

BULLETIN DE LA S. M. F.

PAUL LEVY

L'addition des variables aléatoires définies sur une circonférence

Bulletin de la S. M. F., tome 67 (1939), p. 1-41

http://www.numdam.org/item?id=BSMF_1939__67__1_0

© Bulletin de la S. M. F., 1939, tous droits réservés.

L'accès aux archives de la revue « Bulletin de la S. M. F. » (<http://smf.emath.fr/Publications/Bulletin/Presentation.html>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

BULLETIN

DE LA

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

L'ADDITION DES VARIABLES ALÉATOIRES DÉFINIES SUR UNE CIRCONFÉRENCE;

PAR M. PAUL LÉVY.

1. Introduction. — Soit $F(x)$ une fonction de la variable réelle x , à variation bornée de $-\infty$ à $+\infty$; considérons-la comme définissant une répartition de masses sur une droite D . Posons

$$(1) \quad m = \int_{-\infty}^{+\infty} dF(x), \quad M = \int_{-\infty}^{+\infty} |dF(x)|;$$

c'est-à-dire que m et M représentent respectivement la masse totale et la somme des masses prises en valeur absolue.

L'enroulement de la droite D sur une circonférence Γ de longueur unité conduit à ne pas distinguer des masses situées en des points dont les abscisses aient même partie fractionnaire; x devient une variable définie mod 1. On est ainsi conduit à associer à la répartition sur D une répartition sur Γ , dont la fonction de répartition $\Phi(x)$ sera définie à une constante près par la formule

$$(2) \quad \Delta\Phi(x) = \sum_h \Delta F(x+h),$$

en posant, pour une fonction quelconque $G(x)$,

$$\Delta G(x) = \Delta_a^b G(x) = G(b) - G(a) \quad (1),$$

et désignant par \sum_h une sommation étendue à toutes les valeurs

(1) Pour l'application de cette définition à $F(x+l)$, il doit être entendu que c'est x (et non $x+l$) qui varie de a à b ; il serait plus correct d'employer une notation telle que Δ_x . Dans la suite, pour éviter toute ambiguïté, nous désignerons toujours par x la variable qui doit varier de a à b .

entières de h , de $-\infty$ à $+\infty$. Il résulte évidemment de l'hypothèse faite sur $F(x)$ que la série (2) est toujours absolument convergente; elle définit, à une constante près, une fonction $\Phi(x)$, à variation bornée sur tout intervalle fini, et définissant une répartition périodique, c'est-à-dire que $\Phi(x+l) - \Phi(x)$ est une fonction périodique de l , de période unité; on peut bien la considérer comme définissant une répartition sur Γ . On a d'ailleurs

$$(3) \quad \begin{cases} m' = \int_a^{a+1} d\Phi(x) = \Phi(a+1) - \Phi(a) = m, \\ M' = \int_a^{a+1} d\Phi(x) \leq M, \end{cases}$$

et, si la fonction donnée $F(x)$ est non décroissante (cas du calcul des probabilités)

$$m = M = M'$$

(l'égalité $M = M'$ peut d'ailleurs être réalisée même si $|m| < M$.)

Bien entendu, si la donnée de $F(x)$ détermine $\Phi(x)$ à une constante additive près, la réciproque n'est pas vraie. La donnée de $\Phi(x)$ n'empêche pas que l'on peut choisir arbitrairement, en dehors d'un intervalle semi-ouvert de longueur unité, une répartition de masses sur D ; on peut ensuite, d'une manière et d'une seule, répartir des masses dans cet intervalle de manière que la fonction de répartition $F(x)$ soit liée à $\Phi(x)$ par la formule (2).

La relation entre $F(x)$ et $\Phi(x)$ prend une forme bien simple par la transformation de Fourier. On sait que la fonction $F(x)$ est définie (toujours à une constante près) par sa transformée de Fourier, ou fonction caractéristique

$$(4) \quad \Lambda(t) = \int_0^1 e^{2\pi itx} dF(x),$$

et que la fonction $\Phi(x)$ est de même définie à une constante près par ses coefficients de Fourier, définis par la formule

$$(5) \quad a_n = \int_{\Gamma} e^{2\pi inx} d\Phi(x) \quad (n = \dots, -1, 0, 1, \dots),$$

où l'intégrale étendue à Γ équivaut à une intégrale étendue à n'importe quel intervalle semi-ouvert de longueur unité sur l'axe des x . On a d'ailleurs

$$(6) \quad a_n = \Lambda(n),$$

de sorte que considérer seulement $\Phi(x)$ équivaut à ne tenir compte que des valeurs de $A(t)$ pour les valeurs entières de n . Quoique la transformée de Fourier d'une fonction à variation bornée ne soit pas une fonction quelconque, il est bien clair qu'elle n'est pas déterminée par les valeurs qu'elle prend pour les valeurs entières de t ; cela correspond au fait que la répartition sur D n'est pas déterminée par la répartition sur Γ .

Tout cela est bien connu. Il nous semble pourtant qu'il n'est pas inutile d'attirer l'attention sur les services que peut rendre la relation établie par la formule (1) entre la répartition de masses sur une droite et leur répartition sur une circonférence. Les analystes étudient en général séparément les problèmes relatifs à ces deux modes de répartition; or, il y a souvent intérêt à les étudier parallèlement; grâce à la formule (1), les résultats relatifs aux fonctions $F(x)$ peuvent se transformer en résultats relatifs aux fonctions $\Phi(x)$.

L'objet du présent Mémoire est d'appliquer ces remarques à l'étude simultanée des équations de composition

$$(7) \quad \Delta F(x) = \int_{-\infty}^{+\infty} \Delta F_1(x-y) dF_2(y),$$

$$(8) \quad \Delta \Phi(x) = \int_{\Gamma} \Delta \Phi_1(x-y) d\Phi_2(y),$$

dont on connaît l'importance en calcul des probabilités. Il nous arrivera de retrouver des résultats connus; mais nous obtiendrons aussi un grand nombre de résultats nouveaux, relatifs aux lois stables sur Γ , aux séries à termes aléatoires indépendants sur Γ , et à la théorie des lois indéfiniment divisibles sur Γ .

Il nous arrivera de parler de l'extension des problèmes traités à certains cas où $F(x)$ n'est pas à variation bornée sur tout l'axe réel.

2. Le théorème fondamental. — Nous désignerons ainsi le théorème suivant :

THÉORÈME. — *Si $F(x)$, $F_1(x)$, $F_2(x)$ sont trois fonctions à variations bornées de $-\infty$ à $+\infty$, liées par la relation (7), il leur correspond, par la formule (2), trois fonctions $\Phi(x)$, $\Phi_1(x)$, $\Phi_2(x)$, qui vérifient la relation (8).*

Rappelons d'abord que, si l'on convient, pour les points de discontinuité de $F_1(x)$, d'attribuer à cette fonction la valeur $F_1(x-0)$, la formule (7) conduit à définir une fonction $F(x)$ qui est, comme $F_1(x)$, continue à gauche. On peut faire d'autres conventions. Mais la répartition définie par $F(x)$ est indépendante de ces conventions, et dépend symétriquement des répartitions définies par $F_1(x)$ et $F_2(x)$. Les mêmes remarques s'appliquent à la formule (8).

Nous allons maintenant donner, du théorème énoncé, deux démonstrations différentes.

1° *Démonstration directe.* — D'après (2) et (7), on a

$$\begin{aligned}\Delta\Phi(x) &= \Sigma_h \int_{-\infty}^{+\infty} \Delta F_1(x+h-y) dF_2(y) \\ &= \Sigma_h \Sigma_k \int_0^1 \Delta F_1(x+h-y-k) dF_2(y+k),\end{aligned}$$

d'où, s'il est légitime d'intervertir l'ordre des opérations de sommation et d'intégration

$$\begin{aligned}(9) \quad \Delta\Phi(x) &= \int_0^1 \Sigma_k [\Sigma \Delta F_1(x+l-y)] dF_2(y+k) \\ &= \int_0^1 \Delta\Phi_1(x-y) \Sigma_k dF_2(y+k),\end{aligned}$$

ce qui coïncide avec la formule à démontrer.

Pour légitimer le calcul, il suffit de montrer que la somme triple (9) est absolument convergente. Or, si p est la partie entière de $b-a$, chaque point de l'axe des x appartient à au plus $p+1$ des intervalles $(a+l, b+l)$; en posant

$$M_1 = \int_{-\infty}^{+\infty} |dF_1(x)|, \quad M_2 = \int_{-\infty}^{+\infty} |dF_2(x)|;$$

on a donc

$$\Sigma_l |\Delta F_1(x+l-y)| \leq (p+1) M_1,$$

et la somme triple considérée est majorée par $(p+1)M_1M_2$, ce qui justifie le calcul fait tout à l'heure et termine la démonstration.

2° *Démonstration par les formules de Fourier.* — Cette

démonstration repose sur le fait bien connu que, si $A(t)$, $A_1(t)$ et $A_2(t)$ sont respectivement les transformées de $F(t)$, $F_1(t)$ et $F_2(t)$ par la formule (4), la formule (7) équivaut à

$$(10) \quad A(t) = A_1(t) A_2(t).$$

De même, si a_n , a'_n et a''_n sont les coefficients de Fourier de $\Phi(t)$, $\Phi_1(t)$ et $\Phi_2(t)$, la formule (8) équivaut à

$$(11) \quad a_n = a'_n a''_n \quad (n = \dots, -1, 0, 1, \dots).$$

Or, d'après la formule (6), on a

$$a_n = A(n), \quad a'_n = A_1(n), \quad a''_n = A_2(n),$$

de sorte que (10) entraîne (11). Donc (6) entraîne (7),

C. Q. F. D.

Remarques. — Le théorème précédent s'applique en particulier au cas de répartitions absolument continues, respectivement définies sur D et sur Γ par les densités $f(x)$ et $\varphi(x)$. La fonction $f(x)$ est une fonction sommable de $-\infty$ à $+\infty$, à cela près quelconque; la somme

$$(2') \quad \varphi(x) = \sum_h f(x+h)$$

est alors une fonction périodique, de période unité, sommable dans tout intervalle fini; elle est donc la densité d'une répartition sur Γ , correspondant, dans l'enroulement de D sur Γ , à celle définie par $f(x)$.

En introduisant ces densités, les formules (4) et (5) deviennent

$$(4') \quad A(t) = \int_{-\infty}^{+\infty} e^{2\pi i t x} f(x) dx,$$

$$(5') \quad a_n = \int_{\Gamma} e^{2\pi i n x} \varphi(x) dx,$$

et les formules (7) et (8) deviennent

$$(7') \quad f(x) = \int_{-\infty}^{+\infty} f_1(x-y) f_2(y) dy,$$

$$(8') \quad \varphi(x) = \int_{\Gamma} \varphi_1(x-y) \varphi_2(y) dy.$$

Notre théorème fondamental montre alors que : à trois fonctions $f_1(x)$, $f_2(x)$ et $f(x)$ vérifiant la relation (7') correspondent par la formule (2') trois fonctions $\varphi_1(x)$, $\varphi_2(x)$ et $\varphi(x)$ vérifiant la relation (8').

3. Les répartitions stables. — Nous dirons qu'une répartition est stable si elle est identique à son carré de composition (1). La stabilité s'exprime donc, suivant le cas considéré, par l'une des formules

$$(12) \quad \Delta F(x) = \int_{\mathfrak{D}} \Delta F(x-y) dF(y),$$

$$(12') \quad f(x) = \int_{\mathfrak{D}} f(x-y)f(y) dy,$$

$$(13) \quad \Delta \Phi(x) = \int_{\Gamma} \Delta \Phi(x-y) d\Phi(y),$$

$$(13') \quad \varphi(x) = \int_{\Gamma} \varphi(x-y)\varphi(y) dy,$$

et l'emploi de la transformation de Fourier ramène les deux premières équations à la relation algébrique

$$(14) \quad \Lambda(t) = \Lambda^2(t);$$

les deux dernières se ramènent de même aux formules

$$(15) \quad a_n = a_n^2 \quad (n = \dots, -1, 0, 1, \dots).$$

Étudions les conséquences de ces formules.

