Laurent Schwartz

La vie et l'œuvre d'Andréi Kolmogorov

C. R. Acad. Sci. Paris Sér. Gén. Vie Sci. 6 (6) (1989), p. 573–581

Extrait des *Œuvres de Laurent Schwartz*
publiées par la Société mathématique de France, 2011.

1 27

2

Andréi Nikolaievitch KOLMOGOROV
par Laurent SCHWARTZ

A. N. Kolmogorov est né le 25 avril 1903, à Tambov. Il était membre de l'Académie des Sciences de l'U.R.S.S. depuis 1939 (à l'âge de 36 ans), de la Royal Society of London depuis 1964, et membre associé étranger de notre Académie depuis le 16 mai 1966. Il a eu le prix Wolf en 1980. Il est décédé le 20 octobre 1987, après une vieillesse prolongée et pénible. Ses travaux portent sur de nombreuses branches des mathématiques. Nous nous concentrerons ici essentiellement sur deux : le calcul des probabilités et les systèmes dynamiques (au sens large), où sa contribution a été la plus déterminante.

I) LE CALCUL DES PROBABILITES.

On peut dire que les probabilités modernes ont été fondées avant tout par A. Kolmogorov et Paul Lévy ; c'est Kolmogorov qui en a construit les fondements. Il a d'abord trouvé en quelques années (1925-1935) un grand nombre de très beaux théorèmes.

(I.1) Soit $\sum_{n=1}^{\infty} X_n$ une série de variables aléatoires réelles indépendantes, de moyennes nulles pour simplifier, d'écarts types σ_n. L'écart type de $S_n = X_1 + X_2 + \ldots + X_n$ est alors $\tau_n = (\sigma_1^2 + \ldots + \sigma_n^2)^{1/2}$, et l'inégalité classique (très élèmentaire) de Bienaymé-Tchebyschev dit que $\mathbf{P}\{|S_n| > C\}$ ($\mathbf{P}$ = probabilité) $\leq \tau_n^2/C^2$. Kolmogorov montra une inégalité beaucoup plus fine : $\mathbf{P}\{\operatorname{Sup}_{k \leq n} |S_k| > C\} \leq \tau_n^2/C^2$; on l'appelle aujourd'hui l'inégalité de Kolmogorov, elle a été prolongée à beaucoup d'autres situations, notamment aux inégalités sur les martingales (Doob). Il en déduisit (avec Khintchine) le fameux théorème des 3 séries (1930), condition nécessaire et suffisante pour que la série $\sum_{n=1}^{\infty} X_n$ (où des X_n n'ont plus forcément une moyenne nulle) soit presque sûrement convergente ; c'est que les 3 séries

$$\sum_{n=1}^{\infty} \mathbf{P}\{|X_n| > 1\},$$

$$\sum_{n=1}^{\infty} \mathbf{E}(X_n 1_{\{|X_n| \leq 1\}}),$$

$$\sum_{n=1}^{\infty} (\mathbf{E}(X_n^2 1_{\{|X_n| \leq 1\}}) - (\mathbf{E}(X_n 1_{\{|X_n| \leq 1\}}))^2)$$

convergent, où $\mathbf{E}$ est l'espérance ou valeur moyenne. Si ces conditions ne sont pas réalisées toutes les trois, la série $\sum_{n=1}^{\infty}$ est presque sûrement divergente.

(I.2) Ce fait intervient souvent : il y a des évènements dont on peut dire **a priori** que leur probabilité ne peut être que 0 ou 1, comme la convergence d'une série de variables aléatoires indépendantes, mais bien d'autres encore (par exemple la loi forte des grands nombres : S_n/n a presque sûrement une limite pour $n \to +\infty$ ou presque sûrement n'en a pas, suivant les cas. C'est Kolmogorov qui a trouvé que, pour des variables aléatoires indépendantes équidistribuées, l'existence de la valeur moyenne est nécessaire et suffisante pour que la loi forte des grands nombres soit vérifiée). Grossièrement parlant, ce sont les évènements relatifs à la suite X_n qui, quel que soit k, ne dépendent que des $X_n, n \geq k$. Les processus de Markov utilisent des résultats analogues (pour $t \to 0$ ou $t \to +\infty$, t étant le temps). Ces résultats ont profondément modifié la vision des probabilistes ; on ne cherche presque plus, en probabilités théoriques, des résultats sur des jeux de dés ou de cartes, où un évènement a une certaine probabilité comme 1/2 ou 3/4, on démontre que certains évènements ou leur négation sont "presque sûrs", c'est à dire ont la probabilité 1. On n'écrit même plus aujourd'hui "presque sûrement", mais ps., ou même rien du tout. On dira par exemple : $\sum_{n=1}^{\infty} X_n$ "est convergente", pour dire que la probabilité de sa convergence est 1.

