

Chaînes de Markof
dans les ensembles abstraits
et applications aux processus avec régions
absorbantes et au problème des boucles

par

D. van DANTZIG.

PRÉFACE.

Au moyen de fonctions génératrices généralisées (fonctionnelles) (§7) et de matrices généralisées (§2, 7), on introduit un opérateur linéaire $C_{(A)}$ dans un espace de fonctions mesurables sur un ensemble donné E (contenant, par exemple, toutes les fonctions bornées), dépendant du sous-ensemble A où a lieu « l'absorption » (§3). Une interprétation (§1) donnée antérieurement des variables et des fonctions auxiliaires comme probabilités, par exemple, de la non-apparition d'une certaine « catastrophe » s'avère aussi très utile pour ce type de problèmes. On démontre (§4) que $C_{(A)}$ est un opérateur de projection et que pour $A_1 \supset A_2$, $C_{(A_1)}$ et $C_{(A_2)}$ sont commutables et que leur produit est égal à $C_{(A_2)}$. Cette identité contient comme cas particulier une identité obtenue dans une Thèse de doctorat à Amsterdam par J. H. B. Kemperman. En outre (§5), l'identité fondamentale d'A. Wald est généralisée et l'on montre que la fonction caractéristique d'une distribution peut être considérée comme un cas particulier de valeur propre dans un sens généralisé, où la fonction propre (dans ce cas particulier une fonction exponentielle) doit être proportionnelle à sa transformée par l'opérateur linéaire en dehors de l'ensemble absorbant A seulement. Finalement (§6), on démontre quelques théorèmes relatifs aux boucles dans la trajectoire d'un point mobile et quelques résultats similaires en dérivant la fonctionnelle génératrice par rapport à la fonction dont elle

dépend. La théorie purement mathématique de la multiplication de matrices généralisées est développée dans l'Appendice (§ 7-10). En particulier, on donne des conditions suffisantes pour l'associativité également dans des cas où les matrices ne sont pas bornées (§ 8). L'Appendice contient, en outre, une généralisation d'un processus « à accroissements indépendants » pour le cas d'un ensemble sur lequel, sous certaines conditions, opère un groupe transitif (§ 9) ainsi qu'une extension des résultats du paragraphe 5 à des processus non markoviens (§ 10).

Ce Mémoire a été écrit pendant que l'auteur travaillait en hôte au Statistical Engineering Laboratory du National Bureau of Standards durant l'été 1951, sous le contrat CST-528 entre le National Bureau of Standards et l'Université de la Caroline du Nord. Il fut ensuite révisé avec l'assistance de M. J. J. de Longh, auquel l'auteur est redevable de plusieurs améliorations du texte. Un nombre limité de copies en forme miméographée fut mis en circulation par le National Bureau of Standards à Washington, et, à Amsterdam, par le Centre Mathématique, en l'été 1952 sous le titre : *Time-discrete stochastic processes in arbitrary sets, with applications to processes with absorbing regions and to the problem of loops in Markoff chains*. Depuis lors, le texte a encore été révisé en profitant de l'aide de MM. C. L. Scheffer, R. Doornbos et G. Zoutendijk. Néanmoins, quelques petits détails mis à part, il coïncide avec la version originale et, en ce qui concerne les cinq premiers chapitres, avec l'envoi de l'auteur à la Réunion de l'Institute of Mathematical Statistics tenue à Santa Monica le 15 juin 1951. Le contenu principal de cet article forme d'ailleurs la substance des quatre conférences données par l'auteur à l'Institut Henri Poincaré au mois d'avril 1953. L'auteur tient à remercier MM. Fénon, Fourgeaud, Fuchs et Soulé qui ont eu l'amabilité et la patience de traduire cet article, de même que M. Fréchet, qui a bien voulu surveiller l'œuvre de traduction.

1. La méthode des marques collectives. — Suivant A. Kolmogoroff, un champ de probabilité est défini comme un ensemble Π sur lequel est donnée une fonction d'ensemble complètement additive ⁽¹⁾ [définie,

⁽¹⁾ Les fonctions d'ensemble complètement additives, définies par la condition (1.2), sont quelquefois appelées fonctions d'ensemble totalement additives, absolument additives ou σ -additives.

