ANNALI DELLA
SCUOLA NORMALE SUPERIORE DI PISA
Classe di Scienze

PAUL LÉVY

Sur les intégrales dont les éléments sont des variables aléatoires indépendantes

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 2e série, tome 3, nº 3-4 (1934), p. 337-366

<http://www.numdam.org/item?id=ASNSP_1934_2_3_3-4_337_0>

SUR LES INTÉGRALES DONT LES ÉLÉMENTS SONT DES VARIABLES ALÉATOIRES INDÉPENDANTES

par PAUL LÉVY (Paris) [1].

§ 1. - Introduction.

L'étude des séries dont les termes sont des variables aléatoires indépendantes, en raison de ses nombreuses applications, a été un des problèmes les plus étudiés par les fondateurs du calcul des probabilités, BERNOULLI, LAPLACE, GAUSS, POISSON, CAUCHY, qui se sont attachés à découvrir les propriétés de la loi dont dépend la somme S_n des n premiers termes. M. BOREL a en 1909 renouvelé la question en montrant l'intérêt qu'il y a à ne pas se contenter de considérer chaque valeur de n indépendamment des autres, mais à étudier dans chaque cas particulier la suite des valeurs de S_n, et mettre en évidence ses caractères probables. Ce problème fut l'origine d'importants travaux de MM. CANTELLI, KHINTCHINE, KOLMOGOROFF, que nous avons nous-même complétés sur certains points [2].

---

[1] Les principaux résultats établis dans le présent travail ont été énoncés dans trois Notes présentées à l'Académie des Sciences le 26 février, le 26 mars et le 7 mai 1934. Ils ont en outre fait l'objet d'exposés oraux au Séminaire de M. Hadamard, le 16 mars, à la Société Mathématique du Danemark, le 9 avril, et à la Société Mathématique de France le 25 avril.

[2] PAUL LÉVY: *Sur les séries dont les termes sont des variables éventuelles indépendantes* [Studia Mathematica, t. III (1931), pp. 119-153]. Ce travail sera désigné dans la suite par l'abréviation *Séries*. Nous signalons ci-dessous quelques errata:

p. 123, dernière formule (4)
au lieu de: $\sigma > 0$, lire: $\sigma = 0$.

p. 128, ligne 7 en partant du bas
au lieu de: $P\{x_\nu = \bar{x}_\nu\}$, lire: $P\{x_\nu \neq \bar{x}_\nu\}$.

p. 129, avant dernière ligne
au lieu de: n arbitrairement grand, lire: m arbitrairement grand.

p. 131, § 8, 1°, a)
au lieu de: $\sum E'\{x_\nu^2\}$, lire: $\sum E'\{x_\nu\}$ et $\sum D'\{x_\nu\}$.

p. 148 et 149, Th. XI et XII.

Ces théorèmes supposent l'élimination d'un terme indépendant du hasard. Ainsi l'énoncé correct du th. XI est le suivant: on peut trouver une suite de constantes a_n telle que la loi de probabilité dont dépend la somme $\sum (x_n - a_n)$ soit indépendante de l'ordre des termes.

Les mêmes problèmes se posent si, au lieu d'une variable discontinue n, on introduit une variable continue t; au lieu d'une série, on doit considérer une intégrale de STIELTJES, que nous désignerons par $x(t)$. Or la théorie de ces intégrales dont les éléments sont des variables aléatoires indépendantes n'a guère été étudiée que dans un cas très particulier, celui auquel conduit naturellement l'étude du mouvement brownien ou de la théorie cinétique des gaz: l'accroissement Δx relatif à chaque intervalle Δt dépend de la loi de GAUSS, la valeur probable de $(\Delta x)^2$ étant t. Du point de vue mathématique, le problème avait été considéré par BACHELIER, mais sans être nettement posé; N. WIENER (3) a le premier indiqué et justifié le passage à la limite nécessaire pour permettre l'application des formules fondamentales à des intervalles arbitrairement petits.

Dans un autre ordre d'idées, nous devons mentionner les études sur les probabilités en chaîne de MARKOFF et de MM. HADAMARD, KOLMOGOROFF, HOSTINSKY et FRÉCHET (4); le problème de l'addition de variables aléatoires indépendantes est un cas particulier de celui des probabilités en chaîne, et les auteurs cités ont traité aussi bien le cas discontinu que le cas de la variable continue t. Toutefois les résultats remarquables qu'ils ont obtenus, même complétés sur certains points non encore complètement résolus (5), ne résolvent pas le problème analogue à celui posé par M. BOREL, c'est-à-dire l'étude des caractères probables de la fonction $x(t)$.

C'est à ce problème qu'est consacré le présent travail; indépendamment de la détermination, dans le cas le plus général, de la loi $\mathfrak{L}(t)$ dont dépend $x(t)$, il contient l'étude des caractères probables de $x(t)$, spécialement au point de vue de la continuité de cette fonction. Nous pensons consacrer un autre mémoire à l'étude de questions telles que la détermination de la valeur probable de fonctionnelles de $x(t)$, ou la probabilité que certaines conditions analogues à celle de HÖLDER soient vérifiées.

Le résultat le plus remarquable du présent mémoire nous semble être le rôle que jouent les discontinuités de $x(t)$. Toutes les fois que la loi dont dépend $x(t_2)-x(t_1)$ n'est pas celle de GAUSS, même si la loi $\mathfrak{L}(t)$ varie d'une manière continue avec t de t_1 à t_2, la probabilité que $x(t)$ admette entre t_1 et t_2 des points de discontinuité (nécessairement de première espèce) est positive; elle peut être comprise entre 0 et 1 et dans ce cas le nombre de discontinuités est fini (quoiqu'on ne puisse jamais lui assigner à l'avance une borne supérieure); elle peut être égale

---

(3) N. WIENER: *Differential Space* [Publications of the Massachusetts Institute of Technology, Series II, N.° 60 (juin 1023)].

(4) M. FRÉCHET: *Solution continue la plus générale d'une équation fonctionnelle de la théorie des probabilités en chaîne* [Bulletin de la Société Mathématique de France, t. LX (1932), pp. 242-277 et t. LXI (1933), pp. 182-185].

(5) Dans le travail cité de M. FRÉCHET, il n'est question que d'un système comportant un nombre fini de cas possibles, et de lois de probabilité variant d'une manière continue avec t. La première restriction surtout est essentielle.

à 1, et dans ce cas ce nombre est infini. Les valeur de t correspondant à ces points ne sont d'ailleurs pas données d'avance, mais dépendent du hasard (6).

§ 2. - Résumé et application des résultats connus sur les séries de variables aléatoires; élimination des points de discontinuités fixes.

Nous désignerons par $P\{E\}$ la probabilité d'un évènement E, par $E\{x\}$ la valeur probable de x, par $D\{x\}=\sigma^2\{x\}$ celle de $[x-E\{x\}]^2$. Nous définirons une loi $\mathfrak{L}$, soit par sa *fonction des probabilités totales*

$$F(X)=P\{x<X\} \quad (7),$$

soit par sa fonction caractéristique $\Phi(z)$, ou le logarithme de cette fonction

$$\Psi(z)=\log \Phi(z)=\log E\{e^{izx}\}.$$

La variable z étant réelle, on a $|\Phi(z)|\leq 1$, d'où $R\Psi(z)\leq 0$, le symbole R désignant la partie réelle; on a $\Phi(0)=1$, $\Psi(0)=0$, $\Psi'(0)=iE\{x\}$, $\Psi''(0)=-\sigma^2\{x\}$, sous les réserves respectives, en ce qui concerne ces deux dernières formules, que $E\{|x|\}$ et $E\{x^2\}$ soient finis.

Pour la loi $\mathfrak{L}(t)$, nous écrirons $F(X, t)$, $\Phi(z, t)$, $\Psi(z, t)$.

Nous dirons que la loi $\mathfrak{L}(t)$ tend vers $\mathfrak{L}(t_0)$ si $F(x, t)$ tend vers $F(x, t_0)$, sauf peut-être aux point de discontinuité de cette fonction limite. Cela revient à dire que la *distance des lois* $\mathfrak{L}(t)$, *et* $\mathfrak{L}(t_0)$, définie comme étant le maximum, quand a varie, du segment intercepté par les courbes $y=F(x, t)$ et $y=F(x, t_0)$ sur la droite $x+y=a$, tend vers zéro (8).

Quand on définit la loi $\mathfrak{L}(t)$ par sa fonction caractéristique, la condition nécessaire et suffisante pour qu'elle tende vers la loi $\mathfrak{L}(t_0)$, quand t tend vers t_0, est que $\Phi(z, t)$ tende vers $\Phi(z, t_0)$, et cela uniformément dans tout intervalle fini.

---

(6) [Ajouté à la correction des épreuves]. Le résumé de ma Note du 26 février, rédigé par M. KOLMOGOROFF, a attiré mon attention sur deux Notes de M. B. DE FINETTI et deux autres de M. KOLMOGOROFF lui-même, publiées dans les Atti Accad. naz. Lincei, VI s. [t. 10, pp. 163-168 (1929); t. 12, pp. 278-282 (1930); t. 15, pp. 805-808 et 866-869 (1932)]. Ces dernières notamment contiennent la solution du problème traité dans le présent travail, dans le cas où le processus est homogène, c'est-à-dire que la loi dont dépend $x(t_2)-x(t_1)$ n'est fonction que de t_2-t_1, et où la valeur probable $E\{x^2\}$ est finie. Le résultat fondamental du présent Mémoire apparaît donc comme une extension d'un résultat de M. KOLMOGOROFF.

(7) Une loi de probabilité étant bien définie par la donnée de $F(X)$ aux points où cette fonction est continue, sa définition précise aux points de discontinuité est sans importance. Dans la suite, quand nous parlerons de la courbe $y=F(x)$, nous supposerons la continuité rétablie par un trait vertical en chaque point de discontinuité de la fonction $F(x)$.

(8) Sans introduire l'expression de *distance*, j'ai utilisé cette notion et l'inégalité triangulaire correspondante dans mon *Calcul des Probabilités*, p. 200. Dans la suite, ce livre sera désigné par l'abréviation *Probabilités*.

L'hypothèse essentielle sur laquelle repose la présente étude est que, si

$$t_0 < t_1 < \dots < t_n,$$

les différences $\Delta x_h = x(t_h) - x(t_{h-1})$ *sont des variables aléatoires indépendantes.* Nous supposerons de plus $x(0)=0$; quand t croît à partir de zéro, $x(t)$ est donc la somme d'un nombre de plus en plus grand de variables indépendantes.

D'après les propriétés connues des fonctions caractéristiques, la condition que cette hypothèse impose à la fonction $\Phi(z, t)$ est la suivante: le rapport $\frac{\Phi(z,t')}{\Phi(z,t)}$, si $0 < t < t'$, est une fonction caractéristique, celle de la loi dont dépend $x(t') - x(t)$. En d'autres termes, $\Psi(z, t)$, *considéré comme fonction de* t, *est une intégrale de Stieltjes dont tous les éléments sont des fonctions* $\Psi(z)$ (9).

La variation de la loi $\mathfrak{L}(t)$ ainsi définie a un caractère de monotonie que l'on peut préciser de diverses manières: quand t croît, $|\Phi(z, t)|$, et par suite $R\Psi(z, t)$, ne sont jamais croissants; l'expression $\sigma^2\{x(t)\}$ n'est jamais décroissante, et, si elle est finie, croît même sûrement dans tout intervalle où la loi $\mathfrak{L}(t)$ varie. Mais la forme la plus utile à connaître de ce *principe de monotonie* est *le principe d'augmentation de la dispersion*: la dispersion $l = \omega(\alpha)$, pour une loi $\mathfrak{L}$, et pour chaque valeur de α entre 0 et 1, est la borne inférieure des longueurs des intervalles (x', x'') pour lesquels

$$P\{x' \leqslant x \leqslant x''\} > \alpha.$$

Quand t croît, la dispersion n'est jamais décroissante, et croît même effectivement, au moins pour certaines valeurs de α (notamment pour les valeurs assez voisines de 1), si la loi $\mathfrak{L}(t)$ varie.

