

MÉMORIAL DES SCIENCES MATHÉMATIQUES

PAUL LÉVY

Le mouvement brownien

Mémoires des sciences mathématiques, fascicule 126 (1954)

http://www.numdam.org/item?id=MSM_1954__126__1_0

© Gauthier-Villars, 1954, tous droits réservés.

L'accès aux archives de la collection « Mémoires des sciences mathématiques » 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/>

MÉMORIAL DES SCIENCES MATHÉMATIQUES

PUBLIÉ SOUS LE PATRONAGE DE

L'ACADÉMIE DES SCIENCES DE PARIS,
DES ACADÉMIES DE BELGRADE, BRUXELLES, BUCAREST, COÏMBRE, CRACOVIE, KIEW,
MADRID, PRAGUE, ROME, STOCKHOLM (FONDATION MITTAG-LEFFLER),
DE LA SOCIÉTÉ MATHÉMATIQUE DE FRANCE, AVEC LA COLLABORATION DE NOMBREUX SAVANTS.

DIRECTEUR

Henri VILLAT

Membre de l'Institut,
Professeur à la Sorbonne,
Directeur du « Journal de Mathématiques pures et appliquées ».

FASCICULE CXXVI

LE MOUVEMENT BROWNIEN

Par M. Paul LÉVY



UNIVERSITÉ GRENOBLE 1
CNRS
INSTITUT FOURIER
Laboratoire de Mathématiques

PARIS

GAUTHIER-VILLARS, ÉDITEUR-IMPRIMEUR-LIBRAIRE

Quai des Grands-Augustins, 55

—
1954

Copyright by Gauthier-Villars, 1954.

**Tous droits de traduction, de reproduction et d'adaptation réservés
pour tous pays.**

LE MOUVEMENT BROWNIEN

Par M. Paul LÉVY.

AVERTISSEMENT.

Le présent fascicule a été annoncé au moment de la réédition du fascicule 5 de cette collection (P. Lévy [1]. Entre 1925 et 1951, les travaux relatifs au mouvement brownien étaient devenus si nombreux qu'il n'était plus possible en 1951 de se contenter, comme dans la première édition, d'un bref résumé des deux premiers Mémoires de N. Wiener [1] et [2]. Nous avons alors décidé de faire de cette question l'objet d'un nouveau fascicule.

Par la force des choses, nous avons dû répéter des questions exposées dans deux Ouvrages antérieurs d'une autre collection (P. Lévy [2] et [3]). Mais les dimensions de ce fascicule nous obligent à nous contenter d'un résumé de questions développées dans ces Ouvrages; beaucoup de théorèmes sont énoncés sans démonstration, ou avec une brève esquisse de la méthode à suivre. D'autre part, de nombreux résultats nouveaux ont été obtenus depuis 1948. Pour ces deux raisons, nous pensons que ce fascicule ne sera pas inutile.

Nous avons utilisé le langage du calcul des probabilités. Quelques indications sont données en différents endroits, surtout au début, pour aider le lecteur qui ne serait pas habitué à ce langage. Nous avons supposé connus quelques théorèmes classiques de calcul des probabilités. Nous ne pouvions pas, dans les limites de ce fascicule, donner plus de détails.

CHAPITRE I.

LE MOUVEMENT BROWNIEN LINÉAIRE.

DÉFINITION, CONTINUITÉ, REPRÉSENTATION PAR UNE SÉRIE DE FOURIER.

1. **Définitions.** — Une fonction aléatoire $X(t)$ de la variable réelle t est dite *additive* si ses accroissements successifs sont des variables aléatoires indépendantes. Elle est définie à une constante près si l'on connaît la loi de ces accroissements ⁽¹⁾. Cette loi ne peut d'ailleurs pas être quelconque, puisque, si l'on connaît plusieurs accroissements successifs, l'accroissement total en résulte. Il y a donc une condition de compatibilité, d'un caractère relativement simple. Elle est suffisante pour l'existence de la fonction aléatoire $X(t)$; nous le démontrerons seulement dans le cas du mouvement brownien ⁽²⁾.

La fonction $X(t)$ du mouvement brownien linéaire, qui sera seule considérée maintenant, est la fonction aléatoire additive définie à une constante près par la formule

$$(1) \quad X(t_1) - X(t_0) = \xi \sqrt{t_1 - t_0} \quad (t_1 \geq t_0),$$

où ξ est une *variable gaussienne réduite*, c'est-à-dire que

$$(2) \quad \Pr \{ \xi < x \} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du \quad (3).$$

⁽¹⁾ Précisons bien que *définir une variable aléatoire* (ou une fonction aléatoire), ce n'est pas choisir une de ses déterminations possibles. C'est définir la répartition de la probabilité dans l'ensemble de ses déterminations.

On sait qu'on peut toujours considérer une fonction aléatoire $X(t)$ comme une fonction certaine $f(t, \omega)$ de t et d'une variable auxiliaire ω qui résume tout ce qu'on attend du hasard. L'usage se répand de plus en plus, surtout dans les travaux rédigés en anglais, de faire figurer explicitement cette variable dans les formules. Nous n'avons pas cru devoir nous y conformer; cela alourdit les notations. D'ailleurs ω est un élément étranger, comme le paramètre qu'il faut choisir pour ramener une intégrale curviligne à une intégrale ordinaire, mais qu'on peut choisir d'une infinité de manières; un système de notations intrinsèques devrait l'éliminer.

On reconnaîtra en principe les grandeurs aléatoires à ce qu'elles sont désignées par des majuscules latines (X, Y, Z, M, U), ou, dans un cas qui va être défini avec précision, par les lettres grecques ξ, η (et ζ s'il s'agit de variables complexes).

⁽²⁾ Dans le cas général, cette existence peut se déduire du principe d'augmentation de la dispersion, qui devient, lorsqu'il s'agit d'interpolation, un principe de réduction de la dispersion.

⁽³⁾ Rappelons qu'une variable aléatoire X est *réduite* si $E(X) = 0$, $E(X^2) = 1$;

Les lettres ξ et η , avec ou sans accents ou indices, désigneront toujours de telles variables ; quand plusieurs de ces lettres figureront *dans un même membre* d'une formule, elles seront toujours supposées indépendantes.

Il résulte des propriétés les plus connues de la loi de Gauss que la condition de compatibilité triviale mentionnée à propos de la définition des fonctions additives est vérifiée.

Pour achever la détermination de $X(t)$, et pouvoir déterminer les probabilités absolues relatives à cette fonction, il reste à définir la constante additive. Sauf avis contraire, nous la définirons par la condition $X(0) = 0$, et ne considérerons que les valeurs positives de t ⁽¹⁾. Dans ces conditions, la *covariance* de $X(t)$ est

$$(3) \quad E[X(t)X(u)] = t_0,$$

de sorte que le *coefficient de corrélation* de $X(t)$ et de $X(u)$ est $\sqrt{\frac{t_0}{t_1}}$, t_0 et t_1 désignant respectivement le plus petit et le plus grand des nombres t et u .

Ce coefficient étant symétrique en t et u , et un système de variables gaussiennes réduites étant bien défini par la donnée des coefficients de corrélation, on voit que : *la définition de la fonction réduite $\frac{X(t)}{\sqrt{t}}$ est invariante par le changement de t en $\frac{1}{t}$.*

2. Historique. — Il y a plus d'un siècle que Brown a observé que des particules légères plongées dans un liquide avaient un mouvement qui, en projection sur une droite et au degré de précision des ses observations, était représentable par la forme $x = X(t)$, $X(t)$ étant la fonction définie au n° 1. Il ne peut naturellement s'agir que d'une approximation, acceptable seulement si l'intervalle de temps τ qui sépare deux observations est assez grand. En effet d'après la formule (1), la vitesse moyenne pendant cet intervalle de temps est $\frac{[X]}{\sqrt{\tau}}$, tandis que la vitesse réelle d'une particule a une borne

$E(X)$ désigne sa valeur probable. Dans le cas général, $\sigma = \sigma(X)$ est le nombre non négatif défini par

$$\sigma^2 = E[(X - \mu)^2] = E(X^2) - \mu^2 \quad [\mu = E(X)].$$

⁽¹⁾ Cela n'implique aucune restriction essentielle, $X(-t)$ étant une autre détermination de la même fonction aléatoire.

supérieure finie. La théorie mathématique est donc une schématisation qui étend à l'infiniment petit ce qui physiquement n'est vrai qu'à une échelle finie.

Le premier essai de théorie mathématique est dû à L. Bachelier [1] et [2]. Il n'a guère attiré l'attention, tant à cause d'une erreur étrange dans la définition initiale, que parce que cet auteur n'a pas vu que ses idées ne conduiraient en tout cas qu'à une définition prédictive, qu'il faut compléter par une définition constructive; peut-être aussi parce que ce qu'il y avait d'intéressant était noyé dans des considérations d'un intérêt douteux. Bachelier n'en est pas moins le premier à avoir découvert certaines propriétés intéressantes de la fonction $X(t)$, notamment sa relation avec l'équation de la chaleur, et la loi dont dépend le maximum de $X(t)$ dans un intervalle donné. Ses premiers travaux sont antérieurs à ceux d'Einstein.

La définition correcte de $X(t)$, impliquant une démonstration de l'existence de cette fonction, est due à N. Wiener [1]. Nous en avons donné depuis une démonstration plus simple. Avant de l'exposer, nous allons établir une formule d'interpolation sur laquelle elle repose, et une autre formule qui nous servira dans la suite.

3. Formules d'interpolation. — Nous allons nous placer dans l'hypothèse

$$X(t_0) = x_0, \quad X(t_1) = x_1 \quad (t_0 < t_1),$$

et étudier dans ces conditions la loi de probabilité conditionnelle dont dépend $X(t)$ dans l'intervalle (t_0, t_1) .

1° *On a, dans cet intervalle*

$$(4) \quad X(t) = \mu_t + \sigma_t \eta,$$

η étant une variable gaussienne réduite, et la valeur probable μ_t et l'écart-type σ_t étant définis par les formules

$$(5) \quad \mu_t = \frac{(t_1 - t)x_0 + (t - t_0)x_1}{t_1 - t_0}, \quad \sigma_t = \sqrt{\frac{(t - t_0)(t_1 - t)}{t_1 - t_0}}.$$

On remarque que la valeur probable μ_t résulte d'une interpolation linéaire entre les valeurs connues x_0 et x_1 .

Démonstration. — On sait qu'un système de variables aléatoires gaussiennes est bien défini par ses moments des deux premiers ordres.

Il suffit donc de vérifier qu'en partant de la valeur connue $X(t_0) = x_0$, on obtient les mêmes moments pour $X(t)$ et $X(t_1)$, d'une part en déterminant successivement $X(t)$ et $X(t_1)$ par les formules

$$(6) \quad X(t) - X(t_0) = \xi_0 \sqrt{t - t_0}, \quad X(t_1) - X(t) = \xi_1 \sqrt{t_1 - t},$$

déduites de (1), d'autre part en déterminant d'abord $X(t_1) = x_1$ par la formule (1), puis $X(t)$ par la formule (4). La vérification est immédiate; on peut d'ailleurs simplifier le calcul en supposant $t_0 = x_0 = 0$, ce qui n'est pas restrictif.

On peut aussi utiliser le fait bien connu que la loi dont dépend le point d'un plan dont les coordonnées sont ξ_0 et ξ_1 est invariante dans une rotation ⁽¹⁾. Par suite

$$\xi = \xi_0 \cos \varphi + \xi_1 \sin \varphi, \quad \eta = \xi_0 \sin \varphi - \xi_1 \cos \varphi$$

sont aussi des variables laplaciennes réduites et indépendantes. Si nous posons

$$\cos \varphi = \sqrt{\frac{t - t_0}{t_1 - t_0}}, \quad \sin \varphi = \sqrt{\frac{t_1 - t}{t_1 - t_0}},$$

on déduit des formules (6) que

$$(7) \quad X(t_1) - X(t_0) = (\xi_0 \cos \varphi + \xi_1 \sin \varphi) \sqrt{t_1 - t_0} = \xi \sqrt{t_1 - t_0}.$$

C'est bien la formule (1). On a ensuite

$$\xi_0 = \xi \cos \varphi + \eta \sin \varphi.$$

et, par suite,

$$X(t) - X(t_0) = [X(t_1) - X(t_0)] \cos^2 \varphi + \eta \sin \varphi \sqrt{t - t_0},$$

ce qui est bien identique à la formule (4).

2° Supposant toujours x_0 et x_1 connus, nous allons calculer la *covariance*

$$(8) \quad \Gamma(t, u) = E \{ [X(t) - \mu_t] [X(u) - \mu_u] \}.$$

Comme elle est symétrique en t et u , nous pouvons supposer

$$t_0 \leq t \leq u \leq t_1.$$

(1) La densité de probabilité est en effet $\frac{1}{2\pi} e^{-\frac{r^2}{2}}$ ($r^2 = x^2 + y^2$).

Si alors $X(t)$ est connu, la loi de probabilité conditionnelle de $X(u)$ s'obtient en appliquant dans l'intervalle (t, u) la formule d'interpolation (4), de sorte que la valeur probable conditionnelle de $X(u) - \mu_u$ est

$$\frac{t_1 - u}{t_1 - t} [X(t) - \mu_t].$$

Par suite

$$\Gamma(t, u) = \frac{t_1 - u}{t_1 - t} E\{[X(t) - \mu_t]^2\} = \frac{t_1 - u}{t_1 - t} \sigma_t^2,$$

c'est-à-dire enfin

$$(9) \quad \Gamma(t, u) = \frac{(t - t_0)(t_1 - u)}{t_1 - t_0} \quad (t_0 \leq t \leq u \leq t_1).$$

4. Remarques. — 1° La fonction $\Gamma(t, u)$ n'est autre que la fonction de Green relative à l'équation

$$\frac{d^2 y}{dt^2} = f(t)$$

et à l'intervalle (t_0, t_1) , c'est-à-dire que la solution de cette équation qui s'annule pour t_0 et t_1 est définie dans cet intervalle par la formule

$$(10) \quad y(t) = - \int_{t_0}^{t_1} \Gamma(t, u) f(u) du.$$

2° Le coefficient de corrélation de $X(t)$ et $X(u)$ est

$$(11) \quad r(t, u) = \frac{\Gamma(t, u)}{\sigma_t \sigma_u} = \sqrt{\frac{(t - t_0)(t_1 - u)}{(u - t_0)(t_1 - t)}}.$$

Son carré est un des rapports anharmoniques des quatre nombres t_0, t_1, t, u . Il est donc invariant pour toute transformation homographique qui laisse t et u intérieurs à l'intervalle (t_0, t_1) . Cette invariance subsiste à la limite pour t_1 infini; la formule (9) se réduit dans ce cas à une extension triviale de la formule (3).

La fonction gaussienne réduite $\frac{X(t) - \mu_t}{\sigma_t}$ étant bien définie par le coefficient de corrélation $r(t, u)$ est invariante par les transformations considérées. En termes précis, si une telle transformation transforme t_0, t_1, t en t'_0, t'_1, t' , et si l'on pose

$$\xi(t) = \frac{X(t) - \mu_t}{\sigma_t} = \eta(t'),$$

la fonction $\eta(t')$ peut être définie directement dans l'intervalle (t'_0, t'_1) comme $\xi(t)$ l'est dans (t_0, t_1) (cf. P. Lévy [7]).

3° Du fait qu'une fonction aléatoire gaussienne (non réduite) est bien définie par ses moments des deux premiers ordres, donc, si sa valeur probable est nulle, par sa covariance, et de la formule (9) où n'interviennent pas x_0 et x_1 , résulte que la définition de $X(t) - \mu_t$ est indépendante de $X(t_0)$ et $X(t_1)$.

Cela est même vrai en dehors de l'intervalle (t_0, t_1) , puisque dans ce cas, si pour fixer les idées on n'a sur $X(t)$ aucun autre renseignement que les valeurs x_0 et x_1 de $X(t_0)$ et $X(t_1)$, on a $\mu_t = x_0$ pour $t < t_0$ et $\mu_t = x_1$ pour $t > t_1$.

4° En remplaçant t_0 et t_1 par t et $t + dt$, on déduit des formules (4) et (5) l'équation différentielle stochastique

$$(12) \quad \delta X(t) = \frac{dt}{t_1 - t} [x_1 - X(t)] + \eta \sqrt{dt} \quad (t < t_1, dt > 0 \quad (1)).$$

On remarque que, pour dt très petit, le second terme est prépondérant. Donc, en première approximation, l'allure locale de la courbe $x = X(t)$ n'est pas modifiée par la connaissance de $X(t_1)$. Mais le premier terme n'étant pas aléatoire finit par l'emporter sur les oscillations aléatoires dues au second terme, et (nous en sommes sûrs *a priori*) $X(t)$ tend vers x_1 quand t tend vers t_1 .

L'équation (4), pour des accroissements finis, peut s'écrire

$$(4') \quad X(t) - x_0 = \frac{t - t_0}{t_1 - t_0} (x_1 - x_0) + \eta \sqrt{(t - t_0) \frac{(t_1 - t)}{t_1 - t_0}},$$

et, compte tenu de (9), montre que, pour t_1 infini, pourvu que $x_1 = o(t_1)$, l'influence de la condition $X(t_1) = x_1$ tend vers zéro. A la limite, dans tout intervalle fini, la définition de $X(t)$ est indépendante de cette condition. Nous verrons d'ailleurs que x_1 est presque sûrement $o(t_1)$, de sorte qu'il y a convergence de la définition conditionnelle de $X(t)$ vers sa définition inconditionnelle (2).

(1) La variation $\delta X(t)$ n'est naturellement qu'une approximation de l'accroissement fini : les deux premiers moments de l'erreur sont $o(dt)$. Dans ces conditions, compte tenu de ce qu'il s'agit de variables gaussiennes, cette équation, complétée par une donnée initiale, suffit à définir $X(t)$ [au sens de la note (1) de la page 2].

(2) Il s'agit, bien entendu, de convergence au sens de Bernoulli (ou en réparation), notion qui n'implique aucune hypothèse sur la corrélation entre les variables

5. **Définition constructive de $X(t)$.** — Partons de $X(0) = 0$, et déterminons successivement $X(1)$, $X(2)$, ..., par des applications successives de la formule (1). Revenant ensuite dans chacun des intervalles $(h, h+1)$, nous pouvons y choisir une suite partout dense de nombres $t_n (n=1, 2, \dots)$, et déterminer les $X(t_n)$ par des applications successives de la formule (4). Il s'agit de montrer que les valeurs ainsi obtenues définissent presque sûrement une fonction continue.

Faisons d'abord une remarque qui nous servira au chapitre III : $X(t)$ apparaît, par ce mode de détermination, comme une fonction linéaire de variables gaussiennes indépendantes les unes des autres.

Revenant au problème actuel, plaçons-nous pour fixer les idées dans l'intervalle $(0, 1)$, et prenons pour suite des t_n la suite

$$(13) \quad \frac{1}{2}, \quad \frac{1}{4}, \quad \frac{3}{4}, \quad \frac{1}{8}, \quad \frac{3}{8}, \quad \dots$$

Désignons par $F_p(t)$ la fonction définie dans $(0, 1)$, ayant pour les valeurs $\frac{h}{2^p}$ de t les valeurs obtenues comme il est dit ci-dessus [ce qui revient au même que si on les déterminait dans l'ordre naturel par la formule (1)], et variant linéairement dans chacun des intervalles compris entre ces valeurs. Le passage de $F_p(t)$ à $F_{p+1}(t)$ dépend de 2^p nouveaux choix régis par la formule (4), et pour lesquels σ a la valeur $\sigma_p = \frac{1}{2\sqrt{2^p}}$. Le maximum M_p de $|F_{p+1}(t) - F_p(t)|$ est donc le plus grand de 2^p nombres de la forme $\sigma_p |\eta|$. Par suite

$$(14) \quad \Pr(M_p > \sigma_p \gamma) < 2^p \Pr(|\eta| > \gamma) < \sqrt{\frac{2}{\pi}} \frac{2^p}{\gamma} e^{-\frac{\gamma^2}{2}} \quad (\gamma > 0).$$

Pour $\gamma = \gamma_p = c\sqrt{2^p}$ et $c^2 > \log 2$, cette expression est le terme général d'une série convergente. Donc, compte tenu des lemmes de É. Borel, on a presque sûrement à partir d'une certaine valeur aléatoire de p

$$(15) \quad M_p \leq \sigma_p \gamma_p = c \sqrt{\frac{p}{2^{p+1}}},$$

ou fonctions aléatoires considérées et leur limite. Au contraire, les notions de convergence *en probabilité*, *en moyenne quadratique*, ou *presque sûre*, impliquent une telle hypothèse. Rappelons qu'une loi de probabilité est *voisine* d'une autre si l'on peut passer de l'une à l'autre par une correction sur la variable, très petite sauf dans des cas de probabilité elle-même très petite.

et la série $\Sigma[F_{p+1}(t) - F_p(t)]$ est presque sûrement absolument et uniformément convergente dans $(0, 1)$. Les fonctions continues $F_p(t)$ ont donc presque sûrement une limite continue $X(t)$.

On démontre aisément que cette fonction a bien les propriétés voulues. En supposant d'abord que t_0 et t_1 appartiennent à la suite (13), l'équation (1) est bien vérifiée; nous avons précisément déterminé la formule d'interpolation de manière qu'il en soit ainsi. Le cas où t_0 et t_1 sont quelconques résulte de ce qu'ils sont en tout cas limites de termes de la suite (13).

Au lieu de la suite (13) nous aurions pu prendre n'importe quelle suite $\{t_n\}$ ($n = 1, 2, \dots$) partout dense dans $(0, 1)$. La démonstration de la convergence presque sûre de la suite des fonctions qui remplaceraient $F_n(t)$ peut se faire sans difficulté. Mais il suffit de remarquer *a posteriori* que, l'existence presque sûre de $X(t)$ étant maintenant établie, on retrouve bien les probabilités des différentes déterminations possibles en appliquant la formule (4) aux déterminations successives des $X(t_n)$.

6. La condition de Lipschitz faible. — Il ne saurait être question d'une condition de Lipschitz proprement dite, puisque, d'après la formule (1), quelque petit que soit $t_1 - t_0$, les valeurs possibles de $X(t_1)$ varient de $-\infty$ à $+\infty$. Nous dirons qu'une fonction aléatoire $X(t)$ vérifie la *condition de Lipschitz faible* relative à un intervalle (T, T') et à une fonction $\varphi(h)$ s'il existe presque sûrement un nombre $\eta > 0$ tel que, pour $0 \leq h \leq \eta$, $T \leq t \leq T' - h$, on ait

$$(16) \quad |X(t+h) - X(t)| \leq \varphi(h).$$

La fonction $\varphi(h)$ sera toujours supposée définie au moins dans un petit intervalle $(0, h')$, nulle pour $h = 0$, continue et croissante.

Une telle condition n'est évidemment jamais vérifiée sur tout l'axe des t . Les inégalités

$$|X[(n+1)h] - X(nh)| > \varphi(h) \quad (n = 0, 1, 2, \dots)$$

sont en effet des événements indépendants, ayant une même probabilité positive, et, quelque petit que soit $\varphi(h)$, il est presque sûr qu'une infinité de ces inégalités sont réalisées.

Il ne peut donc s'agir que d'un intervalle fini. Le résultat fonda-

mental est le suivant : *dans tout intervalle fini* (T, T') , *on a presque sûrement*

$$(17) \quad \limsup_{h>0} \max_{T \leq t \leq T'-h} \frac{|X(t+h) - X(t)|}{\sqrt{2h \log \frac{1}{h}}} = 1,$$

formule qu'on peut d'ailleurs appliquer séparément aux accroissements positifs et aux accroissements négatifs. Par suite, pour

$$(18) \quad \varphi(h) = c \sqrt{2h \log \frac{1}{h}},$$

la condition de Lipschitz faible est presque sûrement vérifiée si $c > 1$ et presque sûrement en défaut si $c < 1$.

La démonstration repose sur l'étude préliminaire du cas où l'on ne considère que les valeurs $\tau_p = 2^{-p}$ de h , et les valeurs $n\tau_p$ de t . En nous plaçant pour fixer les idées dans l'intervalle $(0, 1)$, et désignant par m_p le plus grand des nombres

$$|X(n\tau_p) - X[(n-1)\tau_p]| \quad (n = 1, 2, \dots, 2^p),$$

on a la formule

$$(19) \quad \Pr \{ m_p > y \sqrt{\tau_p} \} < \sqrt{\frac{2}{\pi}} \frac{2^p}{y} e^{-\frac{y^2}{2}}$$

analogue à la formule (14). Si d'ailleurs y varie avec p et que le second membre tende vers zéro pour p infini, les deux membres sont des infiniment petits équivalents. On en déduit que

$$\Pr(A_p) = \Pr \left\{ m_p > c \sqrt{2\tau_p \log \frac{1}{\tau_p}} \right\} \quad \left(\tau_p = \frac{1}{2^p} \right)$$

est le terme général d'une série convergente si $c > 1$ et divergente si $c \leq 1$.

Sans entrer dans le détail des raisonnements, disons que les événements A_p sont presque indépendants, en ce sens que, si $p' - p$ est grand, et si l'on sait que A_p n'est pas réalisé, ce renseignement est presque sans influence sur la probabilité de $A_{p'}$. Si alors $c < 1$, on voit aisément que cette faible corrélation n'empêche pas d'appliquer le lemme de Borel relatif au cas de divergence, c'est-à-dire que l'événement A_p est presque sûrement réalisé pour une infinité de fois. Nous arrivons ainsi à une conclusion définitive, puisqu'elle subsiste

a fortiori si l'on ne se limite pas aux valeurs τ_p et $n\tau_p$ de h et t : il est presque sûr que, si $c < 1$, la condition de Lipschitz faible formée avec la fonction (18) n'est pas vérifiée.

