SÉMINAIRE N. BOURBAKI

PAUL ANDRÉ MEYER

Régularité des processus gaussiens

Séminaire N. Bourbaki, 1976, exp. n° 470, p. 256-266

<http://www.numdam.org/item?id=SB_1974-1975__17__256_0>

© Association des collaborateurs de Nicolas Bourbaki, 1976, tous droits réservés.

L'accès aux archives du séminaire Bourbaki (http://www.bourbaki.ens.fr/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI 470-01
27e année, 1974/75, n° 470 Juin 1975

RÉGULARITÉ DES PROCESSUS GAUSSIENS

[d'après X. FERNIQUE]

par Paul André MEYER

Rappel de quelques définitions

Une loi μ sur $\mathbf{R}^n$ est dite gaussienne si sa transformée de Fourier est de la forme

(1) $$\int e^{iu.x}\,\mu(dx) = e^{-q(u)/2}$$

où q est une forme quadratique positive (éventuellement dégénérée). Le mot "gaussienne" est pris ici en un sens restreint : dans un langage plus précis, on dirait "gaussienne centrée". La forme quadratique q peut s'écrire explicitement :

(2) $$q(u) = \int (u.x)^2\,\mu(dx)$$

et la mesure gaussienne μ est donc déterminée par la connaissance des moments du second ordre $\Gamma(i,j) = \int x_i x_j \mu(dx)$, où l'on désigne par x_i les coordonnées.

Soit $(\Omega, \mathbb{F}, P)$ un espace probabilisé, et soit $(X_t)_{t\in T}$ une famille quelconque de variables aléatoires (fonctions mesurables) réelles sur Ω . On dit que (X_t) est un processus gaussien si, pour toute partie finie A de T , la mesure image μ_A de P par l'application $\omega \longmapsto (X_t(\omega))_{t\in A}$ est une mesure gaussienne sur $\mathbf{R}^A$. La covariance du processus gaussien est la fonction sur $T \times T$

(3) $$\Gamma(s,t) = E[X_s X_t] \quad \left(= \int X_s(\omega)X_t(\omega)P(d\omega) \right) .$$

D'après (2), la transformée de Fourier de la loi μ_A est $\exp(-q(u)/2)$, où $q(u) = \sum_{(s,t)\in A\times A} \Gamma(s,t)u_s u_t$. La covariance détermine donc uniquement les lois μ_A , et c'est une fonction de type positif sur $T\times T$. Inversement, étant donnée une fonction de type positif $T\times T$, il existe des processus gaussiens l'admettant pour covariance.

On appelle trajectoire de ω la fonction $t \longmapsto X_t(\omega)$ de T dans $\mathbf{R}$. On rencontre ici, comme partout en théorie des processus, la difficulté suivante : soit (X'_t) une seconde famille de variables aléatoires telle que pour chaque t

on ait $X_t = X'_t$ p.p. sur Ω . Le processus (X'_t) est évidemment un processus gaussien de même covariance que (X_t) , mais, si T n'est pas dénombrable, le passage de (X_t) à (X'_t) - que l'on appellera modification dans la suite - bouleverse complètement les trajectoires. Les probabilistes disposent de moyens techniques ("séparabilité" de Doob) pour choisir de bonnes modifications de processus, mais c'est une question qu'il vaut mieux laisser de côté ici. Nous poserons donc le problème de régularité des processus gaussiens en les termes suivants :

1) Reconnaître sur Γ s'il existe une modification de (X_t) à trajectoires bornées sur T .

2) Reconnaître sur Γ si, une topologie étant donnée sur T , il existe une modification de (X_t) à trajectoires continues.

Ces problèmes ne sont pas sans analogie (ni sans liens) avec les problèmes d'analyse harmonique : reconnaître sur les coefficients de Fourier d'une fonction de L^2 si la fonction est égale p.p. à une fonction bornée, continue... Comme ceux-ci, ils ont suscité des travaux considérables depuis plus de vingt ans, avec des "retombées" parfois plus importantes que le problème lui-même. N'étant pas un spécialiste du sujet, je me garderai bien d'entrer dans des considérations historiques, et je renverrai le lecteur aux exposés d'ensemble de Dudley [1] et de Fernique [2] (je suis fidèlement celui-ci, sans aucune modification de fond ni de forme). Ces exposés comportent des bibliographies.