Quand on étudie, pour n infini, la loi dont dépend la somme x_n de n variables aléatoires indépendantes $u_1, u_2, \dots, u_n$, on sait (Séries, § 4 à 9) que deux cas seulement sont possibles: ou bien $\omega(\alpha)$ augmente indéfiniment pour tout α compris entre 0 et 1 (*cas de divergence*); ou bien $\omega(\alpha)$ tend vers une limite finie, pour tout α compris entre 0 et 1 (*cas de convergence*). Dans ce dernier cas, on peut trouver une suite de constantes A_n, déterminées aux termes près d'une suite convergente, telles que la loi dont dépend $x_n - A_n$ tende vers une limite. Il y a alors une probabilité égale à l'unité pour que la suite des valeurs de $x_n - A_n$ soit convergente, c'est-à-dire, en posant $a_n = A_n - A_{n-1}$, que la série $\sum (u_n - a_n)$ soit convergente. Au contraire, dans le cas de divergence, quelles que soient les constantes a_n, cette probabilité est nulle.

D'une manière plus précise, si $\sum \sigma^2\{u_n\}$ converge, on est dans le cas de convergence, et on peut prendre $a_n = E\{u_n\}$, $A_n = E\{x_n\}$. Mais *on est encore dans le cas de convergence si la convergence de* $\sum \sigma^2\{u_n\}$ *n'est empêchée*

(9) On remarque qu'il aurait suffi, pour aboutir à cette conclusion, de supposer l'hypothèse fondamentale relative aux Δx_h vérifiée pour $n = 2$.

que par des valeurs des u_n trop grandes, mais de probabilité totale arbitrairement petite; cela revient à dire qu'on peut déterminer une variable $\bar{u}_n$, ne différant de u_n que dans des cas de probabilité α_n très petite, de manière que les séries $\sum \sigma^2 \{\bar{u}_n\}$ et $\sum \alpha_n$ soient convergentes (en négligeant un nombre fini de termes, on peut alors rendre arbitrairement petite la probabilité que la suite des $\bar{u}_n$ diffère de la suite des u_n). Cette condition, due à MM. KHINTCHINE et KOLMOGOROFF, est nécessaire et suffisante pour que l'on soit dans le cas de convergence; on peut toujours prendre $a_n = E\{\bar{u}_n\}$.

Il peut être utile de donner à la condition de convergence une forme qui évite d'introduire les variables auxiliaires $\bar{u}_n$; dans le cas de convergence, les valeurs très grandes et très peu probables des u_n, qui empêchent peut-être la convergence de $\sum E\{(u_n - a_n)^2\}$, ne sauraient empêcher celle de $\sum E\left\{\frac{(u_n - a_n)^2}{1+(u_n - a_n)^2}\right\}$. Réciproquement, si cette série converge, il suffit de prendre $\bar{u}_n = u_n$, si $|u_n - a_n| \leqslant 1$, et $\bar{u}_n = a_n$, dans le cas contraire, pour s'assurer qu'on est dans le cas de convergence. Nous obtenons donc une nouvelle forme de la condition de convergence: a_n étant déterminé de manière à rendre minima l'expression $E\left\{\frac{(u_n - a_n)^2}{1+(u_n - a_n)^2}\right\}$, c'est-à-dire que $E\left\{\frac{u_n - a_n}{[1+(u_n - a_n)^2]^2}\right\} = 0$, le minimum obtenu étant désigné par $\sigma'^2\{u_n\}$, *il faut et il suffit, pour qu'on soit dans le cas de convergence, que* $\sum \sigma'^2\{u_n\}$ *converge* ([10]).

On peut évidemment encore donner à cette condition d'autres formes, qui aboutiront à d'autres valeurs de a_n (différant alors de la précédente par le terme général d'une série convergente). Ainsi on peut raisonner sur x_n (au lieu de u_n), et déterminer la constante A_n rendant minima l'expression $E\left\{\frac{(x_n - A_n)^2}{1+(x_n - A_n)^2}\right\}$; cela peut être avantageux de prendre ainsi pour A_n une valeur qui ne dépende que de la loi de probabilité de x_n.

D'après le principe d'augmentation de la dispersion, on ne change pas le cas où l'on se trouve, en changeant l'ordre des termes, en groupant plusieurs termes, ou au contraire en décomposant chaque terme u_n, en une somme de plusieurs termes indépendants. On peut même continuer cette décomposition, et à la limite remplacer la variable discontinue par une variable continue t. Pourvu que l'on détermine comme il vient d'être indiqué les constantes A_n, qui à la limite don-

([10]) L'introduction de $\sigma'^2\{u_n\}$ est à rapprocher d'une idée de M. FRÉCHET, qui définit la *distance* de deux variables aléatoires u et v comme égale à $E\left\{\frac{|u-v|}{1+|u-v|}\right\}$; il donne d'ailleurs d'autres définitions possibles. Nous pourrions introduire *l'écart* (e_1 ou e_2) défini par

$$E\left\{\frac{(u-v)^2}{1+(u-v)^2}\right\} = e_1^2 = \frac{e_2^2}{1+e_2^2};$$

e_1 est plus simple; mais e_2 présente l'avantage d'être une *moyenne*, comprise entre les valeurs extrêmes de $|u-v|$, et de varier de zéro à l'infini. L'expression $\sigma'\{u\}$ est le minimum de neront une fonction $A(t)$, l'*oscillation* de $x_\nu - A_\nu$ (différence de ses valeurs extrêmes) entre n et N peut être bornée supérieurement quand on connaît la dispersion de $x_N - x_n$ ([11]). Cette borne, indépendante du nombre d'indices intermédiaires entre n et N, reste valable à la limite, et par suite la démonstration de la convergence (*Séries*, § 7) subsiste.

Nous pouvons maintenant obtenir des conclusions nettes pour le cas de la variable continue t. Tout d'abord, si le cas de divergence se produit, la dispersion est devenue infinie, et le reste nécessairement quand t croît. Si donc on a une loi $\mathfrak{L}(t)$ bien définie dans un intervalle fini ou infini $(0, T')$, le cas de convergence se rencontre seul tant que $t < T'$; quel que soit l'ensemble d'intervalles dt, extérieurs les uns aux autres, que l'on considère entre 0 et $T < T'$, on est toujours dans le cas de convergence, avec toutes les conséquences qui en résultent.

Donc quand t tend en croissant vers une limite t', la loi $\mathfrak{L}(t)$ a une limite, et il y a une probabilité unité que $x(t) - A(t)$ ait une limite; il en est de même quand t' tend vers t par valeurs plus grandes; mais on peut ne pas avoir la même limite à droite et à gauche.

Il faut dès lors étudier les points de discontinuité et distinguer ceux de $\mathfrak{L}(t)$ de ceux de $x(t)$. L'étude des premiers est simple: si t' est un de ces points, l'indépendance de

$$x(t'-\varepsilon), \qquad x(t'+\varepsilon) - x(t'-\varepsilon), \qquad x(T) - x(t'+\varepsilon), \qquad (t' < T)$$

subsiste à la limite, ε tendant vers zéro, et montre que l'accroissement y de x en ce point est une variable aléatoire indépendante de la variation de $x(t)$ dans le reste de l'intervalle $(0, T)$. Le cas de convergence pour l'addition de ces variables aléatoires y étant seul possible, $\sum \sigma'^2\{y\}$ converge, ce qui implique que les points t' constituent au plus une infinité dénombrable; on peut alors, en faisant rentrer des constantes convenables dans la fonction $A(t)$, supposer

$$E\left\{\frac{y}{(1+y^2)^2}\right\} = 0, \qquad \sigma'^2\{y\} = E\left\{\frac{y^2}{1+y^2}\right\}. \tag{1}$$

Aucune autre restriction n'est évidemment nécessaire; on peut prendre n'importe quel ensemble, fini ou dénombrable, de points t', et associer à chacun d'eux n'importe quelle loi de probabilité (la valeur zéro n'étant pas exclue). Pourvu que la condition (1) soit vérifiée et que $\sum \sigma'^2\{y\}$ converge, non seulement il y a une probabilité unité pour que $\sum y$, les termes de cette série étant rangés dans

---

l'écart e_1 de u et d'une constante. Il y a lieu de noter que l'on a toujours $e_1 \geqslant E\left\{\frac{|u-v|}{\sqrt{1+(u-v)^2}}\right\}$, le second membre étant, d'après les résultats de M. FRÉCHET, une expression possible pour la distance.

([11]) Cf. Séries, § 6. Le résultat obtenu est que l'oscillation et la dispersion sont du même ordre de grandeur, quand on a rendu la première aussi petite que possible par l'addition aux u_n de constantes convenables.

l'ordre adopté pour dénombrer les points t', converge, mais il en est de même de la somme $Y(t)$ des y correspondant aux t' inférieurs à t, et la loi de probabilité dont dépend $Y(t)$ est indépendante de l'ordre considéré (Cf. *Séries*, § 15 et Note 2 ci-dessus).

Cette fonction $Y(t)$ ne varie qu'aux points t', et il y a une probabilité égale à l'unité que $\sum y^2$ soit borné, puisque, compte tenu de (1), $\sum E\{\bar{y}\}$ et $\sum E\{\bar{y}^2\}$ sont bornés, et que les variables auxiliaires $\bar{y}$ ne diffèrent des y que dans des cas de probabilité totale arbitrairement petite ([12]). Si $0<\alpha<2$, la probabilité que $\sum |y|^\alpha$ soit borné est de même 1 ou 0 suivant que la série $\sum E\left\{\frac{|y|^\alpha}{1+|y|^\alpha}\right\}$ converge ou diverge; les deux cas sont possibles suivant les lois dont dépendent les y. De toute façon les problèmes simples qui se posent sur la nature de cette fonction $Y(t)$ sont faciles à résoudre. Elle est bien entendu indépendante de la variation de $x(t)$ en dehors des points t'.

Conclusion. - On peut considérer $x(t)$ comme somme de trois termes indépendants: 1°) *le terme $A(t)$, indépendant du hasard, et qui peut être une fonction quelconque de la variable réelle t*; 2°) *le terme $Y(t)$, qui ne varie qu'en une infinité dénombrable au plus de points donnés d'avance, mais varie en chacun de ces points d'une quantité dépendant du hasard et ayant au moins deux valeurs possibles* [*dont une au moins de chaque signe, si l'on fait l'hypothèse* (1)]; 3°) *un terme, que nous avons maintenant à étudier et supposerons exister seul dans la suite, qui dépend d'une loi variant d'une manière continue avec t. On peut naturellement retrancher de ce terme une fonction continue, choisie de manière à en simplifier l'étude, que l'on supposera ajoutée à $A(t)$.*

Pour chaque t donné d'avance, la probabilité d'une discontinuité de $x(t)$ en ce point est maintenant nulle ([13]). Il n'en résulte pas que la probabilité d'une discontinuité dans l'intervalle $(0, T)$ soit nulle; comme nous l'avons déjà annoncé, cette dernière probabilité n'est nulle que dans le cas de la loi de Gauss.

§ 3. - Choix de $x(t)$; les espaces V et W.

Pour étudier les propriétés probables de $x(t)$, deux méthodes s'offrent à nous, que nous emploierons simultanément.