(I.3) La loi du logarithme itéré s'énonce (dans le cas le plus simple) comme suit : dans un jeu illimité de pile ou face, où les X_n sont une suite de variables indépendantes prenant la valeur $+1$ et la valeur -1 avec des probabilités 1/2, X_n a une moyenne nulle et un écart type 1, donc $S_n = X_1 + X_2 + \ldots + X_n$ a une moyenne nulle et un écart type $\sqrt{n}$: donc, pour un n fixé, $|S_n|$ a une probabilité faible de dépasser de beaucoup $\sqrt{n}$. Mais il s'agit ici de considérer la suite des n. Si alors $0 < C < 1$, il est presque sûr que l'inégalité $|S_n| \geq C\sqrt{2n \log\log n}$ est vérifiée pour une infinité de valeurs de n, et si $C \geq 1$, il est presque sûr qu'elle n'est vérifiée que pour un nombre fini de valeurs de n (encore une probabilité 0 ou 1).

Cette loi remarquable a été trouvée par le mathématicien soviétique Khintchine en 1924. Paul Lévy et Kolmogorov y apportèrent des compléments très fins, en remplaçant aussi $S_n = \sum_{k=1}^{n} X_k$ par B_t, le mouvement brownien. Kolmogorov a défini la classe supérieure et la classe inférieure des fonctions $\varphi > 0$ pour des $t \to 0$, selon que $\mathbf{P}\{\mathrm{Inf}\{t > 0\,; |B_t| > \sqrt{t}\varphi(1/t)\} > 0\} = 1$ ou 0 (i.e. pour $t \to 0$, $|B_t|$ finit par être $\leq \sqrt{t}\varphi(1/t)$ ou est une infinité de fois $> \sqrt{t}\varphi(1/t)$). Alors φ est de la classe supérieure ou inférieure selon que $\int^{\infty} t^{-1}\varphi(t)e^{-\varphi^2(t)/2}dt$ est $< +\infty$ ou $= +\infty$. Ce n'est qu'un cas particulier de nombreuses autres inégalités, notamment pour le module de continuité du mouvement brownien.

Certains de ces résultats de Kolmogorov ont été étroitement mêlés à ceux de Paul Lévy, autour des années avant ou après 30. Ils avaient une très active correspondance, et ils se sont rencontrés à Paris. Il existait une sorte de centre de probabilités Paris-Moscou, malgré la longueur de la correspondance et l'absence de téléphones. Il est arrivé à Paul Lévy de retrouver (sans les connaitre) certains des résultats antérieurs de Kolmogorov (ou de Khintchine), par exemple le théorème des 3 séries, la loi du logarithme itéré, la loi 0 ou 1 ; dans ce cas, il collabora ensuite avec eux sur ces sujets et y apporta d'importants compléments neufs. Ce fut souvent au contraire Paul Lévy qui fut l'initiateur ; dans la théorie des lois indéfiniment divisibles et des lois stables, où travaillèrent Khintchine, Kolmogorov, Paul Lévy, ce sont les résultats de Paul Lévy qui vont le plus loin.