Dans le cas de convergence, on sait que F. P. Cantelli a montré que la conclusion du lemme de É. Borel subsiste indépendamment de toute hypothèse sur l'indépendance ou la presque indépendance des événements considérés. Il est donc presque sûr, si $c > 1$, que l'inégalité

$$m_p \leq c \sqrt{2\tau_p \log \frac{1}{\tau_p}}$$

est constamment vérifiée à partir d'une certaine valeur de p , valeur qui dépend de la fonction $X(t)$.

Il reste à étendre le résultat obtenu au cas des valeurs quelconques de t et h . On y arrive en deux étapes. La première conduit à la conclusion voulue dans le cas où $t = n\tau_p$, $h = \rho\tau_p$, ρ étant entier. On passe de là au cas général en observant que, pour p assez grand, t et h diffèrent arbitrairement peu de multiples de τ_p . Nous renverrons pour les détails à P. Lévy [2], p. 168-172. Le résultat est que : si $c > 1$, la condition de Lipschitz faible formée avec la fonction (18) est presque sûrement vérifiée.

Un résultat tout à fait précis, et que des remarques de K. L. Chung et P. Erdős rendent vraisemblable, est le suivant : si $\varphi(h) = \psi\left(\frac{1}{h}\right)\sqrt{h}$, et si $\psi(u)$ est monotone, la condition de Lipschitz faible relative à $\varphi(h)$ est presque sûrement vérifiée ou non suivant que l'intégrale

$$\int e^{-\frac{\psi^2(u)}{2}} \psi^3(u) du$$

est convergente ou divergente.

Nous dirons, tant dans ce cas que pour d'autres problèmes analogues où l'on arrive à une coupure définie par une condition de convergence, qu'elle répartit les fonctions $\psi(u)$ en une *classe supérieure* et une *classe inférieure* ⁽¹⁾. Ici la classe supérieure, correspondant aux fonctions qui croissent le plus rapidement avec u , correspond au cas de convergence.

(1) Cette terminologie, que l'on m'attribue parfois à tort, vient de l'école de Moscou; j'en ai eu connaissance par une lettre de A. Kolmogoroff, qui l'utilisait pour le problème qui fera l'objet du n° 7, 2°.

Si l'énoncé qui précède est bien exact, un corollaire immédiat est que : *la fonction définie par*

$$\frac{\psi^2(u)}{2} = \log u + \frac{5}{2} \log_2 u + \dots + \log_n u + c \log_{n+1} u$$

(où $\log_1 u = \log u$, $\log_{n+1} u = \log \log_n u$), *appartient à la classe supérieure si $c > 1$ et à la classe inférieure si $c \leq 1$.*

On remarque qu'il n'y a pas de classe intermédiaire. Cela était relativement facile à prévoir en vertu d'un théorème général de A. Kolmogoroff d'après lequel, dans des cas de ce genre, la probabilité cherchée ne peut être que zéro ou un.

7. Propriétés asymptotiques de $X(t)$. — 1° Si $X(0) = 0$, le rapport $\xi(t) = \frac{X(t)}{\sqrt{t}}$ est une variable gaussienne réduite [quelle que soit la valeur de $X(0)$, il le devient asymptotiquement]. Il est alors naturel de se demander si les différentes valeurs possibles pour une telle variable sont réalisées avec des fréquences tendant vers leurs probabilités théoriques. Il ne saurait en être ainsi pour une suite de valeurs de t de la forme $t_n = n\tau$ ($n = 1, 2, \dots$), parce que la corrélation entre les termes voisins est trop forte. Il y a, par exemple, de longues séries de termes positifs et de longue séries de termes négatifs [et ce n'est qu'au cours de ces séries que $|\xi(t)|$ peut atteindre de grandes valeurs], de sorte que la fréquence des termes de chaque signe est parfois très petite et parfois très voisine de l'unité.

Il n'en est plus de même pour la suite des valeurs $t_n = q^n$ ($q > 1$). Alors, en supposant toujours $X(0) = 0$, le coefficient de corrélation entre $\xi_n = \xi(t_n)$ et ξ_{n+1} est $\frac{1}{\sqrt{q}}$; entre ξ_n et ξ_{n+h} il est $q^{-\frac{h}{2}}$, et tend vers zéro pour h infini, et cela d'une manière indépendante de n . La suite des ξ_n est alors une chaîne *stationnaire* ⁽¹⁾ de variables aléatoires presque *indépendantes*, et l'on peut lui appliquer la *loi forte des*

(1) Elle est, en outre, *markovienne*, c'est-à-dire que la loi conditionnelle dont dépend ξ_{n+1} si l'on connaît $\xi_n, \xi_{n-1}, \dots$, ne dépend que de n et ξ_n ; dire alors qu'elle est *stationnaire*, c'est dire qu'elle ne dépend que de ξ_n . Ces définitions relatives aux chaînes s'étendent naturellement aux fonctions $X(t)$, n et $n+1$ étant remplacés par t et $t+\tau$ (τ , constante quelconque), et la suite $\xi_n, \xi_{n-1}, \dots$ par $X(t-u)$ (u variant de zéro à $+\infty$).

grands nombres : il est presque sûr que, quel que soit x , la fréquence des valeurs de n pour lesquelles $\xi_n < x$ tend vers sa probabilité théorique donnée par la formule (2) ⁽¹⁾.

Un corollaire immédiat est que $X(t)$ a presque sûrement indéfiniment des valeurs de deux signes, donc une infinité de racines.

2° Un autre corollaire immédiat est que, presque sûrement, $\xi(t)$ n'est pas borné. Il est alors naturel de se poser la question suivante, analogue à celle dont l'étude a fait l'objet du n° 6 : étant donnée une fonction $\varphi(t) = \psi(t) \sqrt{t}$, quelle est la probabilité pour qu'il existe un nombre (aléatoire) T tel que, pour $t > T$, on ait

$$(20) \quad |X(t)| \leq \varphi(t) = \psi(t) \sqrt{t}.$$

Ici encore, il y a un théorème préliminaire important ⁽²⁾ : cette probabilité ne peut être que zéro ou un. Il y a donc une coupure unique, par rapport à laquelle il existe une classe supérieure, formée de fonctions $\varphi(t)$, pour lesquelles l'inégalité (20) s'applique presque sûrement pour tout t assez grand, et une classe inférieure, formée de fonctions pour lesquelles elle est presque sûrement en défaut pour des valeurs arbitrairement grandes de t .

On voit aussi aisément que la coupure est indépendante de la condition initiale, et qu'elle s'applique aussi séparément aux valeurs positives et aux valeurs négatives, de sorte que, si $\varphi(t)$ appartient à la classe inférieure, $X(t)$ est presque sûrement parfois $> \varphi(t)$ et parfois $< -\varphi(t)$, et cela pour des valeurs arbitrairement grandes de t .

Le premier renseignement assez précis relatif à cette coupure a été la loi du logarithme itéré, découverte en 1924 par A. Khintchine [1] à propos du problème de Bernoulli (donc dans un cas discontinu), mais qui s'applique à notre problème actuel. Elle s'énonce comme suit : on a presque sûrement

$$(21) \quad \limsup_{t \rightarrow \infty} \frac{|X(t)|}{\sqrt{2t \log \log t}} = 1,$$

(1) C'est pour faciliter le langage que nous introduisons des variables discontinues. Mais on peut poser $t = e^{2u}$, $X(t) = \sqrt{t} \Psi(u)$, de sorte que

$$E[\Psi(u) \Psi(u+h)] = e^{-|h|}.$$

La fonction aléatoire $\Psi(u)$ est gaussienne, markovienne et stationnaire.

(2) Pour le cas des suites telles que $\{X(n\tau)\}$, voir P. Lévy [2], p. 147. L'extension au cas général des fonctions additives est immédiate.

ce qui revient à dire que la fonction

$$(22) \quad \varphi(t) = c \sqrt{2t \log \log t}$$

appartient à la classe supérieure si $c > 1$ et à la classe inférieure si $c < 1$.

Nous n'indiquerons que le principe du raisonnement, qui présente l'intérêt de s'appliquer souvent à des problèmes analogues. *Pour une première étude approchée*, on ne considère que la suite des valeurs $t_n = q^n$ de t (q quelconque > 1), et l'on considère les $X_n = X(t_n)$ comme indépendants. D'après les lemmes de É. Borel, on est ainsi conduit à penser que $\varphi(t)$ appartient à la classe inférieure ou à la classe supérieure suivant la nature de la série

$$(23) \quad \sum \frac{1}{\psi(t_n)} e^{-\frac{\psi^2(t_n)}{2}} \quad (t_n = q^n),$$

c'est-à-dire, en supposant $\psi(t)$ monotone, suivant celle de l'intégrale

$$(24) \quad \int_0^\infty e^{-\frac{\psi^2(t)}{2}} \frac{dt}{t \psi(t)} \quad \left[\psi(t) = \frac{\varphi(t)}{\sqrt{t}} \right].$$

Cette règle, appliquée à la fonction (22), donne $c > 1$ comme condition de convergence, donc comme condition pour que $\varphi(t)$ appartienne à la classe supérieure. C'est bien le résultat annoncé, avec une précision supplémentaire pour le cas $c = 1$. Il reste à justifier cette règle, qui n'est qu'à peu près exacte. Notre méthode (P. Lévy [4]) repose sur la remarque que la nature de la série (23) est indépendante de q . On facilite alors le raisonnement en prenant pour q , dans le cas de convergence, une valeur très voisine de l'unité [ce qui permet de négliger les oscillations de $X(t)$ entre t_n et t_{n+1}], et, dans le cas de divergence une valeur très grande (ce qui permet de négliger la corrélation entre X_n et X_{n+1}). Dans les deux cas, l'erreur ne dépasse pas celle due à une petite variation de c . On peut ainsi conclure dans les deux cas, ce qui suffit à établir la formule (21).

Dans notre Mémoire cité de 1930, nous avons commencé l'étude du cas où $c = 1$ en utilisant des suites de nombres t_n qui croissent un peu plus lentement ou un peu plus rapidement qu'une progression géométrique. Peu après, A. Kolmogoroff a résolu définitivement le problème en montrant que : la fonction $\psi(t)$ étant toujours sup-

posée monotone, la condition nécessaire et suffisante pour que $\psi(t)\sqrt{t}$ appartienne à la classe supérieure est la convergence de l'intégrale

$$(25) \quad \int_{t_0}^{\infty} \psi(t) e^{-\frac{\psi_1(t)}{2}} \frac{dt}{t} \quad (1).$$

Ainsi le raisonnement heuristique indiqué d'abord ne donne qu'un résultat approché. Par exemple, pour les fonctions de la forme

$$\psi(t) = \sqrt{2 \log \log t + c' \log \log \log t},$$

l'intégrale (24) converge si $c' > 1$, et l'intégrale (25) converge seulement si $c' > 3$. Le cas où $1 < c' \leq 3$ est précisément le cas difficile qui échappe à notre méthode de 1930. On voit, par le théorème de A. Kolmogoroff, que dans ce cas $\psi(t)$ appartient à la classe inférieure.

Le savant russe avait déduit son théorème de résultats de I. Petrovsky [1] sur l'équation de la chaleur, dont nous verrons au chapitre II la relation avec le mouvement brownien. Plus récemment, W. Feller [1] et [2], l'a retrouvé par une étude plus minutieuse de la méthode directe qui vient d'être esquissée.

8. Étude locale de $X(t)$. — 1° La définition de $X(t)$ étant invariante aussi bien par le changement de x en $x + a$ que par celui de t en $-t$ ou en $t + h$, il suffit d'étudier, dans l'hypothèse $X(0) = 0$, l'allure de la fonction $X(t)$ pour t positif et très petit. Les résultats s'appliqueront aussi bien à droite qu'à gauche de n'importe quel point de la courbe $x = X(t)$.

Nous avons déjà observé que, pour $X(0) = 0$, la définition de la fonction $\xi(t) = \frac{X(t)}{\sqrt{t}}$ ($t > 0$) est invariante par le changement de t en $\frac{1}{t}$. Tous les résultats obtenus au n° 7 se transforment ainsi en résultats locaux, qu'on peut naturellement aussi démontrer directement. On voit ainsi que, presque sûrement, la racine zéro de $X(t)$

(1) Si $\psi(t)$ croît indéfiniment sans être monotone, et si $\psi_1(t) = \min_{t' \geq t} \psi(t')$ appartient à la classe supérieure, il en est *a fortiori* de même de $\psi(t)$. Mais si $\psi(t')$ appartient à la classe inférieure, la nature de $\psi(t)$ dépend de la fréquence des faibles valeurs de $\psi(t) - \psi_1(t)$.

n'est pas isolée, et que la loi du logarithme itéré s'applique en remplaçant $\log \log t$ par $\log \log \frac{1}{t}$ dans la formule (21). Plus généralement, il y a, pour l'étude de l'inégalité (20), une *classe supérieure* et une *classe inférieure*, caractérisées respectivement par la convergence ou la divergence de l'intégrale (25), non plus à l'infini, mais à l'origine.

2° Ces propriétés étant ainsi presque sûrement vraies pour n'importe quelle valeur donnée de t sont par là même presque sûrement vraies dans n'importe quel intervalle, sauf peut-être sur un ensemble de mesure nulle. Il est d'ailleurs à peu près évident que, pour chacune des propriétés considérées, cet ensemble est, ou bien presque sûrement vide, ou bien presque sûrement partout dense.

Avant d'exposer à ce sujet quelques remarques très simples, signalons un problème qui semble n'être pas encore résolu : *quelle condition doit remplir la fonction positive décroissante $\psi(\tau)$ pour qu'il existe presque sûrement une fonction $h = h(t)$, bien définie et positive en tout point d'un intervalle fini donné, et telle que $0 < \tau < h$ entraîne*

$$|X(t + \tau) - X(t)| \leq \sqrt{\tau} \psi(\tau)?$$

Pour le moment, on ne sait rien à ce sujet en ce qui concerne les fonctions $\psi(t)$ qui sont à la fois $\leq \sqrt{2 \log \frac{1}{t}}$ et $\geq c \sqrt{2 \log \log \frac{1}{t}}$ ($c > 1$). Il résulte seulement du n° 6 que, dans ce cas, si la fonction $h(t)$ existe partout, elle ne peut pas, même dans un intervalle fini, avoir une borne inférieure positive.

3° Un corollaire du 1° est qu'en chaque point donné, aussi bien à sa droite qu'à sa gauche, les nombres dérivés de $X(t)$ sont presque sûrement égaux, l'un à $-\infty$, l'autre à $+\infty$. Mais il y a ce point de vue des points exceptionnels; c'est ce que nous allons préciser.

Considérons pour fixer les idées le nombre dérivé inférieur à droite, et montrons qu'il y a presque sûrement, dans n'importe quel intervalle (T_0, T) , une infinité non dénombrable de points où il est ≥ 0 . C'est une conséquence certaine du fait, évidemment presque sûr, que $X(t)$ n'y est pas monotone. On peut alors choisir t_0 et $t_1 > t_0$ de manière que $X(t_1) > X(t_0)$. Alors, quel que soit a entre $X(t_0)$

et $X(t_1)$, la plus grande racine de $X(t) - a$ qui soit $< t_1$ a la propriété indiquée.

Précisons l'allure de la courbe au voisinage d'un tel point, que nous pouvons supposer ramené à l'origine. Donc, par hypothèse, au moins dans un petit intervalle $(0, T)$, on a $X(t) \geq 0$. Dans ces conditions, comme nous le montrerons au chapitre II (n° 14, 4°), la formule (1) est remplacée par la formule

$$(26) \quad \Pr \{ X(t) > x \} = e^{-\frac{x^2}{2t}}.$$

En raisonnant alors comme pour établir la loi du logarithme itéré, on obtient le résultat suivant : *il est presque sûr que la fonction* $\left(\log \frac{1}{t}\right)^{-\alpha} \sqrt{t}$ *borne inférieurement* $X(t)$, *au moins dans un petit intervalle* $(0, \tau)$, *si* $\alpha > 1$, *mais non si* $\alpha < 1$ ⁽¹⁾.

Le résultat subsiste d'ailleurs si la condition $X(t) \geq 0$ est remplacée par $X(t) \geq mt$ ($m > -\infty$). Par suite : *si l'on sait qu'en un point* t *le nombre dérivé inférieur à droite est* $> -\infty$, *il est presque sûrement* $= +\infty$, *de sorte que la courbe* $x = X(t)$ *a une demi-tangente verticale.*

Comme il y a dans tout intervalle une infinité non dénombrable d'occasions d'appliquer cet énoncé, on peut se demander s'il n'y a pas des points exceptionnels où il y aurait un nombre dérivé fini. Notre impression est qu'il n'y en a presque sûrement pas; mais nous ne l'avons pas démontré. S'il en est bien ainsi, les maxima et minima de $X(t)$ sont aussi ceux de $X(t) - mt$ (quel que soit le nombre fini m). Plus généralement, en étudiant l'intersection de la courbe $x = X(t)$ et des droites $x = mt + a$, on n'aurait pas obtenu

(1) Cf. [5], p. 334-336. Le raisonnement est en réalité plus compliqué que pour la loi du logarithme itéré, et cela pour une raison bien simple : si $\frac{X(t)}{\sqrt{t}}$ est exceptionnellement grand pour $t = t'$, cette circonstance subsiste probablement de t' à qt' (q fixe > 1); il n'en est pas de même si ce rapport est exceptionnellement petit. Quelque petit que soit $q - 1$, on ne doit pas s'attendre à trouver à peu près, dans la suite des nombres $q^n X(q^{-2n})$ les valeurs exceptionnellement petites de ce rapport. Nous avons résolu cette difficulté par une étude préliminaire du minimum de $X(t)$ dans l'intervalle $[q^{-n}, q^{1-n}]$.

Nous avons omis en 1939 d'attirer l'attention sur les corollaires très simples du présent théorème qui seront exposés dans la suite du présent paragraphe.

d'autres points exceptionnels que ceux dont nous venons de prouver l'existence presque sûre.

4° Le fait que $X(t)$ ne soit presque sûrement monotone dans aucun intervalle entraîne l'existence presque sûre de maxima et minima qui forment un ensemble dénombrable et partout dense. D'après le résultat du 3° (qui s'applique naturellement, aux signes près, aux quatre nombres dérivés) *chacun de ces points est presque sûrement un point de rebroussement à tangente verticale*. Comme il s'agit ici d'un ensemble dénombrable, *cela est presque sûrement vrai pour tous les maxima et minima*.

Les nombres dérivés à droite d'un point donné étant indépendants de leur valeur à gauche, les points où les quatre nombres dérivés ont une même valeur ($+\infty$ ou $-\infty$) apparaissent comme exactement aussi probables que les points à tangentes de rebroussement verticales. Il est alors naturel de penser qu'ils forment aussi presque sûrement un ensemble dénombrable et partout dense. Mais ce raisonnement heuristique n'est pas une démonstration et nous ne pouvons rien affirmer.

9. La variation totale presque sûre d'ordre deux. — 1° L'ordre de grandeur probable de l'accroissement $\Delta X(t)$ de $X(t)$ étant $\sqrt{\Delta t}$ (Δt désignant l'accroissement de t), on en déduit que, pour n'importe quel intervalle (t_0, t_1) , la variation totale de $X(t)$ est presque sûrement infinie, et il est naturel de considérer la variation d'ordre deux, liée aux sommes de la forme

$$(27) \quad S_D = \Sigma [\Delta X(t)]^2,$$

liées chacune à une division D de (t_0, t_1) en intervalles Δt ; nous désignerons par $\tau = \tau(D)$ le plus grand des Δt .

On voit aisément que, presque sûrement,

$$(28) \quad \lim_{\tau' \rightarrow 0} \inf_{\tau \leq \tau'} S_D = 0, \quad \lim_{\tau' \rightarrow 0} \sup_{\tau \leq \tau'} S_D = \infty,$$

la première de ces formules étant une conséquence certaine de la continuité de $X(t)$. Mais il y a une valeur moyenne qui mérite de retenir notre attention. On a

$$(29) \quad E(S_D) = \Sigma E \{ [\Delta X(t)]^2 \} = \Sigma \Delta t = t_1 - t_0,$$

$$(30) \quad \sigma^2(S_D) = \Sigma \sigma^2 \{ [\Delta X(t)]^2 \} = 2 \Sigma (\Delta t)^2 \leq 2 \tau (t_1 - t_0),$$

de sorte que : τ tendant vers zéro, il y a convergence en moyenne quadratique de S_D vers sa valeur probable $t_1 - t_0$.

2° Le résultat précédent est indépendant de la manière dont D varie quand τ tend vers zéro. Nous allons maintenant supposer que l'on ait choisi une suite partout dense dans (t_0, t_1) de nombres $t_2, t_3, \dots$, et pris pour $D = D_n$ la division de (t_0, t_1) en n intervalles séparés par les points $t_2, t_3, \dots, t_n$. Dans ces conditions $\tau = \tau_n$ tend vers zéro et l'on démontre aisément (cf. P. Lévy [3], p. 205-209) que : S_D tend presque sûrement vers $t_1 - t_0$.

3° Supposons maintenant les points $t_2, t_3, \dots$, choisis au hasard dans (t_0, t_1) indépendamment les uns des autres, d'après une même loi dont la fonction de répartition soit constamment croissante de t_0 à t_1 . Si dans ces conditions, pour une détermination particulière $f(t)$ de $X(t)$, S_D a une limite presque sûre V , nous dirons que V est la *variation totale presque sûre d'ordre deux* de $f(t)$.

Or τ_n tend presque sûrement vers zéro, ce qui permet d'appliquer le résultat du 2°. Donc, compte tenu de la double intervention du hasard [pour le choix des t_n et pour celui de la fonction $X(t)$], il est presque sûr que S_D tend vers $t_1 - t_0$. L'ordre des choix n'ayant pas d'importance, nous voyons que : $X(t)$ a presque sûrement dans tout intervalle (t_0, t_1) une variation presque sûre d'ordre deux égale à $t_1 - t_0$, et cela sans qu'il y ait d'intervalle exceptionnel, puisque la limite considérée est en tout cas monotone en fonction de t_0 comme de t_1 (cf. P. Lévy [3], p. 208).

4° Une généralisation toute naturelle consiste à introduire, au lieu de la somme (27), la somme analogue

$$(31) \quad \Sigma_D \varphi[|\Delta X(t)|],$$

où $\varphi(x)$ est une fonction continue, croissante, s'annulant avec x ; nous supposons de plus $\varphi(x) = x^2 \psi(x)$, $\psi(x)$ étant monotone. Comme dans le cas où $\varphi(x) = x^2$, on en déduit des φ -variations inférieure, supérieure et moyenne, que nous désignerons par $\underline{V}_\varphi, \bar{V}_\varphi, V_\varphi$, analogues à celles définies par les formules (28) et (29).

Compte tenu des résultats obtenus quand $\varphi(x) = x^2$, l'étude de V_φ est triviale; $\psi(0)$ a une valeur déterminée c , nulle, positive, ou infinie, et $V_\varphi = c(t_1 - t_0)$, que cette expression soit définie comme limite en moyenne ou comme limite presque sûre.

En ce qui concerne $\bar{V}_\varphi$, le problème est lié à la loi du logarithme itéré. D'après cette loi, si $\varphi(x) = \frac{x^2}{2 \log \log \frac{1}{x}}$, et si $c = 1$, l'intervalle

(t_0, t_1) est presque sûrement couvert par des intervalles arbitrairement petits pour lesquels $\varphi[|\Delta X(t)|] > c$; mais si $c > 1$, l'ensemble des points ainsi recouverts est de mesure nulle, et, bien que $\varphi[|\Delta X(t)|]$,

pour $\Delta t = \tau$, puisse atteindre $\frac{\log \frac{1}{\tau}}{\log \log \frac{1}{\tau}}$, cette circonstance tout à fait

exceptionnelle ne suffit pas à changer le résultat, et l'on a presque sûrement $\bar{V}_\varphi = t_1 - t_0$.

Le problème relatif à $\underline{V}_\varphi$ est trivial. Nous savons déjà que cette expression est infinie pour $\varphi(x) = x$; pour n'importe quelle fonction continue, elle est nulle si $\varphi(x) = o(x)$. Cela tient à ce qu'à tout instant, ayant mis en évidence par le choix d'un nombre fini de points de division certaines oscillations de $X(t)$, on peut rendre la somme (31) arbitrairement petite en choisissant les nouveaux points de division de manière à ne faire apparaître aucune nouvelle oscillation. L'étude de $\underline{V}_\varphi$ n'apprend donc rien de nouveau au sujet de $X(t)$.

Une notion assez voisine de la précédente est ce que nous appellerons la *mesure de Hausdorff majorée*, surtout importante dans le cas d'un mouvement dans le plan ou dans l'espace (voir n° 24). Disons seulement ici que cette mesure $\bar{H}_\varphi$ est presque sûrement positive et finie pour $\varphi(x) = x^2 \log \log \frac{1}{x}$, ce qui fait apparaître une sorte de symétrie entre $\bar{V}_\varphi$ et $\bar{H}_\varphi$.