Le problème présenté ici ne constitue d'ailleurs qu'une petite partie de la théorie "fine" des processus gaussiens. Lorsque $T = \mathbb{R}$, par exemple, il y a d'innombrables résultats sur des processus à covariance assez régulière : conditions de Hölder, dimension de Hausdorff des trajectoires, temps locaux, nombres de passages par un niveau donné ... C'est toute une branche du calcul des probabilités.

L'étude de T comme espace métrique

Désormais, le processus gaussien $(X_t)_{t\in T}$, de covariance Γ , reste fixé.

L'application $t \mapsto X_t \in L^2(\Omega)$ permet de transporter sur T la métrique de L^2 . Nous poserons

(4) $\qquad d(s,t) = \|X_t - X_s\|_2 = (\Gamma(t,t) + \Gamma(s,s) - 2\Gamma(s,t))^{\frac{1}{2}}$.

Il y a toutefois une petite difficulté, tenant au fait que l'application $t \longmapsto X_t$ ne sépare pas nécessairement T . Il ne suffit pas toujours de dire "on passe au quotient", car cela démolit parfois une structure précieuse sur T . Dans cet exposé élémentaire, nous supposerons que d est une vraie distance sur T .

On doit à Dudley (entre autres choses) d'avoir insisté vigoureusement sur les points suivants : 1) en ce qui concerne la continuité, la topologie naturelle sur T est celle qui est déterminée par d . Le problème une fois résolu pour celle-ci, on la comparera à une autre topologie si besoin est.

2) Les conditions de régularité sont à rechercher sous la forme de caractéristiques géométriques de l'espace métrique (T,d) . Voici le genre de caractéristiques qui interviennent. Nous noterons $B(t,\delta)$ la boule ouverte de centre $t \in T$, de rayon $\delta > 0$ (en abrégé, δ-boule de centre t). S désignant une partie de T ;

(5) $N(S,\delta)$ sera le nombre minimal de δ-boules nécessaires pour recouvrir S (on pose $N(T,\delta) = N(\delta)$) ;

(6) $M(S,\delta)$ sera le nombre maximal de δ-boules disjointes dans T , centrées sur S (on pose $M(T,\delta) = M(\delta)$).

Dudley a montré que si $\sum_n 2^{-n} \sqrt{N(2^{-n})} < \infty$, alors (X_t) admet une modification à trajectoires bornées. Inversement, Sudakov a montré que s'il existe une telle modification, $\sup_n 2^{-n} \sqrt{M(2^{-n})}$ est fini. Pour aller plus loin, Fernique utilise des caractéristiques "locales" de la géométrie de T . Soit S une partie de T ; soit $B(S,\delta) = \{t : d(t,S) < \delta\}$. On définit pour $q \geq 2$

$K(t,\delta,n,q)$ est le nombre maximum de δq^{-n}-boules, disjointes dans T , centrées dans $B(t,\delta q^{-n+1})$

(7) $\qquad K(S,\delta,n,q) = \inf_{t \in B(S,\delta)} K(t,\delta,n,q)$.

Noter que $K(S,\delta,n,q)$ n'est ni une caractéristique du sous-espace S de T , ni une fonction croissante ou décroissante de S .

Une idée des résultats de Fernique

Bien que Fernique ait amélioré les conditions suffisantes de régularité (à la fois en serrant davantage les conditions, et en les rendant plus maniables), sa contribution essentielle concerne les conditions nécessaires. Je vais énoncer, et en grande partie démontrer, son résultat principal sur la minoration des processus gaussiens à trajectoires bornées.

Dans l'énoncé suivant, le symbole [ ] signifie "partie entière de" , le symbole E (espérance) est la manière probabiliste d'écrire l'intégrale sur Ω , enfin, $\sup_t X_t$ est à interpréter comme un sup essentiel des $\sup_{t \in D} X_t$, D dénombrable dans T .