(I.4) Mais sans doute le plus important résultat de Kolmogorov est-il d'avoir donné un fondement mathématique rigoureux aux probabilités, qui auparavant étaient peut-être plus physiques que mathématiques ; c'est de ce fondement qu'on ne s'est depuis plus jamais passé. Il a introduit l'ensemble Ω des épreuves, sa tribu fondamentale $\mathcal{O}$, la probabilité fondamentale $\mathbf{P}$ sur $(\Omega, \mathcal{O})$, comme mesure de masse 1 sur la tribu $\mathcal{O}$. Une épreuve est un point ω de Ω, un évènement est un ensemble A $\mathbf{P}$-mesurable de Ω, sa probabilité est $\mathbf{P}(A)$, une variable aléatoire X est une fonction $\mathbf{P}$-mesurable sur Ω (par exemple réelle ; alors $\mathbf{P}\{X > 0\}$ est $\mathbf{P}\{\omega \in \Omega; X(\omega) > 0\}$). Deux évènements A, B, sont indépendants si $\mathbf{P}(A \cap B) = P(A)P(B)$. D'où les probabilités conditionnelles, espérances conditionnelles, etc..., les probabilités modernes reposent alors sur des bases solides, au fond sur la théorie de la mesure. Beaucoup de gens (non probabilistes) disent : "la science des probabilités n'existe pas, c'est un cas particulier de la théorie de la mesure." Ce n'est pas une bonne manière de voir les choses ; la probabilité commence là où le conditionnement commence, et elle donne alors une foule de résultats profonds, peu en rapport avec les résultats usuels de la théorie de la mesure (avec, le plus souvent, des probabilités 0 ou 1) ; un livre sur la théorie de la mesure ne parlera pas de la loi des grands nombres ou du mouvement brownien. L'article remarquable où Kolmogorov introduisit ces fondements est écrit en allemand : "Grundbegriffe der Wahrscheinlichkeits Rechnung und ihre Anwendungen in der Statistik und theoretischen Physik" paru dans les Ergebnisse, Springer, 1933 ; il fut traduit en anglais en 1950. Il avait été résumé dans un article en 1929 (en français).

Parmi les 23 problèmes de Hilbert, posés par ce dernier au Congrès International des Mathématiciens, Paris, 1900, le sixième est : "Axiomatiser celles des sciences physiques où les mathématiques jouent un rôle important" ; parmi ces sciences physiques, il songeait sans doute aux probabilités. L'axiomatique de Kolmogorov est apprise depuis un certain temps aux éléves des lycées, en tout cas aux étudiants d'université en premier cycle scientifique.

(I.5) Il a apporté des résultats fondamentaux dans les processus de Markov sur un espace E. La probabilité de transition $\mathbf{P}(s,t;x), s \leq t$, loi du processus à l'instant t lorsqu'à l'instant s on sait qu'il est en x, vérifie la relation de **Chapman-Kolmogorov** (1931) : si $r \leq s \leq t$,

$$\mathbf{P}(r,t;x) = \int_E \mathbf{P}(r,s;x)(dy)\mathbf{P}(s,t;y).$$

Grâce à cette relation, on peut construire la loi globale du processus partant de x à l'instant 0. Mais il faut pour cela justement un théorème de Kolmogorov de 1933 : si on se donne un système projectif de probabilités sur $\mathbf{R}^I$, c'est-à-dire un système cohérent de projections sur les produits finis $\mathbf{R}^J$, J fini $\subset I$, il définit une probabilité $\mathbf{P}$ sur $\mathbf{R}^I$ muni de la tribu engendrée par les projections. Ceci reste vrai si on remplace $\mathbf{R}$ par un espace complètement régulier E pourvu que les probabilités projections soient de Radon. Mais, dans ses premiers travaux sur les chaînes de Markov, Kolmogorov n'avait pas encore démontré ces théorèmes, et il a surtout travaillé sur les transitions. Ce théorème introduit des probabilités sur des espaces de dimension infinie (qui ont été utilisées avec un grand succès par Doob et sont constamment utilisées aujourd'hui). D'autre part, il est apparu depuis longtemps (déjà pour le mouvement brownien, qui est markovien) qu'il y a un lien étroit entre le processus et l'équation de la chaleur généralisée ou équation de Focker-Planck ; c'est encore au centre de très nombreuses recherches modernes. Kolmogorov a démontré divers théorèmes à ce sujet ; dans des conditions convenables (qui ont été considérablement étendues depuis), il a montré en 1939 que (pour $E = \mathbf{R}$) $\mathbf{P}(s,t;x)dy = p(s,t;x,y)dy$ et que p satisfait à une équation aux dérivées partielles "rétrograde"