10. Développement de $X(t)$ en série de Fourier. — La fonction $X(t)$, étant presque sûrement continue, est développable dans n'importe quel intervalle fini en une série de Fourier convergente au sens de Cesàro. Il y a d'ailleurs avantage, si l'on suppose $X(0) = 0$ et si l'on se place dans l'intervalle $(0, 2\pi)$, à développer, non $X(t)$, mais la différence

$$(32) \quad X_0(t) = X(t) - \frac{t}{2\pi} X(2\pi) = X(t) - \frac{t\xi}{\sqrt{2\pi}},$$

et, comme elle est indépendante de ξ (n° 4, 3°), le développement de $X_0(t)$ donne immédiatement une expression de $X(t)$.

On simplifie d'ailleurs le calcul en introduisant la *variable gaussienne complexe réduite* $\zeta = \frac{\xi + i\eta}{\sqrt{2}}$, et les fonctions

$$Z(t) = \frac{X(t) + iY(t)}{\sqrt{2}}$$

et $Z_0(t)$ déduites de ζ comme $X(t)$ et $X_0(t)$ le sont de ξ . La plupart des formules des n^{os} 1 à 4 subsistent, en remplaçant ξ et X par ζ et Z , et les moments tels que $E(X^2)$ et $E(XX')$ par $E(|Z|^2)$ et $E(Z\bar{Z}')$ (Z' et $\bar{Z}'$ sont imaginaires conjugués).

La série de Fourier de $Z_0(t)$ dans l'intervalle $(0, 2\pi)$ est alors

$$\sum' A_n (e^{nit} - 1)$$

$\sum'$ désignant une sommation étendue à tous les entiers réels autres que zéro, et les coefficients A_n étant définis par la formule de Fourier

$$A_n = \frac{1}{2\pi} \int_0^{2\pi} Z(t) e^{-nit} dt \quad (n \neq 0).$$

Ce sont des variables gaussiennes complexes, formant un système bien défini par la *covariance*

$$\Gamma_{p,q} = E(A_p \bar{A}_q) = \frac{1}{4\pi^2} \int_0^{2\pi} e^{-pit} \int_0^{2\pi} e^{qiu} \Gamma(t, u) du \quad (p, q \neq 0).$$

D'après la remarque du n^o 4, 1^o, le résultat de l'intégration en u est $\frac{e^{qit} - 1}{q^2}$, de sorte que

$$\Gamma_{p,q} = \frac{1}{4\pi^2 q^2} \int_0^{2\pi} [e^{(q-p)it} - e^{-pit}] dt.$$

Si $p \neq q$, cette covariance est nulle; les A_n sont donc indépendants les uns des autres. Si $p = q = n$, il vient

$$E(|A_n|^2) = \frac{1}{2\pi n^2}, \quad A_n = \frac{\zeta_n}{n\sqrt{2\pi}},$$

et, par suite,

$$(33) \quad Z_0(t) = \sum' \frac{\zeta_n}{n\sqrt{2\pi}} (e^{nit} - 1),$$

$$(34) \quad Z(t) = \frac{1}{\sqrt{2\pi}} \left[\zeta t + \sum' \frac{\zeta_n}{n} (e^{nit} - 1) \right],$$

ζ et tous les ζ_n étant des variables gaussiennes complexes, réduites, toutes indépendantes les unes des autres. Ces séries sont, pour chaque valeur de t , convergentes en moyenne quadratique, donc presque sûrement convergentes.

On peut avoir intérêt à considérer aussi la fonction plus simple

$$(35) \quad Z_1(t) = Z_0(t) - \frac{1}{2\pi} \int_0^{2\pi} Z_0(u) du = \sum' \frac{\zeta_n}{n\sqrt{2\pi}} e^{nit}$$

dont la valeur moyenne dans l'intervalle $(0, 2\pi)$ est nulle.

Il n'y a maintenant aucune difficulté à égaler, dans chacune des formules précédentes, les parties réelles des deux membres. On obtient ainsi la formule

$$(36) \quad X(t) = \frac{\xi't}{\sqrt{2\pi}} + \sum_1^{\infty} \frac{i}{n\sqrt{\pi}} [\xi_n(\cos nt - 1) + \xi'_n \sin nt],$$

et des formules analogues pour $X_0(t)$ (en supprimant le premier terme), et $X_1(t)$ (en supprimant le premier terme et les termes constants). ξ' , les ξ_n et les ξ'_n sont des variables gaussiennes réelles, réduites, indépendantes les unes des autres (nous écrivons ξ' et non ξ parce que cette variable sera dans la suite rapprochée des ξ'_n).

Ces formules, que nous appellerons *formules de Fourier-Wiener*, peuvent servir de définition à $X(t)$ et $Z(t)$. C'est la voie suivie par R. E. A. C. Paley et N. Wiener [1]. La définition initiale de N. Wiener [1] dont nous sommes partis nous paraît plus naturelle. C'est N. Wiener [2] qui a dès 1924 déduit la formule (36) de sa première définition. Notre méthode (P. Lévy [9], p. 184), exposée ici, montre bien plus simplement l'identité des deux définitions.

CHAPITRE II.

ÉTUDE APPROFONDIE DU MOUVEMENT BROWNIEN LINÉAIRE.

11. L'équation de la diffusion. — 1° Il résulte immédiatement de la définition de la fonction $X(t)$ que, si $X(t_0) = u_0$, et $t > t_0$, la loi de probabilité de $X(t)$ est la loi absolument continue dont la densité de probabilité est $\varphi(t - t_0, x - x_0)$, en posant

$$(1) \quad \varphi(t, x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \quad (t > 0).$$

Si alors, à l'instant t_0 , on se donne une loi quelconque, définie par sa fonction de répartition $U_0(x)$, on a, pour $X(t)$, à l'instant $t(t > t_0)$, une loi absolument continue, définie par la densité

$$(2) \quad u(t, x) = \int_{-\infty}^{+\infty} \varphi(t - t_0, x - \xi) dU_0(\xi).$$

On vérifie immédiatement que : $\varphi(t, x)$ et, par suite, $u(t, x)$, sont solutions de l'équation de la chaleur

$$(3) \quad 2 \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

D'une manière plus précise, $u(t, x)$ est la solution du problème de Cauchy. Si $U_0(x)$ est une fonction absolument continue, de dérivée $u_0(x)$, la condition initiale est $u(t_0, x) = u_0(x)$. Dans le cas général, elle s'écrit

$$(4) \quad \lim_{t \downarrow t_0} \int_{-\infty}^x u(t, \xi) d\xi = U_0(x) \quad (t \downarrow t_0).$$

cette équation devant être vérifiée en tout point où $U_0(x)$ est continu. Elle reste d'ailleurs vérifiée en tout point où

$$U_0(x) = \frac{1}{2} [U_0(x-0) + U_0(x+0)].$$

2° La relation entre le mouvement brownien et la théorie de la chaleur a déjà été signalée par L. Bachelier [1]. Elle est un cas particulier d'un phénomène plus général, découvert par l'école russe (1) : si une fonction aléatoire $X(t)$ est solution de l'équation différentielle stochastique

$$(5) \quad dX(t) = A[t, X(t)] dt + B[t, X(t)] \xi \sqrt{dt} \quad (dt > 0),$$

où ξ est toujours une variable gaussienne réduite, indépendante de $X(t)$ et du passé, elle dépend d'une loi absolument continue dont la densité de probabilité dépend de l'équation de la diffusion (ou équation de Fokker-Planck)

$$(6) \quad \frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 (B^2 u)}{\partial x^2} - \frac{\partial (A u)}{\partial x}.$$

(1) A. Kolmogoroff [1] et S. Bernstein [2].

Pour $A = 0$, $B = 1$, $X(t)$ se réduit à la fonction définie au chapitre I, et l'équation (6) à l'équation (3).

3° Revenant au cas particulier du mouvement brownien, nous pouvons préciser par une remarque simple la raison de l'intervention de l'équation (3). On sait que c'est l'équation qui définit la propagation de la chaleur le long d'un fil homogène; u désignant la température, cette équation résulte de ce que le *flux de chaleur* qui traverse la section x pendant le temps dt , compté positivement s'il vient du côté droit, est $c \frac{\partial u}{\partial x} dt$, c désignant une constante. On trouve la même expression, si u a la signification indiquée au 1°, pour la variation de la fonction de répartition $U(t, x)$, primitive de $u(t, x)$; c'est un *flux de probabilité*. Dans les deux cas, on obtient l'équation (3) par dérivation (¹).

On peut rendre cette analogie plus concrète en imaginant un fluide composé de particules qui se déplacent indépendamment les unes des autres, et qui soient assez nombreuses pour qu'à tout instant le nombre de celles qui sont sur un segment donné diffère peu de sa valeur probable. Si pour chaque particule la loi de son mouvement est celle du mouvement brownien, la densité du fluide vérifie l'équation (3).

12. Conditions aux limites. — 1° Nous allons maintenant considérer la probabilité $U(t)$ que

$$(7) \quad f_0(t') \leq X(t') \leq f_1(t') \quad (0 \leq t' \leq t)$$

et la probabilité $u(t, x) dx$ que l'on ait à la fois les inégalités (7) et

$$(8) \quad x \leq X(t) < x + dx.$$

Nous supposons $f_1(t) > f_0(t)$, pour $t \geq t_0$. La seule restriction relative à la continuité de ces fonctions sera la condition de semi-continuité

$$\liminf_{h \rightarrow 0} f_0(t+h) \leq f_0(t) < f_1(t) \leq \liminf_{h \rightarrow 0} f_1(t+h),$$

c'est-à-dire qu'on ne considérera une position comme possible à

(¹) Ce raisonnement, admettant l'existence des dérivées, n'est qu'heuristique. On peut lui donner une forme rigoureuse; mais il est alors plus simple de déduire l'équation (3) de la formule (2).

l'instant t que si elle peut être réalisée pour une trajectoire continue. $X(t_0)$ sera supposé choisi dans l'intervalle $[f_0(t_0), f_1(t_0)]$, suivant une loi quelconque.

La fonction $u(t, x)$ doit être non négative si x appartient à l'intervalle $[f_0(t), f_1(t)]$, et elle y vérifie toujours l'équation (3). La condition (4) subsiste, à cela près que la limite inférieure de l'intervalle d'intégration est $f_0(t)$. Pour achever de définir le problème, il reste à montrer que $u(t, x)$ doit vérifier les *conditions aux limites*

$$(9) \quad u[t, f_0(t)] = u[t, f_1(t)] = 0 \quad (t > t_0),$$

sauf peut-être pour certaines valeurs exceptionnelles de t , qui constituent un *ensemble impropre*.

2° Considérons d'abord le cas où $X(t_0)$ a une valeur donnée $x_0 = f_1(t_0)$; on traiterait de même le cas où $x_0 = f_0(t_0)$. Il faut distinguer deux cas suivant que la fonction

$$\varphi_1(h) = f_1(t_0 + h) - f_1(t_0)$$

appartient à l'une ou l'autre des classes considérées au n° 8, 1°.

Si elle appartient à la classe supérieure, cela revient à dire que la frontière $x = f_1(t)$ fuit assez vite pour n'être presque sûrement pas rattrapée immédiatement par la particule mobile. Donc $U(t_0 + 0) = 1$.

Si au contraire $\varphi_0(h)$ appartient à la classe inférieure, le franchissement immédiat de la borne $f_1(t)$ est presque sûr, et $U(t_0 + 0) = 0$.

Il y a d'ailleurs dans les deux cas une solution positive $u(t, x)$ de l'équation de la chaleur, définie à un facteur constant près par les conditions (9) et la condition

$$U(t, x) = \int_{f_0(t)}^x u(t, y) dy \rightarrow 0 \quad [f_0(t_0) < x < x_0, t \downarrow t_0].$$

Mais, dans le premier cas, quand t tend vers t_0 , $U(t) = U[t, f_1(t)]$ tend vers une limite finie, qu'il faut égaler à l'unité pour obtenir la probabilité définie au 1°. Dans le second cas, $U(t)$ augmente indéfiniment. Au point de vue de la théorie de la chaleur, cela indique qu'il faudrait placer initialement en x_0 une quantité de chaleur infinie. Elle serait presque entièrement absorbée instantanément par la *barrière froide* $x = f_1(t)$, et il resterait une quantité finie et rapidement décroissante.

L'étude de l'équation de la chaleur permet ainsi de distinguer les

deux classes de fonctions $\varphi_1(h)$. C'est sur cette remarque que s'appuie la méthode de Petrovsky et Kolmogoroff mentionnée au n° 7. Naturellement cette méthode implique que le problème local, que nous n'avons étudié qu'au n° 8, soit étudié avant le problème asymptotique.

3° Si $X(t_0)$ a une valeur initiale x_0 comprise entre $f_0(t_0)$ et $f_1(t_0)$, $u(t, x)$, $U(t, x)$ et $U(t)$ deviennent des fonctions $u(t, x|x_0)$, $U(t, x|x_0)$ et $U(t|x_0)$ qui dépendent de x_0 , et si $X(t_0)$ a une fonction de répartition $U_0(x)$, leurs valeurs s'obtiennent par intégration. On a par exemple

$$U(t) = \int_{f_0(t_0)-0}^{f_1(t_0)+0} U(t|x) dU_0(x).$$

Il est important de remarquer que $U(t|x_0)$ est une fonction continue de x_0 . On pourrait à première vue se demander s'il en est bien ainsi quand x_0 tend vers une des valeurs limites, $f_1(t_0)$ par exemple. Mais si la fonction $\varphi_1(h)$ appartient à la classe supérieure, la présence de la barrière $x=f_1(t)$ est sans influence sur $U(t_0+0)$, et *a fortiori* sur $U(t_0+0|x_0)$, toujours égal à un; il n'en résulte aucune discontinuité. Si $\varphi_1(h)$ appartient à la classe inférieure, et si $x_0=f_1(t_0)$, $X(t')-f_1(t')$ a dans l'intervalle (t_0, t) une borne supérieure $M(t)$, presque sûrement positive, dépendant d'une loi continue, et pouvant avec une probabilité positive être arbitrairement petite. Donc

$$U[t|f_1(t_0)-\varepsilon] = \Pr[M(t) \leq \varepsilon]$$

tend vers zéro avec ε , ce qui, dans ce cas encore, établit la continuité.

4° Revenons maintenant à la démonstration des conditions (9), pour n'importe quelle répartition initiale donnée, et pour presque toutes les valeurs de t supérieures à t_0 . La loi du mouvement brownien ne changeant pas si l'on change t en $-t$, nous pouvons appliquer les résultats précédents en faisant décroître t' de t à t_0 . Si $X(t)$ a une valeur connue voisine de $f_1(t)-\varepsilon$, la probabilité que $X(t')-f_1(t')$ ait été positif dans un petit intervalle $(t-\tau, t)$ ne dépend presque pas de la répartition initiale, et, quand ε tend vers zéro, elle tend vers zéro ou un suivant que la fonction

$$\varphi(h) = f_1(t-h) - f_1(t)$$

relative à la valeur t considérée appartient à la classe inférieure ou à la classe supérieure. Cela revient à dire que $u(t, x)$ est le produit de la densité positive obtenue en faisant abstraction des conditions (7) par un facteur qui tend vers zéro dans le premier cas et vers un dans le second.

Le second cas est exceptionnel, ce qui établit le résultat annoncé au 1°. S'il est réalisé en un point t , il cesse aussitôt de l'être. Si

$$f_1(t+h) - f_1(t)$$

appartient à la classe supérieure, c'est parce que la frontière fuit trop vite devant les particules mobiles. Dans l'autre cas, c'est par l'effet d'un flux qui fait que presque toutes les particules voisines ⁽¹⁾ de la frontière la franchissent immédiatement. Il n'est d'ailleurs pas exclu que les points exceptionnels forment un ensemble partout dense. Mais il est exclu que cet ensemble ait une mesure positive; il en résulterait un flux sortant infini, ce qui est impossible.

13. Cas des limites constantes ⁽²⁾. — 1° Supposons d'abord qu'il n'y ait qu'une limite, les valeurs admises pour $X(t)$ variant de $-\infty$ à une valeur constante a . Dans l'hypothèse $X(0) = x_0 < a$, la solution de notre problème est évidemment

$$(10) \quad u(t, x) = \frac{1}{\sqrt{2\pi t}} \left[e^{-\frac{(x-x_0)^2}{2t}} - e^{-\frac{(x_0+x-2a)^2}{2t}} \right].$$

Au point de vue de la théorie de la chaleur, cette formule exprime que le résultat est le même que s'il y avait à l'instant initial une source chaude au point x_0 et une source froide de la même intensité au point *image* $2a - x_0$ (symétrique de x_0 par rapport à a). A cause de la symétrie, au point a , le froid venu de la source froide compense constamment la chaleur venant de la source chaude, et la température y reste nulle.

Par intégration de la formule (10), on trouve aisément

$$U(t) = \int_{-\infty}^a u(t, x) dx = \sqrt{\frac{2}{\pi t}} \int_0^{a-x_0} e^{-\frac{y^2}{2t}} dy,$$

⁽¹⁾ Nous imaginons ici un grand nombre de particules donnant des déterminations indépendantes de la fonction $X(t)$.

⁽²⁾ Il s'agit ici de formules classiques de la théorie de la chaleur. Leur application au calcul des probabilités est due à L. Bachelier.

c'est-à-dire

$$(11) \quad U(t) = \Pr[|X(t)| < a - x_0].$$

Une intégration par rapport à x_0 , les expressions (10) et (11) étant multipliées par $dU_0(x_0)$, donne les nouvelles expressions de $u(t, x)$ et $U(t)$ quand la probabilité est initialement répartie d'une manière quelconque entre $-\infty$ et a .

2° Supposons maintenant que x varie entre deux limites constantes a_0 et $a = a_0 + l$. La solution de notre problème, dans l'hypothèse $X(0) = x_0$ ($a_0 < x_0 < a$), résulte encore aisément de la méthode des images, des parois réfléchissantes étant supposées placées aux points a_0 et a , et chaque image correspondant à une source froide ou chaude suivant la parité du nombre des réflexions. On trouve ainsi

$$(12) \quad u(t, x) = \frac{1}{\sqrt{2\pi t}} \sum_{-\infty}^{+\infty} \left[e^{-\frac{(x-x_0+2nl)^2}{2t}} - e^{-\frac{(x_0+x+2nl-2a)^2}{2t}} \right].$$

Cette expression, et celle de $U(t)$ qui s'en déduit, ne se prêtent pas bien à une étude asymptotique. Pour une telle étude, il vaut mieux utiliser le développement de $u(t, x)$ en série de Fourier. Nous supposerons, pour simplifier l'écriture, $a_0 = 0$, $a = l = \pi$, et une répartition initiale quelconque dans $(0, \pi)$. La solution de l'équation (3), s'annulant pour $x = 0$ et pour $x = \pi$, et à cela près la plus générale, est

$$(13) \quad u(t, x) = \sum_1^{\infty} a_n \sin nx e^{-\frac{n^2 t}{2}},$$

ce qui donne par intégration

$$(14) \quad U(t) = 2 \sum_0^{\infty} \frac{a_{2v+1}}{2v+1} e^{-(2v+1)^2 \frac{t}{2}}$$

La condition initiale donne d'ailleurs

$$(15) \quad a_n = \frac{2}{\pi} \int_0^{\pi} \sin nx dU_0(x),$$

de sorte que a_1 est essentiellement positif. On a donc, pour t infini

$$(16) \quad u(t, x) \sim a_1 \sin x e^{-\frac{t}{2}}, \quad U(t) \sim 2a_1 e^{-\frac{t}{2}}.$$

14. **Les fonctions $M(t)$ et $Y(t)$ (1).** — 1° Nous désignerons respectivement par $M(t)$ et $Y(t)$ le maximum de $X(u) - X(0)$ et celui de $X(u) - X(t)$ dans l'intervalle $(0, t)$. La définition de la fonction $X(u)$ ne changeant pas quand on change le sens de parcours de l'intervalle $(0, t)$, $M(t)$ et $Y(t)$, pour chaque valeur fixe de t , dépendent de la même loi. Mais ce sont des fonctions de t de natures bien différentes. $M(t)$ est monotone, mais non $Y(t)$, qui a presque sûrement une infinité de racines.

Cela revient au même de dire que $M(t) < m$, ou que, $x_0 = X(0)$ étant supposé nul, $X(u)$ n'atteint pas la valeur m dans l'intervalle $(0, t)$. Les formules (10) et (11) s'écrivent donc

$$(10') \quad \frac{d}{dx} \Pr[M(t) < m, X(t) < x] = \frac{1}{\sqrt{2\pi t}} \left[e^{-\frac{x^2}{2t}} - e^{-\frac{(2m-x)^2}{2t}} \right],$$

et

$$(11') \quad \Pr[M(t) < m] = \Pr[|X(t)| < m].$$

Donc : $|X(t)|$, $M(t)$, et, par suite aussi $Y(t)$, dépendent pour chaque t de la même loi de probabilité. Il en est de même naturellement aussi des fonctions $M_1(t)$ et $Y_1(t)$ qui sont à $-X(t)$ ce que $M(t)$ et $Y(t)$ sont à $X(t)$.

2° Les résultats précédents peuvent s'obtenir aisément sans utiliser l'équation de la diffusion. Il n'y a qu'à utiliser un principe de symétrie appliqué par Désiré André à l'étude du gain maximum au cours d'une partie de pile ou face. Dans l'hypothèse $M(t) \geq m$, $X(u)$ atteint pour la première fois la valeur m pour une certaine valeur $T \leq t$ de u , et, quel que soit T supposé connu, la probabilité des valeurs positives de $X(t) - X(T) = X(t) - m$ est égale à la probabilité des valeurs négatives de cette variable, donc à $\frac{1}{2}$ [en écartant l'hypothèse $T = t$, qui est infiniment peu probable, puisqu'elle implique $X(t) = m$]. Cette probabilité étant constante, la probabilité *a priori*, sous la seule condition $M(t) \geq m$, a aussi la même valeur $\frac{1}{2}$, de sorte que

$$\frac{1}{2} \Pr[M(t) \geq m] = \Pr[M(t) \geq m, X(t) > m] = \Pr[X(t) > m],$$

ce qui équivaut à la formule (11').

(1) D'après P. Lévy [5] et [3], chap. VI; de même pour les nos 15 à 18.

On obtient de même la formule (10') en observant que, si $X(T) = m$, les valeurs x et $2m - x$ de $X(t)$ ont la même probabilité.

3° Revenons à la loi à deux variables $M(t)$ et $X(t)$, bien définie par la formule (10'). En la dérivant par rapport à m , on obtient la densité de probabilité

$$(17) \quad \sqrt{\frac{2}{\pi t}} \frac{2m - x}{t} e^{-\frac{(2m - x)^2}{2t}} \quad (m \geq 0, x \leq m).$$

Il suffit de poser $x = m - y$ pour obtenir la densité de probabilité de la loi à deux variables $M(t)$ et $Y(t)$, qui est

$$(18) \quad \sqrt{\frac{2}{\pi t}} \frac{m + y}{t} e^{-\frac{(m + y)^2}{2t}} \quad (m \geq 0, y \geq 0).$$

On vérifie qu'elle est symétrique, comme nous le savions *a priori*. Il résulte de cette symétrie, et du caractère antisymétrique de $x = m - y$, que la densité de la loi à deux variables $X(t)$ et $Y(t)$ se déduit de l'expression (17) en remplaçant $2m - x$ par $2y + x$; on peut aussi le vérifier par un nouveau changement de variables.

4° De la loi à deux variables $M(t)$ et $Y(t)$ on déduit aisément la loi dont dépend $Y(t)$ quand $M(t)$ a une valeur connue m . Si, en particulier, $m = 0$, c'est-à-dire si

$$Y(t) = -X(t) = |X(t)|,$$

il vient

$$(19) \quad \Pr \{ |X(t)| > x / M(t) = 0 \} = e^{-\frac{x^2}{2t}}.$$

Plus généralement, si l'on sait qu'une des valeurs extrêmes de $X(t)$ dans un intervalle (t_0, t_1) est réalisée en l'un des points extrêmes t_0 et t_1 , la formule (1) doit être remplacée par la formule

$$(19') \quad \Pr[|X(t_1) - X(t_0)| > x] = e^{-\frac{x^2}{2(t_1 - t_0)}},$$

formule que nous avons déjà utilisée au n° 8, 3° dans le cas où c'est $X(t_0)$ qui est le minimum de $X(t)$.

5° Une autre conséquence du principe de symétrie qui conduit aux formules (10) et (10') est relative à la probabilité que $X(t)$ s'annule dans un intervalle $[t_0, t_1]$, lorsque $x_0 = X(t_0)$ et $x = X(t_1)$ sont connus. Cette existence étant sûre ou presque sûre si $x_0 x \leq 0$, nous supposons x_0 et x positifs.

Si x_0 est seul connu, on a

$$\Pr[x \leq X(t_1) < x + dx] = \frac{dx}{\sqrt{2\pi(t_1 - t_0)}} e^{-\frac{(x-x_0)^2}{2(t_1-t_0)}},$$

et, en appelant A l'éventualité de l'existence d'une racine de $X(t)$ entre t_0 et t_1 , on déduit de la formule (18), appliquée à $x_0 - X(t)$,

$$\Pr[A, x \leq X(t) < x + dx] = \frac{dx}{\sqrt{2\pi(t_1 - t_0)}} e^{-\frac{(x_0+x)^2}{2(t_1-t_0)}}$$

En divisant ces deux formules l'une par l'autre, on trouve

$$(20) \quad \Pr[A/X(t_1) = x] = e^{-\frac{2x_0x}{t_1-t_0}}.$$

15. **L'inversion de $M(t)$.** — 1° La fonction non décroissante $M(t)$ admet une fonction inverse $T(x)$. En supposant $X(0) = 0$, $T(x-0)$, bien défini pour tout x positif, est la plus petite racine positive de $X(t) = x$.