THÉORÈME 1.- Supposons qu'il existe une modification de (X_t) à trajectoires bornées sur T . On a alors pour tout $\delta > 0$, tout $q \geq 2$, tout $S \subset T$

$$(8) \quad \sqrt{[\log_2 M(S,\delta)]} + \sum_n q^{-n} \sqrt{[\log_2 K(S,\delta,n,q)]} \leq \frac{C}{\delta} E[\sup_t X_t] < \infty$$

(C est une constante universelle).

Le premier terme de (8) correspond au résultat de Sudakov cité plus haut. La convergence de la série (8) est une condition "presque" suffisante pour l'existence de modifications à trajectoires bornées, puisqu'on peut en déduire une condition nécessaire et suffisante dans le cas stationnaire, que nous énonçons maintenant.

Nous supposons que $T = \mathbb{R}^n$, que la covariance $\Gamma(s,t)$ s'écrit sous la forme $\gamma(t - s)$, où γ est une fonction continue. Alors la topologie associée à la distance d est (si elle est séparée) identique à la topologie usuelle, et les notions topologiques de l'énoncé suivant s'entendent au sens usuel (on ne suppose pas, cependant, que l'application $t \mapsto X_t$ sépare $\mathbb{R}^n$). Sous les hypothèses précédentes :

THÉORÈME 2.- Pour qu'il existe une modification de (X_t) à trajectoires localement bornées, il faut et il suffit qu'il existe un voisinage V de 0 tel que la fonction sur $\mathbb{R}_+$

$$\delta \mapsto \sqrt{\log N(V,\delta)}$$

soit intégrable au voisinage de 0 . Alors il existe aussi une modification de (X_t) à trajectoires continues.

On va démontrer presque complètement le théorème 1, mais on ne dira rien du théorème 2.

Démonstration du théorème 1

On ne va pas démontrer exactement (8), mais l'inégalité analogue portant sur les n pairs. Pour avoir celle qui concerne les n impairs, remplacer δ par δq .

Nous posons $p_1 = [\log_2 M(S,\delta)]$

$p_n = [\log_2 K(S,\delta,n,q)]$ pour $n > 1$.

Nous allons raisonner en supposant tous les p_i différents de 0 : cela simplifie le langage - mais la recette pour se passer de cette hypothèse est bien simple : si l'un des p_i est nul, il faut l'oublier dans la construction qui suit.

a) Construction d'un arbre

Considérons l'ensemble $R = R_1 \cup R_2 \cup \ldots \cup R_n \ldots$, où

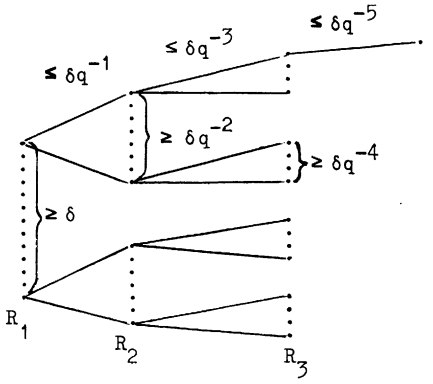
$R_1 = \{0,1\}^{p_1} \quad \ldots \quad R_n = \{0,1\}^{p_1} \times \ldots \times \{0,1\}^{p_n}$.

Nous imaginons R comme un arbre (ou plutôt une forêt). R_1 est l'ensemble des racines, de chaque point de R_1 partent 2^{p_1} branches aboutissant aux points de R_2 , de chaque point de R_2 partent 2^{p_2} branches aboutissant aux points de R_3 , etc.. Nous allons construire une application de R dans T , injective du fait que tous les p_i sont > 0 : cela nous permettra de considérer R comme plongé dans T , et par exemple de parler des variables aléatoires X_r , $r \in R$.

Aux points de R_1 nous associons les centres t_i^1 $(i \in \{0,1\}^{p_1})$ d'une famille de δ-boules disjointes centrées sur S . Une telle famille existe, puisque $M(S,\delta) \geq 2^{p_1}$.