$$\frac{\partial p}{\partial s} = -a(s,x)\frac{\partial^2 p}{\partial x^2} - b(s,x)\frac{\partial p}{\partial x}$$

et une équation directe

$$\frac{\partial p}{\partial t} = \frac{\partial^2}{\partial y^2}(a(t,y)p) - \frac{\partial}{\partial y}(b(t,y)p) \ .$$

C'est aussi lui qui, dans ce cadre, a introduit le semi-groupe de Markov et son générateur infinitésimal (de façon moins précise qu'on ne le fait aujourd'hui). Nous ne donnons ici qu'une partie très petite des résultats très riches de Kolmogorov sur les processus de Markov.

(I.6) Kolmogorov a aussi d'autres résultats très originaux en probabilités : une théorie du risque (étudiée avec (I.4)) avec applications à l'actuariat et aux assurances, un test statistique dit de Smirnov-Kolmogorov (la théorie de la prédiction des processus stationnaires a été inaugurée par Kolmogorov et N. Wiener dans les années 40), une étude des suites aléatoires. Il a publié des suites de nombres choisis au hasard, et, dans les dernières années de sa vie active, il a cherché une définition d'une suite **donnée** de nombres dont on "peut" dire qu'elle a été choisie au hasard. Cette importante recherche rejoint une partie des vieilles idées de von Mises, et se rattache à ses travaux sur les systèmes dynamiques.

(I.7) Terminons par un résultat relativement simple de Kolmogorov, mais utilisé partout en probabilités. Soit Φ une fonction aléatoire, application localement α-holdérienne d'une variété différentiable V de dimension N dans $L^p(\Omega,\mathcal{O},\mathbf{P})$, $1 \leq p \leq +\infty$, $\Phi \in C^\alpha(V;L^p(\Omega,\mathcal{O},\mathbf{P}))$; elle peut être représentée par une fonction Φ presque sûrement localement β-holdérienne, $\Phi \in L^p(\Omega,\mathcal{O},\mathbf{P};C^\beta(V))$, dès que $\frac{\alpha-\beta}{N} > \frac{1}{p}$. Les applications sont très nombreuses. Prenons par exemple le mouvement brownien B ; ici $V = \mathbf{R}_+$, $B_t - B_s$ est gaussienne de paramètre $\sqrt{t-s}$, donc $B \in C^{1/2}(\mathbf{R}_+;L^p(\Omega,\mathcal{O},\mathbf{P}))$, pour $p < +\infty$ arbitraire ; si donc $\beta < \frac{1}{2}$, $\frac{\alpha-\beta}{1} > \frac{1}{p}$ pour p assez grand, donc B est presque sûrement localement β-holdérienne pour tout $\beta < 1/2$, et même $B \in L^p(\Omega,\mathcal{O},\mathbf{P};C^\beta(\mathbf{R}_+))$ pour tout p fini. Par des méthodes directes, on trouve beaucoup mieux (loi du logarithme itéré, (I.3)), mais la méthode de Kolmogorov est sans doute la plus simple pour démontrer la continuité presque sûre du mouvement brownien. On l'utilise dans les travaux les plus récents pour trouver les points multiples de courbes browniennes.

Cet ensemble réussit difficilement à montrer le rôle de très grand pionnier qu'à joué Kolmogorov dans le développement des probabilités modernes.

II LES SYSTEMES DYNAMIQUES.