En un point choisi au hasard, on a presque sûrement $M(t) > X(t)$, donc $Y(t) > 0$. L'ensemble $\mathcal{E}$ des racines de $Y(t)$ est donc presque sûrement de mesure nulle ⁽¹⁾; c'est évidemment un ensemble parfait discontinu, non dénombrable. $M(t)$ est constamment croissant sur $\mathcal{E}$, et constant dans chacun des intervalles i constituant l'ensemble complémentaire $\mathcal{E}'$. Donc $T(x)$ ne peut croître que par sauts (chacun de ces sauts correspond à un intervalle i , et la projection de ces intervalles sur l'axe des t en recouvre la partie positive, à un ensemble de mesure nulle près).

2° On a

$$\Pr[T(x) > t] = \Pr[M(t) < x] = \sqrt{\frac{2}{\pi t}} \int_0^x e^{-\frac{\xi^2}{2t}} d\xi,$$

donc, en posant $\xi^2 = \frac{t}{\tau} x^2$,

$$\Pr[T(x) > t] = \frac{x}{\sqrt{2\pi}} \int_t^\infty \tau^{-\frac{3}{2}} e^{-\frac{x^2}{2\tau}} d\tau.$$

(1) Nous considérons qu'il y a deux interventions indépendantes du hasard, l'une pour choisir la fonction $X(t)$, l'autre pour choisir une valeur particulière t' de t . La condition $X(t') = 0$ définit dans l'espace produit un ensemble mesurable. Il résulte alors du théorème de Fubini qu'on peut intervertir l'ordre des choix; la mesure, qui est nulle, reste nulle.

Donc : $T(x)$ dépend d'une loi absolument continue, dont la densité de probabilité est

$$(21) \quad \frac{x}{\sqrt{2\pi}} t^{-\frac{3}{2}} e^{-\frac{x^2}{2t}}.$$

3° On remarque, ce qui était évident *a priori*, que la loi dont dépend $\frac{T(x)}{x^2}$ est indépendante de x . Une autre propriété presque évidente de $T(x)$ est que

$$(22) \quad T(x_1) - T(x_0) = T_1(x_1 - x_0) \quad (0 < x_0 < x_1),$$

$T_1(x_1 - x_0)$ désignant une expression qui dépend de la même loi que $T(x_1 - x_0)$ et est indépendante de $T(x_0)$. Le point $t_0 = T(x_0)$, x_0 du plan des tx est, en effet, un point de la courbe $x = X(t)$, à partir duquel les propriétés de cette courbe sont les mêmes qu'à partir de l'origine. Le temps $T(x_1) - t_0$ nécessaire pour que $X(t) - x_0$ atteigne la valeur $x_1 - x_0$ est donc indépendant de t_0 , et de la forme $T_1(x_1 - x_0)$, C. Q. F. D.

On peut représenter $T(x_0)$ et $T_1(x_1 - x_0)$ par $a_1\tau_1$ et $a_2\tau_2$, en posant $a_1 = x_0^2$, $a_2 = (x_1 - x_0)^2$, et τ_1 et τ_2 étant deux variables aléatoires indépendantes, dépendant toutes les deux de la loi (20), écrite pour $x = 1$. On peut égaler a_1 et a_2 à deux nombres positifs quelconques; on voit ainsi que : *quels que soient a_1 et a_2 positifs, $a_1\tau_1 + a_2\tau_2$ est de la forme $a\tau$, τ dépendant de la même loi que τ_1 et τ_2* . Le coefficient positif $a = x_1^2$ se déduit de a_1 et a_2 par la formule

$$a^{\frac{1}{2}} = a_1^{\frac{1}{2}} + a_2^{\frac{1}{2}}.$$

Ce résultat peut s'exprimer en disant que : *la loi dont dépend $T(x)$ est une loi stable d'exposant caractéristique $\frac{1}{2}$* . Nous venons de le démontrer indépendamment de l'expression (21) de sa densité de probabilité. On peut le vérifier aussi à l'aide de cette expression, et de la formule

$$f(t) = \int_0^t f_1(u)f_2(t-u) du$$

qui donne la densité $f(t)$ d'une somme de deux variables aléatoires positives et indépendantes l'une de l'autre en fonction des densités $f_1(t)$ et $f_2(t)$ de ces deux variables.

4° Il résulte de la théorie connue des lois stables (cf. par exemple, P. Lévy [2], p. 176-181) que cette propriété de stabilité, compte tenu de ce que $T(x)$ est toujours positif, détermine, à un changement d'unité près, la loi (indépendante de x) dont dépend le rapport $\tau = \frac{T(x)}{x^2}$. La théorie des lois stables conduit à une définition simple de ces lois par leurs fonctions caractéristiques

$$\varphi(z) = E(e^{iz\tau}).$$

Compte tenu de ce que $\Pr(\tau > t)$ est, pour t infini, un infiniment petit équivalent à $\sqrt{\frac{2}{\pi t}}$, ce qui achève de déterminer la loi dont il s'agit, on trouve ainsi

$$(23) \quad \log \varphi(z) = (-1 + i \operatorname{sgn} z) \sqrt{|z|} = \frac{1}{\sqrt{2\pi}} \int_0^\infty (e^{izu} - 1) u^{-\frac{3}{2}} du.$$

5° Dans cette intégrale, le coefficient de $(e^{izu} - 1)$ donne, dans l'intervalle $(0, 1)$, le nombre probable des sauts de la fonction $T(x)$ de hauteurs comprises entre u et $u + du$. Le nombre probable de ceux qui dépassent u , c'est-à-dire le nombre probable des intervalles $\mathbf{i}$, de longueurs $> u$, qui précèdent l'instant $T(1)$, est donc $\sqrt{\frac{2}{\pi u}}$. Leur nombre réel est une variable de Poisson, bien définie par cette valeur probable.

Le nombre total de ces intervalles $\mathbf{i}$ est presque sûrement infini, tandis qu'il n'y en a qu'un nombre fini dépassant une longueur positive donnée. La probabilité qu'une telle longueur soit dépassée pour un intervalle dont on ne connaît que l'origine est donc nulle. Mais on déduit aisément de la formule précédente que : *si l'on sait que la longueur d'un intervalle $\mathbf{i}$ est $\geq u_0$, la probabilité qu'elle soit $\geq u_1$ ($u_1 > u_0$) est $\sqrt{\frac{u_0}{u_1}}$.*

Cette probabilité devient naturellement une fréquence presque sûre sur l'ensemble de l'axe des t .

6° Le résultat précédent est d'ailleurs facile à vérifier indépendamment de la théorie générale des lois stables. Si à un instant donné $Y(t)$ a une valeur y , le temps θ qui s'écoule jusqu'à la prochaine racine est une variable aléatoire du type $T(y) = y^2 \tau$. Or, si

l'on sait qu'à l'instant considéré t l'intervalle $\mathbf{i}$ qui le contient a atteint une longueur u_0 , on a, d'après la formule (19')

$$\Pr \{ Y(t) > y \} = e^{-\frac{y^2}{2u_0}},$$

et la probabilité que sa longueur totale dépasse u_1 est celle de $Y^2(t)\tau > u_1 - u_0$. Un calcul facile redonne bien la valeur $\sqrt{\frac{u_0}{u_1}}$ (cf. P. Lévy [2], p. 213-214).

7° Il résulte du 5°, et de la loi forte des grands nombres que, pour t infini, le nombre des intervalles $\mathbf{i}$ intérieurs à $(0, t)$ et de longueurs $> u$ est un infiniment grand presque sûrement équivalent à $\sqrt{\frac{2}{\pi u}} M(t)$. Mais on peut aussi appliquer la loi forte des grands nombres dans un intervalle fini $(0, t)$, en faisant tendre u vers zéro. On voit ainsi que : *u tendant vers zéro, d'une part le nombre des intervalles $\mathbf{i}$ intérieurs à un intervalle donné $(0, t)$ et de longueur $> u$ est un infiniment grand presque sûrement équivalent à $\sqrt{\frac{2}{\pi u}} M(t)$; d'autre part, la longueur totale de ceux dont la longueur est $< u$ est un infiniment petit presque sûrement équivalent à $\sqrt{\frac{2u}{\pi}} M(t)$ (1).*

Un quelconque de ces énoncés nous montre que : *la connaissance de l'ensemble $\mathcal{S}$ des racines de $Y(t)$ suffit à déterminer $M(t)$* . Cet ensemble n'est d'ailleurs pas quelconque; nous savions déjà que c'est un ensemble parfait discontinu, non dénombrable, de mesure nulle. L'existence des limites qui permettent de définir $M(t)$ constitue une nouvelle restriction.

16. Identité des définitions intrinsèques de $Y(t)$ et de $|X(t)|$. —

1° Nous avons déjà observé que, dans l'hypothèse $X(0) = 0$, $|X(t)|$, $Y(t)$ et $M(t)$ dépendent, pour chaque valeur de t , de la même loi de probabilité. Mais il y a entre les deux premières fonctions une

(1) On voit aussi aisément que l'ensemble des points de $(0, t)$ dont la distance à $\mathcal{S}$ est $< \varepsilon^2$ a une mesure qui est presque sûrement $c\varepsilon M(t) + o(\varepsilon)$ ($\varepsilon \rightarrow 0$, $c = \frac{4}{\sqrt{\pi}}$).

Nous exprimerons ce résultat en disant que $M(t)$ est une *mesure presque sûre du voisinage de $\mathcal{S}$* .

analogie bien plus profonde qu'entre elles et $M(t)$: *considérées comme fonctions aléatoires de t , elles ont la même définition, c'est-à-dire que toutes les deux ont la même probabilité d'appartenir à n'importe quel sous-ensemble donné de l'ensemble des fonctions continues. Il s'agit donc de leur définition intrinsèque, ne faisant pas intervenir leurs relations avec $X(t)$.*

Pour le démontrer, observons d'abord que toutes les deux sont markoviennes. Comme nous partons pour toutes les deux de la même valeur initiale $Y(0) = X(0) = 0$, il suffit de montrer que la loi de probabilité de passage de l'instant t_0 à l'instant $t_1 > t_0$ est la même pour ces deux fonctions.

Si la valeur initiale est nulle, la variation à partir de l'instant initial t_0 est la même qu'à partir de l'origine $t = 0$, et l'on est ramené au résultat appelé ci-dessus.

Si, au contraire, on part d'une valeur initiale $y_0 > 0$, la variation de ces fonctions est d'abord définie par la même équation

$$\delta Y_i(t) = \xi \sqrt{dt} \quad [i = 0, 1; Y_0(t) = Y(t), Y_1(t) = |X(t)|],$$

valable jusqu'à l'instant $t_0 + T(y_0)$ où $Y_i(t)$ s'annule pour la première fois. Il n'est pas nécessaire, pour déterminer $Y_i(t)$ jusqu'à cet instant, de savoir s'il s'agit de $Y(t)$ ou de $|X(t)|$. Si $t_0 + T(y_0) = T_1 \geq t_1$, $Y_i(t_1)$ se trouve ainsi déterminé de la même manière dans les deux cas. Si, au contraire, $T_1 < t_1$, cette valeur constitue un nouveau point de départ, à partir duquel on raisonne comme à partir de l'origine, et $Y(t_1)$ dépend aussi bien que $|X(t_1)|$ de la même loi que $|X(u)|$, pour $u = t_1 - T_1$, quand on sait seulement que $X(0) = 0$. Le résultat annoncé est ainsi démontré.

2° Un corollaire immédiat de ce résultat est que toutes les propriétés de l'ensemble $\mathcal{E}$ des racines de $Y(t)$ s'appliquent sans changement à l'ensemble $\mathcal{E}_0$ des racines de $X(t)$.

Il faut à ce sujet remarquer que la propriété établie au n° 15, 5° au sujet des longueurs des intervalles complémentaires i ne suffit pas à caractériser $\mathcal{E}$. Elle s'appliquerait sans changement si les intervalles i correspondaient, non aux sauts de $T(x)$, mais à ceux d'une fonction telle que $T(x) + cx$. L'ensemble $\mathcal{E}$ aurait alors une mesure positive; l'axe des t serait irrégulièrement dilaté, de manière que les longueurs des intervalles i soient invariantes, mais que $\mathcal{E}$ ait une mesure proportionnelle à $M(t)$. On écarte cette possibilité en préci-

sant que $\mathcal{E}$ a une mesure nulle. Cette propriété, jointe à celle établie au n° 15, 5°, caractérise $\mathcal{E}$, en ce sens qu'elles permettent de définir une succession d'expériences permettant de construire, sans passer par l'intermédiaire de $X(t)$, un ensemble dépendant exactement de la même loi de probabilité que $\mathcal{E}$.

17. **La loi de l'arc sinus.** — 1° Nous serons conduits à cette loi par l'étude du problème suivant : *dans l'hypothèse $X(0) = 0$, quelle est la probabilité pour qu'un intervalle (t_0, t_1) ($0 < t_0 < t_1$) contienne au moins une racine de $X(t)$?*

Ce problème a déjà été résolu en supposant connus, soit $X(t_0)$ seul, soit $X(t_0)$ et $X(t_1)$. Nous supposons maintenant qu'on n'ait pas d'autre renseignement que $X(0) = 0$, de sorte que la densité de probabilité de $|X(t_0)|$ (le signe n'importe pas pour la suite) est

$$\sqrt{\frac{2}{\pi t_0}} e^{-\frac{x^2}{2t_0}}.$$

D'autre part, dans l'hypothèse $X(t_0) = x$, la probabilité de l'existence d'au moins une racine entre t_0 et t_1 est

$$\Pr[T(x) < t_1 - t_0] = \frac{x}{\sqrt{2\pi}} \int_0^{t_1 - t_0} e^{-\frac{x^2}{2\tau}} \tau^{-\frac{3}{2}} d\tau.$$

La probabilité cherchée est donc

$$\alpha = \frac{1}{\pi \sqrt{t_0}} \int_0^{t_1 - t_0} \tau^{-\frac{3}{2}} d\tau \int_0^\infty e^{-\frac{x^2(t_0 + \tau)}{2t_0\tau}} x dx = \frac{\sqrt{t_0}}{\pi} \int_0^{t_1 - t_0} \frac{d\tau}{(t_0 + \tau) \sqrt{\tau}},$$

ou, en posant $\tau = t_0 u^2$,

$$\alpha = \frac{2}{\pi} \int_0^{\sqrt{\frac{t_1 - t_0}{t_0}}} \frac{du}{1 + u^2} = \frac{2}{\pi} \operatorname{Arc} \operatorname{tg} \sqrt{\frac{t_1 - t_0}{t_0}},$$

formule qui peut s'écrire aussi

$$(24) \quad t_0 = t_1 \cos^2 \frac{\pi \alpha}{2} \quad \text{ou} \quad \alpha = \frac{2}{\pi} \operatorname{Arc} \cos \sqrt{\frac{t_0}{t_1}},$$

tandis que pour la probabilité complémentaire $\beta = 1 - \alpha$ on a

$$(24') \quad t_0 = t_1 \sin^2 \frac{\pi \beta}{2} \quad \text{ou} \quad \beta = \frac{2}{\pi} \operatorname{Arc} \sin \sqrt{\frac{t_0}{t_1}}.$$

2° Étant donnée une valeur particulière de t , nous désignerons par T_0 la plus grande racine de $X(t)$ qui soit $< t$, et par T_1 la plus petite qui soit $> t$. L'une et l'autre existent presque sûrement.

La probabilité $\Pr(T_0 < t_0)$ ($0 < t_0 < t$) n'est autre que β , calculé pour $t_1 = t$. Donc

$$(25) \quad \Pr(T_0 < t_0) = \frac{2}{\pi} \text{Arc sin} \sqrt{\frac{t_0}{t}}.$$

Cette formule constitue ce qu'on appelle la *loi de l'arc sinus*. La densité de probabilité correspondante est

$$(26) \quad \frac{2}{\pi} \frac{d}{dt_0} \text{Arc sin} \sqrt{\frac{t_0}{t_1}} = \frac{1}{\pi \sqrt{t_0(t-t_0)}}.$$

Cette loi a une interprétation géométrique simple : si l'on pose

$$T_0 = t \sin^2 \frac{\Phi}{2} = \frac{t}{2} (1 - \cos \Phi) \quad (0 < \Phi < \pi),$$

Φ est une variable uniformément répartie dans l'intervalle $(0, \pi)$. Donc : T_0 est l'abscisse d'un point choisi au hasard sur une demi-circonférence de diamètre $(0, t)$, la probabilité étant répartie uniformément sur cette demi-circonférence (on ne changerait évidemment rien en considérant la circonférence entière).

On remarque qu'il s'agit d'une loi symétrique par rapport au milieu de l'intervalle considéré. Cela n'était peut-être pas évident *a priori*. Mais c'est un corollaire évident du fait que les résultats intrinsèques relatifs à l'ensemble $\mathcal{G}_0$ des racines de $X(t)$ s'appliquent aussi à $\mathcal{G}$. Nous pouvons donc remplacer T_0 par la plus grande racine T'_0 de $Y(t)$ qui soit $< t$, c'est-à-dire la valeur presque sûrement unique qui réalise le maximum de $X(u)$ dans l'intervalle $(0, t)$. La symétrie est alors évidente.

La raison de cette symétrie est la même que pour celle de la loi à deux variables $M(t)$ et $Y(t)$. La loi à trois variables $M(t)$, $Y(t)$ et T'_0 résulte immédiatement de la remarque que, dans l'hypothèse $T'_0 = t'$, on a

$$\Pr[M(t) > m] = e^{-\frac{m^2}{2t'}}, \quad \Pr[Y(t) > y] = e^{-\frac{y^2}{2(t-t')}},$$

ces deux probabilités conditionnelles étant indépendantes.

3° En ce qui concerne T_1 , on a, d'après la formule (24),

$$(27) \quad \Pr(T_1 < t_1) = \frac{2}{\pi} \text{Arc cos} \sqrt{\frac{t}{t_1}} \quad (t_1 > t),$$

de sorte que la densité de probabilité est

$$(28) \quad \frac{2}{\pi} \frac{d}{dt_1} \text{Arc cos} \sqrt{\frac{t}{t_1}} = \frac{1}{\pi t_1} \sqrt{\frac{t}{t_1 - t}}.$$

C'est la *loi de l'arc cosinus*.

4° Pour la loi à deux variables T_0 et T_1 , on a évidemment

$$(29) \quad \text{Pr}(T_0 < t_0, T_1 > t_1) = \frac{2}{\pi} \text{Arc sin} \sqrt{\frac{t_0}{t_1}} \quad (0 < t_0 < t < t_1),$$

et la densité de probabilité est

$$(30) \quad f(t_0, t_1) = \frac{1}{2\pi \sqrt{t_0(t_1 - t_0)}}.$$

On remarque que ces expressions ne contiennent pas t . Cela n'est pas surprenant, puisque nous n'avons fait que récrire la formule (24) avec d'autres notations. Mais il faut noter que cela implique la relation

$$\int_0^t dt_0 \int_t^\infty f(t_0, t_1) dt_1 = 1$$

valable quel que soit t positif.

Remarquons aussi que la densité de probabilité de T_1 lorsque T_0 a une valeur connue t_0 , c'est-à-dire lorsqu'on sait que $X(t_0) = 0$ et qu'il n'y a pas d'autre racine de $X(t')$ entre t_0 et t est, d'après le n° 15, 5°

$$-\frac{d}{dt_1} \sqrt{\frac{t - t_0}{t_1 - t_0}} = \frac{1}{2} \sqrt{\frac{t - t_0}{(t_1 - t_0)^3}},$$

et qu'on peut aussi obtenir la formule (30) en multipliant par cette expression la densité de probabilité relative à T_0 . On peut aussi, inversement, déduire la formule du n° 15, 5° des formules (26) et (30), c'est-à-dire en fin de compte de la formule (24).

5° Pour aller plus loin, il faut utiliser le théorème suivant : *la définition intrinsèque de $\mathcal{E}$ (ou de $\mathcal{E}_0$) entre deux éléments θ_0 et θ_1 de cet ensemble est invariante par une substitution homographique*; on peut en particulier supposer θ_1 infini; l'ensemble $\mathcal{E}$ (ou $\mathcal{E}_0$), qui est fermé, doit en tout cas être considéré comme contenant le point $t = +\infty$.

Ce théorème est une conséquence immédiate de la remarque du n° 4, 2°.

Pour généraliser la formule (24), il n'y a donc qu'à introduire, au lieu de $\frac{t_0}{t_1}$, le rapport anharmonique

$$\varphi = \varphi(t_0, t_1; \theta_0, \theta_1) = \frac{(t_0 - \theta_0)(t_1 - \theta_1)}{(t_1 - \theta_0)(t_0 - \theta_1)},$$

et l'on voit que : *si l'on sait que $\mathcal{E}$ (ou $\mathcal{E}_0$) contient les points θ_0 et θ_1 , et qu'on n'ait aucun autre renseignement concernant l'intervalle (θ_0, θ_1) , et si $\theta_0 < t_0 < t_1 < \theta_1$, la probabilité que cet ensemble contienne au moins un point entre t_0 et t_1 est*

$$(31) \quad \frac{2}{\pi} \text{Arc cos } \sqrt{\varphi} = \frac{2}{\pi} \text{Arc cos } \sqrt{\frac{(t_0 - \theta_0)(\theta_1 - t_1)}{(t_1 - \theta_0)(\theta_1 - t_0)}}.$$

En effet, cet énoncé est invariant par une substitution homographique ne changeant pas l'ordre relatif des quatre points considérés, et se réduit, pour $\theta_0 = 0$ et $\theta_1 = \infty$, à la formule (24).

On en déduit immédiatement des conséquences qui généralisent les formules (25) à (30). Comme on peut évidemment changer le sens positif sur l'axe des t , la loi relative à T_0 , qu'on peut appeler *loi généralisée de l'arc sinus*, et qui s'exprime par la formule

$$\text{Pr}(T_0 < t_0) = \frac{2}{\pi} \text{Arc sin } \sqrt{\varphi(t_0, t; \theta_0, \theta_1)}$$

donne aussi, en échangeant les indices 0 et 1, la probabilité de $T_1 > t_1$.

6° Nous pouvons maintenant donner une définition *intrinsèque* de $\mathcal{E}$ (ou de $\mathcal{E}_0$), c'est-à-dire qu'elle ne fait pas intervenir les relations de cet ensemble avec $X(t)$. Nous nous placerons dans un intervalle $(0, t)$ et supposerons que $0 \in \mathcal{E}$ [s'il s'agit de $\mathcal{E}_0$ et que $X(0) = x_0 \neq 0$, il n'y a qu'à déterminer d'abord le premier point $T(x_0)$ appartenant à $\mathcal{E}_0$, et prendre ce point comme nouvelle origine sur l'axe des t]. Nous utiliserons alors la formule (25) pour déterminer T_0 . L'intervalle (T_0, t) étant sans racine, nous n'avons qu'à considérer l'intervalle $(0, T_0)$, sachant que ses extrémités sont des points de $\mathcal{E}$ (presque sûrement non isolés du côté de cet intervalle).

Choisissons un point t_1 dans cet intervalle (soit $\frac{T_0}{2}$, par exemple, soit, dans une suite formée d'avance, le premier point compris entre 0 et T_0), et déterminons les extrémités T'_0 et T'_1 de l'intervalle **i** sans point de $\mathcal{E}$ qui contient t_1 .

La formule (31), pouvant s'écrire

$$(31') \quad \Pr(T'_0 < t'_0, T'_1 > t'_1) = \frac{2}{\pi} \text{Arc sin } \sqrt{\rho(t'_0, t'_1; 0, T_0)} \quad (0 < t'_0 < t_1 < t'_1 < T_0),$$

définit les conditions de ce choix. On a ainsi un intervalle (T'_0, T'_1) sans racine, et deux nouveaux intervalles $(0, T'_0)$ et (T'_1, T_0) auxquels on appliquera à nouveau la même méthode, et ainsi de suite indéfiniment.

Il est presque sûr que les points de $\mathcal{E}$ ainsi successivement obtenus ne seront jamais aux extrémités de l'intervalle où on les recherche. De plus, si pour fixer les idées le point qui est chaque fois choisi arbitrairement dans un intervalle est le milieu de cet intervalle, l'ensemble $\mathcal{E}$ sera enfermé après n opérations dans la réunion $\mathcal{E}^{(n)}$ de 2^n intervalles de longueurs sûrement inférieures à $\frac{T_0}{2^n}$; on voit d'ailleurs aisément que, T_0 supposé connu, ces longueurs ont une même valeur probable $\left(\frac{\mu}{2}\right)^n T_0$, avec $0 < \mu < 1$. Il en résulte que $\mathcal{E}$ est presque sûrement un ensemble parfait discontinu, sans points isolés à la fois à droite et à gauche, et de mesure nulle (puisqu'il est recouvert par $\mathcal{E}^{(n)}$, dont la mesure probable, $\mu^n T_0$, tend vers zéro). Nous retrouvons bien, par cette construction directe, les propriétés déjà déduites de la définition initiale de $\mathcal{E}$ (ou de $\mathcal{E}_0$).