Etant donné t_i^1 , les $t_{ij}^2 \in R_2$ $(j \in \{0,1\}^{p_2})$ auxquels aboutissent les branches issues de t_i^1 sont les 2^{p_2} centres d'une famille de δq^{-2}-boules disjointes, centrées dans la boule $B(t_i^1, \delta q^{-1})$ - une telle famille existe, puisque $K(S,\delta,2,q) \geq 2^{p_2}$. Noter que ces points ne sont plus dans S ,

mais à distance de S au plus δq^{-1} . Compte tenu de la définition (7) de $K(S,\delta,3,q)$, inf pris sur un voisinage de S , on peut définir ensuite les t^3_{ijk} $(k \in \{0,1\}^{p_3})$, centres d'une famille de δq^{-4}-boules disjointes centrées dans $B(t^2_{ij}, \delta q^{-3})$, etc. Le mieux est de faire un dessin en supposant tous les p_i égaux à 1 :



Tout chemin grimpant dans l'arbre à partir de la racine t^1_i de R_1 reste à une distance de celle-ci inférieure à $\delta q^{-1} + \delta q^{-3} + \ldots \leq \frac{4}{3}\delta q^{-1} \leq \frac{2}{3}\delta$ (on rappelle que $q \geq 2$). Il en résulte que si s et s' sont deux points de R situés sur des chemins issus de racines différentes r et r' , les boules $B(s,\delta/3)$ et $B(s',\delta/3)$ sont contenues dans les boules disjointes $B(r,\delta)$ et $B(r',\delta)$, donc $d(r,r') \geq \delta/3$. Plus généralement - c'est la propriété fondamentale de l'arbre - si s et s' sont deux points de R_{m+n} , $R_{m+n'}$ situés sur des chemins issus de points différents de R_m , leur distance est au moins $\frac{\delta}{3} q^{-2(m-1)}$.

b) Construction d'un processus gaussien $(Y_r)_{r \in R}$

Soient $B_1, B_2, \ldots$ les intervalles $[1,p_1]$, $[p_1+1, p_1+p_2]$,... de $\mathbb{N}$, et A_n l'ensemble $B_1 \cup \ldots \cup B_n = [1, p_1 + \ldots + p_n]$. L'ensemble $R_n = \{0,1\}^{p_1} \times \ldots \times \{0,1\}^{p_n}$ s'identifie à l'ensemble des parties de A_n , et R à l'ensemble des parties finies de $\mathbb{N}$. Cette identification étant faite, donnons-nous sur un espace probabilisé convenable W une suite (λ_k) de variables aléatoires gaussiennes indépendantes réduites (i.e. $E[\lambda_k^2] = 1$). Soit f la fonction sur $\mathbb{N}$ qui vaut $\delta q^{-2(i-1)} / 4\sqrt{p_i}$ sur B_i , et soit pour $r \in R$

$$(9) \qquad Y_r(w) = \sum_{k \in r} f(k)\lambda_k(w) \ .$$

Majorons, pour deux points r et r' de l'arbre, $d_Y^2(r,r') = E[(Y_r - Y_{r'})^2]$. Supposons que les chemins descendant vers la racine à partir de r , r' diffèrent pour la première fois au rang m - ce qui signifie que, en tant que parties finies de $\mathbb{N}$, r et r' ont la même intersection avec $B_1,\dots,B_{m-1}$. Alors

$$E[(Y_r - Y_{r'})^2] \le \sum_{k \in B_m \cup B_{m+1} \cup \dots} f(k)^2 = \frac{\delta^2}{16} \sum_m^\infty q^{-4(i-1)}$$

$$\le \frac{\delta^2}{16} \frac{16}{15} q^{-4(m-1)} .$$

Tandis que nous avons vu en a) que $d_X^2(r,r') = E[(X_r - X_{r'})^2] \ge \frac{\delta^2}{9} q^{-4(m-1)}$.
La conclusion est que

(10) $\quad d_Y(r,r') \le d_X(r,r')$ $\qquad$ quels que soient r , $r' \in R$.

c) Naturellement, ce n'est pas par hasard qu'on est arrivé à (10). Il existe en effet un lemme de comparaison des processus gaussiens, issu d'un travail de Mlle Chevet (et qui est une variante d'un lemme plus ancien, très souvent utilisé, dû à Slepian) qui dit

Si $(Y_s)_{s\in S}$ et $(X_s)_{s\in S}$ sont deux processus gaussiens, où S est un ensemble fini, et si l'on a $d_Y \le d_X$, on a aussi

(11) $\quad E[\sup_s Y_s] \le E[\sup_s X_s]$

(ces intégrales existent du fait que S est fini).