(II.1) La théorie ergodique, née de la mécanique statistique de Boltzman et Gibbs à la fin du 19ème siècle, a commencé son grand développement avec le "théorème ergodique statistique" de John von Neumann et le "théorème ergodique ponctuel" de Georges D. Birkhoff. On se donne un espace mesuré $(X, \mathcal{B}, \mu)$, X ensemble, $\mathcal{B}$ tribu de parties de X, μ mesure ≥ 0 sur $\mathcal{B}$, et on étudie les itérés d'un endomorphisme T de cette structure, $T(\mu) = \mu$. On dit que T est μ-ergodique ou μ T-ergodique si, pour tout $B \in \mathcal{B}$, T-invariant, $\mu(B)$ ou $\mu(X \setminus B) = 0$. Le théorème ergodique de Birkhoff dit que, si f est réelle μ-intégrable, $\frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x))$ converge, pour $n \to +\infty$, vers une limite $f^*(x)$, pour μ-presque tout x ; f^* est T-invariante, μ-intégrable, et, pour tout $B \in \mathcal{B}$ T-invariant, tel que $\mu(B) < +\infty$, $\int_B f^*(x)\mu(dx) = \int_B f(x)\mu(dx)$; si μ est une probabilité, et si T est ergodique, ce que nous supposerons désormais, f^* est la constante $\int_X f(x)\mu(dx)$, et c'est cela qui donne un sens mathématique aux hypothèses ergodiques du siècle dernier. On peut, moyennant une condition de continuité, remplacer le semi-groupe $(T^n)_{n\geq 0}$ par un semi-groupe à temps continu ou flot $(T^t)_{t\geq 0}$, $T^0 = I$, $T^{s+t} = T^t T^s$, et on remplace $\frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x))$ par $\frac{1}{t}\int_0^t f(T^s(x))ds$ pour $t \to +\infty$. De nombreux et très beaux théorèmes ont été démontrés sur ce thème, notamment le théorème ergodique du rapport par Chacon-Ornstein (1960). Des liens étroits avec les processus de Markov ont été trouvés. Un problème important est celui des isomorphismes de systèmes ; deux systèmes

$$(X_1, \mathcal{B}_1, \mu_1, (T_1^t)_{t\geq 0}), (X_2, \mathcal{B}_2, \mu_2, (T_2^t)_{t\geq 0})$$

sont spatialement isomorphes s'il existe une bijection Φ de X_1 sur X_2, telle que $\Phi\mathcal{B}_1 = \mathcal{B}_2, \Phi\mu_1 = \mu_2, \Phi \circ T_1^t = T_2^t \circ \Phi$ ou $T_2^t = \Phi T_1^t \Phi^{-1}$ (tout cela bien sûr, aux ensembles négligeables près) ; un problème central de la théorie ergodique est de classifier les classes d'isomorphismes de systèmes ergodiques.

Pour $(X, \mathcal{B}, \mu)$ donné, un automorphisme de la structure est dit automorphisme de Kolmogorov ou K-automorphisme s'il existe une sous-tribu $\mathcal{B}_0$ de $\mathcal{B}$ telle que $T(\mathcal{B}_0) \not\supseteq \mathcal{B}_0$, $\bigvee_{n\in\mathbf{Z}} T^n(\mathcal{B}_0) = \mathcal{B}$, $\bigwedge_{n\in\mathbf{Z}} T^n(\mathcal{B}_0) = \{\emptyset, X\}$ (tribu triviale ; aux ensembles μ-négligeable près !). Les K-automorphismes jouent un rôle central dans la classification recherchée. Ils ont une propriété de "mélange" (mixing en anglais) : pour tous
$A, B \in \mathcal{B}, \lim_{n\to+\infty} \mu(\Phi^n(A) \cap B)$
$= \mu(A)\mu(B)$. Le "shift" (le décalage) de Bernoulli est défini comme suit : $(Y, \mathcal{A}, \mu)$ est un espace probabilisé, et on pose, $(Y^*, \mathcal{A}^*, \mu^*) = (Y^{\mathbf{Z}}, \otimes^{\mathbf{Z}}\mathcal{A}, \otimes^{\mathbf{Z}}\mu)$; le décalage est la transformation $(y_n)_{n\in\mathbf{Z}} \mapsto (y_{n+1})_{n\in\mathbf{Z}}$; le décalage T est $\otimes^Z\mu$-ergodique (non seulement le semi-groupe $(T^n)_{n\in\mathbf{N}}$ mais le groupe $(T^n)_{n\in\mathbf{Z}}$), et c'est un K-isomorphisme (trivialement, avec $\mathcal{B}_0 = \otimes^{-\mathbf{N}}\mathcal{A}, T\mathcal{B}_0 = \otimes^{-\mathbf{N}+1}\mathcal{A}$).