≥ 0 et ≤ 1 pour tous les sous-ensembles Λ de Π appartenant à un σ -corps σ_Π ⁽²⁾, qui contient Π et chacun de ses éléments]. Désignant cette fonction d'ensemble par P , la valeur qu'elle prend sur un ensemble Λ par $P_{(\Lambda)}$ [au lieu de $P(\Lambda)$], nous avons, en renvoyant le lecteur pour plus de détails concernant les notations et les propriétés à l'Appendice (§ 1),

$$(1.1) \quad 0 \leq P_{(\Lambda)} \leq 1 = P_{(\Pi)},$$

$$(1.2) \quad P_{(\Lambda)} = \sum_1^\infty P_{(\Lambda_n)}, \quad \text{si } \Lambda = \bigcup_1^\infty \Lambda_n \text{ et } \Lambda_m \cap \Lambda_n = 0 \text{ pour } m \neq n.$$

Dans beaucoup de problèmes de la théorie des probabilités, nous avons affaire à plusieurs fonctions d'ensemble complètement additives, toutes définies sur un seul ensemble Π , c'est-à-dire P peut varier sur un autre ensemble Ω , quelquefois appelé l'« ensemble des hypothèses admissibles » ou « espace paramétrique ». Alors, pour tout $\theta \in \Omega$ nous avons un P , désigné par $P^{(\theta)}$, prenant sur Λ la valeur $P_{(\Lambda)}^{(\theta)}$. Si Π est un ensemble dénombrable (c'est-à-dire fini ou infini dénombrable), nous avons $P_{(\Lambda)}^{(\theta)} = \sum_{\lambda \in \Lambda} P_{(\lambda)}^{(\theta)}$ et si Π et Ω sont des ensembles finis, $P_{(\Lambda)}^{(\theta)}$ est

déterminé par la matrice rectangulaire ordinaire des nombres $P_{(\lambda)}^{(\theta)}$, avec $\lambda \in \Pi$. Pour cette raison on appelle matrice (généralisée) le système des valeurs $P_{(\Lambda)}^{(\theta)}$ avec $\theta \in \Omega$, $\Lambda \in \sigma_\Pi$, soumises à la condition d'additivité complète par rapport à Λ . Quelques-unes des propriétés fondamentales de ces matrices seront considérées sous des hypothèses un peu plus générales dans un Appendice (§ 7, 8) ⁽³⁾.

En généralisant une idée de Laplace, selon laquelle un système de probabilités p_n ($n = 0, 1, 2, \dots$) est représenté par sa « fonction génératrice »

$$(1.3) \quad \varphi(z) \stackrel{\text{def}}{=} \sum_0^\infty p_n z^n,$$

z étant une variable auxiliaire, il s'avéra utile (cf. D. van Dantzig, 1941)

⁽²⁾ Un σ -corps est une classe d'ensembles contenant avec chaque ensemble son complément et avec chaque suite (dénombrable) d'ensembles leur union, donc aussi leur intersection.

⁽³⁾ La notation dans l'Appendice (P_X^θ au lieu de P_Λ^θ) est légèrement différente de celle dans ce paragraphe, mais est conforme à celle du paragraphe 2.

de représenter un champ de probabilité P_Λ par la fonctionnelle correspondante d'une fonction auxiliaire U définie sur Π . En admettant que cette fonction est mesurable σ_Π et que l'intégrale (1.4) existe et en désignant la valeur que U prend en un point $\lambda \in \Pi$ par $U^{(\lambda)}$ [au lieu de l'habituel $U(\lambda)$], la fonctionnelle correspondante peut s'écrire comme une intégrale de Stieltjes-Lebesgue-Radon

$$(1.4) \quad C(U) \stackrel{\text{def}}{=} \int P_{(d\lambda)} U^{(\lambda)}.$$

Si $C(U)$ est connue pour une classe suffisamment large de fonctions U , elle détermine P . Dans le cas le plus simple, si elle est connue pour toutes les fonctions U^λ ⁽⁴⁾ bornées et mesurables σ_Π , nous aurons

$$(1.5) \quad P_\Delta = C(\mathbf{1}_\Delta),$$

où $\mathbf{1}_\Delta$ ($\mathbf{1}$ = iota) désigne la fonction caractéristique de l'ensemble Δ , définie par

$$(1.6) \quad \mathbf{1}_\Delta \stackrel{\text{def}}{=} \begin{cases} 1 & \text{si } \lambda \in \Delta, \\ 0 & \text{si } \lambda \notin \Delta \end{cases}$$