La démonstration n'est pas très facile. On commence par se ramener à l'inégalité équivalente

$$E[\sup_{s,t}(Y_s - Y_t)] \le E[\sup_{s,t}(X_s - X_t)]$$

(cela revient à doubler les deux membres de (11)). On se ramène aisément au cas où les matrices de covariance Γ_X et Γ_Y sont inversibles, on réalise X et Y sur un même espace probabilisé, de manière qu'ils soient indépendants, et on pose pour $\lambda \in [0,1]$

$$Z_s^\lambda = \sqrt{\lambda}\, X_s + \sqrt{1-\lambda}\, Y_s$$

processus gaussien de covariance $\lambda\Gamma_X + (1-\lambda)\Gamma_Y$. Après quoi on démontre que

$$\frac{d}{d\lambda} E[\sup_{s,t}(Z_s^\lambda - Z_t^\lambda)] \ge 0$$

et c'est ici que les choses deviennent fatigantes. On n'en dira pas plus.

d) Nous appliquons le lemme précédent à $S = R_n$ et aux deux processus gaussiens X et Y construits plus haut. Grâce à la formule (9), on sait calculer exactement $E[\sup_{r\in R_n} Y_r]$. Rappelons que R_n s'identifie à l'ensemble des parties de A_n . A tout $w \in W$ associons les deux parties $r_+(w)$, $r_-(w)$ de A_n définies par

$$r_\pm(w) = \{k \in A_n : \pm \lambda_k(w) \geq 0\} \qquad \text{(noter que } P\{\lambda_k = 0\} = 0 \text{)}$$

alors (k parcourant A_n , r et r' parcourant $\mathcal{P}(A_n)$)

$$\sum_k f(k)|\lambda_k(w)| \geq \sup_{r,r'} (Y_r(w) - Y_{r'}(w))$$

$$\geq Y_{r_+(w)}(w) - Y_{r_-(w)}(w)$$

$$= \sum_k f(k)|\lambda_k(w)| \quad \text{pour presque tout } w .$$

Intégrons, en tenant compte du fait que $E[|\lambda_k|] = 2/\sqrt{2\pi}$, il vient

$$(12) \qquad \frac{2}{\sqrt{2\pi}} \sum_{k\in A_n} f(k) = 2E[\sup_{r\in R_n} Y_r] .$$

Revenant à l'expression de f donnée en b), on voit que le côté gauche de (12) est, à une constante universelle près $\delta \sum_1^n q^{-2(i-1)}\sqrt{p_i}$. Pour obtenir le côté gauche de (8), il restera à faire tendre n vers $+\infty$.

e) Il n'est évidemment pas question de calculer $E[\sup_{r\in R_n} X_r]$: ce que nous voulons savoir, c'est que le côté gauche de (8) est une série convergente, donc que $E[\sup_{r\in R} X_r]$ converge. Cela résulte d'un lemme bien plus général, dû à Fernique, mais plus ancien que les travaux présentés ici.

Oublions ce qu'est R : soit D un ensemble dénombrable quelconque, soit E l'ensemble $\mathbb{R}^D$ (espace polonais, si l'on veut, muni de sa tribu borélienne), soient ξ_t les applications coordonnées sur E , μ la loi image de P par l'application $\omega \mapsto (X_t(\omega))_{t\in D}$. Soit aussi N la fonction borélienne positive sur E

$$(13) \qquad N(x) = \sup_t |\xi_t(x)|$$

N satisfait aux axiomes d'une semi-norme, à cela près qu'elle n'est pas nécessairement finie - mais l'assertion suivant laquelle (X_t) a des trajectoires p.s. bornées entraîne que N est μ-p.p. finie sur E . Nous voudrions en

déduire que $E[\sup_{t\in D} |X_t|] = \int N(x)\mu(dx) < \infty$. On a beaucoup mieux. En s'appuyant seulement sur le caractère gaussien de μ , et le fait que N est une pseudo-semi-norme mesurable, Fernique démontre