1° Dans le cas où l'on répartit sur D des masses finies ($M < \infty$), la fonction $\Lambda(t)$ est continue. La formule (14) implique donc que la fonction $\Lambda(t)$ soit, ou bien constamment nulle, ou bien constamment égale à l'unité. Dans le premier cas, il n'y a aucune masse; dans le second, il y a une masse unique, égale à l'unité, placée à l'origine. Il n'y a pas d'autre répartition stable.

(1) Cette définition ne coïncide pas avec celle que j'ai introduite en calcul des probabilités; j'ai en effet appelé *loi stable*, ce qu'il eût été plus correct d'appeler *loi d'un type stable*. Cela est sans inconvénient, parce qu'en calcul des probabilités il n'y a pas de loi stable au sens du présent travail (sauf celle qui correspond au cas d'une variable aléatoire qui ne serait pas réellement aléatoire, mais serait sûrement nulle).

Cette conclusion ne subsiste pas si M est infini. Cherchons, par exemple, les répartitions stables, absolument continues, définies par des fonctions $f(x)$ de carrés sommables de $-\infty$ à $+\infty$. On sait que dans ces conditions on peut appliquer la transformation de Fourier sous la forme indiquée par M. Plancherel : la fonction de t

$$\int_{-X}^{+X} e^{2\pi itx} f(x) dx$$

converge en moyenne quadratique, quand X augmente indéfiniment, vers une fonction de carré sommable $A(t)$, qui est la transformée de Fourier-Plancherel de $f(x)$; on déduit inversement $f(x)$ de $A(t)$ par une formule réciproque analogue à la formule de Fourier, mais où la convergence est aussi remplacée par la convergence en moyenne. Cette transformation laisse subsister la correspondance entre les formules (12') et (14), en ce sens que, si l'une est vérifiée presque partout, il en est de même de l'autre; la même remarque s'applique aux formules (7') et (10).

La résolution de la formule (12'), dans le domaine des fonctions de carrés sommables, est alors immédiate. Elle se ramène à l'équation (14). La fonction $A(t)$, devant être de carré sommable et vérifier presque partout cette équation, est égale à 1 sur un ensemble e mesurable et de mesure finie, et presque partout nulle en dehors de e .

On remarque d'ailleurs que, si $f(x)$ est de carré sommable et vérifie presque partout l'équation (12'), on ne change pas la répartition étudiée en supposant cette fonction partout égale au second membre de cette équation, qui est une fonction continue de x [car, dans l'espace des fonctions de carrés sommables de y , le point représentant $f(x - y)$ varie d'une manière continue avec x , c'est-à-dire que $f(x - y)$ tend en moyenne quadratique vers $f(x_0 - y)$ quand x tend vers x_0]. On n'écarte donc aucune solution de notre problème en supposant que la fonction $f(x)$, si elle est de carré sommable, est continue.

Un cas particulier remarquable est celui où l'ensemble e se réduit à l'intervalle $(-c, +c)$. On trouve alors

$$(16) \quad f(x) = \frac{\sin 2\pi cx}{\pi x} \quad (c > 0),$$

et cette fonction, étant continue, vérifie partout l'équation (12'); on peut d'ailleurs facilement le vérifier en appliquant le théorème des résidus au calcul du second membre. De même, si l'on considère une suite finie ou infinie de nombres positifs décroissants c_0, c_1, \dots , la fonction

$$(17) \quad f(x) = \frac{1}{\pi x} \sum (-1)^h \sin 2\pi c_h x$$

est une solution de l'équation (12'); elle correspond à un ensemble e constitué par les intervalles $(c_0, c_1), (c_1, c_2), \dots$, et par leurs symétriques par rapport à l'origine.

2° Occupons-nous maintenant des répartitions stables sur Γ , et étudions-les d'abord indépendamment du problème précédent. Cette étude se ramène à celle de l'équation (15), d'après laquelle chacun des a_n est égal à 0 ou 1.

Si l'on considère d'abord les répartitions absolument continues, définies par des fonctions $\varphi(x)$ sommables et vérifiant presque partout l'équation (13'), a_n , qui est le coefficient de Fourier de $\varphi(x)$, doit tendre vers zéro pour $|n|$ infini. Cela revient à dire qu'un nombre fini de ces coefficients peuvent être seuls égaux à 1, tous les autres étant nuls (1). La solution générale de l'équation (13') est donc

$$(18) \quad \varphi(x) = \sum_{-N}^{+N} a_n e^{2i\pi n x},$$

N étant un entier positif, et chacun des a_n étant égal à 0 ou 1 (2); les solutions réelles sont de la forme

$$(19) \quad \varphi(x) = a_0 + 2 \sum_1^N a_n \cos 2\pi n x;$$

3° Pour les répartitions non absolument continues, ou dont on

(1) Il y a longtemps que j'indique cette résolution de l'équation (13'), comme application de la théorie des séries de Fourier, à mes élèves de l'École Polytechnique (je l'ai proposée comme exercice en 1914, et introduite dans mon cours en 1923). Mais jusqu'ici j'avais toujours supposé la fonction $\varphi(x)$ de carré sommable. On voit que la conclusion subsiste si elle est seulement supposée sommable.

(2) Il résultera de la suite que la seule solution acceptable pour le calcul des probabilités est $\varphi(x) = 1$.

ignore si elles le sont, on ne peut plus affirmer que a_n tend vers zéro. Plaçons-nous d'abord dans le cas du calcul des probabilités, où $m = a_0 = 1$, et où $\Phi(x)$ est non décroissant. Une *valeur possible* est une valeur au voisinage de laquelle $\Phi(x)$ varie; l'ensemble e de ces valeurs est fermé, et a pour complément l'ensemble e' des intervalles ouverts dans chacun desquel $\Phi(x)$ est constant. Pour une répartition stable, c'est-à-dire vérifiant l'équation (13), la somme de deux valeurs possibles est toujours une valeur possible (1). Il en résulte que, ou bien toutes les valeurs sont possibles, ou bien il n'y a qu'un nombre fini p de valeurs possibles, qui sont les multiples de $\frac{1}{p}$ (cela fait bien p points distincts sur Γ).

Supposons d'abord la fonction $\Phi(x)$ continue. On est nécessairement dans le cas où toutes les valeurs de x sont possibles, c'est-à-dire que $\Phi(x)$ est constamment croissant. D'après (13), $\Delta\Phi(x) = \Phi(a+l) - \Phi(a)$ est une moyenne pondérée des différentes valeurs de $\Phi(\xi+l) - \Phi(\xi)$, $\xi = x - \gamma$ décrivant Γ , et le poids $d\Phi(\gamma)$ n'étant jamais nul. Donc, à moins que $\Phi(\xi+l) - \Phi(\xi)$ ne soit constant, $\Phi(a+l) - \Phi(a)$ ne peut pas atteindre les valeurs extrêmes de cette différence; cette conclusion est manifestement absurde, puisque le maximum de cette différence ne change pas si l'on remplace ξ par a (qui est aussi l'abscisse d'un point quelconque de Γ). C'est donc que cette différence est constante; comme cela est vrai quel que soit l , c'est que la répartition est uniforme.

Si au contraire la fonction $\Phi(x)$ est discontinue, il y a des valeurs à probabilités positives; il y en a au plus une infinité dénombrable. Désignons par x_ν une quelconque de ces valeurs, par α_ν sa probabilité, par α la somme $\sum \alpha_\nu$ (on a $\alpha \leq 1$), et par μ le plus grand des α_ν . Si la répartition est stable, la formule (13), appliquée à un intervalle réduit à un point, donne

$$\alpha_h = \sum \alpha_\nu \Pr \{X = x_h - x_\nu\} \leq \mu \sum \alpha_\nu \leq \mu \alpha \leq \mu.$$

Or, il faut bien que α_h atteigne son maximum μ , ce qui exige que $\alpha = 1$, et que, pour au moins une valeur de h , toutes les différences $x_h - x_\nu$ aient même probabilité μ ; leur nombre p est alors

(1) D'une manière générale, si X et Y sont deux variables aléatoires indépendantes, la somme d'une valeur possible de X et d'une valeur possible de Y est une valeur possible de $X + Y$.

fini, et il n'y a pas d'autres valeurs possibles que les x_v . On est donc dans le cas où il y a p valeurs possibles, multiples de $\frac{1}{p}$, et, chaque x_v étant aussi un $x_h - x_v$, elles ont toutes la même probabilité $\mu = \frac{1}{p}$.

Donc : *la stabilité, en calcul des probabilités, implique la répartition uniforme de la probabilité, soit sur toute la circonférence, soit entre les sommets d'un polygone régulier à p côtés ayant un sommet à l'origine* [la valeur $p = 1$ n'est pas exclue, mais correspond au cas où la variable aléatoire définie par $\Phi(x)$, n'ayant qu'une valeur possible, n'est pas réellement aléatoire].

Nous désignerons par $\Phi_p(x)$ la fonction de répartition correspondant à cette répartition uniforme entre les points d'abscisses curvilignes multiples de $\frac{1}{p}$; pour cette répartition, a_n est égal à 1 si n est multiple de p et nul dans le cas contraire.

4° Abandonnons maintenant les conditions restrictives du calcul des probabilités. Nous avons à chercher les fonctions à variations bornées, réelles, et solutions de (13), c'est-à-dire celles pour lesquelles chacun des coefficients de Fourier est égal à 0 ou 1.

Or, en distinguant dans $\Phi(x)$ la partie continue et la partie discontinue, on met a_n sous la forme

$$a_n = a'_n + a''_n \quad (a''_n = \sum \alpha_h e^{2\pi i n x_h}),$$

a'_n tendant vers zéro pour n infini, et a''_n étant une somme finie ou une série absolument convergente ($\sum |\alpha_h| < M'$). Quelque petit que soit ε positif, nous pouvons déterminer N de manière que pour $|n| \geq N$ on ait $|a'_n| < \frac{\varepsilon}{6}$, et réduire a''_n à une somme finie avec une erreur également inférieure à $\frac{\varepsilon}{6}$. On peut ensuite choisir un entier P de manière que, pour les termes conservés dans a''_n , les produits $P x_h$ diffèrent de nombres entiers de moins de $\frac{\varepsilon}{6\pi M'}$. Alors, si les entiers n et $n' = n + P$ sont en module au moins égaux à N , on a

$$|a_n - a_{n'}| < \varepsilon,$$

et, comme a_n et $a_{n'}$ doivent être entiers, il suffit de prendre $\varepsilon < 1$

pour conclure que ces coefficients sont égaux. Donc, compte tenu de ce que a_n et a_{-n} , qui sont toujours imaginaires conjugués, et qui sont ici réels, sont égaux :

la solution la plus générale de l'équation (13) s'obtient en prenant une suite périodique et symétrique de coefficients de Fourier tous égaux à 0 ou 1 et en modifiant un nombre fini de coefficients, sans détruire la symétrie ni introduire d'autres valeurs que 0 ou 1 [bien entendu la symétrie n'est nécessaire que si l'on cherche les solutions réelles de l'équation (13)]. D'ailleurs, toutes les suites ainsi obtenues conviennent bien à notre problème, car une suite périodique et de période P est la suite des coefficients de Fourier d'une répartition ne comportant que P valeurs possibles, multiples de $\frac{1}{P}$, et, pour que cette répartition soit réelle il suffit que a_n et a_{-n} soient imaginaires conjugués (or ils sont ici réels et égaux). La modification d'un nombre fini de coefficients n'empêche pas la suite considérée d'être la suite de Fourier d'une répartition, combinaison finie de celles définies par les fonctions $\Phi_p(x)$ et les fonctions (19); enfin, tous les a_n étant égaux à 0 ou 1, l'équation (13) est vérifiée.