La première est celle employée par Wiener dans l'étude du cas particulier déjà cité. L'intervalle $(0, T)$ est divisé en n intervalles très petits (égaux ou non);

([12]) Nous définissons par exemple $\bar{y}$ par $\bar{y}=y$ si $|y|\leq 1$ et $\bar{y}=0$ si $|y|>1$.

([13]) D'une manière précise, quelque petit que soit $\varepsilon>0$, on a

$$\lim P\{|x(t+\tau)-x(t)|>\varepsilon\}=0, \qquad (\tau\to 0,\ t \text{ fixe}).$$

Bien entendu, le seul fait que la loi $\mathfrak{L}(t)$ varie d'une manière continue avec t n'entraînerait dans chaque intervalle Δt on détermine l'accroissement correspondant Δx d'après la loi de probabilité relative à cet intervalle. Puis on fait augmenter n indéfiniment.

L'avantage de cette méthode est qu'elle se prête bien au calcul des probabilités par un passage à la limite; c'est elle qu'il faut employer par exemple pour déterminer la valeur probable d'une fonctionnelle donnée par une expression analytique simple. Son inconvénient est que, remettant toujours en question les choix déjà faits, elle ne fournit pas un processus aboutissant dans chaque cas particulier à une fonction déterminée, connue avec de plus en plus de précision par l'effet des choix successifs.

La seconde méthode consiste à ne pas remettre les choix en question; les valeurs $x(t_i)$ et $x(t_{i+1})$ une fois choisies, seront conservées; puis l'on choisira $x(t_i')$ pour une valeur t_i' comprise entre t_i et t_{i+1}, et, par des interpolations successives, on arrivera à choisir $x(t)$ pour les points d'un ensemble dénombrable partout dense E.

Or, si l'on détermine convenablement la loi de probabilité dont dépendent les choix de $x(t_i')$ et les choix analogues, les deux méthodes sont équivalentes, au point de vue de la loi dont dépend à un instant quelconque l'ensemble des valeurs de $x(t)$ résultant des choix faits. En effet, dans la première méthode, on choisit indépendamment l'une de l'autre les différences

$$\xi_i' = x(t_i') - x(t_i), \qquad \xi_i'' = x(t_{i+1}) - x(t_i').$$

Dans la seconde, on choisit d'abord la somme $\xi_i = \xi_i' + \xi_i''$, et ensuite ξ_i'. Il n'y a évidemment qu'à choisir ξ_i d'après la loi dont dépend cette somme, puis ξ_i', d'après la loi de probabilité *a posteriori* dont dépend cette variable lorsque la somme ξ_i est connue, pour que l'équivalence des deux méthodes apparaisse.

Mais, au point de vue théorique, la seconde présente l'avantage de montrer nettement un processus conduisant au choix d'une fonction déterminée, de mieux en mieux connue par les choix successifs. La définition de *l'espace fonctionnel pondéré* auquel appartiennent les fonctions $x(t)$ susceptibles d'être obtenues, et que nous appellerons *espace* V, est ainsi très nette. Chaque fois que l'on a à choisir $x(t)$ pour une valeur t_i, l'espace V apparaît ainsi comme divisé en régions, ayant chacune un poids déterminé, et il s'agit d'en choisir une, pour la subdiviser ensuite de la même manière. La probabilité apparaît ainsi comme une *intégrale de Daniell* [14], c'est-à-dire comme une fonctionnelle complètement additive des sous-ensembles de l'espace V.

---

pas cette conclusion si nous n'avions pas supposé en outre que l'accroissement considéré Δx est une variable aléatoire indépendante de $x(t)$. De ces deux hypothèses réunies résulte que $E\left\{e^{iz\Delta x}\right\}$ tend vers l'unité, et cela uniformément par rapport à z dans tout intervalle fini, d'où la conséquence indiquée.

([14]) P. J. Daniell: *A General Form of Integral* [Annals of Mathematics, Series 2, Vol. 19 (1918), pp. 279-294].

La distance de deux points de V, correspondant à deux fonctions $x(t)$ et $y(t)$, sera par définition l'intégrale de $|x(t)-y(t)|$ (*distance* de deux fonctions mesurables bornées). Les régions considérées seront donc très petites pour n assez grand, si les fonctions de V vérifient, sauf dans un ensemble de t de mesure arbitrairement petite, une condition de HÖLDER permettant de limiter la différence de deux fonctions $x(t)$ et $y(t)$ qui ont la même valeur pour les n premiers points de E. S'il en est ainsi, toute fonctionnelle bornée uniformément continue dans V a une valeur probable bien déterminée, qui est une intégrale de DANIELL. Mais en fait ce résultat ne sera que très probablement vrai, c'est-à-dire que l'on pourra trouver pour $|x(t+\tau)-x(t)|$ une borne supérieure tendant vers zéro avec τ et valable sauf pour un ensemble de valeurs de t de mesure arbitrairement petite et pour des fonctions $x(t)$ de probabilité arbitrairement petite. Il est alors seulement très probable que deux points choisis dans un même compartiment de V sont voisins; mais on reste assuré de l'existence de la valeur probable pour les fonctionnelles uniformément continues, bornées dans V, ou même n'augmentant pas trop rapidement dans les régions peu probables de V [15].

Nous pensons revenir dans un autre travail sur le rôle des conditions de HÖLDER dans un espace V. Montrons tout de suite que: *la probabilité que les valeurs choisies pour les points de E définissent une fonction $x(t)$ presque partout continue est égale à* 1.

Choisissons en effet, indépendamment, t entre 0 et T, d'après une loi de probabilité d'ailleurs quelconque, et $x(t)$ dans V. La probabilité que t soit un point de discontinuité est nulle, puisqu'il en est ainsi quel que soit t choisi d'avance. Le résultat subsiste, puisque les choix sont indépendants, si l'on choisit d'abord $x(t)$, puis t. Donc, sauf pour des fonctions $x(t)$ dont la probabililé dans V est nulle, la probabilité qu'un point t choisi entre 0 et T soit point de discontinuité est nulle, c. q. f. d.

Une classe importante d'espaces V est caractérisée par la propriété que la loi dont dépend $x(t+\tau)-x(t)$ ne dépend que de τ, et non de t. En d'autres termes $\Psi(z, t)$ est de la forme $t\Psi(z)$. Dans ce cas l'espace V sera dit espace W [16].

---

(15) Dans son Mémoire sur l'espace différentiel, sans introduire cette notion dès le début comme nous le faisons ici, WIENER a bien indiqué que la moyenne d'une fonctionnelle bornée uniformément continue est une intégrale de DANIELL.

(16) Notons encore que la définition des espaces V (et W) est indépendante du choix de l'ensemble E. On établit en effet aisément, à l'aide du principe de continuité, que l'indépendance des Δx relatifs à des intervalles Δt extérieurs les uns aux autres et limités par des points de E, subsiste si on enlève cette dernière restriction.

Précisons aussi que, si le troisième des termes distingués § 2 existe effectivement, une fonction particulière a toujours une probabilité nulle. Dire qu'une fonction $x_0(t)$ appartient à l'espace V signifie seulement qu'elle a un certain nombre de caractères qui dans cet espace ont une probabilité unité; on ne peut pas dire « *tous les caractères* », car elle n'aura pas celui d'être différente de $x_0(t)$.

§ 4. - Cas où la fonction $x(t)$ est continue.

Lemme. - *La probabilité qu'un nombre donné τ soit module de continuité uniforme relativement à un nombre donné ε est toujours bien définie.*

Dire que τ est module de continuité uniforme équivaut à dire que, pour les points de E (donc pour toutes les valeurs de t), $|t'-t|\leq\tau$ entraîne

$$|x(t')-x(t)|\leq\varepsilon.$$

Or la probabilité que cette condition soit vérifiée pour les n premiers points de E est bien déterminée, et ne peut que décroître quand n croît. Elle a donc une limite bien déterminée, c. q. f. d.

THÉORÈME I. - *Si dans un espace V la probabilité que la fonction $x(t)$ soit continue est égale à l'unité, la loi dont dépend l'accroissement $\Delta x=x(t_2)-x(t_1)$ relative à n'importe quel intervalle (t_1, t_2) est la loi de Gauss, ou cette loi modifiée par l'addition d'une constante. En d'autres termes,*

$$\Psi(z, t)=izf(t)-\frac{z^2}{2}g(t)$$

$f(t)$ étant une fonction continue, et $g(t)$ une fonction continue non décroissante.

Démonstration. - Pour tout ε positif, la probabilité que τ puisse être pris comme module de continuité est bien définie, est une fonction non croissante de τ, et, par hypothèse, tend vers 1 quand τ tend vers zéro. On peut donc choisir τ de manière à rendre cette probabilité supérieure à $1-\varepsilon'$, ε' étant arbitrairement petit. Si on divise l'intervalle (t_1, t_2) en intervalles tous inférieurs à τ, $x(t_2)-x(t_1)$ apparaît comme la somme d'accroissements Δx indépendants et tous inférieurs à ε, sauf dans des cas de probabilité totale inférieure à ε'; ε et ε' étant arbitrairement petits, on sait que dans ces conditions l'expression étudiée, diminuée d'une constante convenable, dépend de la loi de GAUSS ([17]), c. q. f. d.

Nous pouvons bien entendu faire abstraction du terme $f(t)$, en le rattachant à $A(t)$, comme il a été dit à la fin du § 2. Observant qu'alors, quand la fonction $g(t)$ reste constante dans un intervalle, $x(t)$ ne varie pas non plus, et que par suite nous pouvons faire abstraction de cet intervalle, nous voyons que nous pouvons prendre $g(t)$ comme variable, au lieu de t. Par suite:

Corollaire. - *Les cas considérés dans le théorème se ramènent par des réductions élémentaires au cas où*

$$\Psi(z, t)=-t\frac{z^2}{2}.$$

([17]) Si $\varepsilon'=0$, cet énoncé n'est autre que le premier théorème de M. LINDEBERG; or deux lois qui ne diffèrent pas dans des cas de probabilité totale arbitrairement petite sont identiques (Cf. *Séries*, § 10).

Ce cas est celui considéré par WIENER. L'espace V est alors un espace W, que nous désignerons par W_2. On obtient naturellement aussi des espaces W en prenant $f(t)=at$, $g(t)=b^2t$, a et b étant des constantes quelconques.

§ 5. - Considérations générales sur les discontinuités mobiles.

On sait que, si $u_1, u_2, \ldots, u_n, \ldots$ sont des variables aléatoires indépendantes, dépendant d'une même loi de probabilité, cette loi étant telle que $E\{u_n\}=0$ et $E\{u_n^2\}=1$, en posant

$$x_n=u_1+u_2+\ldots+u_n=\xi_n\sqrt{n},$$

la loi de probabilité dont dépend ξ_n tend, pour n infini, vers celle de GAUSS; il en résulte en particulier que, pour chaque valeur de n assez grande, il est très probable que ξ_n n'est ni très petit ni très grand, de sorte que x_n est de l'ordre de grandeur de $\sqrt{n}$.

Ces conclusions sont en défaut si $E\{u_n^2\}$ est infini. Cela tient naturellement au rôle joué par les grandes valeurs des u_n, sans lesquelles le processus que nous venons de rappeler fonctionnerait infailliblement. Si les valeurs supérieures à un certain nombre U ont une probabilité très petite a, la probabilité que cette circonstance se présente au moins une fois au cours de n expériences est sensiblement égale à na, tant que ce produit est petit, et cesse d'être petite avec na. Des valeurs de plus en plus grandes des u_n apparaissent ainsi successivement, leurs probabilités cessant d'être négligeables, quand n augmente. Dans certains cas particuliers chacune de ces valeurs, supérieure à celles qui l'ont précédée, sera pourtant négligeable par rapport à leur somme, et l'on obtiendra encore la loi de GAUSS, mais avec un multiplicateur croissant plus vite que $\sqrt{n}$. Mais en général la somme x_n sera de l'ordre de grandeur du plus grand des u_ν $(\nu\leqslant n)$, et, p étant un nombre suffisamment grand, mais indépendant de n, x_n pourra être représenté avec une erreur relative aussi faible que l'on veut par la somme des p plus grands des u_ν, et dans ce cas la loi dont dépend x_n n'est pas d'un type tendant pour n infini vers celui de la loi de GAUSS ([18]).