(II.2) L'entropie est intervenue au 19ème siècle en thermodynamique, puis dans les travaux mathématiques de Boltzman et Gibbs ; elle n'a pris un sens rigoureux qu'avec l'introduction de l'entropie par Kolmogorov ; on l'appelle aussi entropie de Kolmogorov-Sinai parce que ces deux mathématiciens (Sinai est un élève de Kolmogorov) ont vite travaillé ensemble sur ce sujet, et c'est d'ailleurs tout une école de mathématiciens soviétiques, autour de Kolmogorov, qui l'on fait progresser considérablement. Soit $\xi = (A_k)_{k\in\mathbf{N}}$ une partition finie ou dénombrable, mesurable, de X ; si T est un endomorphisme de $(X, \mathcal{B}, \mu)$, on pose $H(\xi) = -\sum_k \mu(A_k)\log\mu(A_k)$, puis $h(\xi, T) = \lim_{n\to+\infty} \frac{1}{n} H(\vee_{k=0}^{n-1} T^{-k}(\xi))$, qui est finie si $H(\xi) < +\infty$. Alors l'entropie $h(T)$ de T par rapport à μ est $\sup\limits_{\xi \text{finie}} h(\xi, T)$. C'est un invariant par rapport aux isomorphismes spatiaux de $(X, \mathcal{B}, \mu)$. Nous ne pouvons pas citer ici tous les remarquables résultats sur l'entropie. Si T est le décalage de Bernoulli d'un système $(X, \mathcal{A}, \mu)$, où X est dénombrable $= \mathbf{N}$, $\mathcal{A}$ la tribu discrète, μ une probabilité sur $\mathbf{N}$, $\mu\{n\} = p_n, h(T) = -\sum_n p_n \log p_n$, c'est l'opposée de la quantité d'information de Shannon de $(\mathbf{N}, \mu)$. Kolmogorov a toujours souligné combien il avait été inspiré par les travaux de Shannon (1948) en théorie de l'information. Par ce théorème, Kolmogorov a prouvé pour la première fois qu'il y a une infinité continue de décalages de Bernoulli discrets non spatialement isomorphes. Même pour des décalages relatifs à un Y à deux éléments, l'existence de deux décalages non isomorphes était restée inabordable avant l'introduction de l'entropie de Kolmogorov. Un remarquable et très difficile théorème d'Ornstein (1970) dit que deux décalages de Bernoulli discrets sont spatialement isomorphes si et seulement s'ils ont même entropie.

Parmi les découvertes les plus importantes de Kolmogorov en mécanique, il y a son célèbre théorème de la persistance des mouvements quasi-périodiques par des petites perturbations de systèmes hamiltoniens complètement intégrables ; ce fut le premier nouveau pas décisif depuis les "Méthodes nouvelles de la mécanique céleste" de Poincaré. Il fut annoncé dans son exposé au Congrès international des Mathématiciens d'Amsterdam en 1954. Il a été ensuite considérablement développé par V.I. Arnold et J. Moser, il est devenu le théorème KAM (Kolmogorov-Arnold-Moser) dont l'impact ne cesse de croître.

Dans un ordre d'idées voisin, D. Anosov introduisit une classe de difféomorphismes importants sur une variété (on les appelle les difféomorphismes d'Anosov) ; si un tel difféomorphisme a une mesure invariante C^∞, c'est un K-automorphisme.

Tous ces développements très succints montrent combien important fut l'apport de Kolmogorov dans la théorie des systèmes dynamiques ; il a ouvert des voies nouvelles, dont l'exploration n'est pas terminée.

(II.3) Kolmogorov s'est depuis longtemps beaucoup intéressé à la turbulence des flots en dynamiques des fluides. Il a trouvé en particulier (1942), dans le cas d'isotropie locale, pour des nombres de Reynolds très grands, une valeur asymptotique du spectre d'énergie, $E(k) \simeq k^{-5/3}$ pour des grands nombres d'ondes k. Cela correspond très bien avec les valeurs expérimentales. En 1962-63, il a donné de nouvelles contributions à ce sujet ; ses travaux sur la turbulence sont d'un intérêt capital.