Des solutions remarquables sont données par la formule

$$(20) \quad \Phi(x) = \Phi_p(x) - \Phi_q(x),$$

si q est multiple de p ; par

$$(21) \quad \Phi(x) = \Phi_p(x) + \Phi_{p'}(x) - \Phi_q(x),$$

si q est multiple d'un au moins des nombres p et p' et divise leur plus petit commun multiple, par

$$(22) \quad \Phi(x) = \Phi_p(x) + \Phi_{p'}(x) - 2\Phi_q(x),$$

si q est le plus petit commun multiple de p et p' et, plus généralement, par

$$(23) \quad \Phi(x) = \Phi_p(x) - \Phi_q(x) + \Phi_{p'}(x) - \Phi_{q'}(x) + \Phi_{p''}(x) - \dots,$$

le nombre des termes étant fini, et p, q, p', \dots , vérifiant des conditions arithmétiques simples. Le nombre P des termes de la période est le plus petit commun multiple de tous les entiers p, q, p', \dots ; mais, même en introduisant tous les diviseurs d'un nombre

donné P , on n'obtient pas en général la solution générale définie plus haut [cela est évident si P est un nombre premier supérieur à 3; on ne trouve que deux solutions indépendantes, $\Phi_1(x)$ et $\Phi_p(x)$, au lieu de $\frac{P-1}{2}$; c'est seulement pour les valeurs 1, 2, 3, 4, et 6 de P que la solution la plus générale correspondant à une valeur donnée de P peut se déduire de la formule (23)].

5° Occupons-nous maintenant de la liaison que le théorème fondamental du paragraphe 2 établit entre les équations (12) et (13). D'après ce théorème, de toute répartition stable sur D et pour laquelle M soit fini, on déduit une répartition stable sur Γ . Mais on n'obtient ainsi que deux solutions triviales, celle qui correspond à l'absence de toute masse, et celle qui correspond à une masse unité placée à l'origine. Or, nous venons de voir qu'il y en a d'autres.

Celles de ces solutions qui sont continues, et par suite de la forme (18), peuvent cependant être obtenues par application de la formule (2'), mais en partant des répartitions à masse totale infinie, dont la formule (16) nous a donné l'exemple le plus simple. Pour cette détermination (16) de $f(x)$, la formule (2') donne

$$(24) \quad \varphi(x) = \sum_h \frac{\sin 2\pi c(x+h)}{\pi(x+h)} \quad (c > 0),$$

et, pourvu que c ne soit pas entier, cette série, quoique non absolument convergente, est convergente (elle converge séparément à gauche et à droite). La transformée de Fourier $A(t)$ de $f(x)$ est bien définie, pour n'importe quelle valeur de t ; elle est égale à 1, $\frac{1}{2}$, ou zéro, suivant que $|t|$ est inférieur, égal ou supérieur à c . Les coefficients de Fourier de $\varphi(x)$ sont d'ailleurs toujours les valeurs de $A(t)$ pour t entier, de sorte que, si c n'est pas entier, ils sont tous égaux à 0 ou 1; $\varphi(x)$ définit donc une répartition stable sur Γ . D'une manière précise, a_n ayant la valeur un pour $|n| < c$, il vient

$$(25) \quad \varphi(x) = \sum_h \frac{\sin 2\pi c(x+h)}{x+h} = 1 + 2 \Sigma' \cos 2\pi nx,$$

Σ' désignant une sommation étendue à toutes les valeurs de n

positives et inférieures à c . On remarque, ce qui n'est pas évident d'après sa première expression, que $\varphi(x)$ ne dépend que de la partie entière de c .

On peut obtenir de la même manière n'importe laquelle des fonctions (18); il n'y a qu'à choisir des nombres croissants $c_1, c'_1, c_2, c'_2, \dots, c_p, c'_p$, non entiers, et tels que l'ensemble des entiers intérieurs aux intervalles (c_h, c'_h) coïncide avec l'ensemble des valeurs de n pour lesquels $a_n = 1$. La fonction $A(t)$ égale à 1 à l'intérieur de ces intervalles, à $\frac{1}{2}$ à leurs extrémités, et nulle en dehors, est la transformée de Fourier de la fonction

$$f(x) = \sum_1^p \frac{e^{-2\pi i c_h x} - e^{-2\pi i c'_h x}}{2\pi i x},$$

et l'application à cette expression de la formule (2') donne bien la fonction (18) que l'on voulait représenter.

On peut se proposer de retrouver de la même manière les répartitions stables discontinues formées tout à l'heure à l'aide des fonctions $\Phi_p(x)$. Nous n'insisterons pas sur cette question, qui présente surtout l'intérêt de donner un exemple de répartition continue [celle obtenue en limitant la série (2') à un nombre fini de termes] qui devient discontinue à la limite.

4. Premiers éléments de l'arithmétique des lois de probabilité sur Γ . — Les formules (8), (8') ou (11) définissent la *résultante*, ou le *produit* de deux répartitions sur Γ . Si l'on multiplie plus de deux facteurs, la multiplication, d'après la dernière de ces formules, est associative et commutative. On peut étudier l'arithmétique des produits de répartition sur Γ . Nous allons nous placer dans le cas du calcul des probabilités. Cette arithmétique des lois de probabilité sur Γ présente à la fois à certains points de vue de grandes analogies avec celle relative à D et à d'autres des différences importantes. Nous allons d'abord indiquer ses caractères les plus simples. Il est entendu que nous ne tenons pas compte des facteurs unités, c'est-à-dire des lois ne comportant qu'une seule valeur possible.

1° Une loi ne comportant que deux valeurs possibles est indé-

composable si la différence de ces valeurs n'est pas $\frac{1}{2} \pmod{1}$. S'il y a au plus p valeurs possibles, et qu'elles correspondent à des points de Γ situés sur un arc inférieur à une demi-circonférence, la loi considérée ne peut pas être le produit de plus de $p - 1$ facteurs.

Il existe des lois indéfiniment divisibles; on en obtient par l'enroulement des lois indéfiniment divisibles définies sur D . Ces lois seront étudiées au paragraphe 8.

2° Appellons *répartition périodique* celle qui admet une période autre que l'unité. Si une telle répartition n'est pas uniforme, la plus petite période est de la forme $\frac{1}{p}$ (p entier > 1). Toute période d'une répartition composante est aussi une période de la répartition résultante; la démonstration est immédiate, soit d'après la formule (8), soit d'après la formule (11), compte tenu de ce qu'une répartition de période $\frac{1}{p}$ est caractérisée par le fait que tous les a_n d'indices non multiples de p sont nuls.

Mais on remarque que $a_n = a'_n a''_n$ est nul pour tous ces indices si a'_n est nul pour une partie seulement de ces indices et a''_n pour les autres. Une répartition résultante peut donc avoir une période qui n'appartient à aucune répartition composante. Ainsi, P étant le plus petit commun multiple de deux entiers p et q dont aucun ne divise l'autre, en composant deux répartitions ayant respectivement pour plus petites périodes $\frac{1}{p}$ et $\frac{1}{q}$, on obtient une répartition résultante ayant la période $\frac{1}{P}$, que n'a aucune des répartitions composantes.

3° Ces remarques s'appliquent à la répartition uniforme, qu'on peut caractériser par l'existence de périodes arbitrairement petites. Pour cette répartition, tous les a_n autres que a_0 sont nuls. Si une des répartitions composantes est uniforme, il en est de même de la répartition résultante; mais la répartition résultante peut être uniforme sans qu'aucune des composantes le soit, et même sans qu'aucune soit périodique; ainsi les répartitions définies par les

fonctions

$$\varphi_1(x) = 1 + \frac{1}{2}(\cos 2x + \cos 3x), \quad \varphi_2(x) = 1 + \frac{1}{2}(\cos 4x + \cos 5x),$$

dont aucune n'est périodique, ont pour résultante la répartition uniforme.

On obtient des décompositions remarquables de la répartition uniforme sur Γ en la considérant comme résultant de l'enroulement sur Γ de la répartition uniforme sur l'intervalle $(0, 1)$ de la droite D , répartition dont les décompositions en facteurs sont bien connues (¹).

4° Appliquons maintenant les remarques du 2° aux répartitions comportant seulement un nombre fini p de valeurs possibles, égales aux p multiples de $\frac{1}{p}$; elles correspondent donc aux sommets d'un polygone régulier à p côtés ayant un sommet à l'origine. La suite des coefficients de Fourier, pour une telle fonction, est périodique et de période p , et, par suite définie par la donnée de a_1, a_2, \dots, a_{p-1} (a_0 ayant en tout cas la valeur 1); mais a_n et a_{-n} sont imaginaires conjugués; il en résulte que, si $p = 2p'$ ou $2p' + 1$, la répartition est bien définie par les coefficients d'indices $1, 2, \dots, p'$ (ce dernier étant réel si $p = 2p'$). Naturellement, ces coefficients ne sont pas quelconques; mais en les prenant assez petits on est sûr de trouver des probabilités positives pour chacune des p valeurs possibles. S'ils sont nuls, ces p probabilités sont égales; la répartition est périodique; la fonction de répartition est la fonction $\Phi_p(x)$ considérée au paragraphe 3.

Si alors $p' > 1$ (donc $p \geq 4$), les remarques du 2° s'appliquent. *Un produit de deux répartitions entre les p valeurs considérées peut être une répartition uniforme entre ces p valeurs sans qu'aucun des facteurs le soit; il suffit que a'_n soit nul pour certaines des valeurs $1, 2, \dots, p'$ de n , et que a''_n soit nul pour les autres. Si, au contraire, $p' = 1$ (donc $p = 2$ ou 3), la répartition uniforme est caractérisée par $a_1 = 0$, donc $a'_1 a''_1 = 0$, et pour que la répartition résultante soit uniforme, il faut et il suffit qu'une des répartitions composantes le soit.*

(¹) Cf. P. LÉVY, *Arithmétique des lois de probabilité* (*Journal de Mathématiques*, t. XVII, 1938).

Naturellement ces remarques s'appliquant à l'addition de variables aléatoires X_ν , ayant comme valeurs possibles

$$0, \frac{1}{p}, \dots, \frac{p-1}{p},$$

s'appliquent par là même à la multiplication des variables $e^{2\pi i X_\nu}$, qui ont pour valeurs possibles les racines $p^{\text{ièmes}}$ de l'unité.

5° Si l'on compose les deux répartitions définies par $\Phi_p(x)$ et $\Phi_q(x)$, la répartition résultante est celle définie par $\Phi_P(x)$, P étant le plus petit commun multiple de p et q ; cela se voit immédiatement, soit directement, soit par l'intermédiaire des coefficients de Fourier.

5. **Convergence des suites de répartitions sur D et sur Γ .** — Un grand nombre de définitions de la convergence sont possibles; nous allons en considérer deux.

1° *Convergence en moyenne quadratique.* — Il s'agit nécessairement dans ce cas de répartitions absolument continues, à densités de carrés sommables sur D (ou sur Γ), et la convergence dont il s'agit est la convergence en moyenne quadratique des densités $f^{(\nu)}(x)$ [ou $\varphi^{(\nu)}(x)$] vers une limite $f(x)$ [ou $\varphi(x)$]. Par la transformation de Fourier-Plancherel, la convergence en moyenne de $f^{(\nu)}(x)$ vers $f(x)$ équivaut à la convergence en moyenne des fonctions caractéristiques $A^{(\nu)}(t)$ vers la transformée de Fourier $A(t)$ de $f(x)$. Dans le cas des répartitions sur Γ , la convergence en moyenne équivaut à la convergence forte (dans l'espace d'Hilbert) de la suite des coefficients de Fourier $a_{(\nu)}^n$ de $\varphi^{(\nu)}(x)$ vers la suite des coefficients de Fourier a_n de $\varphi(x)$.

Les conditions de convergence d'un produit infini résultent immédiatement de ces définitions, et des formules (10) et (11).

Cherchons par exemple les conditions de convergence de la suite des puissances d'une répartition donnée; la limite ne peut être qu'une répartition stable. S'il s'agit d'une répartition sur D, définie par sa fonction caractéristique $B(t)$, $B^\nu(t)$ doit tendre en moyenne vers une limite $A(t)$, égale à 1 sur un ensemble e mesurable et de mesure finie, et presque partout nulle sur l'ensemble complémentaire e' . Il est donc nécessaire et suffisant

que la fonction $B(t)$, qui, par hypothèse, est mesurable, soit égale à 1 sur e et de module presque partout inférieur à 1 sur e' .