18. La reconstruction de $X(t)$ en partant de $\mathcal{E}$ ou de $\mathcal{E}_0$. — 1° Un ensemble aléatoire étant formé comme il vient d'être indiqué, si l'on veut l'identifier, soit à $\mathcal{E}$, soit à $\mathcal{E}_0$, la détermination de $X(t)$ devient un problème de probabilité conditionnelle dont la résolution comprend deux étapes. La première, qui se pose de la même manière pour les deux problèmes, consiste à déterminer la fonction non négative $Y(t)$ ou $|X(t)|$ qui s'annule sur l'ensemble considéré. La seconde différencie les deux problèmes. Si, en effet, $\mathcal{E}$ et $Y(t)$ sont connus, $M(t)$ et, par suite $X(t) = M(t) - Y(t)$, sont bien définis par les propriétés presque sûres établies au n° 15, 7°. Si, au contraire, $|X(t)|$ est connu, il reste à choisir le signe de $X(t)$ dans chacun des intervalles séparés par les racines de cette fonction, les deux signes possibles étant bien entendu également probables.

2° Revenons à la première opération en supposant, pour fixer les idées, qu'il s'agisse de déterminer $|X(t)|$ entre deux racines

consécutives θ_0 et θ_1 , et que $\mathbf{X}(t)$ soit positif dans (θ_0, θ_1) ; donc $|\mathbf{X}(t)| = \mathbf{X}(t)$. Il faut alors, comme au n° 3, procéder par interpolations, en déterminant successivement $\mathbf{X}(t)$ pour une suite de nombres t_n partout dense dans (θ_0, θ_1) . Il n'y a aucune difficulté. Si $\theta_0 \leq t_0 < t < t_1 \leq \theta_1$, et si $\mathbf{X}(t_0)$ et $\mathbf{X}(t_1)$ ont des valeurs connues x_0 et x_1 , $\mathbf{X}(t)$ dépend d'une loi absolument continue, dont la densité de probabilité est de la forme

$$(32) \quad g(x) = c f(t - t_0, x_0, x) f(t_1 - t, x_1, x),$$

les deux facteurs correspondants respectivement aux renseignements concernant l'intervalle $[t_0, t)$ et à ceux qui concernent $(t, t_1]$. Le coefficient c est ensuite défini par la condition

$$(33) \quad \int_0^{+\infty} g(x) dx = 1.$$

Si l'on n'avait pas la condition $\mathbf{X}(t) > 0$ dans (t_0, t_1) , on retrouverait ainsi les résultats obtenus autrement au n° 3. Cette condition oblige à changer la forme de la fonction f . Comme il suffit de sa valeur à un facteur constant (par rapport à x) près, on n'a qu'à prendre la fonction $u(t - t_0, x_0 - x)$ [notations de la formule (10), α étant ici remplacé par x_0]. Le calcul s'achève sans difficulté. Nous nous contenterons d'indiquer le résultat simple relatif au cas où $x_0 = x_1 = 0$ (ce qui a lieu si et seulement si $t_0 = \theta_0, t_1 = \theta_1$). Comme, d'après la formule (19),

$$f(\tau, 0, x) = \frac{x}{\tau} e^{-\frac{x^2}{2\tau}},$$

on arrive dans ce cas au résultat que $\mathbf{X}(t)$ est de la forme $\sigma(t)U(t)$,

$$(34) \quad \sigma(t) = \sqrt{\frac{(t - t_0)(t_1 - t)}{t_1 - t_0}}$$

ayant la même valeur qu'au n° 3, mais n'étant plus un écart type, et $U = U(t)$ dépendant, quel que soit t , de la loi

$$(35) \quad \Pr(U < u) = \sqrt{\frac{2}{\pi}} \int_0^u v^2 e^{-\frac{v^2}{2}} dv, \quad (1).$$

(1) Non seulement on obtient ainsi une loi indépendante de t , mais la définition intrinsèque de $U(t)$ est invariante par n'importe quelle substitution homographique qui conserve t_0 et t_1 . Ainsi la loi à deux variables $U(t)U(t')[t, t' \in (t_0, t_1)]$ ne dépend

On en déduit notamment que

$$(36) \quad E[X(t)] = 2\sqrt{\frac{2}{\pi}}\sigma(t),$$

et, si T est une variable aléatoire uniformément répartie dans $(t_0 - t_1)$

$$(37) \quad E[X(T)] = \sqrt{\frac{\pi}{8}}(t_1 - t_0).$$

3°. Sans même écrire avec précision les formules qui précèdent, il était évident par raison d'homogénéité que $U(t)$, même au voisinage des points t_0 et t_1 , n'est en général ni très grand, ni très petit (nous supposons toujours $x_0 = x_1 = 0$). $|X(t)|$ est alors en moyenne, et en général, de l'ordre de grandeur de $\sigma(t)$, donc de celui de $\sqrt{\tau}$, τ désignant la différence entre t et la racine de $X(u)$ la plus voisine de cette valeur.

Introduisons alors la fonction $M_0(t)$ qui est à l'ensemble $\mathcal{E}_0$ ce que $M(t)$ est à $\mathcal{E}$ en vertu des théorèmes du n° 15, 7°. On peut dire qu'à l'intérieur de l'intervalle $(0, t)$ cette fonction mesure le voisinage de $\mathcal{E}_0$. Comme entre deux points non consécutifs de $\mathcal{E}_0$ il y en a une infinité d'autres séparés par une infinité dénombrable d'intervalles, les lois des grands nombres s'appliquent, et il est à prévoir que la mesure de l'ensemble défini par $0 < t < T$, $|X(t)| < x$, est, quand x tend vers zéro, un infiniment petit presque sûrement équivalent à $2cxM_0(T)$, c étant un coefficient constant. Si l'on remplace la condition $|X(t)| < x$ par $0 \leq X(t) < x$, le coefficient numérique 2 disparaît.

Il n'est pas très difficile de démontrer qu'il en est bien ainsi, et que $c = 1$. Le lecteur trouvera la démonstration dans notre Mémoire de 1939 [5], p. 330-331 et 337-338. Comme le résultat obtenu s'applique à l'ensemble $\mathcal{E}_x$ des racines de $X(t) = x$, et à son voisinage défini par $x \leq X(t) < x + dx$, on obtient le résultat suivant : $\mu(T, x)$ désignant la mesure de l'ensemble défini par $0 < t < T$,

que du rapport anharmonique de t_0, t_1, t, t' . C'est une conséquence immédiate du théorème analogue établi au n° 4, 2°. En effet la propriété d'invariance établie à cet endroit s'étend *ipso facto* à la loi conditionnelle obtenue en ajoutant à la condition $X(t_0) = X(t_1) = 0$ la condition invariante $X(t) > 0$ dans (t_0, t_1) .

$X(t) < x$, et $M_x(t)$ étant lié à $\mathcal{E}_x$ comme $M(t)$ l'est à $\mathcal{E}$ d'après le n° 15, 7°, on a presque sûrement

$$\mu(T, x) = \int_{-\infty}^x M_{\xi}(T) d\xi.$$

19. **Le passage du fini à l'infini.** — 1° Nous avons au n° 1 introduit explicitement la loi de Gauss dans la définition de $X(t)$. On sait qu'un des caractères les plus remarquables de cette loi est qu'elle s'introduit asymptotiquement dans l'étude des sommes de variables aléatoires indépendantes. Sous la seule condition que le plus grand des termes soit très probablement très petit par rapport à la somme supposée centrée sur sa médiane, la somme diffère très peu d'une variable gaussienne. Dans la théorie générale des fonctions aléatoires additives, ce théorème classique (quoique énoncé ici sous une forme moderne), prend la forme suivante : *une fonction aléatoire additive et presque sûrement continue est à accroissements gaussiens; par une réduction triviale, elle se ramène à $X(t)$.*

2° Il résulte des théorèmes que nous venons de rappeler que l'on peut définir $X(t)$ de la manière suivante ; S_n désignant le gain après n coups de pile ou face, avec enjeux tous égaux à l'unité, on définit $X_{\tau}(t)$ par les conditions : *a.* $X_{\tau}(n\tau) = S_n \sqrt{\tau}$; *b.* $X_{\tau}(t)$ est continu et varie linéairement dans chacun des intervalles $[n\tau, (n+1)\tau]$. On peut considérer $X(t)$ comme limite de $X_{\tau}(t)$, pour τ tendant vers zéro, mais seulement *au point de vue de Bernoulli*; cela signifie que, pour n'importe quel domaine Ω dans l'espace des fonctions continues on a

$$\lim \Pr[X_{\tau}(t) \in \Omega] = \Pr[X(t) \in \Omega].$$

L'existence d'une telle limite n'est pas évidente, et c'est pour cela que la théorie de Bachelier mentionnée au n° 2 n'était pas satisfaisante. On peut la déduire du fait que les lois considérées appartiennent à un ensemble compact. Notre démonstration des n°s 3 à 5 est plus élémentaire, et offre en outre l'avantage d'être constructive. Mais, cette existence une fois établie, la méthode d'induction qui précède donne une méthode possible pour l'étude de $X(t)$. Chaque théorème relatif à la suite des S_n (qu'on peut souvent généraliser en ne supposant pas que chaque accroissement $S_{n+1} - S_n$ soit de la forme ± 1), donne à la limite un théorème relatif à $X(t)$. C'est la

méthode du passage du fini à l'infini. L'étude directe de $X(t)$ nous a, en principe, paru préférable. Mais l'autre méthode peut être utile aussi. Nous allons donner, à propos de deux problèmes qui n'ont pas été considérés dans les précédents paragraphes, quelques indications qui permettent de s'en rendre compte.

3° Le premier problème est relatif à la mesure $\mu(T)$ de l'ensemble des t de l'intervalle $(0, T)$ pour lesquels $X(t) > 0$; c'est donc $T - \mu(T, 0)$, d'après les notations du n° 18, 3°. Dans notre Mémoire de 1939 [p. 325, form. (54)], nous avons démontré que, si $X(0) = 0$, $\mu(T)$ se répartit entre 0 et T suivant la loi de l'arc sinus

$$(38) \quad \Pr[\mu(T) < \lambda T] = \frac{2}{\pi} \text{Arc sin } \sqrt{\lambda} \quad (0 \leq \lambda \leq 1).$$

Naturellement, si $X(T)$ est connu, on a une formule différente, et, dans l'hypothèse $X(0) = X(T) = 0$, nous avons obtenu une formule plus simple encore : *la répartition de $\mu(T)$ est uniforme dans l'intervalle $(0, T)$* [loc. cit., p. 323, form. (47)].

Il arrive que les formules asymptotiques soient plus simples que les formules finies; aussi n'avions-nous pas pensé que ce résultat remarquablement simple se retrouverait dans le domaine fini. Or il en est ainsi, du moins dans le cas du jeu de pile ou face. K. L. Chung et W. Feller [1] ont en effet démontré un théorème équivalent au suivant : *pour ce jeu, et dans l'hypothèse $S_{2n} = 0$, les $n + 1$ valeurs possibles pour le nombre des termes positifs dans la suite $S_1, S_3, \dots, S_{2n-1}$, sont également probables.* On aurait naturellement un résultat moins simple en considérant les sommes de rangs pairs; qui peuvent s'annuler.

Ce résultat simple est très remarquable, et, par la méthode du passage du fini à l'infini, il conduit à notre théorème de 1939 rappelé ci-dessus.

4° Désignons par $M^*(T)$ le maximum de $|X(t)|$ dans l'intervalle $[0, T]$. La loi du logarithme itéré et le théorème de Kolmogoroff qui la précisent s'appliquent indifféremment à $X(t)$, $|X(t)|$, $M(t)$, $M^*(t)$; il s'agit là de la recherche des valeurs exceptionnellement grandes de ces fonctions. Le problème de la recherche des valeurs exceptionnellement petites ne se pose que pour les deux dernières, et n'est pas le même pour les deux. La probabilité des valeurs inférieures à $m\sqrt{t}$, pour m très petit, est, pour $M(t)$, $\sim \sqrt{\frac{2}{\pi}} m$,

et, pour $M^*(t)$, il résulte de la formule (16) qu'elle est $\sim c e^{-\frac{\pi^2}{8m^2}}$, c étant constant. L'ordre de grandeur n'est pas le même dans les deux cas, ce qui était à prévoir; les valeurs exceptionnellement petites de $\frac{M(t)}{\sqrt{t}}$ ont surtout chance d'être réalisées dans les intervalles où $X(t)$ est resté longtemps négatif, et $M^*(t)$ y est en général au moins de l'ordre de grandeur de $\sqrt{t}$.

Un raisonnement analogue à celui qui conduit à la loi du logarithme itéré donne alors les résultats suivants :

a. La fonction $\sqrt{t}(\log t)^{-\alpha}$ est presque sûrement, pour t assez grand, une borne inférieure de $M(t)$ si $\alpha > 1$; elle ne l'est pas si $\alpha < 1$.

b. La fonction $c \frac{\pi}{2} \sqrt{\frac{t}{\log \log t}}$ est presque sûrement, pour t assez grand, une borne inférieure de $M^(t)$ si $c < 1$, mais non si $c > 1$.*

Ces résultats sont ainsi des conséquences relativement simples de nos formules de 1939. En ce qui concerne les problèmes analogues pour les sommes finies, le premier a été résolu il y a quelques années par Warren Hirsch (*Thèse*, Univ. de New-York). Le second, sur l'intérêt duquel W. Feller avait attiré l'attention, a été complètement résolu par K. L. Chung [1], non seulement dans le cas du jeu de pile ou face, mais dans un cas beaucoup plus général. Par le passage du fini à l'infini, on déduit de son résultat le théorème suivant, beaucoup plus précis que l'énoncé *b* ci-dessus.

b'. Si $\psi(t)$ est une fonction non décroissante, $\frac{\pi \sqrt{t}}{2 \sqrt{\psi(t)}}$ est presque sûrement, pour t assez grand, une borne inférieure de $M^(t)$ si l'intégrale*

$$\int_0^\infty \psi^*(t) e^{-\psi^*(t)} \frac{dt}{t}$$

est convergente; elle ne l'est presque sûrement pas si cette intégrale est divergente ⁽¹⁾.

La démonstration directe de ce résultat serait sans doute plus

(1) Ce théorème relatif à $M^*(t)$ s'applique évidemment aussi à la fonction aléatoire

$$Y^*(t) = \max_{u \in [0, t]} Y(u).$$

simple. Mais il faut remarquer que, si K. L. Chung n'avait eu en vue que le passage du fini à l'infini, il aurait pu se borner au cas du jeu de pile ou face, et simplifier sa démonstration. En tout cas il y a lieu de remarquer que, contrairement à ce qui a eu lieu pour le problème traité au 3°, le problème relatif aux sommes finies a été posé et résolu avant de l'être pour la fonction $X(t)$.

20. Fonctionnelles de $X(t)$. — 1° Les fonctionnelles considérées jusqu'ici, telles que $M(t)$, $M^*(t)$, $Y(t)$, dépendaient surtout de points particuliers. Pour une étude générale, il y a lieu de mentionner d'abord que l'étude des fonctionnelles linéaires est particulièrement simple; une telle fonctionnelle est toujours une variable aléatoire gaussienne.

Considérons, par exemple, l'intégrale

$$(39) \quad U_1(T) = \int_0^T f(t) dX(t),$$

où $f(t)$ est supposé continu dans $(0, T)$. Cette condition ne suffit pas pour que U soit défini comme intégrale de Stieltjes-Young. Les sommes de Riemann-Stieltjes

$$(40) \quad \sum_1^n f(t_v) [X(t_v) - X(t_{v-1})] = \sum_1^n f(t_v) \xi_v \sqrt{t_v - t_{v-1}}$$

n'ont pas toujours une limite sous les seules conditions

$$0 = t_0 < t_1 < \dots < t_n = T, \quad \text{Max}(t_v - t_{v-1}) \rightarrow 0.$$

Mais elles sont toujours de la forme $\sigma_n \xi$, σ_n ayant une limite σ définie par

$$(41) \quad \sigma^2 = \int_0^T f^2(t) dt,$$

de sorte qu'en répartition elles tendent vers la forme $\sigma \xi$.

Nous reviendrons au chapitre III sur les intégrales de la forme (39), à propos d'une application pour laquelle nous chercherons à dépasser le point de vue de Bernoulli. Mais à ce point de vue il est bien évident que $U_1(T)$ est de la forme $\sigma \xi$, σ étant défini par la formule (41). Si d'ailleurs nous remplaçons $f(t_v)$ dans la formule (40) par la

moyenne de $f(t)$ dans (t_{v-1}, t_v) , nous voyons que le résultat subsiste sous la seule condition que $f(t)$ soit de carré sommable.

2° La théorie des fonctionnelles quadratiques de $X(t)$ repose sur une remarque bien simple, que nous aurions pu déjà utiliser pour l'étude des fonctionnelles linéaires. Nous savons, par la formule de Fourier-Wiener ou plus simplement par une remarque du n° 5 (2° alinéa), que $X(t)$ dépend linéairement d'une suite de variables gaussiennes ξ_n , indépendantes les unes des autres. Donc une fonctionnelle quadratique U_2 de $X(t)$ est une fonction quadratique des ξ_n . Or les ξ_n peuvent être considérés comme les coordonnées d'un point dans l'espace de Hilbert. Il en résulte que la forme quadratique U_2 peut être orthogonalisée, c'est-à-dire mise sous la forme

$$(42) \quad U_2 = \sum \frac{\lambda_n}{2} \chi_n^2,$$

cette somme étant finie ou infinie, et les χ_n étant, comme les ξ_n , des variables gaussiennes, réduites et indépendantes les unes des autres ⁽¹⁾.

Il est bien connu que chaque terme $\frac{\chi_n^2}{2}$ a pour fonction caractéristique $(1 - iz)^{-\frac{1}{2}}$. Celle de U_2 est donc

$$(43) \quad E(e^{izU_2}) = \prod \frac{1}{\sqrt{1 - \lambda_n iz}}.$$

La condition nécessaire et suffisante pour la convergence presque sûre de la série (42) est la convergence du produit (43). Donc inversement n'importe quelle fonction de la forme (43) est la fonction caractéristique de fonctionnelles quadratiques de $X(t)$.

L'ensemble de lois ainsi caractérisé est manifestement un sous-semi-groupe du semi-groupe des lois indéfiniment divisibles.

3° Comme exemple simple, considérons l'expression

$$(44) \quad S = \frac{\xi\eta' - \eta\xi'}{2} = \frac{1}{4}(\chi_1^2 + \chi_2^2 - \chi_3^2 - \chi_4^2),$$

⁽¹⁾ Cf. M. Kac et A. J. F. Siegert [1]. Ces auteurs font intervenir la théorie des équations intégrales, tandis que nous nous contentons de faire appel à un théorème classique de la théorie de l'espace de Hilbert. Mais l'application concrète de la théorie générale peut obliger tout de même à introduire une équation intégrale, comme nous allons le faire au 4° de ce paragraphe.

qui représente l'aire du triangle inscrit dans la courbe du mouvement brownien plan qui sera étudiée au chapitre III (les sommets correspondants aux points $t = 0, 1, 2$). La fonction caractéristique de $U = 2S$ est $\frac{1}{1+z^2}$; la densité de probabilité correspondante est $\frac{1}{2}e^{-|u|}$.

Plus généralement, toutes les fois que les λ_n sont deux à deux égaux, le radical disparaît, et la fonction caractéristique (43) est méromorphe. S'ils sont deux à deux égaux et opposés, elle devient réelle et paire. Si ces deux circonstances sont réunies, elle prend la forme

$$\prod \frac{1}{1 + \lambda_n^2 z^2}.$$

Nous venons de voir l'exemple le plus simple de cette circonstance.

4° Un exemple moins élémentaire est celui de l'intégrale

$$(45) \quad I = \int_0^1 X^2(t) dt \quad (2).$$

On peut écrire

$$I = \int_0^1 dt \int_{1-t}^1 \xi_u \sqrt{d(1-u)} \int_{1-t}^1 \xi_v \sqrt{d(1-v)} = \int_0^1 \int_0^1 \xi_u \xi_v \text{Min}(u, v) \sqrt{du dv}.$$

D'après la théorie connue des formes quadratiques définies dans l'espace de Hilbert, les λ_n (*a priori* tous positifs) sont les constantes fondamentales de l'équation de Fredholm

$$(46) \quad 2 \int_0^1 \text{Min}(t, u) f(u) du = \lambda f(t),$$

qui peut s'écrire

$$(47) \quad 2 \int_0^t u f(u) du + 2t \int_t^1 f(u) du = \lambda f(t).$$

Par deux dérivations successives, il vient

$$(48) \quad 2 \int_t^1 f(u) du = \lambda f'(t),$$

$$(49) \quad -2f(t) = \lambda f''(t).$$

(2) Nous avons introduit cette intégrale en 1940, à propos d'une application qui sera indiquée au chapitre III. La loi dont elle dépend a été formée par R. H. Cameron et W. T. Martin [1]. Nous reproduisons ici notre exposé de Berkeley (P. Lévy [9]), qui utilisait lui-même, en le simplifiant, un exposé de R. Fortet.

Cette équation différentielle, compte tenu des conditions

$$f(0) = f'(1) = 0$$

qui résultent de (47) et (48), donne

$$f(t) = c \sin \left(n + \frac{1}{2} \right) \pi t$$

et, par suite,

$$(50) \quad \lambda_n = \frac{8}{(2n+1)^2 \pi^2}.$$

La formule (43) se réduit alors à la formule de Cameron et Martin

$$(51) \quad E(e^{t|z|}) = \frac{1}{\sqrt{\cos \sqrt{2iz}}}.$$

Comme $\cos \sqrt{2iz}$ est une fonction entière, et ne s'annule pas sur l'axe réel, il n'y a qu'à choisir la détermination du radical égale à un pour $z = 0$; elle est bien définie par continuité.

5° D'une manière beaucoup plus générale, M. Kac [1] a étudié les fonctionnelles de la forme

$$(52) \quad U(T) = \int_0^T \Phi[X(t)] dt,$$

la fonction $\Phi(x)$ étant supposée continue et non négative, et démontré le théorème suivant, que nous nous contenterons d'énoncer : si $F(t, u)$ est la fonction de répartition de $U(t)$, on peut la définir par sa transformée de Laplace-Stieltjes, définie par la formule

$$(53) \quad \int_0^\infty \int_0^\infty e^{-(\alpha t + \beta u)} dt d_u F(t, u) = \int_{-\infty}^{+\infty} f(x, \alpha, \beta) dx,$$

où $f = f(x, \alpha, \beta)$ est la solution de l'équation différentielle

$$(54) \quad \frac{d^2 f}{dx^2} = 2[\alpha \Phi(x) + \beta] f(x)$$

qui vérifie les conditions suivantes : $f(\pm \infty) = 0$; $f'(x)$ est borné et continu sauf à l'origine où

$$(55) \quad f(+0) - f(-0) = -2.$$

Le cas où $\Phi(x) = x$ est trivial. M. Kac en a considéré deux autres.

Pour $\Phi(x) = x^2$, il retrouve la formule de Cameron et Martin. Pour $\Phi(x) = \frac{1 + \operatorname{sgn} x}{2}$, observant que la discontinuité à l'origine n'empêche pas l'application de son théorème, il retrouve notre loi de l'Arc sinus rappelée au n° 19, 3°.

6° Dans un autre ordre d'idées, signalons nos recherches, exposées en 1950 au Symposium de Berkeley (P. Lévy [9]), sur les équations différentielles stochastiques de la forme

$$\delta U(t) = dt \int_0^t f(t, \tau) dU(\tau) + \sigma(t) dZ(t),$$

où $Z(t)$ est la fonction de Wiener complexe du n° 10. Si l'on suppose $Z(t)$ préalablement formé, c'est une équation de Volterra. Si $Z(t)$ n'est pas connu, $dZ(t)$ devant alors être remplacé par $\zeta \sqrt{dt}$, on peut, par un changement de la forme $U(t) = \lambda t + \mu + U_1(t)$, ramener $U(t)$, dans l'intervalle $(0, 2\pi)$, à une fonction $U_1(t)$ qu'on peut supposer définie de $-\infty$ à $+\infty$, continue, périodique de période 2π , et nulle en moyenne. On peut représenter cette fonction par une série de Fourier à coefficients gaussiens, et, pour que ces coefficients soient indépendants, il faut et il suffit qu'elle soit stationnaire. On le vérifie aisément dans le cas particulier de la série de Fourier-Wiener.

D'autres types d'équations intégrales ou différentielles stochastiques liées à $X(t)$ ont été étudiés par divers savants, notamment K. Itô [1] à [3].

CHAPITRE III.

LE MOUVEMENT BROWNIEN DANS LE PLAN ET DANS L'ESPACE ⁽¹⁾.

21. Notions générales. — D'une manière générale, le mouvement brownien dans l'espace E_N à N dimensions est défini par les formules

$$(1) \quad x_v = X_v(t) \quad (v = 1, 2, \dots, N),$$

⁽¹⁾ Nous nous bornerons au cas d'un espace euclidien à N dimensions. Rappelons que le cas de la surface d'une sphère a été dès 1928 étudié par F. Perrin [1]; plus récemment, le cas d'espaces plus généraux a été considéré par K. Yosida.

Sauf lorsqu'une source différente sera précisée, les résultats que nous exposons dans ce Chapitre sont tirés de notre Mémoire de 1940 ou de notre livre de 1948 (P. Lévy [6] et [3], chap. VII).

où les X_v sont des déterminations indépendantes les unes des autres de la fonction $X(t)$ étudiée dans les chapitres I et II.

Dans le cas du mouvement plan, nous écrirons $X(t)$ et $Y(t)$ au lieu de $X_1(t)$ et $X_2(t)$.