(cf. D. van Dantzig, 1935). Dans le cas plus général, où P varie sur un ensemble Ω , nous pouvons aussi introduire une fonction d'ensemble auxiliaire, complètement additive F sur un σ -corps σ_Ω de sous-ensembles de Ω , en supposant que pour tout Δ fixé P_Δ^0 est mesurable σ_Ω , et définir la fonctionnelle correspondante

$$(1.7) \quad C(F, U) \stackrel{\text{def}}{=} \int F_{d\theta} \int P_{d\lambda}^0 U^\lambda.$$

Ces fonctionnelles furent introduites dans nos cours à Amsterdam en 1947 et dans un exposé fait à Lyon en 1948 (publié en 1949) et quelques applications de la méthode y furent données. Les fonctions auxiliaires furent désignées par « marques », leur fonctionnelle par « marque collective ». Il s'avérera utile d'employer une notation et une terminologie, introduite à une autre occasion (1935), à savoir d'appeler des fonctions de points et des fonctions d'ensemble complètement additives (soumises

⁽⁴⁾ Dorénavant nous omettrons les parenthèses et écrirons P_Λ , P^0 et P_Λ^0 au lieu de $P_{(\Lambda)}$, $P^{(0)}$ et $P_{(\Lambda)}^{(0)}$ respectivement, parce qu'en ce qui suit, il n'y aura pas d'ambiguïté dans les notations.

à certaines conditions légèrement restrictives spécifiées dans l'Appendice) respectivement *fonctions de la première et de la seconde espèce* et de désigner systématiquement les arguments (points) des premières par des lettres minuscules écrites en indices supérieurs et ceux (ensembles) des secondes par des majuscules écrites en indices inférieurs. Exceptionnellement, dans le cas où l'on effectue une intégrale, les limites d'ensembles « élémentaires » ou « petits » sont représentées par le symbole désignant le point variable correspondant, précédé de la lettre d , au lieu d'une lettre majuscule (par exemple $d\lambda$ au lieu de Λ).

Des applications antérieures ont montré que l'application de la méthode était souvent considérablement facilitée si l'on interprétait les variables auxiliaires et collectives, fonctions et fonctionnelles comme des *probabilités*, en restreignant temporairement leur domaine de valeurs à l'intervalle réel $(0,1)$. Ainsi, dans (1.7) nous pouvons restreindre $F_\theta (\theta \in \sigma_\Omega)$ à $0 \leq F_\theta \leq 1 = F_\Omega$, et interpréter F_θ comme la probabilité que, au moyen d'un certain mécanisme aléatoire, on ait choisi un $\theta \in \Theta$. De même, en supposant $0 \leq U^\lambda \leq 1$, nous pouvons considérer un événement auxiliaire $\mathcal{C}$ (qui n'a rien à voir avec le problème de probabilités que l'on considère) et supposer que, chaque fois que le résultat de ce dernier est $\lambda \in \Pi$, un mécanisme aléatoire, dépendant de λ , détermine avec une probabilité $1-U^\lambda$ contre U^λ , si $\mathcal{C}$ a lieu ou non. Alors $C(F, U)$ est la probabilité totale que $\mathcal{C}$ n'ait pas lieu. Il est essentiel que le mécanisme aléatoire demeure (au moins partiellement) *indéterminé*, de façon à garantir la variabilité de F et de U sur des ensembles suffisamment larges de fonctions. En gros, plus la classe de fonctions sur lesquelles F et U peuvent varier est large, plus la classe de problèmes pour lesquels leur introduction est utile est étendue. L'interprétation d'un F auxiliaire comme une distribution de probabilités est due à J. von Neumann (1928) et fut largement utilisée dans : J. von Neumann et O. Morgenstern, *Theory of Games and Economic Behavior* (1947) et dans A. Wald, *Statistical Decision Functions* (1950).