De même, sur Γ , la condition nécessaire et suffisante pour que la suite des puissances d'une répartition définie par ses coefficients de Fourier b_n ait une limite est que, un nombre fini de ces coefficients ayant la valeur 1, les modules des autres aient une borne supérieure inférieure à l'unité.

Il faut maintenant observer que, en faisant correspondre à chaque répartition sur D la répartition sur Γ définie par la formule (2') (du moins si cette formule a un sens), la convergence en moyenne sur D n'entraîne pas nécessairement la convergence en moyenne sur Γ . Cela est bien évident, puisque la formule même qui définit la répartition enroulée peut n'avoir pas de sens si on l'applique à la répartition limite.

Prenons, par exemple, pour $e^{(\nu)}$ l'ensemble constitué par l'intervalle $(\nu - \frac{1}{\nu}, \nu + \frac{1}{\nu})$ et son symétrique par rapport à l'origine, et pour $A^{(\nu)}(t)$ la fonction égale à 1 dans $e^{(\nu)}$ et nulle en dehors; pour ν infini, elle tend en moyenne vers zéro. Les coefficients $a_n^{(\nu)}$ qui s'en déduisent par la formule (6) sont égaux à 1 si $n = \pm \nu$ et nuls pour les autres valeurs de n ; donc $\varphi^{(\nu)}(x) = 2 \cos \nu x$; il n'y a donc pas convergence de la répartition enroulée vers une limite.

Il est facile de varier cet exemple et de faire en sorte qu'il n'y ait même pas convergence faible; tel est le cas si l'on remplace dans ce qui précède l'intervalle $(\nu - \frac{1}{\nu}, \nu + \frac{1}{\nu})$ par $(\nu' - \frac{1}{\nu}, \nu' + \frac{1}{\nu})$, ν' étant un entier qui varie périodiquement avec ν .

2° *Convergence au sens du calcul des probabilités.* — Nous appellerons ainsi le mode de convergence caractérisé par les deux conditions suivantes :

Condition préliminaire a, sur D. — Les intégrales

$$M^{(\nu)} = \int_{-\infty}^{+\infty} |dF^{(\nu)}(x)|$$

sont bornées dans leur ensemble, et également convergentes;

Condition a sur Γ . — Les intégrales

$$\int_{-\infty}^{+\infty} d\Phi^{(v)}(x)$$

sont bornées dans leur ensemble.

Condition de convergence b (sur D ou sur Γ). — On peut définir la constante indéterminée dont dépend la fonction de répartition $F^{(v)}(x)$ [ou $\Phi^{(v)}(x)$] de manière qu'elle tende vers une limite $F(x)$ [ou $\Phi(x)$], sauf peut-être aux points de discontinuité de la fonction limite ⁽¹⁾.

D'après la condition préliminaire, les répartitions considérées appartiennent à un ensemble compact. On en déduit aisément ⁽²⁾ que : *compte tenu de a, la condition b équivaut à la suivante*, qui fait intervenir les transformées de Fourier.

Condition c sur D. — $A^{(v)}(t)$ tend vers une limite $A(t)$ qui est continue à l'origine; la convergence est d'ailleurs nécessairement uniforme dans tout intervalle fini ⁽³⁾.

Condition c sur Γ . — $a_n^{(v)}$ tend vers a_n ⁽⁴⁾.

L'équivalence des conditions *b* et *c* ne subsiste pas si l'on ne tient pas compte de *a*.

Si maintenant nous considérons, en même temps que des répartitions définies sur D, celles qui leur correspondent sur Γ dans l'enroulement de D sur Γ , il résulte de $M' < M$ [formule (3)] que

⁽¹⁾ Dans le cas des répartitions sur D, il résulte de la condition préliminaire *a* que $F^{(v)}(-\infty)$ est bien défini, et l'on supprime toute indétermination dans la définition de la fonction $F^{(v)}(x)$ correspondant à une fonction donnée en supposant $F^{(v)}(-\infty) = 0$.

⁽²⁾ Cf. Notre récent Mémoire du *Bulletin des Sciences Mathématiques*, t. LXII, 1938.

⁽³⁾ Nous utilisons ici un résultat de M. Cramer, qui permet de ne pas introduire l'uniformité de la convergence dans l'énoncé de la condition *c*; la convergence est nécessairement uniforme dans tout intervalle fini si la fonction limite est continue à l'origine.

⁽⁴⁾ Rappelons la démonstration, dans le cas des répartitions sur Γ . Il est évident que *a* et *b* entraîne *c*. Réciproquement, si *a* et *c* sont vérifiés, et si *b* ne l'était pas, les répartitions considérées, qui appartiennent à un ensemble compact, auraient un élément d'accumulation distinct de la loi défini par les a_n ; il résulte de la première partie du théorème que c'est impossible.

si la condition a est vérifiée pour les répartitions initiales, elle l'est pour les répartitions enroulées. Il en est de même de la condition c , d'après la formule (6). Donc, à l'inverse de ce qui avait lieu pour la convergence en moyenne, la convergence vers une limite d'une suite de répartitions définies sur D entraîne la convergence de celles qui leur correspondent dans l'enroulement de D sur Γ .

6. Cas du calcul des probabilités; convergence des séries à termes aléatoires indépendants. — C'est, comme nous l'avons déjà dit, celui où $m = M = 1$; alors M' a aussi la valeur 1, et $F(x)$ et $\Phi(x)$ sont des fonctions non décroissantes. La condition a du paragraphe 4, 2°, étant vérifiée, la convergence d'une suite est définie indifféremment par la condition b ou par la condition c . Les produits infinis que nous étudions correspondent à des séries

$$(26) \quad X_1 + X_2 + \dots + X_\nu + \dots,$$

à termes aléatoires indépendants les uns des autres, $F^{(\nu)}(x)$ étant la fonction de répartition de la somme S des ν premiers termes de cette série; on sait que, sur D , la convergence de $F^{(\nu)}(x)$ vers une limite est liée à la convergence presque sûre de cette série; sa somme dépend alors de la loi définie par la limite $F(x)$ de $F^{(\nu)}(x)$.

Commençons par une remarque qui s'applique aussi bien sur Γ que sur D ; considérons la formule (8), d'après laquelle $\Delta\Phi(x)$, pour un intervalle (a, b) de longueur l , est une moyenne pondérée des valeurs de $\Delta\Phi_1(x)$ pour l'ensemble des intervalles ayant cette longueur. Si donc nous désignons respectivement par $\omega(l)$ et par $\Omega(l)$ la borne inférieure et la borne supérieure de $\Phi(x+l) - \Phi(x)$, la première de ces expressions ne peut qu'augmenter, et la seconde ne peut que diminuer, quand on passe d'une répartition composante à la répartition résultante. La même remarque, d'après la formule (7), s'applique aux accroissements de $F(x)$. Dans les deux cas, $\omega^{(\nu)}(l)$ et $\Omega^{(\nu)}(l)$ désignant les déterminations de $\omega(l)$ et $\Omega(l)$ relatives à la somme $S^{(\nu)}$, ces deux fonctions ne peuvent que se rapprocher l'une de l'autre. Mais il y a une différence importante entre les deux cas : sur D , la densité moyenne est nulle, de sorte que l'on a nécessairement $\omega(l) = 0$; sur Γ , la densité moyenne est 1, la moyenne non pondérée des

valeurs de $\Phi(x+l) - \Phi(x)$ est donc l , et $\Omega^{(\nu)}(l)$ et $\omega^{(\nu)}(l)$ ne peuvent devenir égaux qu'en prenant la valeur commune l .

1° Rappelons d'abord les résultats relatifs aux répartitions sur D . La *concentration minima* $\omega^{(\nu)}(l)$ (ou probabilité minima pour la longueur l) étant nulle, il n'y a à étudier que la décroissance de la *concentration maxima* de $\Omega^{(\nu)}(l)$, qui a nécessairement une limite non négative $\Omega(l)$. Or, $l \leq l' \leq 2l$ entraîne évidemment

$$\Omega^{(\nu)}(l) \leq \Omega^{(\nu)}(l') \leq 2\Omega^{(\nu)}(l),$$

formule qui subsiste à la limite. Donc $\Omega(l)$ ne peut s'annuler qu'en étant identiquement nul. Il y a donc seulement deux cas possibles.

Dans le premier cas, $\Omega(l)$ est nul pour tout l positif; la concentration maxima relative à $S^{(\nu)}$ tend vers zéro pour ν infini, et inversement la *dispersion* de cette somme (longueur du plus petit intervalle fermé auquel correspond une probabilité au moins égale à un nombre positif α) augmente indéfiniment, quel que soit α . La série (26) est alors presque sûrement divergente (en effet, dans le cas contraire, elle aurait une somme finie, dans des cas de probabilité positive; la dispersion, pour α assez petit, ne serait pas infinie); il en est de même de toutes les séries $\Sigma(X_\nu - a_\nu)$, les a_ν étant des constantes (non aléatoires); c'est ce qu'on exprime en disant qu'il y a *divergence essentielle*.

Dans le second cas, $\Omega(l)$ est positif pour tout l positif, et peut-être même pour $l=0$. La dispersion, comme la somme d'une série convergente à termes positifs, augmente, mais reste finie; c'est le cas de *quasi-convergence*; on peut, dans ce cas, déterminer une suite de constantes νx_ν , de manière que la série $\Sigma(X_\nu - x_\nu)$ soit presque sûrement convergente.

Une règle simple, due à M. Kolmogoroff, permet de reconnaître si l'on est dans ce cas. Elle repose sur le fait que, dans le cas des séries à termes aléatoires indépendants les uns des autres, et bornés, la convergence presque sûre est liée à la convergence en moyenne quadratique, et caractérisée par la convergence de $\Sigma \sigma_\nu^2$, en posant

$$(27) \quad \mu_\nu = \mathfrak{M}\{X_\nu\}, \quad \sigma_\nu^2 = \mathfrak{M}\{X_\nu - \mu_\nu\}^2.$$

Pour les séries à termes non bornés, il est nécessaire que l'on puisse remplacer X_ν par une variable X'_ν , égale à X_ν sauf dans des cas dont la probabilité totale α_ν soit le terme général d'une série convergente, et toujours située dans un intervalle de longueur positive donnée L (la valeur choisie pour L est sans importance; l'essentiel est qu'elle ne dépende pas de ν); si cette condition est remplie, la quasi-convergence de ΣX_ν est liée à celle de $\Sigma X'_\nu$, caractérisée par la convergence de $\Sigma \sigma_\nu^2$ (σ_ν étant défini en partant de X'_ν comme σ_ν en partant de X_ν).

2° Étudions maintenant les conséquences des résultats précédents au point de vue de l'enroulement de D sur Γ . Nous dirons qu'une série Σx_ν converge sur Γ si les points représentant ses sommes successives ont une limite; cela revient à dire qu'on peut déterminer une suite d'entiers γ_ν tels que la série $\Sigma(x_\nu + \gamma_\nu)$ soit convergente. De même la convergence sur Γ d'une série aléatoire ΣX_ν est liée à celle de $\Sigma(X_\nu + Y_\nu)$, Y_ν étant un entier aléatoire convenablement déterminé en fonction de X_ν .

Nous allons montrer que, comme sur D , il n'y a pas d'autre circonstance possible que la quasi-convergence et la divergence essentielle, et que la quasi-convergence se reconnaît à la convergence d'une série $\Sigma \sigma_\nu^2$ tout à fait analogue à celle dont dépend la convergence presque sûre dans le cas des variables définies sur D et bornées.