Si le point mobile, que nous désignerons par $A(t)$ est initialement à l'origine O , et si $R(t)$ désigne la distance $OA(t)$, on a

$$(2) \quad \Pr[R(t) < r\sqrt{t}] = \frac{2}{\sqrt{2N}\Gamma\left(\frac{N}{2}\right)} \int_0^r \rho^{N-1} e^{-\frac{\rho^2}{2}} d\rho,$$

ce qui, pour $N = 2$, donne

$$(2') \quad \Pr[R(t) > r\sqrt{t}] = e^{-\frac{r^2}{2}} \quad (r > 0).$$

On remarque que

$$(3) \quad E[R^2(t)] = NE[X^2(t)] = Nt,$$

ce qui en particulier, si $N = 2$, donne

$$(3') \quad E[|X(t) + iY(t)|^2] = 2t.$$

Ces formules sont en relation avec celles du n° 10, où nous avons préféré introduire la variable $\frac{X(t) + iY(t)}{\sqrt{2}}$, qui est réduite pour $t = 1$.

On pourrait évidemment, ici aussi, dans le cas général, introduire le vecteur de composantes $\frac{X_v(t)}{\sqrt{N}}$, qui, pour $t = 1$, aurait un caractère unitaire. Il nous a semblé qu'il n'y avait pas d'avantage à le faire, et c'est le vecteur $\xi', \xi'', \dots, \xi^{(n)}$ (notations du n° 1) que nous considérerons comme *réduit*.

Si l'on introduit des notations vectorielles, le produit de deux vecteurs étant bien entendu leur produit scalaire, toutes les formules des n°s 2 et 3 comportent des généralisations évidentes, sur lesquelles il semble inutile d'insister; N figurera naturellement en facteur dans les expressions de la covariance.

22. Propriétés locales et étude des points exceptionnels. —

1° Nous désignerons par C la trajectoire du point $A(t)$. Il suffit évidemment d'étudier son allure au voisinage de l'origine, et pour les valeurs positives de t . Nous poserons $t_n = q^n$ ($0 < q < 1$)

et $A_n = A(t_n)$. Comme les ξ_n au n° 7, les points A_n forment une suite de variables aléatoires asymptotiquement indépendantes : l'influence de A_n sur le vecteur réduit $\frac{OA_{n+\nu}}{\sqrt{t_{n+\nu}}}$ tend vers zéro pour ν infini, et cela d'une manière uniforme par rapport à n . On en déduit, comme pour les ξ_n , que les extrémités des vecteurs réduits

$$O a_n = \frac{OA_n}{\sqrt{t_n}}$$

se répartissent dans les différentes régions de l'espace avec des fréquences presque sûrement proportionnelles aux probabilités théoriques.

Naturellement l'irrégularité de chacune des fonctions $X_\nu(t)$ entraîne une irrégularité aussi grande du mouvement du point $A(t)$, et en projection sur n'importe quelle direction le mouvement n'est presque sûrement monotone dans aucun intervalle. Il en est alors évidemment de même de la variation de $R(t)$. Quand t tend vers zéro, $R(t)$ décroît en moyenne comme $\sqrt{Nt}$, mais avec des irrégularités très grandes.

2° Si nous considérons les angles $\theta_n = A_n OA_{n+1}$, il y a évidemment aussi entre eux indépendance asymptotique, mais asymptotique seulement, puisqu'à une grande valeur de $R(t_n)$ correspondent probablement de petites valeurs de θ_{n-1} et de θ_n ; il y a une corrélation positive entre les θ_n d'indices peu différents. Mais, si $N > 2$, l'orientation du plan $OA_n A_{n+1}$ autour de OA_n est absolument indépendante du point A_{n-1} .

Supposons maintenant $N = 2$. Nous pouvons alors attribuer un signe à θ_n , qui variera de $-\pi$ à $+\pi$ et l'on a $E(\theta_n) = 0$. Posons $E(\theta_n^2) = \sigma^2$. En raison de l'indépendance asymptotique des θ_n , et compte tenu de théorèmes connus (P. Lévy [2], p. 233-246), l'angle polaire

$$\Theta_n = \theta_1 + \theta_2 + \dots + \theta_n$$

est asymptotiquement de la forme $\sigma \xi \sqrt{n}$, et a presque sûrement des valeurs arbitrairement grandes des deux signes. Il en résulte que la ligne polygonale $A_0 A_1, \dots$, et *a fortiori* la courbe C , n'atteignent presque sûrement le point O qu'après une extraordinaire variation de l'angle polaire Θ , qui a des valeurs arbitrairement grandes des deux signes.

3° Si l'on tient compte, en outre, de l'irrégularité de la décroissance du rayon vecteur, il est bien clair que la courbe C doit se recouper elle-même, d'autant plus que les mêmes circonstances se retrouvent au voisinage de n'importe quel point, et qu'il n'est pas à prévoir que l'arc de retrait réussisse à éviter de recouper l'arc d'approche. Ces remarques ne constituent, bien entendu, pas une démonstration; mais on démontre aisément que : *les valeurs de t qui correspondent aux points doubles constituent presque sûrement un ensemble partout-dense, non dénombrable, de mesure nulle* ⁽¹⁾.

Le problème des points doubles, dans le cas où $N > 2$, qui est bien plus difficile, a été résolu par A. Dvoretzky, P. Erdős et S. Kakutani [1]. Le résultat est le suivant : *si $N = 3$, il est encore presque sûr qu'il y a des points doubles sur n'importe quel arc de C ; si $N \geq 4$, il est presque sûr qu'il n'y en a pas.*

4° Il existe d'autres points exceptionnels que les points doubles. En nous bornant au cas du plan, considérons ceux qui correspondent aux maxima et minima de $Y(t)$. Ce sont des points au voisinage desquels la courbe est tout entière du même côté d'une parallèle à l'axe des x . Comme au voisinage d'un tel point t_0 il arrive une infinité de fois que $X(t) - X(t_0)$ [et aussi $X(t_0) - X(t)$] soit égal à

$$\sqrt{2\tau \log \log \frac{1}{\tau}} \quad (\tau = t - t_0),$$

les valeurs correspondantes de $Y(t) - Y(t_0)$ étant presque toutes $O(\sqrt{\tau})$, la courbe, tout en étant du même côté d'une droite, n'est presque sûrement pas à l'intérieur d'un angle d'ouverture $\omega < \pi$.

Cette remarque est importante; il en résulte qu'une rotation des axes, si faible soit-elle, fait apparaître d'autres points exceptionnels.

(1) Brièvement : n'importe quel arc de longueur finie partant de $A(t_0)$ est avec une probabilité positive β un tracé approché de l'arc (t_0, t_1) de C . S'il a un point double, l'arc de C en a un aussi, avec une probabilité $\alpha \geq \beta > 0$. Or il résulte de théorèmes généraux connus que α ne peut être que zéro ou un; donc $\alpha = 1$.

D'autre part, l'intersection des arcs (t_0, t_1) et (t_1, t_2) ($t_0 < t_1 < t_2$) a presque sûrement la structure d'un ensemble parfait discontinu, non dénombrable. L'ensemble des points doubles de l'arc (t_0, t_2) est *a fortiori* non dénombrable.

Enfin le fait que ces points doubles restent exceptionnels résulte du fait, qui sera démontré au n° 24, que la mesure superficielle de la courbe C est presque sûrement nulle. Elle n'a donc aucune chance de passer exactement en un point donné, ni de repasser par le point atteint à un instant t donné.

Il y a ainsi presque sûrement une infinité non dénombrable et partout dense de points au voisinage desquels la courbe est tout entière du même côté d'une droite. Ces points restent exceptionnels, puisqu'il s'agit d'une propriété qui n'est presque sûrement pas réalisée pour une valeur donnée de t ; sur l'axe des t , ils correspondent à un ensemble de mesure nulle.

5° Considérons la plus petite aire convexe, soit $\mathcal{R}(t_0, t_1)$, ou plus simplement $\mathcal{R}$, qui contienne l'arc (t_0, t_1) de la courbe. Les points $A(t_0)$ et $A(t_1)$ sont presque sûrement à son intérieur. Il résulte du 4° que son contour n'a presque sûrement pas de points anguleux, de sorte que sur ce contour la direction de la tangente varie d'une manière continue. Mais les points où la courbe C atteint ce contour sont exceptionnels, et ne peuvent pas constituer un arc de ce contour. Ce contour comprend donc une infinité de segments rectilignes, et la tangente ne varie qu'aux points d'un ensemble parfait discontinu. On démontre d'ailleurs aisément qu'en un point de cet ensemble la courbure est presque sûrement infinie, et il en résulte qu'il est presque sûrement de mesure nulle (aussi bien sur le contour que sur l'axe des t).

D'autre part, lorsque t_1 croît, la croissance de $\mathcal{R}$ est évidemment continue, et ne se produit que pour des valeurs exceptionnelles de t_1 . La région $\mathcal{R}$, et son aire, sont donc des fonctions singulières de t_1 (c'est-à-dire continues, et ne variant que sur un ensemble de mesure nulle) ⁽¹⁾.

6° L'aire $\mathcal{R}'(t_0, t_1)$ comprenant l'arc (t_0, t_1) et l'ensemble des points du plan entouré par cet arc, a une variation tout à fait différente. Quand t_1 croît, elle ne peut varier qu'au moment où une boucle se ferme. Elle varie alors brusquement, et le point $A(t_1)$, au moment de cette variation, est un point double. Naturellement il ne peut y avoir qu'une infinité dénombrable de tels sauts. S'il y a une

(¹) Les remarques précédentes s'étendent aisément au cas où $N > 2$. Si, par exemple, $N = 3$, la surface convexe S limitant la région $\mathcal{R}$ n'a presque sûrement ni arête, ni point conique, mais a partout un plan tangent bien déterminé. D'autre part, un plan d'appui de l'arc (t_0, t_1) étant déterminé par trois points, il est infiniment peu probable que cet arc atteigne une quatrième fois un tel plan sans le dépasser. On en déduit aisément que, presque sûrement, la surface S est presque entièrement composée par des triangles d'appui, l'ensemble complémentaire de la réunion de ces triangles étant de mesure nulle; cet ensemble comprend des segments rectilignes dont les extrémités seules appartiennent à l'arc (t_0, t_1) .

infinité non dénombrable de points doubles, il n'y en a qu'une infinité dénombrable qui ne soient pas complètement entourés par l'arc séparant les deux passages au point considéré.

23. Loi du logarithme itéré et condition de Lipschitz. — 1°
THÉOREME. — *Si $c > 1$, il existe presque sûrement un nombre fini T tel que $t > T$ entraîne*

$$(4) \quad R(t) < c \sqrt{2t \log \log t}.$$

Si $c \leq 1$, il est presque sûr que T n'existe pas.

La seconde partie de ce théorème est évidente. Vraie pour chaque projection du vecteur $OA(t)$, elle l'est pour sa longueur.

Supposons donc $c > 1$, et choisissons un angle aigu α assez petit pour que $c' = c \cos \alpha > 1$. Considérons, d'autre part, n droites D_v ; $Y_v(t)$ désignant la projection de $OA(t)$ sur D_v , il existe presque sûrement, d'après le n° 7, 2°, un nombre fini T_v tel que $t > T_v$ entraîne

$$(5) \quad |Y_v(t)| < c' \sqrt{2t \log \log t}.$$

Si T est le plus grand des T_v , $t > T$ entraîne à la fois les n inégalités (5). Or on peut évidemment choisir n assez grand, et les directions des droites D_v convenablement réparties, de manière que n'importe quelle droite fasse avec au moins une des droites D_v un angle $< \alpha$. Alors l'ensemble des n inégalités (5) entraîne l'inégalité (4), qui est ainsi vérifiée pour $t > T$,

C. Q. F. D.

Ainsi, *en première approximation*, on a un énoncé indépendant de N . Les grandes valeurs de $R(t)$, par exemple celles qui dépassent $\sqrt{2t \log \log t}$, sont seulement d'autant moins rares que N est plus grand et, si $c > 1$, l'ordre de grandeur probable du nombre T croît avec N .

Mais cela n'est vrai qu'en première approximation. Il n'est guère douteux que, pour généraliser le théorème de A. Kolmogoroff énoncé au n° 7, il faille remplacer l'intégrale qui y figure par une intégrale contenant N , qui est peut-être

$$(6) \quad \int \psi^N(t) e^{-\frac{\psi^N(t)}{2}} dt.$$

2° Les remarques qui précèdent s'étendent immédiatement à la loi locale du logarithme itéré, et aussi, par la même méthode, à la condition de Lipschitz du n° 8. On voit ainsi que, si $c > 1$, il existe presque sûrement, pour tout arc fini de la courbe C, un nombre positif τ tel que $|t_1 - t_0| < \tau$ entraîne

$$r[\Lambda(t_0)\Lambda(t_1)] < c\sqrt{2\tau \log \frac{1}{\tau}}.$$

[$r(A, B)$ désignant la distance AB]. Il est seulement d'autant plus petit en probabilité que N est plus grand.

24. La mesure superficielle de la courbe C. — 1° En nous plaçant dans le cas du plan, nous allons montrer que :

THÉORÈME. — *La courbe C constitue un ensemble dont la mesure superficielle est presque sûrement nulle.*

Il suffit de le démontrer pour un arc fini $(0, t)$. Par raison d'homogénéité, la mesure superficielle de cet arc (sûrement bien définie, puisque c'est un ensemble fermé), a une valeur probable de la forme μt , μ étant une constante absolue. Cette mesure étant majorée par celle du carré circonscrit, on voit aisément que μ est fini.

Aux arcs $(0, t)$, $(t, 2t)$, et à leur réunion $(0, 2t)$, correspondent ainsi des mesures de valeurs probables μt , μt et $2\mu t$. Il en résulte que celle de leur intersection a pour valeur probable

$$(7) \quad \mu t + \mu t - 2\mu t = 0.$$

Elle est donc presque sûrement nulle.

Désignons maintenant par $\Phi\left(\frac{r}{\sqrt{t}}\right)$ la probabilité que l'arc $(0, t)$ contienne un point donné M situé à la distance r de l'origine; c'est une fonction monotone de t , donc aussi de r . On a la même probabilité pour les deux arcs $(0, t)$ et $(t, 2t)$, si r désigne maintenant la distance $MA(t)$. Ces deux probabilités étant indépendantes, la probabilité que M appartienne à la fois à ces deux arcs, en supposant $t = 1$, est $\Phi^2(r)$.

Or μ , et la mesure probable de l'intersection de ces deux arcs, sont respectivement les intégrales, étendues au plan entier, de $\Phi(r)$ et de la probabilité dont nous venons de calculer la valeur $\Phi^2(r)$.

Elles sont donc toutes les deux positives ou toutes les deux nulles. D'après la formule (7), la seconde est nulle. Donc $\mu = 0$, et n'importe quel arc de C a une mesure superficielle presque sûrement nulle.

2°. Dans le cas de l'espace, on ne peut pas parler de mesure superficielle. Mais Hausdorff a défini des mesures, dépendant d'un paramètre p , et qui pour $p = 2$ et pour les ensembles plans se réduit à la mesure superficielle de Lebesgue. La question se pose donc, pour $N > 2$, de savoir si la mesure de Hausdorff d'ordre 2 de la courbe C (qui *a priori* ne peut qu'augmenter avec N , la mesure d'une courbe majorant celles de ses projections) est aussi presque sûrement nulle.

La réponse est affirmative. Contentons-nous de dire que la démonstration repose sur la définition d'une mesure de Hausdorff majorée, dont l'étude, sans être triviale, est plus simple que celle de la mesure de Hausdorff elle-même (1).

25. La fermeture de la courbe C . — 1°. Chaque arc fini de C est un ensemble fermé qui, d'après le n° 23, ne saurait remplir une aire, ni *a fortiori* un volume. Mais il n'en est pas de même de la courbe entière, et l'on peut se demander si sa fermeture $\bar{C}$ ne remplit pas un volume, ou même tout l'espace. Il résulte d'ailleurs aisément de théorèmes généraux du calcul des probabilités qu'on ne peut trouver dans ces questions que des probabilités égales à zéro ou un, de sorte que, *ou bien il est presque sûr que $\bar{C}$ remplit tout l'espace, ou bien il est presque sûr que $\bar{C}$ a une mesure nulle*. En fait : la première de ces circonstances est réalisée dans le plan, et la seconde si $N > 2$.

Un problème de ce genre a été considéré pour la première fois par G. Pólya [1] et [2]. Au lieu du mouvement brownien, il considérait un mouvement discontinu ; à chaque saut, une des coordonnées variait de $\pm a$, les autres restant constantes ; il y avait ainsi $2N$ possibilités, également probables. Pólya a démontré que, si $N = 2$, le point mobile passe presque sûrement une infinité de fois à chacune de ses positions possibles (qui sont les sommets d'un quadrillage rectangulaire), tandis que si $N > 2$, il s'éloigne presque sûrement

(1) Cf. P. Lévy [13]. Ce travail pose aussi un problème plus général qui n'est pas complètement résolu.

indéfiniment de n'importe quel point donné. Par la méthode du passage du fini à l'infini, cet énoncé conduit à celui indiqué ci-dessus.

Pour ne pas reproduire exactement les raisonnements de G. Pólya, ou ceux de notre livre de 1948 (P. Lévy [3], p. 259-262), nous allons nous proposer de chercher une borne inférieure presque sûre de $R(t)$. Si $N > 2$, elle augmentera indéfiniment avec t , et le théorème énoncé en résultera.

2° Considérons d'abord la suite des points A_n qui correspondent à des valeurs t_n , indéfiniment croissantes, de t , et posons $R_n = OA_n = R(t_n)$. D'après la formule (2), la probabilité de $R_n \leq cn^{-\alpha} \sqrt{t_n}$ est le terme général d'une série convergente si $N\alpha > 1$. A cette condition il existe presque sûrement un $\bar{n}'$ tel que, pour $n > \bar{n}'$, on ait

$$(8) \quad R_n \geq cn^{-\alpha} \sqrt{t_n}.$$

Soit, d'autre part, ρ_n le maximum de la distance $A_n M$ quand M décrit l'arc $A_n A_{n+1}$. Il est borné supérieurement par

$$\sqrt{N} \text{Max} |X_v(t) - X_v(t_n)| \quad (t_n \leq t \leq t_{n+1}; v = 1, 2, \dots, N),$$

d'où, compte tenu du théorème de Boole et des formules du n° 14, 1°,

$$\text{Pr}[\rho_n > x_n \sqrt{N(t_{n+1} - t_n)}] \leq 4N \text{Pr}(\xi > x_n)$$

En prenant $x_n = c' \sqrt{2 \log n}$, et $c' > 1$, cette probabilité est le terme général d'une série convergente. On a donc presque sûrement, pour n assez grand (soit $n > \bar{n}''$)

$$(9) \quad \rho_n \leq c' \sqrt{2N(t_{n+1} - t_n) \log n}.$$

Or $R(t)$, sur l'arc $A_n A_{n+1}$, est borné inférieurement par $R_n - \rho_n$, et les formules (8) et (9) donnent ainsi une borne inférieure, qui n'est utile qu'à condition d'être positive. Si, pour fixer les idées, nous prenons $t_n = n^\beta$, on a $\rho_n = o(R_n)$ si $\alpha < \frac{1}{2}$, tandis que dans le cas contraire $R_n = o(\rho_n)$ et les formules ci-dessus ne donnent aucune borne inférieure positive pour $R(t)$.

Nous avons déjà supposé $\alpha > \frac{1}{N}$. Nous n'arrivons donc au résultat cherché que si $N > 2$. A cette condition, nous pouvons choisir α

entre $\frac{1}{2}$ et $\frac{1}{N}$; observant alors que le terme $-\rho_n$ est sans importance puisque c est une constante positive quelconque, nous voyons que, pour n assez grand, on a

$$R(t) \geq ct_n^{\frac{1}{2} - \frac{\alpha}{\beta}} \quad (t_n \leq t \leq t_{n+1}),$$

et par suite, pour t assez grand

$$R(t) \geq ct^{\frac{1}{2} - \varepsilon} \quad \left(\varepsilon = \frac{\alpha}{\beta} \right).$$

Or β peut être arbitrairement grand, donc ε arbitrairement petit. Donc enfin : *quelque petit que soit ε positif, si $N > 2$, il existe presque sûrement un nombre T tel que, pour $t \geq T$, on ait*

$$(10) \quad R(t) \geq ct^{\frac{1}{2} - \varepsilon}$$

(c étant un nombre positif quelconque; T dépend de c et de ε).

Il est, par contre, bien évident que, pour $\varepsilon = 0$, l'inégalité (10) est presque sûrement en défaut pour des valeurs arbitrairement grandes de t .

3° Dans le même ordre d'idées, et *toujours pour $N > 2$* , A. Dvoretzky et P. Erdős[1] ont obtenu un résultat tout à fait précis, que nous nous contenterons d'énoncer : *la fonction $g(t)$ étant supposée non croissante, la probabilité que l'on ait*

$$(11) \quad R(t) > g(t) \sqrt{t}$$

pour tout t assez grand est zéro ou un suivant que l'intégrale

$$\int g^{N-2}(t) \frac{dt}{t}$$

est divergente ou convergente à l'infini.

Ainsi, pour $g(t) = (\log t)^{-\alpha}$, il y a divergence, et $g(t)$ appartient à la *classe supérieure*, si $\alpha(N-2) \leq 1$, et à la *classe inférieure* si $\alpha(N-2) > 1$.

4° Le cas $N = 2$ semble plus difficile à traiter. Il y a lieu de penser que, pour tout ε positif, l'inégalité $R(t) > t^{-\varepsilon}$ est dans ce cas presque sûrement vérifiée pour tout t assez grand.

La difficulté consiste en ceci : en désignant par $m(t_0, t_1)$ le minimum de $R(t)$ dans l'intervalle (t_0, t_1) , il faudrait avoir une valeur

approchée de la probabilité $\text{Pr}[m(t_0, kt_0) < \rho \sqrt{t_0}]$ pour les petites valeurs de ρ . Dvoretzky et Erdős ont pu, pour $N > 2$, et $k \geq 4$, borner inférieurement cette probabilité, et cela leur a suffi. Pour $N = 2$, une double inégalité est nécessaire pour arriver au résultat par une méthode analogue à la leur.

5° Nous avons vu (n° 1) que la définition de la fonction $\frac{X(t)}{\sqrt{t}}$, est invariante par le changement de t en $\frac{t}{t}$. Il en est évidemment de même de $\frac{R(t)}{\sqrt{t}}$, et les résultats précédents se transforment en résultats relatifs à l'allure de la courbe au voisinage de l'origine (ou d'un point quelconque). Ainsi : *dans le cas $N > 2$, si et seulement si $\alpha(N-2) > 1$, l'inégalité $R(t) > \sqrt{t} \left(\log \frac{1}{t}\right)^{-\alpha}$ est presque sûrement vérifiée pour tout t positif assez petit.*

Si $N = 2$, en admettant que l'énoncé suggéré au 4° ci-dessus soit exact, $R(t)$ est presque sûrement $> t^{1+\varepsilon}$ pour tout t assez petit (ε positif), mais non $> ct$ (quelque petit que soit c).

26. L'équation de la diffusion. — 1° La densité de probabilité de la loi dont dépend $A(t)$ dans l'hypothèse où $A(t_0)$ a une position connue A_0 , est

$$(12) \quad u(t - t_0, A_0, M) = [2\pi(t - t_0)]^{-\frac{N}{2}} e^{-\frac{r^2}{2(t-t_0)}},$$

t étant $> t_0$, et r désignant la distance A_0M . Cette fonction vérifie l'équation de la chaleur

$$(13) \quad \Delta u = 2 \frac{\partial u}{\partial t}$$

qui devient ainsi, comme dans le cas linéaire, l'équation de la diffusion de la probabilité ⁽¹⁾. Comme dans le cas linéaire, il est facile, soit de le vérifier directement, soit de le déduire de l'évaluation du flux de probabilité à travers un élément de surface dS ; ce flux est, en un temps dt , de la forme $c \frac{du}{dn} dS dt$.

⁽¹⁾ Il s'agit d'une généralisation bien naturelle de l'équation de Bachelier pour le cas linéaire. L'équation analogue relative à la sphère ayant été utilisée en 1928 par F. Perrin [1], il y a lieu de penser qu'à cette date celle relative au plan était familière aux physiciens.

2° La question des problèmes aux limites se pose aussi tout naturellement comme dans le cas linéaire. Désignons par $u(t, M) dV$ la probabilité que le point mobile soit à l'instant t dans un élément de volume dV entourant le point M , et cela sans avoir atteint, depuis l'instant initial t_0 , une surface fermée S . On se donne, à l'instant t_0 , une répartition quelconque de la probabilité dans le volume V intérieur à S .

Analytiquement, il s'agit de chercher la solution de l'équation (13), définie dans V pour $t > t_0$, s'annulant sur S , et définissant une répartition qui, t tendant vers t_0 , tend vers la répartition initiale donnée.

On peut représenter cette fonction par un développement de la forme

$$(14) \quad u(t, M) = \sum_1^{\infty} c_n v_n(M) e^{-\lambda_n t},$$

v_n et λ_n ($n = 1, 2, \dots$) représentant les fonctions et les constantes fondamentales relatives à la région intérieure à S et à l'équation

$$(15) \quad \Delta v + 2\lambda v = 0.$$

Les c_n dépendent des données initiales.