Lemme.- Si $N(x) < \infty$ μ-p.p., il existe $\alpha > 0$ tel que

$$(14) \qquad \int \exp[\alpha N^2(x)]\mu(dx) < +\infty \ .$$

La démonstration est si simple qu'il faut la donner. Formons l'espace $E \times E$, applications coordonnées X et Y , loi $P = \mu \otimes \mu$. Une propriété des mesures gaussiennes, déjà connue des probabilistes babyloniens, affirme que l'application $(X,Y) \mapsto \left(\frac{X-Y}{\sqrt{2}}, \frac{X+Y}{\sqrt{2}}\right)$ de $E\times E$ dans $E\times E$ préserve P (c'est l'identité $q(x) + q(y) = q\left(\frac{x-y}{\sqrt{2}}\right) + q\left(\frac{x+y}{\sqrt{2}}\right)$ pour les formes quadratiques). Nous avons donc si $s \le t$

$$(15) \qquad P\{N(X) \le s\}P\{N(Y) > t\} = P\{N\left(\frac{X-Y}{\sqrt{2}}\right) \le s \ , \ N\left(\frac{X+Y}{\sqrt{2}}\right) > t\}$$

mais les conditions au second membre ne peuvent être réalisées - du fait que N est une pseudo-semi-norme - que si

$$N(X) > \frac{t-s}{\sqrt{2}} \ , \qquad N(Y) > \frac{t-s}{\sqrt{2}} \ .$$

En effet, supposons par exemple $N(X) \le \frac{t-s}{\sqrt{2}}$. Alors $N\left(\frac{X+Y}{\sqrt{2}}\right) \le \frac{t-s}{2} + N\left(\frac{Y}{\sqrt{2}}\right)$ de sorte que $N\left(\frac{X+Y}{\sqrt{2}}\right) > t$ entraîne $N\left(\frac{Y}{\sqrt{2}}\right) > \frac{t+s}{2}$. Alors $N\left(\frac{X-Y}{\sqrt{2}}\right) + N\left(\frac{X}{\sqrt{2}}\right) > \frac{t+s}{2}$, soit $N\left(\frac{X-Y}{\sqrt{2}}\right) > \frac{t+s}{2} - \frac{t-s}{2} = s$, de sorte que (15) n'est pas réalisé. Finalement, il vient pour la fonction $n(t) = P\{N(X) > t\}$,

$$(16) \qquad (1 - n(s))n(t) \le n^2\left(\frac{t-s}{\sqrt{2}}\right) \ .$$

Le fait que N est μ-p.s. finie signifie que $n(+\infty) = 0$. Prenons s tel que $1 - n(s) = q > \frac{1}{2}$, et posons

$$t_k = s(\sqrt{2} + 1)(2^{(k+1)/2} - 1) \qquad (t_o = s)$$

suite à croissance bien rapide. Nous avons $t_{k+1} = s + t_k\sqrt{2}$, donc d'après (16)

$$qn(t_{k+1}) \leq n^2(t_k) \qquad n(t_o) = 1 - q \ ,$$

d'où par récurrence

$$n(t_k) \leq q\left(\frac{1-q}{q}\right)^{2^n} .$$

Comme $\frac{1-q}{q} < 1$, c'est une décroissance très rapide, et le reste est facile. La démonstration du théorème 1 est achevée.

(Il est bon de savoir que la probabilité pour que $N(x) = +\infty$ ne peut être que 0 ou 1 - résultat également dû à Fernique.)

BIBLIOGRAPHIE

[1] R. M. DUDLEY - Sample functions of the Gaussian process, Ann. Prob., 1 (1973), p. 66-103.

[2] X. FERNIQUE - Régularité des trajectoires des fonctions aléatoires gaussiennes, Cours de l'Ecole d'été de St Flour, 1974. A paraître dans les Lecture Notes in Math., Springer.

[3] X. FERNIQUE - Intégrabilité des vecteurs gaussiens, C.R.Acad. Sc., 270 (1970), p. 1698-1699.