(II.4) Si X est un espace métrique borné, $A \subset X, N_\varepsilon(A)$ est le nombre minimum de boules de rayon $\varepsilon > 0$ de centres dans X, permettant de recouvrir A ; l'ε-entropie de A est $H_\varepsilon(A) = \log N_\varepsilon(A)$, et on s'intéresse à sa croissance quand $\varepsilon \to 0$; l'exposant d'entropie de A est $\lim . \sup ._{\varepsilon \to 0} \log H_\varepsilon(A)/\log \frac{1}{\varepsilon}$. Il est lié à l'épaisseur et à la complexité de $A \subset X$; d'ailleurs Kolmogorov a dit là qu'il cherchait des nombres exprimant la complexité d'un ensemble. Si X est la boule unité d'un espace euclidien de dimension n, A une boule concentrique plus petite, $H_\varepsilon(A) = n$. En raisonnant ensuite sur un espace de Banach et des sous-espaces de dimension $n \to +\infty$, on examine la croissance de suites d'entropies en fonction de n, qui définit la dimension approximative, introduite indépendamment par A. Pelczynski. On en déduit des critères de continuité de certains processus gaussiens (très dépassés depuis), ou des critères de compacité ou de nucléarité d'opérateurs (théorèmes de Mitjagin, 1961). On utilise aussi l'ε-entropie dans des codages de messages.

III. AUTRES TRAVAUX

Ils sont très nombreux, et touchent des sujets très divers.

(III.1) Séries de Fourier.

Ce sont là ses premiers travaux (1926), il a trouvé le premier une fonction intégrable périodique dont la série de Fourier est partout divergente (alors que si la fonction est dans L^p, $1 < p < +\infty$, sa série de Fourier converge presque partout (Carleson 1966)). La démonstration remplit une note aux C.R.A.S. Dans le même ordre d'idées, il a indiqué comment trouver la fonction conjuguée (au sens de l'analyse harmonique) d'une fonction non intégrable. La série de Fourier est assez généralement de convergence mauvaise ; cette convergence a suscité une énorme quantité de travaux. L'exemple de Kolmogorov est un des premiers cas très "méchants" signalés.

(III.2) A peu près en même temps que A. Gorny et H. Cartan, il a étudié les majorations des dérivées successives d'une fonction, en liaison avec la quasi-analyticité (1939).

(III.3) Il n'a pas beaucoup travaillé en topologie algébrique, mais il a introduit des notions fondamentales. L'homologie était définie de différentes manières ; après l'homologie de Vietoris (1927) vint l'homologie d'Alexander-Kolmogorov pour des espaces localement compacts (1936) ; mais surtout il introduit dans des cas très généraux l'anneau de cohomologie, cohomologie de Kolmogorov-Spanier (1948). En fait la cohomologie de Čech et celle des faisceaux (Leray) ont joué ensuite un rôle plus important.

(III.4) Le 13ème problème de Hilbert (parmi les 23 posés par Hilbert au Congrès international des mathématiciens, Paris, 1900) est le suivant : montrer l'impossibilité de résoudre l'équation algébrique du 7ème degré

$$x^7 + ax^3 + bx^2 + cx + 1 = 0,$$

par des composées de fonctions continues de 2 variables, à partir des coefficients a, b, c. A la suite de très intéressants travaux de A.G. Vitushkin (1954-55), Kolmogorov s'est intéressé à ce sujet. En 1956-57, Kolmogorov et son élève (très jeune à l'époque) V.I. Arnold ont montré, de manière très compliquée, que toute fonction continue de n variables peut être représentée par des composées de fonctions continues de 2 variables (si $n = 3$, on utilise 9 composées, $F(x_1, x_2, x_3) = \sum_{i=1}^{9} h_i(g_i(x_1, x_2), x_3))$. Puis Kolmogorov a montré élémentairement qu'on peut prendre des composées de fonctions d'une variable, au nombre de 7 pour $n = 3$, avec la forme canonique, $F(x_1, x_2, x_3) = \sum_{i=1}^{7} h_i(g_{i,1}(x_1) + g_{i,2}(x_2) + g_{i,3}(x_3))$, et les $g_{i,j}$ peuvent être choisies une fois pour toutes, indépendamment de F (1963).