On arrive à une conclusion tout à fait analogue si la loi dont dépend u_n varie avec n.

L'inversion de ce résultat fait bien comprendre le rôle des discontinuités mobiles. Pour chaque valeur donnée de t, $x(t)$ apparaît comme la somme d'un grand nombre n de variables Δx indépendantes et toutes très petites, en ce sens que, pour chacune d'elles, quelque petit que soit $\varepsilon>0$, $a=P\{|\Delta x|>\varepsilon\}$ tend vers zéro

---

([18]) Je me propose de revenir ailleurs sur cette question, pour compléter et préciser les résultats que j'ai déjà publiés antérieurement. Il m'a semblé que des développements trop longs sur ce sujet risqueraient ici de détourner l'attention des lecteurs des points essentiels pour l'étude des espaces V.

avec Δt, donc avec $\frac{1}{n}$ (Cf. Note 11). Si la loi dont dépend $x(t)$ n'est pas celle de GAUSS, c'est alors nécessairement que certains des Δx ne sont pas très petits; en termes plus précis, $\sum \alpha$ ne sera pas très petit, de sorte que la probabilité que le plus grand des $|\Delta x|$ soit $\leqslant \varepsilon$ ne sera pas voisine de l'unité. Si nous passons à la limite, les intervalles Δt devenant infiniment nombreux et infiniment petits, on trouvera des points dans le voisinage desquels n'existe aucun module de continuité relatif à ε; du moins la probabilité de l'existence de ces points sera positive. Mais leurs places dépendent du hasard, et la probabilité qu'un tel point occupe une position dònnée d'avance est toujours nulle (les discontinuités fixes ayant été éliminées).

§ 6. - La loi des petites probabilités de Poisson.

Considérons un évènement A, pouvant être réalisé, ou non, pour certaines valeurs de t et supposons que: 1°) la probabilité de sa réalisation dans l'intervalle (t, t') tend vers zéro quand t' tend vers t (quel que soit le signe de $t'-t$), 2°) les probabilités de sa réalisation relatives à différents intervalles extérieurs les uns aux autres soient indépendantes.

Alors la probabilité de la non-réalisation de A dans l'intervalle $(0, t)$ est une fonction continue de t, décroissante, ou du moins non croissante; désignons-la par $e^{-\theta}$; à chaque intervalle Δt correspond, d'après l'hypothèse 2°, une probabilité $e^{-\Delta\theta}=1-\alpha$ pour la non-réalisation de A, et, d'après l'hypothèse 1°, θ ne peut augmenter indéfiniment pour aucune valeur de t. Comme, dans tout intervalle où θ est constant, la probabilité de la réalisation de A est nulle, nous pouvons étudier en fonction de θ le nombre de réalisations de A.

Dans tous les cas où le nombre des réalisations de A, θ variant de zéro à θ', est p, la plus petite des variations de θ entre deux de ces réalisations a une valeur déterminée, supérieure à $\frac{\theta'}{n}$, si n est assez grand. La probabilité que A soit réalisé plusieurs fois dans un même intervalle d'étendue $\frac{\theta'}{n}$ tend donc vers zéro avec $\frac{1}{n}$. Par suite, la probabilité que le nombre de réalisations de A ait précisément la valeur p est la limite, pour n infini, de la probabilité que, si l'on divise l'intervalle $(0, \theta')$ en n intervalles égaux, A soit réalisé dans p de ces intervalles, et ne le soit pas dans les autres; elle est donc, en posant $\Delta\theta=\frac{\theta'}{n}$, $e^{-\Delta\theta}=1-\alpha$ $\left(\text{d'où } \alpha \sim \frac{\theta'}{n}\right)$,

$$\lim \frac{n(n-1)\ldots.(n-p+1)}{p!}\alpha^p e^{-(n-p)\Delta\theta}=e^{-\theta'}\frac{\theta'^p}{p!}.$$

De même, pour un intervalle $(0, \theta)$, cette probabilité sera $e^{-\theta}\frac{\theta^p}{p!}$. On reconnaît la loi de POISSON. On sait (on vérifie d'ailleurs aisément) que, pour cette loi, on a

$$(2) \qquad \left\{\begin{array}{l} E\{p\}=\theta, \qquad \sigma^2\{p\}=E\{(p-\theta)^2\}=\theta, \\ \log E\{e^{ipz}\}=\theta(e^{iz}-1). \end{array}\right.$$

On peut alors concevoir que l'expérience dont dépend la réalisation de A dans un intervalle $(0, T)$ soit réalisée de la manière suivante: on choisira un entier p d'après la loi de POISSON, le paramètre θ qui intervient dans cette loi étant le nombre probable de réalisations de A entre 0 et T. On choisira ensuite les abscisses $t_1, t_2, \ldots, t_p$ des p points de réalisation de A par de nouveaux tirages au sort, dans chacun desquels chaque intervalle dt corresponde à une probabilité proportionnelle à $d\theta$.

Le résultat est le même, au point de vue de la probabilité des différents cas possibles, que si, pour chaque intervalle élémentaire dt, la réalisation de A dépendait d'un tirage indépendant des autres lui donnant la probabilité $d\theta$.

§ 7. - Exemples simples d'espaces V et W à discontinuités mobiles.

Pour former un tel espace, il suffit de supposer que $x(t)$ ne varie que par sauts. L'existence d'un ou plusieurs sauts sera l'évènement A considéré au § 6, et chaque fois qu'un saut se produira en un point t, sa *hauteur* u (c'est-à-dire la variation de x en ce point) dépendra d'une loi $\mathfrak{L}'$ pouvant ou non varier avec t; bien entendu la valeur $u=0$ aura une probabilité nulle, si l'on veut que les sauts existent effectivement; à cela près la loi $\mathfrak{L}'$ sera quelconque.

Il est clair que l'on obtient ainsi un espace V, qui devient un espace W si θ est proportionnel à t et si $\mathfrak{L}'$ est indépendant de t.

Plaçons-nous dans ce dernier cas, et définissons la loi $\mathfrak{L}'$ par sa fonction des probabilités totales $F(u)$. Le nombre probable des sauts de hauteur égale à u (à du près), entre 0 et t, est $\theta dF(u)$. Comme $\theta = at$, la somme $\xi = pu$ de ces sauts dépend d'une loi définie par

$$\log E\{e^{iz\xi}\} = \log E\{e^{ipuz}\} = at(e^{iuz}-1)dF(u),$$

et, comme $x(t)$ est la somme des variables indépendantes ξ correspondant aux différents intervalles du, la loi $\mathfrak{L}(t)$ dont dépend $x(t)$ est définie par

$$\log E\{e^{izx(t)}\} = at\int_{-\infty}^{+\infty}(e^{iuz}-1)dF(u). \tag{3}$$

Indiquons quelques cas particuliers:

1°) La loi $\mathfrak{L}'$ est telle que l'on ait sûrement $u=1$. Alors $x(t)$ devient égal au nombre des discontinuités entre 0 et t et dépend précisément de la loi de POISSON.

2°) Si u admet deux valeurs possibles et également probables, 1 et -1, on obtient une loi de probabilité liée à une remarque de M. LINDEBERG et que j'ai moi-même déjà étudiée (*Probabilités*, p. 241).

3°) Un cas assez remarquable est celui où la loi $\mathfrak{L}'$ est la loi de GAUSS; l'espace W ainsi obtenu présente certaines analogies avec celui lié au mouvement d'une particule dans le mouvement brownien ou dans la théorie cinétique des gaz. La théorie de N. WIENER est déjà en relation avec ces applications; mais elle est obtenue en étendant à l'infiniment petit ce qui dans ces applications n'est vrai qu'à une échelle très petite; cela cesse d'être vrai pour des intervalles dt plus petits encore, et comparables à l'intervalle qui sépare deux chocs. Ces chocs ont les caractères de l'évènement A du § 6, et leur nombre dépend de la loi de POISSON, qui permet de calculer l'intervalle moyen de deux chocs et le libre parcours moyen.

Mais il importe de remarquer que notre exemple actuel ne donne pas non plus une image exacte du mouvement d'une molécule. Si l'on représente par $x(t)$ l'abscisse d'une molécule, l'hypothèse de la constance de la vitesse entre deux chocs est en contradiction avec les hypothèses fondamentales de la théorie des espaces V et W, d'après lesquelles la vitesse ne saurait rester constante dans un intervalle fini, si petit soit-il. Si au contraire on prend pour $x(t)$ une composante de la vitesse, cela donne une idée exacte de la fréquence des chocs; mais dans la théorie cinétique des gaz, les vitesses avant et après le choc sont indépendantes, et $E\{x^2\}$ reste constant; au contraire dans l'espace W que nous considérons actuellement, c'est la valeur de $x(t)$ avant la discontinuité et l'accroissement Δx qui sont indépendants, et $E\{x^2\}$ augmente proportionnellement à t. Ainsi d'aucune manière un espace W ne peut donner une représentation tout à fait exacte du mouvement d'une particule dans le mouvement brownien ou d'une molécule dans un gaz.

Revenons au cas général; il résulte de (3) que $E\{x\}$ et $\sigma^2\{x\}$ sont finis en même temps que $E\{u\}$ et $\sigma^2\{u\}$, et que

$$(4) \qquad \left\{\begin{aligned} E\{x\} &= at\int_{-\infty}^{+\infty} u\,dF(u) = atE\{u\}, \\ \sigma^2\{x\} &= at\int_{-\infty}^{+\infty} u^2 dF(u) = atE\{u^2\}. \end{aligned}\right.$$

Lorsque ces expressions sont finies, il est bien évident que la loi de probabilité dont dépend la *variable réduite* $\frac{x-E\{x\}}{\sigma\{x\}}$ tend pour t infini vers la loi de GAUSS; mais pour chaque valeur finie de t, on a une loi différente, bien que $x(t)$ soit la somme d'un nombre arbitrairement grand de variables Δx, indépendantes, et à valeur quadratique moyenne très petite. Mais la probabilité que les $|\Delta x|$ soient tous simultanément très petits n'est pas voisine de l'unité (puisqu'elle est $e^{-\theta}$), et cela suffit pour que la loi dont dépend $x(t)$ puisse n'être pas celle de GAUSS.

§ 8. - Extension au cas où le nombre des discontinuités mobiles est infini.

Chaque saut de $x(t)$ est caractérisé par deux nombres, l'abscisse t et la hauteur u du saut. Le nombre probable des sauts correspondant à une région donnée S du plan des tu est une fonctionnelle additive, toujours positive ou nulle, que nous pouvons représenter par

$$\Phi(S)=\iint_S d_t d_u N(u,t), \tag{5}$$

l'élément $d_t d_u N(u, t)$ correspondant à chaque rectangle $dtdu$ étant toujours positif ou nul. Nous pouvons supposer $N(u, 0)=0$, et, les discontinuités fixes étant supposées éliminées, $N(u, t)$ varie d'une manière continue avec t. Nous verrons dans un instant que l'intégrale précédente est nécessairement bornée dans toute région, finie ou infinie, sur laquelle $|u|$ est borné inférieurement; nous pouvons alors supposer $N(u, t)$ nul pour u infini; quand u franchit la valeur zéro, $N(u, t)$ varie d'une quantité, finie ou infinie, égale au nombre probable des sauts d'abscisses inférieures à t et de hauteurs quelconques. Il faudra alors, si une aire S traverse l'axe des t, faire abstraction pour le calcul de l'intégrale (5) de cette discontinuité, c'est-à-dire considérer cette intégrale comme la somme des intégrales relatives aux deux parties de l'aire S séparées par l'axe des t, cet axe étant considéré comme n'appartenant à aucune des deux parties.