Cependant on donne souvent une interprétation économique de « gains » et « pertes » aux fonctions de la première espèce, quand elles interviennent dans ces textes. Ceci, sans aucun doute, a l'avantage qu'elles peuvent varier sur tout l'ensemble des nombres réels, sans restriction à l'intervalle $(0,1)$, mais aussi l'inconvénient que les produits et les puissances de gains ou de pertes n'ont pas d'interprétation évidente,

alors que les produits de probabilité peuvent directement être interprétés comme des probabilités.

Pour la même raison C et U^λ ont été interprétées comme les probabilités totales et conditionnelles pour qu'un événement *n'ait pas lieu* (au lieu de : *ait lieu*). Nous appellerons cet événement C une « catastrophe ». L'avantage de terminologie de prendre la réalisation de non- C au lieu de celle de C réside dans le fait que la conjonction de plusieurs non-réalisations de C peut être décrite comme une non-réalisation (totale) de C alors que la conjonction de plusieurs réalisations de C ne peut être décrite aussi simplement comme une réalisation. Quoi qu'il en soit, l'interprétation des quantités auxiliaires n'est pas de première importance et, de plus, rien ne nous empêche, tant que n'apparaissent pas des difficultés de convergence, d'appliquer la terminologie probabiliste également à des quantités négatives ou complexes, de la même manière qu'on l'a fait pour la terminologie géométrique [cf. aussi Bartlett (1944)].

Le but de ce Mémoire est d'obtenir certains résultats concernant les processus stochastiques par cette méthode.

2. Processus stochastiques. — Nous considérons des variables aléatoires, c'est-à-dire des fonctions sur Π , dont les « valeurs » sont des éléments x, y, z , d'un ensemble arbitraire E . Des variables aléatoires (ou des événements aléatoires) seront désignées en omettant l'argument $\lambda \in \Pi$ et en soulignant le symbole de la fonction : par exemple, $\underline{x}, \underline{y}, \underline{z}_n, \dots$ ⁽⁵⁾. Nous supposons encore que sur l'ensemble E et sur ses produits directs ($E \times E, E \times E \times E, \dots$) on ait donné des σ -corps de sous-ensembles $\sigma_E, \sigma_{E^2}, \sigma_{E^3}$ (des conditions plus générales sont considérées dans l'Appendice (§ 7) ⁽⁶⁾).

Nous définirons un processus stochastique (discret dans le temps) comme une suite de variables aléatoires $x_0, x_1, x_2, \dots$ dans E , telles

⁽⁵⁾ On désigne souvent des variables aléatoires par des lettres majuscules. Ceci est rarement fait d'une façon cohérente, vu que les majuscules sont également utilisées à d'autres fins et que certaines variables aléatoires sont aussi désignées par d'autres symboles. Notre système de notations nous permet d'utiliser à d'autres fins tout un alphabet (lettres majuscules).

⁽⁶⁾ Pour obtenir une définition de la distribution de probabilité sur E , nous devons restreindre les variables aléatoires aux fonctions sur Π qui sont mesurables (σ_E, σ_Π) , c'est-à-dire pour lesquelles les originaux d'ensembles $\in \sigma_E$ sont toujours des ensembles $\in \sigma_\Pi$.

que la distribution des probabilités conditionnelles de chacune d'elles, les précédentes étant données, existe

$$(2.1) \quad P_{(n)}^{x_0, \dots, x_{n-1}} = P[\underline{x}_n \in X \mid \underline{x}_0 = x_0, \dots, \underline{x}_{n-1} = x_{n-1}],$$

où $x_0, \dots, x_{n-1}$ sont des éléments arbitraires de E et X est un élément de σ_E . On suppose que ces distributions de probabilité sont mesurables- σ_{E^n} (en $x_0, \dots, x_{n-1}$) et complètement additives en X et qu'elles satisfont aux relations

$$(2.2) \quad 0 \leq P_{(n)}^{x_0, \dots, x_{n-1}} \leq 1 = P_{(n)}^{x_0, \dots, x_{n-1}}.$$