Les λ_n étant rangés par ordre de grandeur croissante, $c_1 v_1(M)$ est toujours positif. Le premier terme de la série (14) donne donc la partie principale de $u(t, M)$ pour t infini, et la probabilité que le point mobile n'ait pas atteint la surface S avant l'instant t est un infiniment petit équivalent à

$$c_1 e^{-\lambda_1 t} \int_V v_1(M) dV.$$

3° La probabilité perdue (ou si l'on préfère, la chaleur perdue) à travers un élément dS de surface pendant le temps dt est $\frac{1}{2} \frac{du}{dn} dS dt$. La perte totale, t variant de t_0 à l'infini, est donc

$$(16) \quad \frac{1}{2} \frac{dU}{dn} dS \quad \left[U(M) = \int_{t_0}^{\infty} u(t, M) dt \right].$$

Or il est presque évident que, à un facteur constant près, U est la fonction de Green relative au point A_0 . On en déduit aisément ce

théorème de G. Pólya [2], note 11 ⁽¹⁾, le point mobile partant de A_0 finit presque sûrement par atteindre S ; *la probabilité qu'il atteigne S pour la première fois sur une portion ouverte s de cette surface est la mesure harmonique de s relative à A_0* , c'est-à-dire la valeur en A_0 d'une fonction harmonique $\Phi(A)$, régulière à l'intérieur de S , et tendant respectivement vers un ou vers zéro quand A tend vers un point de s ou de $S - \bar{s}$; pour les points frontières de $\bar{s}$, elle est discontinue mais bornée. Cette mesure est, en effet, l'intégrale de l'expression (16) dans s .

Pólya, pour démontrer son théorème, avait employé la méthode de passage du fini à l'infini, en étudiant d'abord le cas d'une promenade au hasard dans un réseau de rues. Mais on peut appliquer directement sa méthode au mouvement brownien, et l'on obtient sans doute ainsi la démonstration la plus simple de son théorème : $P(A_0)$ désignant la probabilité cherchée, et σ désignant la surface d'une sphère de centre A_0 intérieure à S , le point A où le point mobile atteint σ pour la première fois est une variable aléatoire uniformément répartie sur cette surface. Donc $P(A_0)$ est la moyenne de $P(A)$ sur σ . Donc $P(A)$ est une fonction harmonique. Les autres propriétés qui caractérisent la fonction $\Phi(A)$ étant manifestement vérifiées par $P(A)$, on a $P(A) = \Phi(A)$,
C. Q. F. D.

5° Dans le cas du plan, une autre démonstration, que nous avons donnée en 1945 (P. Lévy [8]) en même temps que celle du 3°, est basée sur un théorème intéressant par lui-même : *les propriétés intrinsèques de la courbe C (c'est-à-dire celles qui sont indépendantes de sa représentation paramétrique) sont invariantes par n'importe quelle représentation conforme*.

C'est une conséquence immédiate des remarques suivantes, elles-mêmes triviales : *a.* la définition de la courbe C est caractérisée par les propriétés élémentaires de ses différents arcs, qu'on peut considérer indépendamment les uns des autres; *b.* localement, toute représentation conforme se réduit à une similitude; *c.* la définition de C n'est pas changée par une similitude.

On peut alors ramener la *surface* S , devenue ici une courbe, et le

⁽¹⁾ Voir aussi S. Kakutani [1]. Nos travaux, dont les circonstances n'ont pas permis la publication avant 1945, remontent à 1941. Ils sont donc indépendants de ceux de Kakutani.

point intérieur A_0 , à une circonférence et son centre, et le théorème de Pólya devient évident.

27. Intégrales stochastiques ⁽¹⁾. — 1° Nous allons donner une définition généralisée de l'intégrale

$$(17) \quad I = \int_{T_0}^{T_1} f(t) dg(t),$$

qui s'applique dans certains cas où la définition classique n'a pas de sens. Nous considérerons les sommes de Riemann

$$(18) \quad S_n = \sum_1^n \frac{f(t_{v-1}) + f(t_v)}{2} [g(t_v) - g(t_{v-1})] \quad (t_0 = T_0 < t_1 < \dots < t_n = T_1),$$

qui correspondent à des aires limitées à des lignes polygonales inscrites dans la courbe $x = f(t)$, $y = g(t)$. Au point de vue classique, on exige que S_n ait une limite, qui sera l'intégrale, sous la seule condition que, n augmentant indéfiniment, la plus grande des différences $t_v - t_{v-1}$ tende vers zéro. On sait que, même si, ce que nous supposons, $f(t)$ et $g(t)$ sont continus, cette limite peut ne pas exister. Alors l'intégrale, au point de vue classique, n'existe pas.

⁽¹⁾ Nous avons introduit ces intégrales, à propos de l'application qui fera l'objet du n° 28, dans notre Mémoire de 1940 [6], et avons en 1940 et 1941 donné des indications nouvelles dans diverses Notes (*C. R. Acad. Sc.*; *Bull. Soc. Math.*; *Ann. Fac. Sc. de Lyon*). Un exposé d'ensemble de nos idées sur ces questions a été publié plus récemment (P. Lévy [11]).

Nos intégrales stochastiques n'ont rien de commun avec ce que, depuis E. Slutsky, plusieurs auteurs désignent par le même nom. Il s'agit ici d'une définition applicable à des fonctions individuelles, que nous appliquons en particulier à chaque détermination possible d'une fonction aléatoire. Une telle définition, sauf erreur, ne se trouve, avant 1940, que dans Paley et Wiener [1], chap. IX. Mais ces auteurs partent de la représentation de la fonction complexe $g(t) = Z(t)$ par la formule de Fourier-Wiener. Nous utiliserons cette formule au n° 30; mais, pour la théorie générale, il y a intérêt à partir de la définition intrinsèque de $f(t)$ et $g(t)$ et de sommes de Riemann-Stieltjes.

Depuis 1944, une théorie qui utilise aussi des sommes riemanniennes a été développée par K. Itô [1] à [3]; il remplace notre somme (18) par

$$\sum_1^n f(t_{v-1}) [g(t_v) - g(t_{v-1})],$$

et ne cherche pas à obtenir une définition applicable à des fonctions individuelles.

Nous introduisons alors une suite de nombres aléatoires T_n ($n = 2, 3, \dots$), indépendants les uns des autres, chacun étant choisi entre T_0 et T_1 avec répartition uniforme de la probabilité. Pour définir S_n , nous prendrons pour $t_1, t_2, \dots, t_{n-1}$ les nombres $T_2, T_3, \dots, T_n$ rangés par ordre de grandeur croissante; S_n devient alors une variable aléatoire, sûrement bornée pour chaque valeur de n . La suite des T_n étant presque sûrement partout dense dans (T_0, T_1) , si l'intégrale classique existe, il y a convergence presque sûre de S_n vers I . Mais il peut arriver que l'intégrale classique n'existe pas, et que S_n ait une limite presque sûre, qui ne saurait être aléatoire; cette limite sera alors l'*intégrale stochastique*.

On démontre aisément, que, pour n infini, la suite S_n a presque sûrement une borne inférieure et une borne supérieure respectivement égales à deux nombres I et $\bar{I}$ non aléatoires. Si ces deux nombres sont égaux, l'expression $f(t) dg(t)$ est *stochastiquement intégrable*, et leur valeur commune est I .

2° On peut considérer d'autres modes de convergence que la convergence presque sûre, notamment la convergence en probabilité et la convergence en moyenne quadratique. On peut donc donner d'autres définitions de l'intégrale stochastique, qui peuvent s'appliquer lorsque celle donnée ci-dessus est en défaut.

3° Une généralisation plus importante s'obtient en remarquant que S_n a toujours une valeur probable bien définie μ_n , et la convergence presque sûre de S_n vers I (de même que la convergence en moyenne quadratique) implique qu'il y ait à la fois convergence de μ_n vers I et convergence de $S_n - \mu_n$ vers zéro. Si la première de ces conditions n'est pas réalisée, il y a des *oscillations forcées*; si la seconde ne l'est pas, il y a des *oscillations fortuites*. Ces deux sortes d'oscillations peuvent exister indépendamment l'une de l'autre.

La généralisation dont nous voulons parler s'obtient en négligeant les oscillations fortuites; $S_n - \mu_n$ est *nul en moyenne*, et il suffit que μ_n ait une limite I pour qu'il soit naturel de considérer I comme une *limite en moyenne généralisée* de la suite des S_n . Nous dirons alors que I est une *intégrale stochastique généralisée*.

Il est facile de généraliser encore en exigeant seulement que la suite des μ_n ait une limite généralisée. Il est facile d'indiquer des exemples de courbes définies de manière à faire apparaître ainsi à

une certaine échelle une prépondérance de boucles décrites dans un sens, et à une échelle plus fine une prépondérance de boucles décrites dans l'autre sens, cette alternance se reproduisant indéfiniment. Si elle est régulière, la suite des μ_n peut avoir une limite généralisée.

Tel est le cas pour la courbe connue qui remplit l'aire d'un triangle rectangle isocèle, chacune de ses moitiés étant semblable à la courbe entière, mais avec retournement du sens des angles. Alors μ_n n'a pas de limite, mais la moyenne

$$\frac{a_1 \mu_1 + a_2 \mu_2 + \dots + a_n \mu_n}{a_1 + a_2 + \dots + a_n}$$

en a une si l'on prend par exemple

$$a_n = \frac{1}{n}.$$

4° La définition qui précède, par cela même qu'elle implique aussi peu de restrictions que possible, présente un inconvénient. Il peut arriver que l'intégrale de $f(t) dg(t)$ ait un sens de $-T$ à $+T$, sans en avoir dans chacun des intervalles partiels $(-T, 0)$ et $(0, T)$. Tel sera le cas si

$$f(t) + f(-t) = g(t) + g(-t) = 0.$$

A cause de la symétrie, μ_n sera toujours nul. Cela n'implique pas l'existence des intégrales partielles.

On peut alors introduire une condition restrictive : on ne considérera l'intégrale (17) comme ayant un sens pour un intervalle donné que si elle en a pour chaque intervalle partiel.

On peut remarquer qu'en ce qui concerne les oscillations fortuites, une condition de ce genre ne serait pas restrictive. Si, pour un intervalle donné, elles tendent vers zéro (presque sûrement ou en moyenne quadratique), il en est de même pour n'importe quel intervalle partiel.

5° Terminons ces notions générales en disant que de nombreux problèmes relatifs aux intégrales stochastiques restent actuellement posés, notamment : la recherche des conditions précises d'application des différentes définitions possibles ; l'influence des changements de variables ; et surtout l'extension aux intégrales multiples.

28. Application à la courbe du mouvement brownien plan ⁽¹⁾. —
 1° Supposant $X(0) = Y(0)$, nous considérerons l'intégrale

$$(19) \quad S(T) = \frac{1}{2} \int_0^T [X(t) dY(t) - Y(t) dX(t)]$$

qui représente, avec les conventions usuelles concernant les signes, l'aire comprise entre l'arc $(0, T)$ de la courbe C et sa corde, l'arc étant décrit dans le sens des t croissants. Par raison d'homogénéité, $S(T)$ est, pour chaque valeur de T , de la forme ST . La loi dont dépend S est indépendante de T ; elle est symétrique. Pour l'étudier, on peut supposer $S = S(1)$.

Remarquons d'abord que, au point de vue classique, S n'a aucune signification. Désignons, en effet, par L_n la ligne polygonale inscrite dans l'arc $(0, 1)$ de C , ayant pour sommets les points $t = 2^{-n}v$ ($v = 0, 1, \dots, 2^n$). La région comprise entre L_n et L_{n+1} se compose de 2^n triangles $T_n^{(v)}$, indépendants les uns des autres; leurs aires sont, de la forme $2^{-n}U$, la loi dont dépend U vérifiant les conditions

$$(20) \quad E(U) = 0, \quad E(|U|) = \mu = \frac{1}{4}, \quad E(U^2) = \sigma^2 = \frac{1}{8} \quad (2).$$

La somme de ces aires, prises en valeur absolue, tend presque sûrement vers sa valeur probable μ , et la somme de celles qui sont positives tend presque sûrement, pour n infini, vers $\frac{\mu}{2}$. Il en résulte que, presque sûrement, l'intégrale n'existe pas, au point de vue classique.

2° Nous allons maintenant montrer que : *l'intégrale stochastique, telle qu'elle a été définie au n° 27, 1°, existe presque sûrement.*

Si s_n désigne l'aire comprise entre L_n et L_{n+1} , qui est la somme algébrique de celles des 2^n triangles $T_n^{(v)}$, on a

$$(21) \quad E(s_n) = 0, \quad E(s_n^2) = 2^{-n} \sigma^2.$$

La série $\sum s_n$ est d'ailleurs ce que J. Ville a appelé une martingale,

⁽¹⁾ Il est à peine besoin de remarquer qu'on peut supposer qu'une seule des fonctions $f(t)$ et $g(t)$ est aléatoire. C'est ce qu'ont fait Paley et Wiener, en supposant $f(t)$ connu, mesurable, et de carré sommable, et $g(t) = Z(t)$. Comme $X(t)$ est presque sûrement continu, leur théorie s'applique à notre intégrale (19).

⁽²⁾ Ces formules résultent immédiatement, soit du n° 20, 3°, soit de $4U = \xi R(1)$, $R(t)$ étant défini par la formule (2'). La seule chose qui importe ici est que σ soit fini.

c'est-à-dire que la valeur probable de s_n est indépendante des valeurs de $s_1, s_2, \dots, s_{n-1}$, supposées connues. Dans ces conditions, les formules (21) entraînent à la fois la convergence en moyenne quadratique et la convergence presque sûre de cette série.

Le lecteur vérifiera aisément que ce résultat subsiste si l'on modifie la décomposition de l'intervalle d'intégration en intervalles partiels, en se donnant une suite de nombres T_n qui y soit partout dense, et si l'on prend les $n - 1$ premiers points $A(T_v)$, rangés dans leur ordre naturel sur C , comme sommets d'une ligne inscrite L_n . Alors chaque aire s_n ne comprend qu'un seul triangle, et la série Σs_n reste presque sûrement convergente.

Pour arriver au résultat énoncé, il n'y a plus qu'à supposer les T_n aléatoires, comme il a été dit au n° 27, 1° ; ils forment presque sûrement une suite partout dense dans l'intervalle d'intégration, et l'on peut appliquer le résultat précédent. Donc, compte tenu de ce que le hasard intervient, d'abord dans le choix des T_n , ensuite dans celui de la courbe C , la somme Σs_n converge presque sûrement vers une limite qui ne dépend que de C . D'après le théorème de Fubini le résultat subsiste si l'on intervertit l'ordre des choix. Cela signifie que la courbe C , choisie d'abord, a presque sûrement une propriété, qui est précisément l'existence de l'intégrale stochastique, telle qu'elle a été définie au n° 27.

On peut démontrer que le résultat subsiste si l'on définit l'intégrale stochastique par la convergence en moyenne quadratique de la suite des S_n . La courbe C a presque sûrement la propriété que

$$(22) \quad \lim_{n \rightarrow \infty} E[(S - S_n)^2] = 0.$$

29. Étude de la loi dont dépend S . — 1° En décomposant l'aire $S(t + dt)$ en trois aires partielles séparées par les cordes $OA(t)$ et $A(t)A(t + dt)$, on a rigoureusement

$$S(t + dt) - S(t) = \frac{1}{2} \xi_t R(t) \sqrt{dt} + S_1 dt,$$

S_1 dépendant de la même loi que $S(1)$; ξ_t est toujours une variable gaussienne réduite.

Le dernier terme peut être négligé dans l'intégration, et comme

en le négligeant on remplace l'aire étudiée par une aire limitée à une ligne polygonale inscrite, l'intégrale

$$(23) \quad S = S(1) = \frac{1}{2} \int_0^1 R(t) \xi_t \sqrt{dt}$$

ainsi obtenue est bien l'intégrale stochastique considérée au n° 27.

On peut, pour déterminer l'arc $(0, 1)$, choisir d'abord $R(t)$, et ensuite les déplacements transversaux $\xi_t \sqrt{dt}$. Après le premier choix, S est en distribution la limite d'une somme riemannienne dont les termes, toujours indépendants, très petits, et à valeurs probables nulles, ne sont plus exactement gaussiens et réduits. Mais ils le sont d'autant plus exactement que les dt sont plus petits, et à la limite, S apparaît comme une variable gaussienne d'écart type $\frac{1}{2} \sqrt{J}$, en posant

$$(24) \quad \begin{cases} J_1 = \int_0^1 X^2(t) dt, & J_2 = \int_0^1 Y^2(t) dt, \\ J = J_1 + J_2 = \int_0^1 R^2(t) dt. \end{cases}$$

On a donc

$$(25) \quad S = \frac{\xi}{2} \sqrt{J},$$

ξ et J étant indépendants ⁽¹⁾.

2° Utilisons maintenant la formule de R. H. Cameron et W. T. Martin (n° 20, 4°). Elle définit la loi dont dépend J_1 par sa fonction caractéristique

$$(26) \quad E(e^{izJ_1}) = \frac{1}{\sqrt{\cos \sqrt{2} iz}}.$$

Comme J_1 et J_2 sont indépendants, on en déduit

$$(27) \quad E(e^{izJ}) = \frac{1}{\cos \sqrt{2} iz}.$$

Posons

$$(28) \quad \frac{1}{\cos z} = \sum_0^\infty c_n \frac{z^{2n}}{(2n)!}.$$

⁽¹⁾ Cette formule est déjà dans notre Mémoire de 1940. L'application de la formule de Cameron et Martin est indiquée dans notre communication à Berkeley [9], et dans un Mémoire plus développé [12].

Il vient

$$E(e^{izJ}) = \sum_0^{\infty} c_n \frac{(iz)^n}{(2n)!},$$

et par suite

$$E(J^n) = \frac{2^n n!}{(2n)!} c_n,$$

On sait d'ailleurs que

$$E(\xi^{2n}) = 1.3 \dots (2n-1) = \frac{(2n)!}{2^n n!},$$

et il résulte de la formule (25) que

$$E[(2S)^{2n}] = c_n,$$

et, les moments d'ordres impairs étant nuls, il vient

$$(29) \quad E(e^{izS}) = \sum_0^{\infty} (-1)^n c_n \frac{z^{2n}}{(2n)!} = \frac{1}{\operatorname{ch} z}.$$

La loi dont dépend $2S$ est ainsi bien définie par sa fonction caractéristique. C'est une loi connue, évidemment absolument continue ⁽¹⁾.

La densité de probabilité correspondante est $\frac{1}{2\operatorname{ch} \frac{\pi x}{2}}$. Pour la variable S

elle-même, la fonction caractéristique et la densité de probabilité sont donc

$$(30) \quad \frac{1}{\operatorname{ch} \frac{z}{2}} \quad \text{et} \quad \frac{1}{\operatorname{ch} \pi x}.$$

30. Application de la série de Fourier-Wiener. — 1° En nous plaçant maintenant pour simplifier les formules dans l'intervalle $(0, 2\pi)$, représentons $X(t)$ et $Y(t)$ par leurs séries de Fourier-Wiener

$$X(t) = \frac{\xi't}{\sqrt{2\pi}} + \sum_1^{\infty} \frac{1}{n\sqrt{\pi}} [\xi_n(\cos nt - 1) + \xi'_n \sin nt],$$

$$Y(t) = \frac{\eta't}{\sqrt{2\pi}} + \sum_1^{\infty} \frac{1}{n\sqrt{\pi}} [\eta_n(\cos nt - 1) + \eta'_n \sin nt].$$

⁽¹⁾ Elle a été considérée depuis longtemps, à titre d'exemple, par Hausdorff, et est mentionnée en 1934 dans la table de transformées de Fourier, de Campbell et Foster. Mais le présent problème semble être le premier où elle s'introduise naturellement. Pour le calcul de la densité de probabilité, nous renvoyons à notre Mémoire cité, ou à la thèse de G. Kuntz (*Thèse d'Université*, Paris, 1937).

Un simple calcul formel donne alors

$$(31) \quad S(2\pi) = \sum_1^{\infty} \frac{1}{n} [\xi_n(\eta'_n - \eta' \sqrt{2}) - \eta_n(\xi'_n - \xi' \sqrt{2})].$$

Cette série, étant presque sûrement convergente, peut être prise comme définition de $S(2\pi)$. On est ainsi conduit à une définition des intégrales stochastiques différente de celle du n° 27, et qui se ramène aisément à celle de Paley et Wiener. Au lieu d'approcher de la courbe C par une ligne polygonale, on approche de $f(t)$ et $g(t)$ par des polynômes trigonométriques. Nous avons démontré ailleurs que, dans le cas particulier qui nous intéresse ici, cette définition équivaut à l'autre. Nous nous contenterons de montrer qu'elle conduit à la même loi de probabilité pour $S(2\pi) = 2\pi S$, donc pour S .

2° Nous supposons d'abord ξ' et η' connus et désignerons par E' les valeurs probables calculées dans ces conditions. La série (31) étant alors une série à termes indépendants, il est facile de calculer sa fonction caractéristique en partant de la formule

$$\begin{aligned} E[e^{i\eta(\xi-h)z}] &= E\left[e^{-\frac{(\xi-h)^2 z^2}{2}}\right] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}[(x-h)^2 z^2 + x^2]} dx = \frac{1}{\sqrt{1+z^2}} e^{-\frac{h^2 z^2}{2(1+z^2)}}. \end{aligned}$$

On en déduit

$$E'[e^{iz[\xi_n(\eta'_n - \eta' \sqrt{2}) - \eta_n(\xi'_n - \xi' \sqrt{2})]}] = \frac{1}{1+z^2} e^{-\frac{\rho^2 z^2}{1+z^2}} \quad (\rho^2 = \xi'^2 + \eta'^2),$$

et, en remplaçant z par $\frac{z}{\pi n}$ et faisant le produit des formules ainsi écrites

$$(32) \quad E'(e^{2izS}) = \Phi(z, \rho) = \frac{z}{\operatorname{sh} z} e^{\frac{\rho^2}{2} \left[1 - \frac{z \operatorname{ch} z}{\operatorname{sh} z}\right]}.$$

Cette formule définit la loi dont dépend $S(t) = St$ dans l'hypothèse $R(t) = \rho \sqrt{t}$ ⁽¹⁾.

⁽¹⁾ Elle définit donc aussi la loi à deux variables $R(t)$ et $S(t)$. Dans nos premiers travaux sur cette question ([6] et [3], p. 266), nous n'avions réussi qu'à ramener la recherche de cette loi à l'intégration d'une équation aux dérivées partielles du type

Pour obtenir la loi inconditionnelle de S , il n'y a qu'à observer que

$$E(e^{2izS}) = E\{\Phi[z, R(1)]\} = \int_0^\infty \Phi(z, \rho) e^{-\frac{\rho^2}{2}} d\frac{\rho^2}{2},$$

et l'on retrouve bien l'expression $\frac{1}{\operatorname{ch} z}$ obtenue au n° 28.

3° Revenons à la formule (32). Si $\rho = 0$, le second membre se réduit à $\frac{z}{\operatorname{sh} z}$. Donc, si l'on sait que la boucle $(0, t)$ est fermée, l'aire $S(t) = St$ qu'elle entoure dépend aussi d'une loi très simple ; S a alors pour fonction caractéristique $\frac{z}{2 \operatorname{sh} \frac{z}{2}}$ et pour densité de probabilité $\frac{\pi}{2 \operatorname{ch}^2 \pi x}$.

CHAPITRE IV.

LE MOUVEMENT BROWNIEN A PLUSIEURS PARAMÈTRES ⁽¹⁾.

31. Notions générales. — 1° Il s'agira dans ce chapitre de l'étude d'une fonction aléatoire $X(A)$ d'un point de l'espace euclidien E_N à N dimensions, définie, à une constante près, par la formule

$$(1) \quad X(B) - X(A) = \xi \sqrt{r(A, B)},$$

où $r(A, B)$ désigne la distance AB . Nous déterminerons la constante additive par la condition $X(O) = 0$; il suffira de remplacer $X(M)$ par $X(M) - X(O)$ pour obtenir des formules indépendantes de cette hypothèse.

La fonction $X(M)$ est ainsi une fonction aléatoire gaussienne, bien définie par sa covariance

$$(2) \quad E[X(A)X(B)] = \frac{1}{2}[r(O, A) + r(O, B) - r(A, B)].$$

elliptique. On voit maintenant que, pour $t = 1$, la densité de probabilité, solution de cette équation, est

$$\frac{\rho}{\pi} e^{-\frac{\rho^2}{2}} \int_0^\infty \cos sz \Phi(z, \rho) dz.$$

⁽¹⁾ D'après P. Lévy [8] et [3], chap. VIII. Il n'existe, à notre connaissance, aucun autre travail que les nôtres sur le mouvement brownien à plusieurs paramètres, qui nous semble cependant constituer un exemple remarquable de fonction aléatoire de plusieurs variables.

Or il est bien connu qu'une fonction quelconque $\Gamma(A, B)$ n'est pas une covariance. Dans le cas réel, la condition nécessaire et suffisante pour qu'elle le soit (et dans ce cas elle sera en particulier la covariance d'une fonction aléatoire gaussienne), est qu'elle soit symétrique et que, quel que soit l'entier n et quels que soient les points $A_1, A_2, \dots, A_n$, la forme quadratique

$$\sum_{h=1}^n \sum_{k=1}^n \Gamma(A_h, A_k) x_h x_k$$

soit non négative. Il s'agit donc d'étudier la forme

$$(3) \quad \sum_{h=1}^n \sum_{k=1}^n [r(O, A_h) + r(O, A_k) - r(A_h, A_k)] x_h x_k,$$

Or, d'après un théorème de I. J. Schönberg ⁽²⁾, cette forme est définie positive. Il n'y a donc pas d'incompatibilité au point de vue de Bernoulli. Si $\{A_n\}$ est une suite de points partout dense dans E_N , on peut déterminer successivement tous les $X(A_n)$, chacun d'eux étant une fonction linéaire des précédents, et d'une nouvelle variable gaussienne réduite ξ_n .