Divisons alors la région $0<t<T$ en aires élémentaires s, telles que u varie peu pour chacune d'elles; elles pourront être très petites dans tous les sens, ou chacune peut comprendre tout l'intervalle de variation de t; les remarques qui vont suivre s'appliquent dans tous les cas. Désignons par n l'intégrale (5) étendue à chaque région s; le nombre p des sauts correspondant à cette région est une variable aléatoire, de valeur probable n, et dépendant de la loi de POISSON (bien définie par la donnée de n); la somme ξ de ces sauts est égale à pu, et si la somme des variables aléatoires ξ correspondant à toutes les aires s converge (c'est-à-dire si la probabilité de sa convergence est l'unité), et cela pour tout T dans un certain intervalle $(0, T')$, cette somme $x(T)$ vérifie évidemment toutes les conditions requises dans la définition des espaces W. Nous avons ainsi défini un espace V, dans lequel $x(t)$ est réduit à la fonction des sauts, et bien défini par la donnée de la fonctionnelle additive (5). Des formules (2) relatives à la loi de POISSON, et de la formule

$$\log E\{e^{izx(T)}\}=\sum \log E\{e^{iz\xi}\}$$

qui résulte de ce que $x(T)$ est la somme des variables aléatoires indépendantes ξ, résulte que

$$\log E\{e^{izx(T)}\}=\int_{-\infty}^{+\infty}(e^{izu}-1)d_u N(u,T), \tag{6}$$

formule qui définit parfaitement la loi dont dépend $x(T)$.

Mais nous pouvons aller plus loin. Il n'est pas nécessaire que la somme $\sum \xi$ converge pour qu'à la fonctionnelle additive (5) corresponde un espace V; il suffit qu'à chaque aire s on puisse faire correspondre un nombre a, indépendant du hasard, et tel que la convergence de $\sum (\xi - a)$ ait une probabilité égale à l'unité. Cette condition est d'ailleurs aussi nécessaire, car sans elle nous savons (§ 2) qu'il serait impossible, par l'addition d'autres termes, d'obtenir une somme $x(T)$ comprenant tous les termes ξ, et ayant un sens; si d'ailleurs cette condition est vérifiée pour toute la région $0 < t < T$, elle l'est aussi pour toute région partielle. Pour la former, nous pouvons grouper les aires s de manière que chacune d'elles comprenne tout l'intervalle de variation de t, et soit définie par le fait que u appartienne à un intervalle du. Cette condition est alors que l'intégrale

$$\sum \sigma^2 \{\xi\} = \int_{-\infty}^{+\infty} u^2 d_u N(u, T), \tag{7}$$

ou bien converge, ou bien puisse être rendue convergente par des modifications portant sur des cas de probabilité totale arbitrairement petite.

Cette dernière restriction montre que la condition qui nous intéresse ne peut être empêchée par la divergence à l'infini de l'intégrale précédente. Cela est d'ailleurs bien évident *a priori*: pour toute région $|u| > \varepsilon$, l'intégrale (5) étant finie, le nombre des sauts est fini (sauf dans des cas de probabilité nulle), et la question de convergence ne se pose pas; la contribution plus ou moins grande des grandes valeurs de u à l'intégrale (5) ne peut que rendre plus ou moins grande la probabilité des grandes valeurs de la somme $x(T)$, et en particulier, si l'intégrale (7) diverge à l'infini, $E\{x^2(T)\}$ est infini.

Au contraire la divergence de cette intégrale dans un intervalle fini empêche nécessairement la convergence de $\sum (\xi - a)$. Pour le montrer, nous pouvons supposer le plan divisé en régions s telles que pour chacune d'elles l'intégrale (5) soit bornée; la divergence de $\sum \sigma^2 \{\xi\}$ subsiste si l'on remplace ξ par une variable $\bar{\xi}$ égale à 0 si $\xi = 0$ et à u dans le cas contraire; en effet

$$\sigma^2 \{\xi\} = nu^2 \quad \text{et} \quad \sigma^2 \{\bar{\xi}\} = u^2 e^{-n}(1 - e^{-n})$$

sont du même ordre de grandeur. Chaque terme $\bar{\xi}$ étant alors borné, la convergence de $\sum \sigma^2 \{\bar{\xi}\}$ ne peut être modifiée par une modification ne portant que sur des cas de probabilité totale arbitrairement petite; cela est vrai *a fortiori* pour $\sum \sigma^2 \{\xi\}$. Il faut donc que cette intégrale converge dans tout intervalle fini.

Le résultat obtenu sur l'intégrale (6) entraîne celui annoncé plus haut sur l'intégrale (5). Toutes les deux sont en effet en même temps bornées dans toute aire où $|u|$ est borné inférieurement et supérieurement. À l'infini, les valeurs très grandes et très peu probable de u qui peuvent empêcher la convergence de l'intégrale (7) ne peuvent empêcher celle de l'intégrale (5); dire qu'elles sont très peu probables revient en effet à dire que leur contribution à cette intégrale est très petite. Nous pouvons donc énoncer le théorème suivant:

THÉORÈME II. - *Si $\Phi(S)$ est une fonctionnelle bien définie pour toute aire S du plan des t, u pour laquelle $0<t<T$, la condition nécessaire et suffisante pour qu'il existe des espaces V tels que le nombre probable des discontinuités mobiles correspondant à des points de l'aire S soit $\Phi(S)$, est que $\Phi(S)$ soit une fonctionnelle additive, non négative, nulle pour toute aire réduite à une droite $t=$*const., *finie pour toute aire sur laquelle $|u|$ est borné inférieurement, et telle que l'intégrale*

$$\iint_S u^2 d\Phi \tag{8}$$

soit finie pour toute aire finie.

Un des espaces V correspondant à $\Phi(S)$ étant ainsi défini, on obtiendra naturellement tous les autres en ajoutant une fonction indépendante du hasard, un terme gaussien, et un terme qui soit somme de discontinuités fixes.

Pour former un de ces espaces, on ne pourra pas toujours définir $x(t)$ par $\sum \xi$, mais par $\sum (\xi - a)$, en prenant $a=0$ pour $|u|$ supérieur à un nombre U choisi arbitrairement, et $a=E\{\xi\}$ dans le cas contraire. En désignant par $\omega(u)$ une fonction égale à u pour $|u|<U$ et à zéro pour $|u|>U$, la loi dont dépend $x(t)$ sera alors définie par la formule

$$\log E\{e^{izx(t)}\} = \int_{-\infty}^{+\infty} [e^{izu} - 1 - iz\omega(u)] d_u N(u, t). \tag{9}$$

Par une définition différente de $\omega(u)$, on pourra aisément éviter l'inconvénient résultant de la possibilité de l'existence de discontinuités simultanées de $\omega(u)$ et $N(u, t)$ pour $u=U$. Ainsi, on pourra dans tous les cas prendre $\omega(u)=\frac{u}{1+u^2}$. Les différents espaces V ainsi obtenus pour une même fonction $N(u, t)$ se déduiront les uns des autres par l'addition à $x(t)$ d'une fonction continue.

La possibilité de prendre $\omega(u)=u$ pour tout u dépend évidemment de la nature à l'infini de l'intégrale

$$\int u d_u N(u, T). \tag{10}$$

Cela est possible si cette intégrale converge à l'infini, même en cas de divergence à l'origine, et dans ce cas l'on aura, pour la fonction $x(t)$ correspondant à ce choix de $\omega(u)$,

$$E\{x\}=0.$$

Naturellement, $E\{x^2\}$ sera alors égal à $\sigma^2\{x\}$, donc à l'intégrale (7), qu'elle soit finie ou infinie.

La possibilité de prendre $u=0$, et d'avoir ainsi une fonction $x(t)$ qui représente la somme des sauts sans l'addition d'aucun terme indépendant du hasard, dépend au contraire de l'allure de $N(u, t)$ pour $u=0$. Il y a trois cas à distinguer.

1°) *Cas où l'intégrale* (10) *converge pour* $u=0$ (*donc dans tout intervalle fini.*

Dans ce cas on peut prendre $\omega(u)=0$; l'intégrale (9) reste convergente, et, comme on peut dans cette intégrale séparer les valeurs positives et les valeurs négatives de u, on peut considérer $x(t)$ comme la somme de deux termes représentant respectivement la somme des sauts positifs et la somme des sauts négatifs. La fonction $x(t)$ est donc une fonction à variation bornée, ne variant que par sauts, et, sauf dans des cas de probabilité nulle, le nombre des sauts est dans tout intervalle (t_1, t_2) fini ou infini en même temps que sa valeur probable.

Pour préciser un peu la nature de cette fonction, observons que la somme $\sum |u|^\alpha$, étendue à tous les sauts de $x(t)$ est évidemment, sauf dans des cas de probabilité nulle, finie pour $\alpha>1$, et, pour α compris entre 0 et 1, finie ou infinie en même temps que l'intégrale

$$\int_{-1}^{+1} |u|^\alpha d_u N(u, T). \tag{11}$$

Il n'y a pour vérifier ce résultat qu'à raisonner sur $\sum |u|^\alpha$ comme nous venons de le faire pour $\sum u = x(T)$. La nature de cette somme dépend donc de l'allure de $N(u, T)$ pour $u=0$, tandis que la nature de $E\{\sum |u|^\alpha\}$ dépend en outre de l'allure de l'intégrale précédente à l'infini; on a d'ailleurs

$$E\{\sum |u|^\alpha\} = \int_{-\infty}^{+\infty} |u|^\alpha d_u N(u, T), \tag{11'}$$

les deux membres étant en même temps finis ou infinis.

Quant à l'expression $E\{|x(T)|^\alpha\}$, on voit aisément que, la probabilité que les grandes valeurs de $x(T)$ soient obtenues par l'addition de termes bornés étant très petite, sa nature ne dépend que de celle de l'intégrale (11') à l'infini: cette expression est finie ou infinie suivant que cette intégrale converge ou diverge à l'infini.

2°) *Cas où l'intégrale* (10) *diverge pour* $u=0$, *et où les deux intégrales*

$$\int_{-1}^{0} u d_u N(u, T), \qquad \int_{0}^{1} u d_u N(u, T), \tag{12}$$

sont de natures différentes [19].

---

[19] Les conclusions relatives à ce cas n'ont pas été exactement indiquées dans ma Note du 26 mars. Le premier exposé net de ce cas, et de son application à la théorie des lois stables, a été fait le 25 avril 1934 à la Société Mathématique de France.

Supposons pour fixer les idées la première infinie et la seconde finie. En retranchant un terme du type étudié au premier cas, on peut supposer que $N(u, t)$ ne varie avec u (quel que soit t fixe) que de -1 à 0. Dans la formule (9), l'intervalle d'intégration étant réduit à $(-1, 0)$ nous prendrons $\omega(u)=u$, c'est-à-dire $u=E(\xi)$. Les caractères de la fonction $x(t)$ sont alors les suivants : elle admet une infinité de sauts, par l'effet desquels elle décroît d'une quantité infinie ; bien que la place exacte de ces sauts dépende du hasard, l'addition d'une terme indépendant du hasard conduira à une somme finie.