La distribution de probabilité de l'ensemble $\underline{x}_0, \dots, \underline{x}_n$ est donnée par

$$(2.3) \quad P_{X_0, \dots, X_n}^{(n)} = P[\underline{x}_0 \in X_0, \dots, \underline{x}_n \in X_n].$$

Nous nous restreindrons au cas où elle est complètement additive, non seulement par rapport à chaque X_k séparément, mais aussi par rapport à l'ensemble des X_k , donc par rapport à $\sigma_{E^{n+1}}$. En admettant, en outre, qu'une intégrale multiple est égale au résultat d'intégrations successives, elle est alors liée à (2.1) par la relation récursive pour $n \geq 1$

$$(2.4) \quad P_{X_0, \dots, X_n}^{(n)} = \int_{X_0} \dots \int_{X_{n-1}} P_{d x_0, \dots, d x_{n-1}}^{(n-1)} P_{(n)}^{x_0, \dots, x_{n-1}}.$$

D'une façon descriptive nous appellerons cette suite de variables aléatoires un « cheminement aléatoire » dans E ou un point « cheminant » ou « sautant » à travers ou sur E , les éléments $x \in E$ les « états » ou les « positions » où il se trouve et $P_{(n)}^{x_0, \dots, x_{n-1}}$ la « probabilité de passage » d'un point ayant « passé par » $x_0, \dots, x_{n-1}$ successivement pour arriver à un état quelconque dans X . La suite des états $x_0, x_1, \dots$ par lesquels passe le point cheminant s'appelle sa trajectoire, la suite $x_0, \dots, x_n$ son segment de trajectoire d'ordre n , x_0 son état initial, la transition de x_{n-1} à x_n le $n^{\text{ième}}$ échelon (ou saut).

Un processus stochastique est un *processus de Markof* (simple) si les probabilités de passage pour $n \geq 1$ dépendent seulement du dernier état par lequel a passé le point cheminant

$$(2.5) \quad P_{(n)}^{x_0, \dots, x_{n-1}} = P_{(n)}^{x_{n-1}} \quad (n \geq 1);$$

$P_{(n)}$ est appelée la matrice de transition de la $n^{\text{ième}}$ transition; $P_{(n)X}^x$ est la probabilité pour que le point cheminant, si x est son $(n-1)^{\text{ième}}$ état (c'est-à-dire $\underline{x}_{n-1} = x$), saute dans un élément de X (c'est-à-dire $\underline{x}_n \in X$). La probabilité, sous la condition $\underline{x}_{n-1} = x$, que $\underline{x}_{n+1} \in X$ est $\int P_{(n)dy}^x P_{(n+1)X}^y$, que nous écrirons, en utilisant la notation matricielle discutée en plus de détails dans l'Appendice [définition (8.37)], $(P_{(n)} P_{(n+1)})_X^x$. D'une façon générale, nous avons pour $n \geq 1$

$$(2.6) \quad P[x_{n-1+k} \in X \mid x_{n-1} = x] = (P_{(n)} \dots P_{(n-1+k)})_X^x.$$

la multiplication matricielle étant associative (cf. Appendice, § 8, lemme 5).

Le processus de Markof est *stationnaire* si toutes les matrices de transition $P_{(n)}$ sont égales à une seule et même matrice P

$$(2.7) \quad P_{(n)X}^x = P_X^x \quad (n \geq 1).$$

La matrice définie par (2.6) est alors simplement la $k^{\text{ième}}$ puissance P^k de P , indépendante de n . Les propriétés asymptotiques des processus de Markof stationnaires généraux ont été étudiés dans un excellent Mémoire de W. Doeblin (1940).