Pour une variable aléatoire X définie sur Γ , nous désignerons par X' celle des abscisses curvilignes du point X qui est située dans un intervalle semi-ouvert donné $(\xi, \xi + 1)$; nous appellerons écart quadratique moyen et désignerons par $\sigma\{X\}$ le nombre non négatif défini par

$$(27') \quad \sigma^2\{X\} = \text{Min } \mathcal{N}\{(X' - x)^2\}$$

(\mathcal{N} désignant la valeur probable, et x et ξ variant de $-\infty$ à $+\infty$). On ne change pas la définition de σ en supposant $x = \xi + \frac{1}{2}$, car à chaque valeur donnée de x correspond un minimum obtenu pour $\xi = x - \frac{1}{2}$; si, d'ailleurs, cette valeur $x - \frac{1}{2}$ appartient à un intervalle qui correspond à un arc de Γ ne contenant aucune valeur

possible de X , on peut aussi choisir ξ arbitrairement dans cet intervalle sans changer $\mathcal{M}\{(X' - x)^2\}$. Inversement, si ξ est donné, le minimum de cette expression est obtenu pour la valeur bien déterminée $\mu := \mathcal{M}\{X'\}$ de x .

Le minimum σ est un minimum effectivement atteint pour au moins un point de x de Γ ; celle de ses abscisses curvilignes que l'on choisit est indifférente. On vérifie aisément que le point diamétralement opposé sur Γ à un tel point x ne peut pas être pour X une valeur à probabilité positive [de sorte qu'il est indifférent, pour la définition de σ , de considérer l'intervalle $(\xi, \xi + 1)$ comme ouvert à droite ou à gauche].

Nous désignerons par $\mu = \mu\{X\}$ une des abscisses d'un tel point minimisant. Il peut y en avoir plusieurs; tel est le cas évidemment pour une répartition périodique, et même, dans le cas de la répartition uniforme, μ est complètement indéterminé. Mais, dans ces cas, σ ne peut pas être très petit (il admet un minimum égal à $\frac{1}{4}$ réalisé si X admet seulement deux valeurs possibles, également probables, correspondant à deux points diamétralement opposés).

Nous désignerons par μ_ν et σ_ν les expressions $\mu\{X_\nu\}$ et $\sigma\{X_\nu\}$. On a alors le théorème suivant :

TÉORÈME. — Si la série $\Sigma \sigma_\nu^2$ est convergente, la série $(\Sigma X_\nu - \mu_\nu)$ est presque sûrement convergente sur Γ . Si la série $\Sigma \sigma_\nu^2$ est divergente, la série ΣX_ν est essentiellement divergente sur Γ .

La première partie est évidente : si la série $\Sigma \sigma_\nu^2$ est convergente, la série $\Sigma (X'_\nu - \mu_\nu)$ est convergente en moyenne, donc presque sûrement convergente, X'_ν étant celle des abscisses curvilignes du point X_ν qui appartient à l'intervalle $(\mu_\nu - \frac{1}{2}, \mu_\nu + \frac{1}{2})$.

Pour démontrer la réciproque, raisonnons par l'absurde. Si la série ΣX_ν n'est pas essentiellement divergente, on peut déterminer des constantes x_ν telles que la probabilité de la convergence de $\Sigma (X'_\nu - x_\nu)$ soit positive, X'_ν étant celle des abscisses curvilignes du point X_ν qui appartient à l'intervalle $(x_\nu - \frac{1}{2}, x_\nu + \frac{1}{2})$. Il s'agit d'une série définie sur D , à termes aléatoires indépendants et bornés; il en résulte que la série $\Sigma \mathcal{M}\{(X'_\nu - \mu'_\nu)^2\}$ [en

posant $\mu'_v = \mathfrak{N}\{X'_v\}$ est convergente [car autrement la série $\Sigma X'_v$ serait essentiellement divergente, et la série $\Sigma(X'_v - x_v)$ serait presque sûrement divergente]. Comme cette série est une majorante de $\Sigma\sigma_v^2$, la convergence de cette dernière série en résulte.

C. Q. F. D.

7. Cas du calcul des probabilités (suite). Convergence et quasi-convergence des produits infinis sur Γ . — Nous venons d'étudier la probabilité de la convergence de la série (26). Nous allons maintenant étudier la convergence du produit infini de lois de probabilité qui lui est associé, c'est-à-dire la convergence (au sens rappelé au début du paragraphe 5) de la suite des lois L_v , dont dépendent les sommes S_v . Il est évident que, comme sur D , la convergence presque sûre de la série (26) implique la convergence de la suite des lois L_v vers une limite L , qui est la loi dont dépend la somme de la série. La même remarque s'applique à la quasi-convergence [ou convergence presque sûre d'une des séries $\Sigma(X_v - \xi_v)$]. Mais nous verrons que, sur Γ , la réciproque n'est pas vraie; la série (26) peut être essentiellement divergente, et pourtant il y a toujours quasi-convergence de la suite des lois L_v vers une limite L .

Nous utiliserons plusieurs modes de raisonnement qui se complètent.

1° L'ensemble des lois de probabilité définies sur Γ est compact. La suite des L_v a donc, ou bien une limite L , ou bien des éléments d'accumulation L, L', \dots , qui constituent un ensemble fermé.

Si alors on démontre que toutes les lois L, L', \dots , appartiennent à un ensemble fermé \mathcal{E} , il en résulte que L_v tend vers \mathcal{E} , c'est-à-dire qu'il existe une loi L' infiniment peu différente de L_v et appartenant à \mathcal{E} . Nous dirons dans ce cas que L_v appartient à la limite \mathcal{E} ; cette manière de parler ne préjuge pas de l'existence d'une limite unique L .

2° Nous avons remarqué plus haut que $\omega^{(v)}(l)$ et $\Omega^{(v)}(l)$, valeurs extrêmes de $\Phi^{(v)}(x+l) - \Phi^{(v)}(x)$ quand x décrit Γ [$\Phi^{(v)}(x)$ désignant la fonction de répartition de S_v], tendent d'une manière monotone vers des limites $\omega(l)$ et $\Omega(l)$. On a toujours

$$(28) \quad \omega^{(v)}(l) \leq \omega(l) \leq l \leq \Omega(l) \leq \Omega^{(v)}(l),$$

de sorte que, si

$$\Omega^{(v)}(l) = \omega^{(v)}(l),$$

qui est l'oscillation de

$$\Phi^{(v)}(x + l) - \Phi^{(v)}(x)$$

quand x décrit Γ , tend vers zéro, on a

$$(29) \quad \Omega(l) = \omega(l) = l.$$

Contrairement à ce qui se passe sur D pour l'égalité $\Omega(l) = 0$, cette égalité peut être réalisée pour certaines valeurs de l et ne pas l'être pour d'autres. Elle indique en effet que, pour n'importe laquelle des trois limites L, L', \dots , la répartition est périodique et de période l ; si l est rationnel, cela est possible sans que la répartition soit uniforme.

Il y a donc trois cas à distinguer :

Premier cas. — L'égalité (29) est vérifiée quel que soit l ; alors L_v tend vers la répartition uniforme.

Deuxième cas. — Il existe un entier tel que l'égalité (29) soit vérifiée pour les valeurs de l multiples de $\frac{1}{p}$, et pour celle-là seulement. La répartition définie par L_v , sans tendre vers l'uniformité, tend à devenir périodique et de période $\frac{1}{p}$, c'est-à-dire qu'il existe une loi L'_v ayant la période indiquée et différant infiniment peu de L_v .

Troisième cas. — L'égalité (29) n'est vérifiée pour aucune valeur non entière de l ; alors les répartitions limites L, L', \dots , ne sont ni uniformes, ni périodiques.

3° Supposons les lois composantes définies par leurs coefficients de Fourier $b_n^{(v)}$; les lois L_v sont alors définies par les coefficients

$$(30) \quad a_n^{(v)} = b_n^{(1)} b_n^{(2)} \dots b_n^{(v)}.$$

Comme $|b_n^{(v)}| \leq 1$, $|a_n^{(v)}|$ tend d'une manière monotone vers une limite $|a_n|$, et toutes les valeurs limites de $a_n^{(v)}$ (a_n, a'_n, \dots) ont même module $|a_n|$.

On sait alors immédiatement dans lequel des cas énumérés tout à l'heure on se trouve. Le premier cas est caractérisé par $a_n = 0$

pour tout n non nul (comme $a_n^{(v)}$ et $a_{-n}^{(v)}$ sont imaginaires conjugués, il suffit de considérer les valeurs positives de n). Le second cas est caractérisé par $a_n = 0$ pour tout n non multiple de p , $|a_n|$ étant positif pour certaines valeurs dont le p. g. c. d. est p . Le troisième cas est caractérisé par le fait que $|a_n|$ soit positif pour des valeurs de n dont le p. g. c. d. est 1. En effet ces différentes propriétés, d'après des remarques déjà utilisées paragraphe 4, caractérisent respectivement sur Γ la répartition uniforme, les répartitions de période $\frac{1}{p} < 1$, et les répartitions non périodiques.

Si nous convenons de considérer le produit $\prod |b_n^{(v)}|$ comme convergent si

$$(31) \quad \lim_{v \rightarrow \infty} |b_n^{(v)} b_n^{(v+1)} \dots| = 1,$$

et comme divergent si cette limite est zéro, on remarque que a_n peut s'annuler dans deux cas bien différents, soit si un (au moins) des $b_n^{(v)}$ est nul, soit si le produit infini $\prod b_n^{(v)}$ est divergent. Il importe peu, dans ce qui précède, que a_n soit nul pour une de ces raisons ou pour l'autre, et la divergence des produits infinis $\prod |b_n^{(v)}|$ n'est pas nécessaire pour obtenir les coefficients nuls qui caractérisent les répartitions limites uniformes ou périodiques. Nous avons indiqué au paragraphe 4 des exemples qui précisent cette remarque; nous avons vu qu'un produit fini peut correspondre à une répartition uniforme ou périodique sans qu'aucun des facteurs ait cette propriété, et, une fois obtenue, cette propriété subsiste indéfiniment, que les produits $\prod |b_n^{(v)}|$ soient convergents ou divergents.

4° Nous allons maintenant montrer que toutes les lois limites L, L', \dots , appartiennent à l'ensemble fermé des lois déduites de l'une d'elles par addition d'une constante. D'après une remarque faite plus haut (1° du présent paragraphe), il existe donc dans cet ensemble une loi L'_v infiniment peu différente de L_v , ce qui revient à dire qu'il y a quasi-convergence de la suite des lois L_v .

S'il n'y a qu'une loi limite L (et il en est nécessairement ainsi dans le premier des cas énumérés tout à l'heure), il n'y a rien à démontrer. Supposons donc qu'il y ait au moins deux lois limites,

et désignons par L et L' deux telles lois; on peut déterminer des entiers croissants

$$\nu_1, \nu'_1, \nu_2, \nu'_2, \dots, \nu_h, \nu'_h, \dots,$$

auxquels correspondent des lois L_ν tendant vers L pour les indices ν_h et vers L' pour les indices ν'_h . Considérons alors les sommes

$$X_{\nu_{h+1}} + X_{\nu_{h+2}} + \dots + X_{\nu'_h};$$

elles dépendent de lois ayant au moins un élément d'accumulation L'' , et, par la manière même dont cette loi L'' est définie, on a $L' = LL''$; donc, a_n, a'_n, a''_n désignant respectivement les coefficients de Fourier de L, L', L'' , on a $a'_n = a_n a''_n$. Comme $|a_n| = |a'_n|$, si $|a_n| > 0$, on a

$$(32) \quad |a''_n| = 1.$$

D'après la définition des coefficients de Fourier, cela implique que, pour toutes les valeurs de x possibles pour la loi L'' , $e^{2\pi i n x}$ ait la même valeur. Elles correspondent donc sur Γ (s'il y en a plusieurs) aux sommets d'un polygone régulier à n' côtés, n' divisant n . Or ce raisonnement s'applique à des valeurs de n dont le p. g. c. d. est p dans le second cas et 1 dans le troisième. Il n'y a donc, pour la loi L'' , qu'une détermination possible dans le troisième cas; il peut y en avoir p dans le second; de toute façon $p x$ ($p = 1$ dans le troisième cas) correspond à un point bien déterminé de Γ ; et la formule (32) est vérifiée, non seulement pour certaines valeurs de n dont le p. g. c. d. est p , mais pour tous les multiples de p .