Pour rendre ceci plus net par un exemple, supposons que les sauts se produisent sur la droite $u=t-T$; en termes plus précis, supposons que la fonctionnelle $\Phi(S)$ se réduise pour toute aire S à une intégrale étendue à la portion de la droite $u=t-T$ qui traverse cette aire ; chaque élément de cette intégrale sera alors $d_u N(u, T)$. La fonction $N(u, T)$ augmentera indéfiniment quand u tendra vers zéro par valeur négative, mais pour $t<T$, on aura

$$
\begin{aligned}
&N(u, t)=N(u, T) && \text{pour } -1<u\leqq t-T, \\
&N(u, t)=N(t-T, T) && \text{pour } t-T\leqq u<0.
\end{aligned}
$$

La fonction $x(t)$ sera alors la somme des sauts, tous négatifs, augmentés du terme indépendant du hasard

$$
-\int_{-1}^{t-T} u d_u N(u, T);
$$

t tendant vers T, elle apparaîtra comme une différence de deux termes, l'un aléatoire et l'autre indépendant du hasard, augmentant indéfiniment ; mais il y aura une probabilité unité pour l'existence d'une limite finie.

D'une manière générale, il peut arriver que, pour n'importe quel intervalle (t_1, t_2), $x(t_2)-x(t_1)$ se présente, comme dans le cas précédent pour l'intervalle $(0, T)$, sous la forme $\infty-\infty$. Tel sera le cas si l'intégrale

$$
\int_{-1}^{0} u d_u[N(u, t_2)-N(u, t_1)]
$$

est toujours infinie ; il en sera ainsi nécessairement si l'espace V est un espace W, la probabilité de toutes les circonstances possibles se répartissant alors uniformément sur l'axe des t. La nature de la fonction $x(t)$ est alors moins facile à imaginer d'une manière intuitive. Mais la loi dont dépend $x(T)$, pour chaque valeur particulière T de t, peut être définie comme dans l'exemple précédent. On ne la modifie pas en effet en déplaçant horizontalement, c'est-à-dire sur le segment $0<t<T$ de n'importe quelle droite $u=$const., les masses dont la somme dans l'aire S définit $\Phi(S)$. On peut ainsi les amener sur la droite $u=t-T$, et on est ramené à l'exemple précédent.

Les remarques sur la somme $\sum |u|^{\alpha}$, sur ses relations avec les intégrales (11) et (11′), et sur $E\{|x|^\alpha\}$, que nous avons faites dans le premier cas, subsistent ici sans modification. Il en sera d'ailleurs de même dans le troisième cas.

Nous rattacherons d'ailleurs au second cas celui où les circonstances qui le définissent sont réalisées au moins dans un intervalle (t_1, t_2), sans l'être pour l'ensemble de l'intervalle $(0, T)$.

3°) *Les deux intégrales* (12) *sont divergentes, et, pour n'importe quel intervalle partiel* (t_1, t_2), *les deux intégrales déduites des précédentes en remplaçant* $N(u, T)$ *par* $N(u, t_2)-N(u, t_1)$ *sont de même nature.*

On peut alors prendre $\omega(u)=0$; l'intégrale (9) sera semi-convergente, et on peut aisément lui attribuer une valeur déterminée. Il suffit de considérer deux courbes C' et C'', de part et d'autre de l'axe $u=0$, et tendant vers cet axe. Pour chaque intervalle dt, les ordonnées u' et u'' de ces deux courbes seront liées par une relation telle que

$$\left(\int_{-1}^{u''}+\int_{u'}^{1}\right) d_u d_t N(u, t)$$

ait une limite. Si d'ailleurs on ne cherche qu'à définir la loi dont dépend $x(T)$, ou si l'espace V étudié est un espace W, on peut prendre pour ces courbes des parallèles à l'axe des t.

On peut ainsi dire que $x(t)$ représente bien la somme des sauts; mais ce n'est pas une fonction à variation bornée; cette somme est seulement semi-convergente, et on peut la définir de plusieurs manières différentes en faisant varier la relation entre les courbes C' et C''. Passer d'une définition à l'autre revient à modifier le terme en iz dans la formule (9), c'est-à-dire à ajouter à $x(t)$ une fonction indépendante du hasard. L'indétermination qui existe sur $x(t)$ est donc la même que dans le cas précédent, et la différence entre ces deux cas n'est pas très profonde. Celle entre le premier et les deux autres est au contraire essentielle.

§ 9. - Application à l'étude des lois stables.

Rappelons qu'une loi est dite stable si, x_1 et x_2 étant deux variables indépendantes obéissant à cette loi et a_1 et a_2 deux constantes positives quelconques, on peut trouver une constante positive A, fonction de a_1 et a_2, telle que

$$x=\frac{a_1x_1+a_2x_2}{A}$$

dépende de la même loi. J'ai montré (*Probabilités*, pp. 252-263) qu'il existe des types de lois stables dépendant de deux paramètres α et β (un type de lois est l'ensemble des lois déduites de l'une d'elles par le changement de x en cx). Elles sont définies par la formule

$$\log E\{e^{izx}\}=-\frac{|z|^\alpha}{\Gamma(\alpha+1)}\left(1+i\frac{z}{|z|}\beta \operatorname{tg}\frac{\pi\alpha}{2}\right), \tag{13}$$

où $0<\alpha\leqslant 2$, $|\beta|\leqslant 1$; la relation entre a_1, a_2, et A est

(14) $$A^\alpha = a_1^\alpha + a_2^\alpha.$$

J'ai appelé ces lois $L_{\alpha,\beta}$ (ou L_α, si $\beta=0$); L_2 est la loi de GAUSS; pour $\alpha=1$, $\beta \operatorname{tg}\frac{\pi\alpha}{2}$ est une constante quelconque, qu'on peut ramener à zéro par l'addition à x d'une constante convenable; on est alors ramené à la loi L_1, ou loi de CAUCHY, pour laquelle à chaque intervalle dx correspond la probabilité $\frac{dx}{\pi(1+x^2)}$. Les autres lois $L_{\alpha,\beta}$ sont des lois continues pour lesquelles la densité de probabilité est une fonction transcendante.

CAUCHY avait vu dès 1853 que la formule (13), dans le cas où $\beta=0$, entraîne la propriété de stabilité; mais il n'avait pas établi que le second membre est une valeur acceptable pour le logarithme d'une fonction caractéristique; il s'agissait de démontrer que la densité de probabilité, qu'on déduit de la fonction caractéristique par la formule de FOURIER, n'est jamais négative. M. G. PÓLYA a signalé la difficulté et l'a résolue pour $\alpha<1$; je l'ai résolue dans le cas général et ai montré en même temps l'existence des lois stables dissymétriques (celles pour lesquelles $\beta\neq 0$).

Au point de vue qui nous occupe maintenant, il est évident que la formule

(15) $$\log E\{e^{izx(t)}\} = -t\frac{|z|^\alpha}{\Gamma(\alpha+1)}\left(1+i\frac{z}{|z|}\beta\operatorname{tg}\frac{\pi\alpha}{2}\right),$$

où $0<\alpha\leqslant 2$, $|\beta|\leqslant 1$, définit des espaces W, que nous désignerons par $W_{\alpha,\beta}$ (ou W_α, si $\beta=0$); pour ces espaces, $(\Delta t)^{-1/\alpha}\Delta x$ dépend d'une loi indépendante, non seulement de t, mais de Δt; ce sont les seuls pour lesquels le type de la loi dont dépend Δx ne dépende pas de t ni de Δt.

L'espace W_2, correspondant à la loi de GAUSS, n'est autre que l'espace différentiel de N. WIENER; les autres ne peuvent être que des espaces à discontinuités mobiles. Nous allons montrer qu'on peut les définir en partant des formules du § 8, et obtenir par là une nouvelle démonstration très simple de l'existence des lois stables.

Nous n'avons pour cela qu'à prendre

(16) $$d_u N(u,t) = \frac{ctdu}{|u|^{\alpha+1}},$$

c étant une constante, qui peut, toutefois, dépendre du signe de u. D'après le théorème II; cette expression sera acceptable si $0<\alpha<2$.

Premier cas: $0<\alpha<1$. Dans ce cas la formule (9) donne, en prenant d'abord $c=0$ pour $u<0$,

(17) $$\log E\{e^{izx(t)}\} = ct\int_0^\infty (e^{izu}-1)\frac{du}{u^{\alpha+1}}.$$

Le changement de variable $|z|u=v$ montre que le second membre est proportionnel, pour chaque signe de z, à $|z|^\alpha$; il est donc de la forme (13). La théorie des fonctions euleriennes (*Probabilités*, p. 259), montre d'ailleurs que

$$\int_0^\infty (e^{iu}-1)\frac{du}{u^{\alpha+1}}=\Gamma(-\alpha)\left(\cos\frac{\pi\alpha}{2}-i\sin\frac{\pi\alpha}{2}\right)$$
$$=\frac{-\pi}{2\Gamma(\alpha+1)\sin\frac{\pi\alpha}{2}}\left(1-i\operatorname{tg}\frac{\pi\alpha}{2}\right),$$

ce qui permet d'identifier exactement les lois définies par les formules (15) et (17), pour

$$\pi c=2\sin\frac{\pi\alpha}{2},\qquad \beta=-1. \tag{18}$$

La fonction $x(t)$, somme de sauts positifs, est donc non décroissante, et même croissante (le nombre des sauts étant infini dans toute intervalle), et par suite positive (pour $t>0$); cette propriété remarquable de la loi $L_{\alpha,-1}$ est déjà indiquée dans notre ouvrage cité (p. 261).

Naturellement, il suffit de considérer deux variables indépendantes x_1 et x_2, dépendant de la loi $L_{\alpha,-1}$, et de poser

$$x=\frac{1-\beta}{2}x_1-\frac{1+\beta}{2}x_2$$

pour que x dépend de la loi $L_{\alpha,\beta}$, dont l'existence est bien ainsi établie pour $|\beta|\leq 1$.

Deuxième cas: $1<\alpha<2$. La formule (17) doit ici être remplacée par

$$\log E\{e^{izx(t)}\}=ct\int_0^\infty (e^{izu}-1-izu)\frac{du}{u^{\alpha+1}}, \tag{19}$$

et, la fonction $x(t)$, ne se réduisant plus à la somme des sauts, n'est plus nécessairement croissante et positive. À cela près, les résultats obtenus dans le cas précédent, y compris la formule (18), subsistent. L'existence de la loi $L_{\alpha,\beta}$ est encore bien établie dans ce cas.

Troisième cas: $\alpha=1$. Il suffit alors de considérer le cas symétrique $\beta=0$. Il vient

$$\log E\{e^{izx(t)}\}=ct\int_0^\infty (\cos zu-1)\frac{du}{u^2} \tag{20}$$
$$=ct|z|\int_0^\infty(\cos u-1)\frac{du}{u^2}=-c\frac{\pi}{2}t|z|,$$

de sorte qu'on retrouve bien la loi de Cauchy, la formule (18) étant toujours applicable.

Les propriétés des lois $L_{\alpha,\beta}$ et des espaces $W_{\alpha,\beta}$ sont ainsi établies par application des résultats du § 8. Nous pouvons aisément les obtenir aussi par une méthode en quelque sorte inverse de la précédente, en partant de la formule (15) et étudiant la probabilité des grandes valeurs des variables. Ainsi (*Probabilités*, pp. 259-263), pour la loi définie par la formule (15), avec $\beta = -1$, la probabilité des valeurs de x supérieures à X est, pour X infiniment grand positif, ou, ce qui revient au même, pour t infiniment petit

$$P\{x(t) > X\} \sim \frac{Ct}{X^\alpha}, \qquad \left(C = \frac{2}{\pi\alpha} \sin\frac{\pi\alpha}{2}\right), \tag{21}$$

et on a la même formule en remplaçant t par Δt et $x(t)$ par Δx. Le nombre probable de réalisations de l'inégalité $\Delta x < X$, si l'on divise l'intervalle $(0, t)$ en intervalles très petits Δt, est donc équivalent, pour Δt très petit, à

$$\frac{Ct}{X^\alpha} = \frac{2}{\pi\alpha} \sin\frac{\pi\alpha}{2} \frac{t}{X^\alpha},$$

et à la limite, on obtient le nombre probable de sauts de hauteur $u > X$. Ce résultat concorde bien avec les formules (16) et (18).