2° On peut donner une forme géométrique au résultat précédent. On sait que des variables aléatoires $X_0, X_1, \dots, X_n$ à variances finies peuvent toujours être représentées dans E_n ou dans l'espace de Hilbert Ω par $n+1$ points $a_0, a_1, \dots, a_n$, de manière que

$$(4) \quad E[(X_h - X_k)^2] = r^0(a_h, a_k),$$

et que réciproquement, à n points donnés, on peut toujours faire correspondre des variables aléatoires, qu'on peut en particulier supposer gaussiennes, de manière que la relation (4) soit vérifiée.

Par suite, le résultat obtenu au 1° est exactement équivalent au suivant : étant donnés $n+1$ points $A_0, A_1, \dots, A_n$, dans E_n (ou dans Ω), on peut toujours leur faire correspondre $n+1$ points $a_0, a_1, \dots, a_n$ de manière que

$$(5) \quad r^0(a_h, a_k) = r(A_h, A_k) \quad (h, k = 0, 1, \dots, n).$$

Il importe peu que les points $A_0, A_1, \dots, A_n$ soient, ou non, dans une même variété linéaire à moins de n dimensions ; si ces points

(2) I. J. Schönberg [1] et P. Lévy [3], p. 276.

sont distincts, $\alpha_0, \alpha_1, \dots, \alpha_n$ ne seront pas dans une telle variété, c'est-à-dire qu'il n'y a pas de relation certaine entre les $X(A_v)$.

3° Les $X(A_n)$ une fois obtenus pour les points d'un ensemble dénombrable et partout dense, il reste à démontrer qu'ils définissent presque sûrement une fonction continue. Contentons-nous de dire qu'on y arrive par la même méthode que dans le cas où $N = 1$ (n° 5). On arrive aussi à une condition de Lipschitz faible, qui s'exprime par la formule, *valable dans toute région finie*,

$$(6) \quad \Pr \left\{ \lim_{\rho \rightarrow 0} \limsup_{r < \rho} \frac{[X(B) - X(A)]'}{2Nr \log 1/r} = 1 \right\} = 1 \quad r = r(A, B) \quad (1).$$

32. Deux formules simples. Le cas de l'espace de Hilbert. —

1° Prenons pour A_n le point d'ordonnée $+1$ sur l'axe des x_n , et posons $X_n = X(A_n)$. Si $X(O) = 0$, on a

$$E \left[\left(\frac{X_1 + X_2 + \dots + X_n}{n} \right)^2 \right] = \frac{1}{n} E(X_1^2) + \frac{n-1}{n} E(X_1 X_n) = 1 - \frac{n-1}{n\sqrt{2}}.$$

On en déduit qu'inversement, si $X_1, X_2, \dots, X_n$ sont connus, et $X(O)$ inconnu, on a

$$(7) \quad X(O) = \mu_n + \sigma_n \xi_n,$$

formule où

$$(8) \quad \mu_n = \frac{X_1 + X_2 + \dots + X_n}{n}, \quad \sigma_n^2 = \frac{\sqrt{2}-1}{\sqrt{2}} + \frac{1}{n\sqrt{2}}.$$

On remarque que, pour n infini, σ_n tend *en décroissant* vers une limite *positive* σ (le premier point était évident *a priori*, mais non le second).

Par un calcul presque aussi simple, en supposant toujours $X_1, X_2, \dots, X_n$ connus, on arrive à

$$(9) \quad X_{n+1} = \mu_n + \sigma_n \xi'_n \quad \left(\bar{\sigma}_n^2 = \frac{1}{\sqrt{2}} + \frac{1}{n\sqrt{2}} \right).$$

On obtient ainsi, pour n infini, une limite de $\bar{\sigma}_n$ supérieure à celle de σ_n . Cela était à prévoir. Si H_n est le centre de gravité des points $A_1, A_2, \dots, A_n$, les directions OH_n et $H_n A_{n+1}$ sont perpendiculaires

(1) La loi du logarithme itéré se généralise de même en remplaçant r par Nr . Cela permet de prévoir un fait sur lequel nous reviendrons plus loin : pour N infini, c'est-à-dire dans le cas de l'espace de Hilbert, $X(A)$ n'est pas continu.

au plan P_n de ces n points; mais OH_n tend vers zéro, tandis que $H_n A_{n+1}$ tend vers un. La part du hasard augmente quand le point considéré s'éloigne de P_n dans une direction perpendiculaire.

Le fait qu'elle ne tende pas vers zéro entraîne une conséquence importante. Si l'on fait augmenter N indéfiniment, E_N devenant donc l'espace de Hilbert, et qu'on détermine successivement tous les X_n par la formule (9) [en partant de $X_1 = 0$; si l'on se donne $X(O) = 0$, il faut remplacer cette formule par une autre analogue], la suite des ξ_n n'est presque sûrement pas bornée; celle des X_n ne l'est donc pas non plus. Comme ce résultat subsiste si l'on remplace la figure considérée par une figure homothétique arbitrairement petite, il est presque sûr qu'au voisinage de n'importe quel point, $X(A)$ n'est ni continu, ni même borné.

2° La remarque qui précède conduit à se demander s'il est possible de définir $X(A)$ dans l'espace de Hilbert⁽¹⁾. La réponse est affirmative. Cela résulte de ce qu'on peut définir une suite de points B_n qui soit partout dense dans cet espace. Si l'on désigne par $F_n(A)$ la valeur probable conditionnelle de $X(A)$, quand on connaît $X(B_1)$, $X(B_2)$, ..., $X(B_n)$, on démontre aisément que, pour chaque point A , la suite des $F_n(A)$ est convergente en moyenne quadratique. Les différences $F_{n+1}(A) - F_n(A)$ étant indépendantes, la convergence presque sûre en résulte, et $X(A)$ est la limite ainsi obtenue, c'est-à-dire qu'il n'y a pas lieu de distinguer $X(A)$ de sa valeur probable quand on connaît tous les $X(B_n)$. Bien entendu, la convergence n'est pas uniforme, puisque les $F_n(A)$ sont continus, mais non $X(A)$.

Précisons bien qu'il y a peut-être des points où la suite des $F_n(A)$ n'a pas de limite. La seule chose qui résulte du raisonnement précédent est que cette circonstance n'a aucune chance d'être réalisée en un point donné d'avance.

33. La droite ou le plan et un point extérieur. — 1° Commençons par rappeler une propriété connue des systèmes gaussiens de variables aléatoires X_ν ($\nu = 0, 1, \dots$): si l'on se donne $X_1, X_2, \dots, X_n$,

(1) Rappelons que c'est A qui décrit l'espace de Hilbert, et non X . Il est naturellement possible remplacer X par un vecteur à une infinité de dimensions. Mais ce ne serait pas un vecteur de l'espace de Hilbert; ses composantes ne seraient même pas bornées dans leur ensemble.

on a pour X_0 une expression de la forme (7), $E(\mu_n)^2$ ne pouvant que croître, et σ_n^2 que décroître, quand n croît; chaque nouvelle information fait décroître σ_n , si elle consiste dans la donnée d'une variable qui n'était pas, immédiatement avant, orthogonale à X_0 .

Supposons alors $X(A)$ connu dans le plan des $x_1 x_2 \dots x_n$ (plan P_n), et prenons pour A_ν ($\nu = 0$ ou $\nu > n$) le point $x_\nu = 1$ sur l'axe des x_ν . En posant $X(A_\nu) = X_\nu$, on a

$$(10) \quad X_0 = \mu_n + \sigma'_n \xi_n, \quad X_{n+1} = \mu_n + \sigma'_n \eta_n,$$

et de même, au point A'_0 symétrique de A_0 par rapport à O ,

$$(10') \quad X'_0 = \mu_n + \sigma'_n \xi'_n,$$

avec $\sigma'_n < \sigma_n$, et aussi, d'après le même principe général, $\sigma'_{n+1} < \sigma'_n$; donc σ'_n tend en décroissant vers une limite $\sigma' \leq \sigma$ [σ_n^2 ayant la valeur (8) et $\sigma^2 = 1 - \frac{1}{\sqrt{2}}$].

Il est d'ailleurs facile de borner inférieurement σ'_n . On a en effet

$$(11) \quad E[(X'_0 - X_0)^2] = 2 = \sigma_n'^2 E[(\xi'_n - \xi_n)^2] \leq 4 \sigma_n'^2,$$

et, par suite,

$$(12) \quad \sigma_n'^2 > \sigma'^2 \geq \frac{1}{2}.$$

2° D'autre part, le coefficient de corrélation ρ_n de X_0 et X_ν (ou de X_ν et $X_{\nu'}$, $\nu' > \nu > n$), déduit des formules (10), donc de

$$E[(X_\nu - X_0)^2] = \sqrt{2} = \sigma_n'^2 E[(\eta_n - \xi_n)^2] = 2 \sigma_n'^2 (1 - \rho_n)$$

a la valeur

$$(13) \quad \rho_n = 1 - \frac{1}{\sigma_n'^2 \sqrt{2}}.$$

Or ρ_n est non négatif. En effet, $X_{n+1}, X_{n+2}, \dots, X_{n+p}$ ont deux à deux le même coefficient de corrélation ρ_n ; en désignant par $\eta_{n,1}, \eta_{n,2}, \dots, \eta_{n,p}$ les variables réduites qui leur correspondent par la formule (10), on a

$$E[\eta_{n,1} + \eta_{n,2} + \dots + \eta_{n,p}]^2 = p + p(p-1)\rho_n \geq 0,$$

et, comme p peut être pris arbitrairement grand, $\rho_n \geq 0$, d'où

$$(14) \quad \sigma_n'^2 \geq \frac{1}{\sqrt{2}}.$$

Cette borne inférieure coïncide d'ailleurs avec la limite σ'^2 de $\sigma_n'^2$. Cela résulte de ce que l'hypothèse $\rho_n > \varepsilon$ permet de borner inférieurement la valeur, pour la détermination de X_0 , des renseignements supplémentaires fournis par X_{n+1} , et *a fortiori* celle des renseignements obtenus lorsqu'on se donne $X(A)$ dans tout le plan P_{n+1} . Compte tenu de la formule (14), il en résulte une borne inférieure positive pour $\sigma_n'^2 - \sigma_{n+1}'^2$; $\sigma_n'^2$ ayant une limite, l'hypothèse $\rho_n > \varepsilon$ ne peut être vérifiée qu'un nombre fini de fois, et ρ_n tend vers zéro ⁽¹⁾; d'après (13), $\sigma_n'^2$ tend donc vers $\sigma'^2 = \frac{1}{\sqrt{2}}$.

Rappelons d'ailleurs que, d'après notre livre de 1948 (p. 289), $\sigma_1'^2 = \frac{\pi}{4} = 0,785 \dots$; comme $\sigma'^2 = 0,707 \dots$, on voit que la décroissance de σ_n' est singulièrement faible. Une fois que $X(A)$ est connu sur l'axe des x_1 , les renseignements supplémentaires qui résultent pour la connaissance de X_0 de la connaissance de $X(A)$ en tous les points du plan $x_0 = 0$ n'ajoutent pas grand chose (quelque grand que soit N , et même s'il s'agit de l'espace de Hilbert).

De même ρ_n décroît en partant d'une valeur initiale déjà très petite.

3° Il résulte de la formule (11), et de la formule évidente

$$\sigma_n'^2 = 1 - E(\mu_n^2) < 1,$$

que $E(\xi_n \xi_n') < 0$. Donc : *si $X(A)$ est connu sur un plan, il y a une corrélation négative entre les valeurs de cette fonction en deux points symétriques par rapport à ce plan.* Dans cet énoncé, s'il s'agit de l'espace E_N , *plan* signifie « variété linéaire à au plus $N - 1$ dimensions ».

En particulier, s'il s'agit d'un plan à $N - 1$ dimensions, cet énoncé implique que : *si $N > 1$, le mouvement brownien à N paramètres n'est pas un processus markovien.*

Ce résultat assez surprenant nous a conduit à nous demander s'il existait des processus markoviens dans un espace E_N ($N > 1$), sauf

⁽¹⁾ Si, sur l'intersection du plan $x_1 = x_2 = \dots = x_n = 0$ et de la sphère $\Sigma x_i^2 = 1$, nous marquons deux points A et B , le coefficient de corrélation de $X(A)$ et $X(B)$ [$X(M)$ étant connu dans P_n] décroît d'une manière continue quand l'angle AOB croît de 0 à π . Pour n infini, la valeur de cet angle pour laquelle il s'annule tend en décroissant vers $\frac{\pi}{2}$.

dans des cas triviaux comme celui où $X(A)$ ne dépendrait que d'une des coordonnées. J. Kampé de Fériet, puis nous-même, avons répondu par l'affirmative, mais en donnant des exemples pour lesquels il y avait toujours ce que nous appelions des caractères de dégénérescence ; par exemple, dans le plan, l'existence de lignes de discontinuité. Récemment, Ottaviani [1] a défini, dans le cas du plan, un processus markovien n'ayant aucun caractère de dégénérescence. Mais sa méthode repose sur le fait que, dans le plan, la fermeture de la courbe du mouvement brownien est partout dense. Cette propriété ne s'étend pas au cas d'un espace à plus de deux dimensions. La question reste donc posée pour $N > 2$: *existe-t-il un processus markovien n'ayant aucun caractère de dégénérescence?* ⁽¹⁾

34. La sphère et son centre. — Si μ désigne la moyenne de $X(A)$ sur la surface Σ de la sphère de rayon unité et de centre O , un calcul facile donne

$$(15) \quad E \{ [\mu - X(O)]^2 \} = 1 - \frac{1}{2} \rho_N,$$

ρ_N désignant la distance moyenne d'un point M de Σ à un point fixe A de cette surface. Si donc $X(A)$ est connu sur Σ , μ est évidemment la valeur probable de $X(O)$, et il vient

$$(16) \quad X(O) = \mu_N + \sigma_N \xi \quad \left(\sigma_N^2 = 1 - \frac{1}{2} \rho_N \right).$$

Ici encore, il était évident que, l'augmentation de N équivalant à un accroissement de l'information, σ_N ne peut que décroître quand N croît ; on le vérifie aisément d'après la formule (15). Il est évident aussi que σ_n est inférieur aux nombres σ_n , σ'_n et $\bar{\sigma}_n$ des nos 32 et 33. Ce qu'on pouvait en conséquence se demander, c'est si la limite σ de σ_n , pour n infini, est positive ou nulle. Or un théorème connu de E. Borel apprend que la distance x de M au plan diamétral perpendiculaire tend en mesure vers zéro ; c'est une conséquence

(1) Nous ne proposons pas de définition précise de ce que nous appelons un caractère de dégénérescence. Il s'agit d'un de ces problèmes dont parle Poincaré, qu'on ne peut bien poser que quand on les a résolus. Ici, c'est seulement dans l'hypothèse d'une réponse négative à notre question qu'on serait conduit à une définition précise.

immédiate de ce que la moyenne de x^2 est $\frac{1}{p}$. Il en résulte immédiatement que la distance AM tend en mesure vers $\sqrt{2}$. Donc

$$\sigma^2 = 1 - \frac{1}{\sqrt{2}} > 0 \quad (1).$$

BIBLIOGRAPHIE.

BACHELIER (L.) :

- [1] Trois Mémoires, publiés dans les *Ann. Éc. norm. sup.*, t. 17, 1900, p. 21 86; t. 18, 1901, p. 143-210; t. 30, 1913, p. 77 119.
- [2] *Calcul des probabilités* (Paris, Gauthier-Villars, 1912, p. 1-516).

BERNSTEIN (S.) :

- [1] *Principes de la théorie des équations différentielles stochastiques* (*Trav. math. Inst. phys. math. Stekloff*, t. 5, 1933, p. 95-124).

CAMERON (R. H.) et MARTIN (W. T.) :

- [1] *The Wiener measure of Hilbert's neighbourhood in the space of real continuous functions* (*J. Math. Phys.*, Massachusett's Inst. of Technology, t. 23, 1944, p. 195-209).

CHUNG (K. L.) :

- [1] *On the maximum partial sums of sequences of independent random variables* (*Trans. Amer. Math. Soc.*, t. 64, 1948, p. 205-233).

CHUNG (K. L.) et FELLER (W.) :

- [1] *On fluctuations in coin tossing* (*Proc. Nat. Acad. Sc.*, t. 35, n° 10, 1949, p. 605-610).

DVORETSKY (A.) et ERDÖS (P.) :

- [1] *Some problems on random walk in space* (*Proc. of the Second Symposium on math. Statistics and Probability*, Berkeley, 1950, p. 353-368).

(¹) Cette formule constitue une rectification de celle que, par suite d'une erreur de calcul, j'avais indiquée autrefois (P. Lévy [3], p. 290, formule (38)). Les conséquences que j'en avais déduites aux p. 295 297 sont évidemment aussi inexactes. Malgré le caractère discontinu de la fonction $X(A)$ dans l'espace de Hilbert, la dégénérescence dont j'avais parlé n'existe pas. La part du hasard subsiste en un point A_0 tant qu'on ne sait rien sur les valeurs de $X(A)$ au voisinage immédiat de ce point.

DVORETSKY (A.), ERDOS (P.) et KAKUTANI (S.) .

- [1] *Double points of paths of Brownian motion in n -space* (*Acta Scientiarum mathematicarum*, Szeged, t. 12, 1950, p. 75-81).

FELLER (W.) :

- [1] *The general form of the so-called law of the iterated logarithm* (*Trans. Amer. Math. Soc.*, t. 54, 1943, p. 373-402).
 [2] *The law of the iterated logarithm for identically distributed random variables* (*Ann. of Math.*, t. 47, 1946, p. 631-638).

FORTET (R.) .

- [1] *Quelques travaux récents sur le mouvement brownien* (*Ann. Inst. Henri Poincaré*, t. 11, 1947, p. 438-442).

ITÔ (K.)

- [1] *Stochastic integral et On a stochastic integral equation* (*Proc. Japan Acad.*, t. 20, 1944, p. 519-522, t. 22, 1946, p. 32-35).
 [2] *Stochastic differential equations in a differentiable manifold* (*Nagoya Math. J.*, t. 1, 1950, p. 35-47, voir aussi t. 3, 1951, p. 55-66).
 [3] *On stochastic differential equations* (*Memoirs of the Amer. Math. Soc.*, t. 4, 1951, p. 1-51).

KAC (M.) .

- [1] *Random walk and the theory of brownian motion* (*Amer. Math. Monthly*, t. 54, 1947, p. 369-391).
 [2] *On some connections between probability theory and differential and integral equations* (*Symposium on Math. Statistics and Probability*, Berkeley, 1950, p. 189-215).

KAC (M.) et SIEGERT (A. J. F.) .

- [1] *An explicit representation of a stationary gaussian process* (*Ann. of Math. Statistics*, t. 18, 1947, p. 438-442).

KAKUTANI (S.) :

- [1] *Two dimensional Brownian motion and harmonic functions* (*Proc. Math. Inst.*, Osaka, t. 20, 1944, p. 700-708).
 [2] *Two dimensional Brownian motion and the type problem of Riemann surfaces* (*Proc. Math. Inst.*, Osaka, t. 21, 1945, p. 138-140).

KHINTCHINE (A.) :

- [1] *Ein Satz der Wahrscheinlichkeitsrechnung* (*Fundam. Math.*, t. 6, 1924, p. 9-20).
 [2] *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung* (*Ergebn. der Math. und ihrer Grenzgebiete*, Springer, t. 2, 4, 1933, p. 1-77).

KOLMOGOROFF (A.) :

- [1] *Ueber die analytischen Methoden in der Wahrscheinlichkeitsrechnung* (*Ergebn. der Math. und ihrer Grenzgebiete*, Springer t. 2, 3, 1933, p. 1-62).

LÉVY (P.)

- [1] *Analyse fonctionnelle* (*Mém. Sc. Math.*, fasc. 5, 2^e éd., 1951, 58 p.).
 [2] *Théorie de l'addition des variables aléatoires* (Gauthier Villars, 1^{re} éd., 1937, xvii, 329 p., 2^e éd., 1954, 387 p.)
 [3] *Processus stochastiques et mouvement brownien* (Gauthier-Villars, 1948, 365 p.).
 [4] *Sur un théorème de M. Khintchine* (*Bull. Sc. math.*, t. 55, 1931, p. 145-160),
 [5] *Sur certains processus stochastiques homogènes* (*Compositio math.*, t. 7, 1939, p. 283-339).
 [6] *Le mouvement brownien plan* (*Amer. J. Math.*, t. 62, 1940, p. 487-550).
 [7] *Un théorème d'invariance projective relatif au mouvement brownien* (*Comm. math. helv.*, t. 16, 1943, p. 242-248).
 [8] *Trois théorèmes sur le mouvement brownien* (*Congrès de l'Assoc. franç. pour l'Avanc. des Sciences*, Paris, oct. 1945, et *Intermédiaire des recherches math.*, suppl. au fasc., janv. 1947, p. 124-126).
 [9] *Wiener's random function and other laplacian random function* (*Proc. of the second Berkeley Symposium on math. Statistics and Probability*, Berkeley, août 1950, p. 171-187).
 [10] *Problèmes concrets d'analyse fonctionnelle* (Gauthier Villars, 1951, 484 p.).
 [11] *Intégrales de Stieltjes généralisées* (*Ann. Soc. polonaise de math.*, t. 25, 1952 p. 17-26).
 [12] *Random functions : general theory with special reference to laplacian random functions* (*Univ. of California public in Statistics*, t. 1, 1953, p. 331-390).
 [13] *La mesure de Hausdorff de la courbe du mouvement brownien* (*Giornale dell'Istituto ital. d. attuari*, t. 16, 1953, p. 1-37).

OTTAVIANI (G.) :

- [1] *Sulle catene doppie di Markoff* (*Giornale dell'Istituto ital. d. attuari*, t. 14, 1951, p. 7-15).

PALEY (R. E. A. C.) et WIENER (N.) :

- [1] *Fourier transforms in the complex domain* (New York, 1934, 184 p.; notamment chap. IX, p. 140-162).

PERRIN (F.)

- [1] *Étude mathématique du mouvement brownien de rotation* (*Ann. Ec. norm. sup.*, 3^e série, t. 45, 1928, p. 1-51).

PETROWSKY (I.) :

- [1] *Ueber das Irrfahrtproblem* (*Math. Ann.*, t. 109, 1933-1934, p. 425-434).

- [2] *Zur ersten Randwertaufgabe der Wärmeleitungsgleichung* (*Compositio math.*, t. 1, 1935, p. 383-419).

PÓLYA (G) .

- [1] *Ueber eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Strassennetz* (*Math. Ann.*, t. 84, 1921, p. 149-160).
[2] *Sur la promenade au hasard dans un réseau de rues* (*Actualités scientifiques, Hermann*, n° 734, 1938, p. 25-44).

SCHONBERG (I. J.) .

- [1] *Metric spaces and positive definite functions* (*Trans. Amer. Math. Soc.*, t. 44, 1938, p. 522-536).

WIENER (N.) .

- [1] *Differential space* (*J. Math. Phys.*, Massachusetts Inst. of Technology, t. 2, 1923, p. 131-174)
[2] *Sur un problème de probabilités dénombrables* (*Bull. Soc. math. Fr.* t. 52, 1924, p. 569-578).

YOSIDA (K.) .

- [1] *Brownian motion on the surface of the 3 sphere* (*Ann. of math. Statistics*, t. 20, 2, 1949).



TABLE DES MATIÈRES

CHAPITRE I

Le mouvement brownien linéaire.

Définition, continuité, représentation par une série de Fourier.

1. Définitions.	2
2. Historique	3
3. Formules d'interpolation	4
4 Remarques	6
5. Définition constructive de $X(t)$	8
6 La condition de Lipschitz faible.	9
7 Propriétés asymptotiques de $X(t)$	12
8. Étude locale de $X(t)$	15
9. La variation totale presque sûre d'ordre deux	18
10 Développement de $X(t)$ en série de Fourier.	20

CHAPITRE II.

Étude approfondie du mouvement brownien linéaire.

11 L'équation de la diffusion	22
12. Conditions aux limites.	24
13 Cas des limites constantes.	27
14. Les fonctions $M(t)$ et $Y(t)$	29
15 L'inversion de $M(t)$	31
16 Identité des définitions intrinsèques de $Y(t)$ et de $ X(t) $	34
17. La loi de l'arc sinus.	36
18. La reconstruction de $X(t)$ en partant de $\mathcal{E}$ ou de $\mathcal{E}_0$	40
19 Le passage du fini à l'infini.	43
20. Fonctionnelles de $X(t)$	46

CHAPITRE III

Le mouvement brownien dans le plan et dans l'espace.

21. Notions générales.	50
22. Propriétés locales et étude des points exceptionnels	51

23. Loi du logarithme itéré et condition de Lipschitz.....	55
24. La mesure superficielle de la courbe C.....	56
25. La fermeture de la courbe C.....	57
26. L'équation de la diffusion.....	60
27. Intégrales stochastiques.....	63
28. Application à la courbe du mouvement brownien plan.....	66
29. Étude de la loi dont dépend S.....	67
30. Application de la série de Fourier Wiener.....	69

CHAPITRE IV.

Le mouvement brownien à plusieurs paramètres.

31. Notions générales.....	71
32. Deux formules simples. Le cas de l'espace de Hilbert.....	73
33. La droite ou le plan et un point extérieur.....	74
34. La sphère et son centre.....	77
BIBLIOGRAPHIE.....	78