Or, dans le second cas, la loi L correspondant à une répartition périodique et de période $\frac{1}{p}$, l'addition d'un multiple de $\frac{1}{p}$ est indifférent, et la loi L' résulte, aussi bien dans ce cas que dans le troisième cas, de l'addition d'une constante non aléatoire à la variable aléatoire définie par la loi L .

Nous avons donc démontré le résultat annoncé : *il y a, dans les trois cas, convergence ou quasi-convergence vers une répartition L (uniforme dans le premier cas, périodique et non uniforme dans le second, non périodique dans le troisième).*

5° Nous allons maintenant étudier les relations qui existent entre

la convergence de la série (26) et celle du produit infini de lois de probabilité qui lui correspond. Il est évident qu'il faut pour cela donner de cette dernière convergence une définition qui ne soit pas modifiée par la suppression d'un nombre fini de facteurs.

Le cas de divergence essentielle du produit infini sera alors le cas où il y a convergence vers la répartition uniforme (premier cas ci-dessus), et où de plus cette circonstance ne peut pas être modifiée par la suppression d'un nombre fini de facteurs. En faisant intervenir les coefficients de Fourier, ce cas est caractérisé par le fait que, quels que soient n différent de zéro et ν , on a

$$(33) \quad b_n^{(\nu+1)} b_n^{(\nu+2)} \dots = 0.$$

Le *cas mixte* est celui où il y a quasi-convergence vers une répartition périodique (de période $\frac{1}{p}$) et peut-être uniforme; si l'on supprime un nombre fini assez grand de facteurs, l'uniformité disparaît et la périodicité subsiste. Ce cas est caractérisé par le fait que la formule (33) est vérifiée pour tout n non multiple de p , tandis que la formule (31) l'est au contraire pour des valeurs de n dont le p. g. c. d. est p ; elle l'est par suite, pour les raisons indiquées à propos de la formule (32), pour tout n multiple de p .

Le *cas de quasi-convergence* est alors celui où la formule (31) est vérifiée quel que soit n . Si l'on supprime un nombre suffisamment grand de lois composantes, il y a alors quasi-convergence vers une loi qui, non seulement n'est ni uniforme ni périodique, mais est arbitrairement voisine de l'unité (1).

On remarque qu'on peut définir le cas mixte comme étant celui où il y a convergence par rapport au module p , mais non par rapport au module 1. On peut alors prévoir le théorème suivant :

(1) C'est-à-dire la loi ne comportant qu'une seule valeur possible; on ne restreint rien en supposant cette valeur nulle.

On remarque que, pour la suite des sommes S_n , la convergence de leurs lois de probabilité vers une répartition non périodique implique la formule (31) pour toutes les valeurs de n , donc la convergence en probabilité vers zéro de $S_{\nu'} - S_\nu$ pour ν indéfiniment croissant et $\nu' > \nu$, donc la convergence en probabilité de la suite des S_n vers une variable aléatoire S . Il n'en serait pas ainsi pour une suite de variables aléatoires quelconques; la convergence des lois de probabilité, n'excluant pas l'indépendance des variables considérées, n'entraîne pas la convergence en probabilité.

Théorème. — *Le cas de quasi-convergence est réalisé si la série (26) est quasi-convergente sur Γ ; le cas mixte est réalisé si elle est quasi-convergente par rapport au module $\frac{1}{p}$ sans l'être par rapport au module 1; le cas de divergence essentielle est celui où elle est essentiellement divergente par rapport à n'importe quel module.*

La première partie est évidente. Si la série (26) est quasi-convergente, on peut, par l'addition à chaque terme d'une constante convenable, la rendre presque sûrement convergente. Alors, si ν est assez grand, et quel que soit $\nu' > \nu$, la somme

$$(34) \quad S_{\nu'} - S_{\nu} = X_{\nu+1} + X_{\nu+2} + \dots + X_{\nu'}$$

dépend d'une loi arbitrairement voisine de l'unité, et cette circonstance subsiste pour ν' infini [la somme (34) étant définie pour ν' infini, sauf dans des cas dont la probabilité est nulle].

Montrons maintenant que la quasi-convergence du produit infini entraîne celle de la série (26). D'après la définition de la quasi-convergence, on peut, par addition aux X_{ν} de constantes convenables, rendre convergent le produit considéré. Comme nous l'avons vu, il y a alors, non seulement convergence de la suite des lois dont dépendent les S_{ν} , mais convergence en probabilité, donc aussi convergence mutuelle en probabilité, de la suite des S_{ν} ; c'est-à-dire que, ε étant arbitrairement petit, on a, pour ν assez grand [$\nu > N(\varepsilon)$] et $\nu' \geq \nu$,

$$(35) \quad \Pr \{ |S_{\nu'} - S_{\nu}| < \varepsilon \pmod{1} \} > 1 - \varepsilon;$$

il faut montrer que la convergence presque sûre en résulte. Nous allons utiliser à cet effet un mode de raisonnement souvent employé dans la théorie des séries aléatoires. La variable X_{ν} , définie seulement sur Γ , dépendant d'un entier arbitraire, nous pouvons choisir cet entier de manière que $|X_{\nu}| \leq \frac{1}{2}$. Donnons-nous une longueur $l < \frac{1}{2}$. Désignons par e_h (pour $h > \nu$) l'hypothèse

$$\begin{aligned} |S_k - S_{\nu}| &\leq l & (k = \nu + 1, \nu + 2, \dots, h - 1), \\ |S_h - S_{\nu}| &> l, \end{aligned}$$

et par α_h la probabilité de cette hypothèse; posons

$$(36) \quad \alpha = \alpha_{\nu+1} + \alpha_{\nu+2} + \dots = \Pr \left\{ \limsup_{h > \nu} |S_h - S_\nu| > l \right\},$$

et prenons pour ν' un nombre assez grand pour que

$$\alpha_{\nu+1} + \alpha_{\nu+2} + \dots + \alpha_{\nu'} > \frac{\alpha}{2}.$$

Pour h compris entre ν et ν' [$\nu' \geq h > \nu > N(\varepsilon)$], on a, d'après la formule (35),

$$\Pr \left\{ |S_{\nu'} - S_h| < \varepsilon \pmod{1} \right\} > 1 - \varepsilon,$$

et, la différence $S_{\nu'} - S_h$ ne contenant que des termes X_k d'indices $k > h$, donc indépendants de ceux qui figurent dans l'hypothèse e_h , cette probabilité est indépendante de cette hypothèse. On a donc

$$\begin{aligned} & \Pr \left\{ e_h, |S_{\nu'} - S_h| > l - \varepsilon \pmod{1} \right\} \\ & \geq \Pr \left\{ e_h, |S_{\nu'} - S_h| < \varepsilon \pmod{1} \right\} \geq \alpha_h (1 - \varepsilon), \end{aligned}$$

d'où, par une sommation par rapport à h ($h = \nu + 1, \nu + 2, \dots, \nu'$)

$$\Pr \left\{ |S_{\nu'} - S_\nu| > l - \varepsilon \pmod{1} \right\} > \frac{1 - \varepsilon}{2} \alpha,$$

et compte tenu de (35), et si $\varepsilon < \frac{l}{2}$,

$$\frac{1 - \varepsilon}{2} \alpha < \varepsilon, \quad \alpha < \frac{2\varepsilon}{1 - \varepsilon}.$$

Donc, quelque petit que soit l , ε étant inférieur à la fois à $\frac{l}{2}$ et à un nombre arbitrairement petit ε' , et $\nu > N(\varepsilon)$, on a

$$\alpha < \frac{2\varepsilon'}{1 - \varepsilon'},$$

c'est-à-dire que α tend vers zéro pour ν infini. Il y a donc bien convergence presque sûre de la série (26).

Nous avons ainsi établi que la quasi-convergence de la série (26) et celle du produit infini sont liées (de même pour la convergence). Or on ne change rien à nos raisonnements en remplaçant Γ par une circonférence de rayon $\frac{1}{p}$; la convergence presque sûre ou la quasi-

convergence de la série (26) par rapport au module $\frac{1}{p}$ sont donc liées à la convergence ou à la quasi-convergence du produit infini dans les conditions indiquées dans la définition du cas mixte. Par suite, la divergence essentielle de la série entraîne celle du produit infini. La démonstration du théorème est ainsi complète.

Rappelons que : *dans le cas de divergence la répartition limite est uniforme; elle peut l'être dans les autres cas; dans le cas mixte, elle est périodique et de période $\frac{1}{p}$; elle peut l'être dans le cas de quasi-convergence. Dans ce dernier cas, la répartition limite peut donc être uniforme, ou périodique et non uniforme, ou encore non périodique* (1).

6° Terminons ce paragraphe par quelques remarques qui sont des applications évidentes des résultats précédents.

Considérons d'abord le cas où toutes les valeurs possibles de X_v sont multiples de $\frac{1}{p}$ (p indépendant de v); le cas de quasi-convergence et le cas mixte sont seuls possibles. Le premier est caractérisé par la convergence de la série $\Sigma \alpha_v$, en désignant par $1 - \alpha_v$ la probabilité de la plus probable des valeurs de X_v . Si cette série diverge, il y a, soit convergence vers une répartition uniforme entre les p valeurs possibles, soit quasi-convergence vers une répartition de période $\frac{1}{p}$ (p' étant un diviseur de p).

Considérons maintenant les puissances successives d'une loi quelconque. S'il y a au moins deux valeurs possibles, le cas de quasi-convergence est exclu, et il n'y a que deux cas à distinguer. Dans le premier, les valeurs possibles correspondent à des sommets d'un polygone régulier de p côtés; tous les sommets ne sont peut-être pas des valeurs possibles; mais nous pouvons, d'une part ramener une des valeurs possibles à zéro par addition d'une constante, d'autre part supposer qu'après cette réduction $\frac{1}{p}$ est le p. g. c. d.

(1) L'exemple de la série

$$\pm \frac{1}{4} \pm \frac{1}{8} \pm \frac{1}{16} \pm \dots$$

(les signes + et - étant, pour chaque terme, également probables) montre que, même dans le cas de convergence sûre, la répartition peut devenir uniforme même si elle ne l'est pour aucune somme finie.

des valeurs possibles (autrement on remplacerait p par un nombre plus petit, p' , diviseur de p). Dans ce cas, les puissances de la répartition étudiée ont pour limite une répartition uniforme entre les p valeurs possibles.

Dans tous les autres cas, les puissances de la loi donnée ont pour limite la répartition uniforme sur Γ .

Ce résultat s'applique même s'il n'y a que deux valeurs possibles, pourvu que la différence de ces valeurs soit irrationnelle. Comme les $\nu + 1$ valeurs possibles de S_ν ont des probabilités proportionnelles aux coefficients du binôme (pour l'exposant ν), le résultat ainsi obtenu peut se rattacher à des propriétés de ces coefficients.

8. Les chaînes continues et les lois indéfiniment divisibles sur Γ . — Considérons, au lieu de la suite des sommes S_ν , une fonction $X(t)$, à accroissements aléatoires indépendants, c'est-à-dire que, si $t_2 > t_1$, $X(t_2) - X(t_1)$ est indépendant, aussi bien de $X(t_1)$ que de toutes les valeurs de $X(t)$ pour $t < t_1$; les valeurs de la fonction sont, soit des variables réelles (représentées par des points de D), soit des variables définies module 1 (représentées par des points de Γ); $X(t - 0)$ et $X(t + 0)$ peuvent ne pas être égaux.