Les méthodes qui précèdent s'étendent aisément à l'étude *des lois semi-stables* (Cf. *Probabilités*, pp. 270-277). Il suffit, dans les formules (17), (19) et (20), de multiplier l'élément différentiel par une fonction périodique (de période $\log q$) et non négative de $\log u$; si l'on ne compare que des valeur de z constituant une progression géométrique de raison q, la proportionnalité des seconds membres de ces formules à $|z|^\alpha$ subsiste; si alors on considère des valeurs de Δt constituant une progression géométrique de raison q^α, la loi dont dépend $(\Delta t)^{-1/\alpha}\Delta x$ sera la même pour toutes ces valeurs. Si donc on ajoute un grand nombre n de variables indépendantes obéissant à une telle loi, le produit de leur somme par $n^{-1/\alpha}$ dépendra d'une loi qui peut être considérée comme une fonction périodique, de période $\alpha \log q$ et continue dans une période, de $\log n$; si q^α est un entier q', pour des valeurs de n en progression géométrique de raison q', on retrouve la même loi de probabilité.

Naturellement, les lois que nous venons de définir en partant des formules (17), (19) et (20) sont des lois continues; il faudrait introduire des intégrales de STIELTJES pour obtenir des lois discontinues et avoir ainsi les lois semi-stables les plus générales.

§ 10. - Les espaces V les plus généraux.

D'après ce qui précède, on se rend compte que *les espaces V les plus généraux s'obtiennent en considérant $x(t)$ comme somme de quatre termes*:

1°) *un terme indépendant du hasard, qui est une fonction quelconque de t*;

2°) *un terme gaussien*;

3°) *la somme des discontinuités fixes;*

4°) *le terme correspondant aux discontinuités mobiles.*

Ces termes sont distincts, et indépendants, aux deux restrictions près qui suivent:

1°) le terme correspondant à chaque discontinuité fixe n'est défini qu'à une constante près, qu'on peut ajouter à ce terme et retrancher de la partie indépendante du hasard;

2°) Le somme des discontinuités mobiles peut n'être pas convergente; on ne peut alors donner un sens au dernier terme ci-dessus que par addition d'un terme indépendant du hasard, qui sera nécessairement retranché du premier; dans ce cas, la séparation entre le premier et le quatrième terme est arbitraire, une fonction continue de t pouvant être ajoutée à l'un et retranchée de l'autre; mais dans des cas particuliers, comme celui des lois stables, des considérations autres que celles qui interviennent dans la définition générale des espaces V, peuvent conduire à choisir pour le quatrième terme une forme plus simple que les autres.

Toutefois, les résultats qui précèdent ne sont pas encore établis; nous avons seulement défini par synthèse le terme correspondant aux discontinuités mobiles. Il nous reste à établir qu'après avoir éliminé les discontinuités fixes comme il a été indiqué au § 2, on peut éliminer les discontinuités mobiles en les groupant en un terme du type étudié au § 8, et qu'on se trouve ainsi ramené au cas où le théorème I s'applique.

Cette démonstration repose sur le lemme suivant:

Lemme. - *Les discontinuités fixes étant éliminées*:

1°) *les seules discontinuités possibles sont des sauts*;

2°) $x(t)$ *peut être considéré comme la somme des sauts de hauteurs supérieures en valeur absolue à n'importe quel nombre positif donné* U, *et d'un terme indépendant de cette somme.*

La première partie de ce lemme n'est nullement évidente; nous avons déjà démontré que pour chaque point t la probabilité d'une discontinuité de nature quelconque est nulle, lorsqu'on a éliminé les discontinuités fixes. C'est une question différente dont il s'agit ici: dans un intervalle $(0, t)$, la probabilité de l'existence de discontinuités est devenue positive, et il s'agit de montrer que la probabilité qu'il y ait d'autres discontinuités que des sauts reste nulle.

Admettons d'abord le lemme pour montrer que le résultat énoncé sur l'élimination des discontinuités mobiles en résulte. Considérons à cet effet une suite de nombres positifs U_1, U_2,...., U_n,...., décroissants et tendant vers zéro. Désignons par $\xi_n(t)$ la somme des sauts de hauteur u telle que

$$U_{n-1} \geqslant |u| > U_n.$$

Il est nécessaire que la somme $\sum \xi_n(t)$, ou bien soit convergente (sauf dans des cas de probabilité nulle), ou bien puisse être rendue convergente par addition de termes indépendants du hasard, de manière à conduire à une somme $\sum [\xi_n(t) - a_n(t)]$ du type étudié au § 8. Dans le cas contraire, en effet, la dispersion relative à la somme $\sum_1^n \xi_\nu(t)$, d'après les résultats rappelés au § 2, augmenterait indéfiniment; cela est impossible, puisqu'elle est inférieure à la dispersion relative à $x(t)$.

Posons alors

$$x(t) = \sum [\xi_n(t) - a_n(t)] + y_n(t) = S_n(t) + y_n(t). \tag{21}$$

Le premier terme $S_n(t)$ tend vers une limite $S(t)$ (sauf dans des cas de probabilité nulle); $y_n(t)$ tend de même vers une limite $y(t)$. Ces limites $S(t)$ et $y(t)$, sommes de séries convergentes de variables aléatoires, dépendent de lois déterminées, limites de celles dont dépendent $S_n(t)$ et $y_n(t)$. Ces lois sont d'ailleurs indépendantes, puisque, ε et ε' étant arbitrairement petits, on a, pour chaque t, pour n assez grand,

$$P\{|S(t) - S_n(t)| > \varepsilon'\} < \varepsilon,$$

ce qui revient à dire qu'en faisant une erreur arbitrairement petite sur $S(t)$ et $y(t)$, et négligeant des cas de probabilité arbitrairement petite, on peut confondre $S(t)$ et $y(t)$ avec les variables indépendantes $S_n(t)$ et $y_n(t)$; ce sont donc bien des variables indépendantes.

La fonction $y(t)$ ainsi obtenue est alors le résultat de l'élimination des discontinuités mobiles, et est, d'après le § 4, somme d'un terme indépendant du hasard et d'un terme dépendant de la loi de GAUSS.

Il reste à démontrer le lemme, ce qui est peut-être le point le plus délicat de la présente étude.

D'abord, un point de discontinuité autre qu'un saut implique l'existence, d'un même côté de ce point, de valeurs différentes pour les limites inférieure et supérieure de $x(t)$, et par suite d'une infinité d'oscillations entre ces deux valeurs. Il s'agit de montrer l'impossibilité de cette circonstance. La difficulté de la démonstration provient de ce que, pour choisir $x(t)$ comme il a été dit § 3, il faut choisir un procédé déterminé de division de l'intervalle $(0, T)$ en intervalles Δt très petits; on n'est pas sûr que les points de division choisis soient ceux qui mettent le mieux en évidence les oscillations de $x(t)$, et l'on doit se demander si le fait que le nombre des $|\Delta x|$ dépassant un nombre donné U, à chaque instant de l'opération, est borné, n'est pas compatible avec l'existence d'une infinité d'oscillations.

Plaçons-nous à un instant déterminé de l'opération; désignons par a, pour chaque intervalle Δt, la probabilité $P\{|\Delta x| \geqslant U\}$; $\sum a$ est alors le nombre probable des intervalles pour lesquels $|\Delta x| \geqslant U$. D'après l'hypothèse que la loi $\mathfrak{L}(t)$ varie d'une manière continue avec t (en raison de l'élimination des discontinuités fixes), et d'après le théorème de BOREL-LEBESGUE, non seulement chaque a est très petit, mais le plus grand des α (et même, ce qui nous servira plus loin, la probabilité $P\{|\delta x| \geqslant U\}$ pour tout δt intérieur à un même Δt) est inférieur à un nombre α' arbitrairement petit, si tous les Δt sont assez petits. Le nombre des intervalles pour lesquels $|\Delta x| \geqslant U$ dépendra naturellement à la limite de la loi de Poisson; mais il s'agit de montrer que sa valeur probable $\sum \alpha$ reste finie, quand les Δt deviennent très nombreux et très petits.

Soit à chaque instant n le nombre de ces intervalles; il est de l'ordre de grandeur de sa valeur probable $\nu = \sum \alpha$, et d'après l'expression de la loi de Poisson) la valeur la plus probable a une probabilité de l'ordre de $\frac{1}{\sqrt{\nu}}$. La dispersion est donc de l'ordre de grandeur $\sqrt{\nu}$ [indiqué autrement par le fait que $\sigma(n) = \sqrt{\nu}$]: à un intervalle (n_1, n_2) petit par rapport à $\sqrt{\nu}$ correspond une probabilité très petite. Si alors on considère la somme $x_1(T)$ des Δx au moins égaux à U en valeur absolue, la dispersion correspondante est au moins de l'ordre de $U\sqrt{\nu}$. Il en est ainsi même dans le cas le plus défavorable où les valeurs très voisines respectivement de U et $-U$ sont seules possibles (pour chaque Δx) et également probables; la compensation entre les termes positifs et les termes négatifs ne se produit qu'à une erreur près de l'ordre de grandeur indiqué. Dans aucun cas d'ailleurs le choix de la valeur exacte de Δx ne peut réduire la dispersion due à l'incertitude sur n. La dispersion de $x(T)$ ne pouvant qu'être supérieure à celle de $x_1(T)$, et étant bornée, $\nu = \sum \alpha$ est bien borné, pour chaque valeur fixe de U.

Dans ces conditions $\sum \alpha^2$ tend vers zéro; il en est de même de $\sum \beta$, si β est petit par rapport à α: nous pouvons négliger la probabilité totale de toute circonstance dont la probabilité β pour chaque intervalle partiel est petite par rapport à α. Nous allons montrer qu'il en est ainsi de la probabilité pour que l'on ait à la fois $|\Delta x| < U$, et, à l'intérieur de l'intervalle considéré $\Delta t = t'' - t'$, une oscillation supérieure à $4U$. Cette circonstance implique l'existence d'au moins une valeur t de cet intervalle pour laquelle $|x(t) - x(t')| \geqslant 2U$. Soit t la plus petite valeur pour laquelle il en soit ainsi (ou la borne inférieure de ces valeurs); l'hypothèse faite ne modifiant en rien la loi dont dépend $x(t'') - x(t)$, on a

$$P\{|x(t'') - x(t)| \geqslant U\} \leqslant \alpha',$$

de sorte que, s'il y a une probabilité β' que la valeur considérée de t existe, il y a au moins une probabilité $\beta'(1-\alpha')$ que $|\Delta x| \geqslant U$; donc $\beta'(1-\alpha') \leqslant \alpha$, et β' est au plus de l'ordre de grandeur de α. La probabilité que cette valeur t existe, et que l'on ait tout de même $|\Delta x| < U$, est au plus $\alpha'\beta'$; elle est donc très petite par rapport à α.