(III.5) Il y a d'autres travaux divers, notamment en logique, théorie de la dimension des espaces topologiques, dynamique des populations,... Il s'est aussi beaucoup intéressé à l'enseignement élémentaire et à la pédagogie, et a écrit plusieurs articles à l'usage des enseignants et des parents. C'est lui qui a fondé les "écoles spéciales", pour des jeunes de l'enseignement secondaire, doués en mathématiques, y enseignant lui-même ainsi que d'autres universitaires. Ces écoles, qui ont groupé plusieurs milliers d'élèves, ont donné à l'URSS des centaines de mathématiciens. Elles ont malheureusement été fermées, parce qu'elles rassemblaient trop d'enfants de familles juives ou libérales ; cependant que cet exemple était imité et a subsisté dans des pays socialistes, mais aussi à Taiwan et aux Etats-Unis, à une très grande échelle.

Dans les années 1985-87, on a édité à Moscou ses oeuvres complètes en 3 volumes, avec des commentaires. Souvent ses découvertes les plus importantes sont décrites dans de très courts articles, mais de manière très claire : l'ensemble des volumes et relativement peu épais ; c'est un des traits de son caractère scientifique. Il a été un des très grands formateurs de la jeunesse mathématique : parmi ses éléves, citons A.I. Malcev, S.M. Nikolskij, B.V. Gredenko, F.M. Gelfand, G.E. Shilov, A.M. Obukhov, M.S. Pinsker, J.V. Prokhorov, R.L. Dobrushin, V.A. Uspenskij, V.M. Alexeiev, V.D. Ierokhin, J.G. Sinai, V.M. Tikhomirov et V.I. Arnold (dans l'ordre plus ou moins chronologique).

Kolmogorov était un homme particulièrement souriant, ouvert, généreux. Il s'est spécialement donné du mal pour que les jeunes mathématiciens doués puissent arriver vite à des positions de responsabilité. Il a toujours travaillé à l'Université de Moscou depuis 1925, il a eu la chaire de probabilités de 1938 à 1966, il a été de 1966 à 1976 le Directeur du laboratoire interdépartemental de statistique (qu'il avait créé) et de 1976 à 1980 il a eu la chaire de statistique, ensuite enfin celle de logique. Ayant reçu le prix Bolzano, il en a consacré une part substantielle à l'achat de livres pour la Bibliothèque de statistique. Après un voyage en France, il y a laissé de l'argent personnel pour qu'un de ses élèves puisse venir quelques mois en France. Il a traversé de manière variée la période terrible de la dictature soviétique. On ne peut que regretter une faiblesse insigne sous Brejnev : il a publié le 4 février 1974, dans la Pravda, une lettre exprimant sa satisfaction de l'exil de Soljénitsyne et de sa perte de la citoyenneté. Par contre, sous Staline, il a oeuvré au maximum pour protéger la mathématique de la répression idéologique. Et, aussi sous Staline, il a publié dans les comptes rendus de l'Académie des Sciences de l'URSS, en 1940, une note sur "la nouvelle confirmation des lois de Mendel" (volume 27, pages 38-42) où, en pleine montée de Lissenko, il montra ouvertement que les expériences de ce dernier ne faisaient que confirmer les lois de Mendel.

On a souvent dit qu'il était un "prince des mathématiques", sûrement parmi les plus grands. Pionnier de première grandeur dans des domaines variés (mais présentant presque toujours une certaine unité), il a été connu et admiré pendant près de 60 ans dans le monde entier. Notre Académie a perdu en lui un de ses membres associés les plus grands et les plus prestigieux.

Laurent SCHWARTZ

Je remercie Jean-Marie Strelcyn, qui m'a donné de précieux conseils et renseignements.