La fonction de répartition de $X(t) - X(s)$ (pour $t > s$) étant désignée par $F(s, t, x)$ si X est défini sur D et par $\Phi(s, t, x)$ si X est défini sur Γ , les propriétés énoncées s'expriment par les équations intégrales

$$(37) \quad \Delta F(s, u, x) = \int_D \Delta F(s, t, x - y) d_y F(t, u, y)$$

$$(38) \quad \Delta \Phi(s, u, x) = \int_\Gamma \Delta \Phi(s, t, x - y) d_y \Phi(t, u, y) \quad (s < t < u).$$

Pour chaque système de valeurs des paramètres s, t et u , ces équations sont des formes (7) et (8); le théorème fondamental du paragraphe 2 s'applique donc (que l'on soit ou non dans le cas du calcul des probabilités) et donne le moyen de déduire de chaque solution de l'équation (37) une solution de l'équation (38). Mais toutes les solutions de cette dernière équation ne peuvent pas être obtenues de cette manière.

Nous allons nous placer dans le cas du calcul des probabilités,

et montrer les caractères qui distinguent les solutions spéciales à l'équation (38) de celles qui se déduisent de solutions de l'équation (37).

1° La plus simple des solutions de l'équation (38) est obtenue en prenant $\Phi(s, t, x) = x$: quels que soient s et $t > x$, $X(t) - X(s)$ est une variable aléatoire choisie sur Γ avec répartition uniforme de la probabilité. La valeur de $X(t)$, pour chaque valeur de t , est donc choisie au hasard sur Γ , chacun de ces choix étant indépendant de tous les autres; il n'est restreint par aucune condition de continuité.

Si t désigne le temps, les fonctions $X(t)$ correspondent à un mouvement d'une particule infiniment mobile, dont chaque position n'aurait aucun lien avec ses positions antérieures. Un tel mouvement ne peut pas être utile à considérer dans les sciences appliquées. De plus, au point de vue du calcul des probabilités, le choix de $X(t)$ implique l'idée d'une loi définie dans un ensemble ayant une puissance supérieure à celle du continu; nous avons déjà montré ailleurs la difficulté qu'il y a à raisonner sur une pareille loi, à moins de se contenter d'étudier ce qu'on peut appeler ses projections (1).

Des remarques tout à fait analogues peuvent être faites au sujet du cas où $X(t) = x(t) + X_1(t)$, $x(t)$ étant une fonction donnée, et $X_1(t)$ ayant seulement p valeurs possibles, multiples de $\frac{1}{p}$, et également probables, les choix correspondant à différentes valeurs de t étant indépendants les uns des autres.

2° Pour une étude plus générale, nous dirons que t est un *point de divergence essentielle à gauche* si, $t_1, t_2, \dots, t_n, \dots$ étant une suite de valeurs croissantes et tendant vers t , la série $\Sigma[X(t_{n+1}) - X(t_n)]$ est, par rapport à n'importe quel module rationnel, essentiellement divergente. Il en résulte évidemment que, quel que soit $s < t$, la différence $X(t) - X(s)$ est une variable aléatoire choisie sur Γ avec une répartition uniforme de

(1) Les remarques que nous avons faites sur ce sujet dans notre théorie de l'addition des variables aléatoires ont été depuis développées par J. L. Doob, dont les conclusions confirment absolument ce que nous avons dit d'une manière plus concise.

la probabilité. Cette conclusion ne subsistant pas si la série considérée est quasi-convergente par rapport à un module rationnel (car elle est alors en défaut pour $s = t_n$), il en résulte que la définition considérée est indépendante du choix des points t_n .

Tout point t , qui est point d'accumulation de points de divergence essentielle à gauche, situés à sa gauche, est lui-même un point de divergence essentielle à gauche.

On définit de même les points de divergence essentielle à droite. Un point de divergence essentielle est un point t qui a cette propriété au moins d'un côté. Alors, si $s < t < u$, la différence $X(u) - X(s)$ est une variable aléatoire choisie sur Γ avec une répartition uniforme de la probabilité.

L'ensemble de ces points est fermé. L'ensemble des points pour lesquels il y a quasi-convergence des deux côtés, par rapport à au moins un module, est donc ouvert.

Des remarques analogues s'appliquent aux *points mixtes*, qui sont des points de divergence, au moins d'un côté, par rapport au module 1, et qui sont des points de convergence par rapport à au moins un module $\frac{1}{p}$ (donc aussi par rapport à $\frac{1}{p'}$, si p' est multiple de p). On en déduit que l'ensemble des points de *quasi-convergence*, c'est-à-dire de ceux qui le sont des deux côtés par rapport au module 1 (donc aussi par rapport à $\frac{1}{p}$, quel que soit l'entier p), est un ensemble ouvert.

3° Nous allons maintenant étudier $X(t)$ dans un intervalle fermé, intérieur à un des intervalles ouverts dont tous les points sont des points de quasi-convergence (1). Dans ce cas, la fonction aléatoire $X(t)$ peut être déduite d'une fonction définie sur D , en faisant abstraction de sa partie entière. On le montre en recommençant la théorie des fonctions à accroissements aléatoires indépendants, bien connue dans le cas des fonctions définies sur D . Il n'y a aucun changement essentiel. Nous nous contenterons d'en rappeler les grandes lignes.

(1) Les résultats qui suivent, ne dépendant que des valeurs de $X(t)$ dans e , s'appliquent naturellement aussi si l'extrémité de gauche de e est un point de divergence à gauche, et si son extrémité de droite est un point de divergence à gauche.

En ajoutant d'abord à $X(t)$ une fonction non aléatoire convenablement déterminée, on peut réduire le cas général à celui où il y a en tout point convergence presque sûre à gauche et à droite, et pas seulement quasi-convergence. Il peut ensuite exister des points τ (*points de discontinuité fixes*) pour lesquels la différence $\Delta X(\tau) = X(\tau + 0) - X(\tau - 0)$ n'est pas presque sûrement nulle; la somme de ces discontinuités $\Delta X(\tau)$, pour l'intervalle e , et par suite pour n'importe quel intervalle e' intérieur à e , est quasi-convergente. En retranchant la somme (ou quasi-somme) de ces discontinuités, on effectue une seconde réduction, après laquelle $X(t)$ dépend d'une loi qui varie d'une manière continue avec t .

De plus, nous étudierons $X(t)$ de gauche à droite, à partir d'une valeur t_0 pour laquelle nous supposerons $X(t_0) = 0$ (¹) Pour tout $t > t_0$, $X(t)$ est alors une somme de termes aléatoires indépendants les uns des autres, dépendant de lois très voisines de l'unité. La loi dont dépend $X(t)$ est alors appelée *loi infiniment divisible sur Γ* .

Il peut y avoir encore des *discontinuités mobiles*. On peut naturellement attribuer à chaque saut une hauteur $u = \Delta X(\tau)$ comprise entre 0 et 1. Le nombre probable de tous les sauts peut être infini; mais pour chaque intervalle (u_1, u_2) intérieur à l'intervalle $(0, 1)$, et pour chaque intervalle e' intérieur à e , le nombre probable des sauts pour lesquels τ appartient à e' et u à l'intervalle (u_1, u_2) est fini, et tend vers zéro avec la longueur de e (u_1 et u_2 étant fixes). On en déduit que pour chaque intervalle fixe e' , ce nombre probable étant de la forme $N(u_2) - N(u_1)$, le nombre effectif de ces sauts est une variable aléatoire dépendant de la loi de Poisson, et leur somme est une variable aléatoire dont la nature est bien définie par la donnée de la fonction de répartition des sauts $N(u)$. Quand u_1 et $1 - u_2$ tendent vers zéro, la somme considérée devient une série à termes aléatoires indépendants les uns des autres, et qui doit être quasi-convergente. Les conditions à imposer à la fonction $N(u)$ pour qu'il en soit ainsi sont les mêmes que dans la théorie des lois indéfiniment divisibles sur D , pour les

(¹) Sans cette hypothèse, les résultats qui suivent s'appliqueraient, non à $X(t)$, mais à $X(t) - X(t_0)$.

valeurs très petites de u . Si elles sont vérifiées pour e , elles le sont aussi pour n'importe quel intervalle e' intérieur à e . On définit ainsi une variable aléatoire $X_1(t)$ qui est la quasi-somme des sauts effectués en des points de e d'abscisses inférieures à t ; elle est définie à une fonction non aléatoire près, que l'on peut ajouter; nous la supposons bien entendu définie de manière que $X_1(t)$ soit nul pour $t = t_0$ et dépende d'une loi qui varie d'une manière continue avec t .

Ce terme $X_1(t)$ peut être considéré comme somme ou quasi-somme de variables aléatoires indépendantes les unes des autres, et dont chacune est la quasi-somme des sauts ayant une même valeur u . Naturellement, les valeurs de u qu'il faut considérer peuvent constituer un ensemble fini, ou infini dénombrable, ou non dénombrable. Dans ce dernier cas qui est le cas général, la quasi-somme qui représente $X_1(t)$ se présente sous la forme d'une intégrale de Stieltjes. De toute façon, comme dans l'étude des lois indéfiniment divisibles sur D , il y a un théorème d'unicité : *il n'y a qu'une seule manière de représenter $X_1(t)$ par une somme, une série ou une intégrale (quasi-convergentes) dont chaque élément corresponde à une valeur donnée de u et représente une variable aléatoire dépendant de la loi de Poisson enroulée* (1).

4° Il nous reste à étudier le terme $\xi(t) = X(t) - X_1(t)$.

Il est nul pour $t = t_0$, indépendant des autres termes, est toujours une fonction de t à accroissements aléatoires indépendants, et est de plus une fonction de t presque sûrement continue. Cela veut dire que, ε et ε' étant des nombres positifs arbitrairement petits, on peut diviser e en intervalles partiels e_1, e_2, \dots, e_n , de manière que, sauf dans des cas de probabilité totale inférieure à ε' , les accroissements de $\xi(t)$ dans les intervalles e_i soient tous en modules inférieurs à ε . Naturellement, sur Γ , la partie entière de $\xi(t)$ est indifférente. Mais on peut

(1) Rappelons le principe de la démonstration. Un produit de lois de Poisson peut être obtenu par un processus homogène, c'est-à-dire que $\log b_n$ varie (de zéro à sa valeur finale) proportionnellement à t . Le coefficient qui correspond, dans la somme ou intégrale considérée, à chaque intervalle Δu (fini ou indéfiniment petit, mais ne comprenant pas l'origine), est alors le nombre probable des sauts pour lesquels u appartient à cet intervalle. Cette interprétation exclut toute indétermination.

définir cette fonction sur D de manière que l'accroissement de $\xi(t)$ pour chacun des e_i soit inférieur en module à ε , sauf dans des cas de probabilité inférieure à ε' . Comme ε et ε' sont arbitrairement petits, l'accroissement de $\xi(t)$ (sur e ou sur tout intervalle intérieur à e) est, d'après le second théorème fondamental du calcul des probabilités, la somme d'une constante et d'une variable gaussienne. Alors $\xi(t)$ est, sur D, la somme d'une fonction continue et non aléatoire de t (qu'on peut rattacher à la fonction non aléatoire considérée dans la première réduction; on peut donc la négliger dans ce qui suit) et d'une variable gaussienne, dont la valeur quadratique moyenne est une fonction continue et non décroissante de t . Donc enfin, sur Γ , $\xi(t)$ obéit à la *loi de Laplace-Gauss enroulée sur Γ* , c'est-à-dire celle dont la densité de probabilité est, pour chaque valeur de t , de la forme

$$(39) \quad \varphi(x) = \frac{1}{m\sqrt{2\pi}} \sum_{-\infty}^{+\infty} e^{-\frac{(x+h)^2}{2m^2}},$$

m étant un paramètre, fonction non décroissante de t .

Les coefficients de Fourier, qui se déduisent par la formule (6) de la fonction caractéristique de la loi de Laplace-Gauss, sont donc

$$a_n = e^{-2\pi^2 m^2 n^2}.$$

Lorsque m est petit, la variable aléatoire définie par ces formules est très petite, sauf dans des cas très peu probables, et la loi de Laplace-Gauss enroulée ne diffère guère de la loi correspondante définie sur D. La différence s'accroît quand m croît; la densité maxima, réalisée pour $x = 0$ (1), ne peut que décroître; de même la densité minima ne peut que croître (2); pour m infini, la répartition devient uniforme.

