta$ est une valeur propre de l'action de D sur ν. Or on sait (voir par exemple [4]) que l'ensemble des valeurs propres d'un système non singulier conservatif est de mesure nulle pour la mesure de Lebesgue : Θ est donc de mesure nulle dans $\mathbb{Z}_p$, ce qui contredit le lemme 1. On a donc $\int e(ax)\,d\mu(x) = 0$ pour tout $a\neq 0$. ∎

References

[1] G. Brown, W. Moran and C. Pearce, *Riesz products and normal numbers*, Journal of the London Mathematical Society **32** (1985), 12–18.

[2] J. Feldman and M. Smorodinsky, *Normal numbers from independant processes*, Ergodic Theory and Dynamical Systems **12** (1992), 707–712.

[3] J. Feldman, *A generalization of a result of Lyons about measures in* $[0,1]$, Israel Journal of Mathematics **81** (1993), 281–287.

[4] B. Host, J. F. Méla and F. Parreau, *Non singular transformations and spectral analysis of measures*, Bulletin de la Société Mathématique de France **119** (1991), 33–90.

[5] A. Johnson, *Measures on the circle invariant under multiplication by a nonlacunary subgroup of the integers*, Israel Journal of Mathematics **77** (1992), 211–240.

[6] K. Petersen, *Ergodic Theory*, Cambridge Studies in Advanced Mathematics, 2, Cambridge University Press, Cambridge, 1983.

[7] D. J. Rudolph, $\times 2$ *and* $\times 3$ *invariant measures and entropy*, Ergodic Theory and Dynamical Systems **10** (1990), 395–406.

[8] W. M. Schmidt, *On normal numbers*, Pacific Journal of Mathematics **10** (1960), 661–672.