Nous pouvons donc négliger la possibilité que des intervalles pour lesquels $|\Delta x| < U$ donnent une oscillation supérieure à $4U$. Sans doute cette circonstance peut se produire au début de l'opération. Mais lorsqu'on subdivise chaque Δt en intervalles plus petits, et ainsi de suite indéfiniment, il viendra un moment où cette circonstance cessera de se produire; on peut fixer le moment à partir duquel sa probabilité sera devenue inférieure à un nombre arbitrairement petit ε; c'est donc bien que la probabilité de sa répétition indéfinie est nulle.

Chaque point de discontinuité au voisinage duquel l'oscillation totale dépasse un nombre arbitrairement petite $4U$ implique donc l'existence d'une suite d'intervalles Δt intérieurs les uns aux autres, relatifs aux phases successives de la division de l'intervalle $(0, T)$ en intervalles partiels, et pour lesquels on aura à partir d'un certain moment $|\Delta x| \geqslant U$. Le nombre de ces suites est donc fini (puisqu'en négligeant une probabilité arbitrairement petite on peut borner à chaque instant le nombre n des intervalles pour lesquels $|\Delta x| \geqslant U$). Dans ces conditions, toute circonstance qui pour un de ces points a une probabilité nulle a une probabilité totale nulle. Nous allons montrer que c'est le cas pour l'existence d'une discontinuité autre qu'un saut et pour laquelle les oscillations dépassent un nombre fixe.

Cela résulte immédiatement de ce que l'existence d'une telle discontinuité n'est pas rigoureusement locale. Il viendra un moment, en subdivisant chaque Δt en intervalles plus petits δt, où les oscillations considérées devront exister et être très nombreuses, non seulement dans l'intervalle δt contenant le point des discontinuité, mais dans δt et à l'extérieur de δt. Or ce δt est distingué des autres uniquement par des circonstances intérieures à cet intervalle, nécessairement sans influence sur l'allure de $x(t)$ à l'extérieur. La probabilité de l'existence de ces oscillations dans Δt à l'extérieur de δt est donc très petite, comme elle le serait pour tout autre intervalle δt, ce qui achève la démonstration de la première partie du lemme.

Le seconde n'offre plus aucune difficulté. La somme $x_1(T)$ tend vers la somme des sauts supérieurs à U en valeur absolue (et de peut-être une partie de ceux égaux à $\pm U$, circonstance sans importance, puisqu'elle n'intervient avec une probabilité positive que pour des valeurs de U constituant au plus une infinité dénombrable). Il reste à démontrer que la somme $x_2(T) = x(T) - x_1(T)$ peut à la limite être considérée comme une variable aléatoire indépendante de T. Nous montrerons à cet effet qu'on peut définir une variable aléatoire $\bar{x}_2(T)$, indépendante de $x_1(T)$, et telle que

$$P\{|\bar{x}_2(T) - x_2(T)| > \varepsilon'\} < 2\varepsilon,$$

ε et ε' étant arbitrairement petits.

Choisissons d'abord n assez grand pour que, les Δt étant tous très petits, le nombre des intervalles pour lesquels $|\Delta x| \geqslant U$ soit inférieur à n, sauf dans des cas de probabilité inférieure à ε; cela est possible, puisque ce nombre a une valeur probable bornée $\sum \alpha$; il suffit que $n\varepsilon > \text{Max} \sum \alpha$. Prenons ensuite les Δt assez petits pour que, dans chacun de ces intervalles

$$P\left\{|\Delta x| > \frac{\varepsilon'}{n}\right\} < \frac{\varepsilon}{(1-\alpha)n};$$

le maximum du premier membre tendant vers zéro et le second vers $\frac{\varepsilon}{n}$, cela est possible; $x(T)$ étant la somme des Δx choisis dans les différents intervalles Δt, prenons enfin pour $\bar{x}_2(T)=\bar{x}(T)$ la somme des termes $\Delta\bar{x}$ égaux à Δx si $|\Delta x|<U$ et obtenus en recommençant l'expérience, si $|\Delta x|\geqslant U$; toutefois il faut que, dans l'expérience ainsi recommencée, l'éventualité $|\Delta x|\geqslant U$ soit exclue, la probabilité de chaque valeur inférieure à U étant alors multipliée par $\frac{1}{1-\alpha}$. Dans ces conditions, $\bar{x}_2(T)$ est indépendant aussi bien du nombre et des rangs des intervalles pour lesquels on aura eu d'abord $|\Delta x|\geqslant U$, que des valeurs de Δx dans ces intervalles; donc $\bar{x}_2(T)$ et $x_1(T)$ sont bien indépendants. D'autre part on passe de $\bar{x}_2(T)$ à $x_2(T)$ en remplaçant dans n intervalles au plus par zéro une valeur initiale de Δx telle que $|\Delta x|\leqslant\frac{\varepsilon'}{n}$, c'est-à-dire en faisant une erreur au plus égale à ε', sauf dans des cas de probabilité totale au plus égale à

$$\varepsilon+n(1-\alpha)\frac{\varepsilon}{(1-\alpha)n}=2\varepsilon, \qquad \text{c. q. f. d.}$$

Le lemme, et par suite le fait que les espaces V les plus généraux sont bien définis comme il a été dit au début du présent paragraphe, sont bien démontrés.

§ 10. - L'arithmétique des lois de probabilité.

Si deux variables indépendantes et leur somme dépendent de lois $\mathfrak{L}'$, $\mathfrak{L}''$, et $\mathfrak{L}$, nous considérerons $\mathfrak{L}$ comme le produit de $\mathfrak{L}'$ et $\mathfrak{L}''$; la multiplication de deux lois équivaut donc à celle de leurs fonctions caractéristiques.

Désignons par L la loi de GAUSS, par L'_u la loi définie par

$$\log E\{e^{izx}\}=e^{izu}-1,$$

c'est-à-dire la loi de POISSON avec le coefficient u; en ne distinguant pas deux lois déduites l'une de l'autre par addition à x d'un terme indépendant du hasard, les résultats obtenus expriment que, après élimination des discontinuités fixes, la loi $\mathfrak{L}$ dont dépend $x(T)$ est nécessairement de la forme

$$\mathfrak{L}=L^{\alpha}\Pi(L'_u)^{\varphi(u)},$$

Π désignant un produit généralisé (en ce sens que son logarithme est une intégrale de STIELTJES). Cette formule constitue une décomposition de $\mathfrak{L}$ en facteurs premiers, les éléments premiers autres que L dépendant d'un paramètre continu u, et l'exposant étant, pour chaque élément, un nombre positif ou nul quelconque.

Nous allons montrer que cette décomposition est unique.

Observons à cet effet que la loi de probabilité dont dépendent les sauts de $x(t)$ est une loi à deux variables. Mais on ne modifie pas la loi dont dépend $x(T)$, ni la fonction $N(u, T)$, en déplaçant la probabilité parallèlement à l'axe des t

(t restant compris entre 0 et T) et la répartissant d'une manière uniforme sur chaque parallèle à Ot. Cela revient à dire que, sans changer (quels que soient u_1 et u_2 de même signe) la loi de probabilité dont dépend le nombre des sauts pour lesquels $0<t<T$, $u_1<u<u_2$, on peut supposer $N(u, t)$, et par suite $\log E\{e^{izx(t)}\}$, proportionnels à t. La loi dont dépend chaque Δx est ainsi bien définie par la formule

$$\log E\{e^{iz\Delta x}\} = \frac{\Delta t}{T} \log E\{e^{izx(T)}\},$$

quand on connaît celle dont dépend $x(T)$. On peut donc déterminer

$$\sum P\{u_1 < \Delta x < u_2\}$$

c'est-à-dire à la limite le nombre probable $N(u_2, T) - N(u_1, T)$ de sauts pour lesquels $u_1 < u < u_2$ (sauf peut-être pour les valeurs où ce nombre varie brusquement, ce qui est sans importance).

Ainsi chaque loi ne peut être obtenue que d'une seule manière; si la répartition de la probabilité dans l'intervalle $(0, T)$ peut être modifiée, le nombre probable des sauts de chaque espèce est bien déterminé par la donnée de la loi dont dépend $x(T)$, et l'exposant du facteur gaussien obtenu après élimination des discontinuités mobiles est bien déterminé aussi, c. q. f. d.

Ce théorème d'unicité de la décomposition entraîne naturellement toutes les conséquences connues en arithmétique élémentaire: théorème de GAUSS, définition et propriétés des plus grands communs diviseurs.

Naturellement, cette arithmétique ne s'applique qu'aux lois $\mathfrak{L}$ susceptibles d'être rencontrées dans la théorie des espaces V où $\mathfrak{L}(t)$ varie d'une manière continue avec t. Leurs fonctions caractéristiques sont définies par

$$(22) \qquad \log \Phi(z) = -a\frac{z^2}{2} + \int_{-\infty}^{+\infty} [e^{izu} - 1 - iz\omega(u)]dN(u),$$

et, indépendamment de cette représentation, se reconnaissent, d'après ce qui précède, à ce que $\Phi^{\varepsilon}(z)$ est une fonction caractéristique, quelque petit que soit ε. Si cette condition est vérifiée, la théorie précédente donne le moyen de résoudre l'équation intégrale (22), en en tirant a et $N(u)$; $\omega(u)$ reste naturellement indéterminé.

Pour d'autres lois que celles que nous venons d'étudier, il est facile de montrer que l'arithmétique considérée ne s'applique pas. Ainsi une variable ξ ayant comme valeurs possibles 0 et 1, ces deux valeurs étant également probables, ne peut être somme de deux variables aléatoires indépendantes (à moins que l'une ne soit constante). La loi correspondante a le caractère d'un facteur premier. Posons alors

$$x = \frac{\xi_1}{2} + \frac{\xi_2}{4} + \dots + \frac{\xi_p}{2^p} + \dots,$$

$\xi_1, \xi_2, \dots, \xi_p, \dots$, étant des variables indépendantes dont chacune obéit a cette loi; x étant une variable choisie au hasard entre 0 et 1 (chaque intervalle ayant une probabilité égale à sa longueur), la loi $\mathfrak{L}$ dont dépend x se trouve ainsi décomposée en facteurs premiers. Or, en faisant intervenir la numération à base 3, on a de même pour x une décomposition de la forme

$$x = \frac{\eta_1}{3} + \frac{\eta_2}{9} + \dots + \frac{\eta_p}{3^p} + \dots,$$

et pour la loi $\mathfrak{L}$ une autre décomposition en facteurs premiers. Si nous considérons un des termes $3^{-p}\eta_p$, ou la somme de plusieurs termes, la différence de deux valeurs possibles d'une telle variable n'étant jamais divisible par 2^{-p}, la loi dont elle dépend est un diviseur de la loi $\mathfrak{L}$ sans contenir aucun des facteurs premiers obtenus dans la première décomposition.

Naturellement, dans ce qui précède, on pourrait remplacer 2 et 3 par d'autres entiers positifs; pourvu que le second ne divise aucune puissance du premier, le raisonnement subsiste.

Une question se pose encore, au sujet des lois représentables par la formule (22). *L'unicité de leur décomposition subsiste-t-elle si l'on admet d'autres facteurs composants que ceux qui interviennent dans cette formule?*

Nous pensons que la réponse est affirmative; en ce qui concerne la loi de GAUSS, c'est d'ailleurs un résultat que depuis 1931 nous considérons comme vraisemblable, sans avoir pu le démontrer. Il faut remarquer qu'en tout cas la loi de GAUSS est la seule loi pouvant jouir de cette propriété d'être décomposable en facteurs élémentaires, la décomposition étant unique et les facteurs étant du type qu'elle-même; les autres lois stables sont en effet un produit de facteurs élémentaires dépendant de la loi de POISSON; et la loi de POISSON comprend une infinité de lois, dont l'ensemble constitue un groupe, mais dont chacune prise isolément n'est pas stable.