(1) C'est une conséquence évidente de ce que les lois considérées sont des carrés de lois absolument continues, et symétriques; si $\varphi(x)$ est la densité de probabilité, on a $\varphi(y) = \varphi(-y)$, d'où

$$\int_{\Gamma} \varphi(y) \varphi(-y) dy - \int_{\Gamma} \varphi(y) \varphi(x-y) dy = \frac{1}{2} \int_{\Gamma} [\varphi(y) - \varphi(x-y)]^2 dy \geq 0.$$

(2) C'est une conséquence évidente de propriétés des fonctions $\Omega(t)$ et $\omega(t)$ définies paragraphe 6, qui ne peuvent que croître quand on passe d'une loi résultante à la loi composante; or, $\omega'(0)$ et $\Omega'(0)$ sont, pour une loi absolument continue, les valeurs extrêmes de la densité.

Pour chaque valeur de m , la densité $\varphi(x)$ décroît quand x croît de zéro à $\frac{1}{2}$. Cela est évident, si m est petit, d'après la formule (39); et l'on déduit de la formule (8'), qui s'applique ici aux fonctions $\varphi(x)$ correspondant aux valeurs m, m_1 et m_2 de m liées par la relation

$$m^2 = m_1^2 + m_2^2$$

(en prenant m_2 très petit), que la propriété indiquée ne cesse pas d'être exacte quand m croît. D'ailleurs, pour m assez grand, la propriété indiquée se vérifie immédiatement par le développement

$$\varphi(x) = 1 + 2 \sum_1^{\infty} e^{-2\pi^2 m^2 n^2} \cos 2\pi n x$$

de $\varphi(x)$ en série de Fourier.

Cette propriété est d'autre part rendue intuitive par l'interprétation de la fonction (39) que donne la théorie de la chaleur. Cette fonction représente, pour une valeur du temps proportionnelle à m^2 , la température d'une circonférence homogène initialement chauffée au point d'abscisse zéro.

De cette propriété, on déduit $\mu = 0$ (μ étant le nombre défini paragraphe 6, 2^o). L'écart quadratique moyen se calcule alors aisément par la formule

$$(41) \quad \tau^2 = \int_{-\frac{1}{2}}^{+\frac{1}{2}} x^2 d\Phi(x) = \frac{1}{12} + \frac{1}{\pi^2} \sum_1^{\infty} \frac{(-1)^n}{n^2} \alpha_n$$

évidemment valable pour toute loi symétrique et telle que $\mu = 0$ (²); on trouve ainsi

$$(42) \quad \tau^2 = \frac{1}{12} + \frac{1}{\pi^2} \sum_1^{\infty} \frac{(-1)^n}{n^2} e^{-2\pi^2 m^2 n^2},$$

(²) Cela est du moins évident pour une loi absolument continue, et dont la densité $\varphi(x)$ soit représentable par une série de Fourier uniformément convergente. Or on peut modifier n'importe quelle loi donnée de manière à réaliser cette hypothèse, et cela en modifiant arbitrairement peu les deux membres de la formule à démontrer.

expression qui est bien nulle pour $m = 0$ et égale à $\frac{1}{12}$ pour m infini.

5° Revenons à la somme des sauts pour étudier quelques cas particuliers en commençant par la loi de Poisson elle-même, pour laquelle tous les sauts ont une même valeur u . Les valeurs à probabilités positives sont les valeurs hu ($h = 0, 1, 2, \dots$).

Si u est irrationnel, ces valeurs sont toutes distinctes et forment un ensemble partout dense sur Γ (alors tous les points de Γ correspondent à des valeurs possibles de X). On sait que chaque valeur hu a la probabilité $e^{-\theta} \frac{\theta^h}{h!}$, θ étant le nombre probable des sauts.

Si au contraire $u = \frac{q}{p}$ (p et q étant des entiers premiers entre eux), il n'y a que p valeurs distinctes possibles, $0, u, \dots, (p-1)u$, et il vient

$$(43) \quad \Pr \{X = hu\} = \sum_{k=0}^{\infty} \frac{e^{-\theta} \theta^{h+kp}}{(h+kp)!} \quad (h = 0, 1, 2, \dots, p-1).$$

En donnant à u les valeurs $\frac{1}{p}, \frac{2}{p}, \dots, \frac{p-1}{p}$, on obtient $p-1$ lois, qui sont les éléments d'un groupe de lois indéfiniment divisibles pour lesquelles toutes les valeurs possibles sont multiples de $\frac{1}{p}$; une loi de ce groupe est bien définie par les $(p-1)$ paramètres $\theta_1, \theta_2, \dots, \theta_{p-1}$, chacun des θ_h étant le nombre probable des sauts de hauteur $\frac{h}{p}$. La décomposition d'un tel produit en facteurs étant unique, on obtient des lois dépendant de $p-1$ paramètres; si l'on identifie ces lois avec une répartition arbitraire de la probabilité entre les p valeurs de x considérées, il peut y avoir des conditions de possibilité imposées aux $p-1$ probabilités indéterminées dont dépend une telle loi; mais il ne peut s'agir que de conditions d'inégalité. Il existe effectivement de telles conditions; il est évident notamment que, s'il y a deux valeurs à probabilités positives dont la différence soit $\frac{q}{p}$, q étant premier avec p , les p valeurs considérées ont des probabilités positives.

Supposons d'abord $p = 2$. Les probabilités des deux valeurs

possibles sont

$$(44) \quad \begin{cases} e^{-\theta} \left(1 + \frac{\theta^2}{2!} + \dots \right) = e^{-\theta} \operatorname{ch} \theta, \\ e^{-\theta} \left(\theta + \frac{\theta^3}{3!} + \dots \right) = e^{-\theta} \operatorname{sh} \theta. \end{cases}$$

Leur rapport est $\operatorname{th} \theta$. Comme l'addition d'une constante égale à 0 ou $\frac{1}{2}$ permet en tout cas de supposer que zéro est la plus probable des deux valeurs, on voit que, si les deux valeurs possibles sont inégalement probables, et dans ce cas seulement, on a une loi de Poisson. L'égalité des probabilités, quoiqu'elle s'obtienne asymptotiquement, ne définit pas une loi indéfiniment divisible (nous avons déjà remarqué qu'elle ne peut s'obtenir par l'addition d'un nombre fini de termes indépendants que si elle est réalisée par l'un d'eux). Si $p = 3$, les probabilités des trois valeurs possibles, pour la loi de Poisson, sont faciles à calculer; on trouve

$$(45) \quad \frac{1}{3} \left(1 + 2e^{-\frac{\theta}{2}} \cos \frac{\theta\sqrt{3}}{2} \right), \quad \frac{1}{3} \left[1 - e^{-\frac{\theta}{2}} \left(\cos \frac{\theta\sqrt{3}}{2} \pm \sqrt{3} \sin \frac{\theta\sqrt{3}}{2} \right) \right].$$

Les calculs nécessaires pour savoir si une répartition donnée de la probabilité entre ces trois valeurs est un produit de lois de Poisson, par la méthode directe, sont assez compliqués. En introduisant les coefficients de Fourier, ils deviennent très simples, et peuvent être effectués facilement pour une valeur quelconque de p .

On sait que, pour la loi de Poisson correspondant à une valeur donnée de u , on a

$$\log A(t) = \theta(e^{2\pi i u t} - 1),$$

et par suite, d'après (6),

$$(46) \quad \log a_n = \theta(e^{2n\pi i u} - 1).$$

Pour un produit de $p - 1$ lois de Poisson qui correspondent aux valeurs $\frac{1}{p}, \frac{2}{p}, \dots, \frac{p-1}{p}$ de u , on a donc

$$(47) \quad \log a_n = \sum_{h=1}^{p-1} \theta_h \left(e^{\frac{2n-h}{p} \pi i} - 1 \right),$$

et, comme la suite des a_n est périodique, l'identification d'un tel produit avec une répartition donnée de la probabilité entre les p valeurs possibles dépend de la résolution d'un système de $p - 1$ équations linéaires (obtenues en donnant à n les valeurs $1, 2, \dots, p - 1$), et le déterminant de ces équations n'est pas nul. Si en effet il était nul, le système homogène aurait au moins une solution non identiquement nulle $\theta'_1, \theta'_2, \dots, \theta'_{p-1}$. Pour λ assez petit, si $\theta_1, \theta_2, \dots, \theta_p$ sont positifs, tous les $\theta_h + \lambda \theta'_h$ seraient positifs, et le produit de lois de Poisson défini par ces paramètres serait indépendant de λ , ce qui est en contradiction avec le théorème d'unicité énoncé à la fin du 3° du présent paragraphe.

Si alors une répartition de la probabilité entre les p valeurs multiples de $\frac{1}{p}$ est donnée, pour savoir si c'est un produit de lois de Poisson, il suffit de calculer ses coefficients de Fourier, a_1, a_2, \dots, a_{p-1} , et d'en déduire les θ_h comme il vient d'être dit; la condition nécessaire et suffisante est que les valeurs obtenues des θ_h (nécessairement réelles, puisque a_k et a_{-k} sont imaginaires conjugués) soient positives. La condition nécessaire et suffisante pour que la loi donnée soit indéfiniment divisible est alors que cette condition soit réalisée, soit pour cette loi, soit pour une des $p - 1$ autres lois que l'on en déduit par l'addition d'une constante multiple de $\frac{1}{p}$ (1).

6° Terminons ce travail par l'étude d'un autre cas particulier, celui des lois obtenues par l'enroulement sur Γ des lois définies sur D et que nous avons jusqu'ici appelées lois stables, mais que l'emploi de ce mot fait plus haut nous oblige à appeler ici *lois de types stables*. On les obtient, en excluant seulement la loi de Laplace-Gauss, en supposant que, pour chaque signe de u , le nombre probable des sauts dont la hauteur prise en valeur absolue dépasse u est de la forme $cu^{-\gamma}$, avec $\gamma < 2$ (restriction imposée par la condition de quasi-convergence de la somme des sauts, pour

(1) Il serait intéressant d'avoir une règle simple pour déterminer cette constante, si elle existe. Pour $p = 2$, il suffit de ramener à zéro la valeur la plus probable; pour p assez grand, cette règle ne s'applique évidemment plus; on voit aisément qu'elle est déjà en défaut pour $p = 3$.

les petites valeurs de u). Par l'enroulement sur Γ , on obtient une loi pour laquelle, à chaque intervalle élémentaire du , correspondent des sauts dont le nombre probable est

$$(48) \quad \sum_{-\infty}^{+\infty} \frac{[c + c' \operatorname{sgn}(u+h)]}{(u+h)^{\gamma+1}} du \quad (0 \leq c' \leq c).$$

La loi de Cauchy est la loi symétrique pour laquelle $\gamma = 1$. Pour la loi de Cauchy enroulée, l'expression précédente devient

$$(49) \quad c du \sum_{-\infty}^{+\infty} \frac{1}{(u+h)^2} = \frac{c \pi^2 du}{\sin^2 \pi u},$$

de sorte que, de 0 à 1, on peut prendre $-\pi c \cot \pi u$ comme fonction de répartition du nombre probable des sauts en fonction de leur hauteur. C'est (comme les autres lois stables enroulées) une loi absolument continue. La densité de probabilité est, d'après (2'),

$$(50) \quad \varphi(x) = \frac{1}{\pi c} \sum_{-\infty}^{+\infty} \frac{1}{1 + \frac{(x+h)^2}{c^2}}.$$

La fonction caractéristique de la loi de Cauchy étant d'ailleurs $e^{-2\pi c|x|}$, les coefficients de Fourier de $\varphi(x)$, d'après la formule (6), sont $e^{-2\pi c|n|}$, et l'on a

$$(51) \quad \varphi(x) = 1 + 2 \sum_1^{\infty} e^{-2\pi n c} \cos 2\pi n x,$$

d'où l'on déduit pour le carré de l'écart quadratique moyen la valeur

$$(52) \quad \sigma^2 = \frac{1}{12} + \frac{1}{\pi^2} \sum_1^{\infty} \frac{(-1)^n}{n^2} e^{-2\pi n c}.$$

Le 1^{er} août 1938.