Un cas particulier important est celui où E est l'ensemble de tous les nombres réels et pour lequel le processus de Markof est invariant par rapport à une translation, c'est-à-dire

$$(2.8) \quad P_{(n)X+\nu}^{x+\nu} = P_{(n)X}^x$$

pour tous les nombres réels ν . En posant $P_{(n)X}^{\text{def}} = P_{(n)X}^0$ nous avons alors

$$(2.9) \quad P_{(n)X}^x = P_{(n)X-x}^0 = P_{(n)X-x} = \int P_{(n)d\nu} \mathbf{I}_X^{x+\nu}.$$

Un tel processus est quelquefois appelé « processus à accroissements indépendants », puisque la coordonnée du $(n+1)^{\text{ième}}$ état est la somme de $(n+1)$ variables stochastiquement indépendantes

$$(2.10) \quad \underline{x}_n = \underline{x}_0 + \underline{v}_1 + \dots + \underline{v}_n$$

$\underline{v}_k$ étant distribuée suivant $P_{(k)X}$; nous préférons le terme « processus invariant ».

Plus généralement nous pouvons considérer le cas où E est un espace euclidien à r dimensions. Alors, en définissant un processus invariant par (2.8) où x et ν sont des vecteurs à r dimensions, (2.9) reste valable.

La même chose est valable dans des cas encore plus généraux où E est un groupe abélien arbitraire, écrit sous forme additive, x et y étant alors des éléments du groupe.

Le cas où E est un groupe non abélien peut être considéré comme une spécialisation du cas encore plus général où il existe un groupe de transformations de E dans lui-même, qui est transitif sur E et par rapport auquel P est invariant (cf. Appendice, § 9).

3. La matrice collective d'un processus de Markof stationnaire dans un milieu absorbant. — Nous considérons un processus de Markof stationnaire dans un ensemble E déterminé par la matrice de transition P_x^y . Nous supposons que le point cheminant part d'un état donné quelconque x , que s'il est dans un état y , il existe une probabilité A^y pour que le processus ne se continue pas (nous disons alors que le point cheminant est « absorbé » en y), auquel cas il existe une probabilité $1 - U^y$ pour que la catastrophe $\mathcal{C}$ arrive; et une probabilité $B^y = 1 - A^y$ qu'il se continue, auquel cas il existe une probabilité $1 - T^y$ pour que $\mathcal{C}$ arrive (*).

Au lieu de la probabilité totale C que $\mathcal{C}$ n'arrive pas, nous introduisons la probabilité conditionnelle C^x pour que le point cheminant, s'il part de x , soit éventuellement absorbé *sans qu'aucune catastrophe ne soit arrivée*.

Il est aisé de calculer C^x . Le point étant en x , il est ou bien absorbé immédiatement (probabilité A^x) auquel cas non- $\mathcal{C}$ a la probabilité U^x , ou non (probabilité B^x) auquel cas la probabilité de non- $\mathcal{C}$ est T^x . Mais alors il saute en un certain point y (probabilité de transition P_{xy}^x si dy représente un ensemble « élémentaire » contenant y , à savoir un ensemble sur lequel la variation de la quantité à intégrer tend vers zéro) et alors la probabilité d'une absorption éventuelle sans $\mathcal{C}$ est C^y , d'où

$$(3.1) \quad C^x = A^x U^x + B^x T^x \int P_{dy}^x C^y.$$

Pour obtenir une solution de cette équation, nous introduisons la

(*) On peut interpréter un processus de Markof avec absorption comme un processus dans un ensemble $E \cup \omega$, où ω consiste d'un seul élément seulement, n'appartenant pas à E . Pour la matrice Q du processus dans $E \cup \omega$ on aura, avec $x \in E$, $X \in \sigma_E$ arbitraires :

$$Q_X^x = B^x P_X^x, \quad Q_\omega^x = A^x, \quad Q_X^\omega = 0, \quad Q_\omega^\omega = 1.$$

Une interprétation similaire se laisse donner à l'événement d'une catastrophe.

« matrice collective » C_X^x qui désigne la probabilité pour qu'un point, partant de x , soit éventuellement absorbé quelque part dans X , sans qu'une catastrophe arrive. L'équation pour C_X^x , correspondant à (3.1) est alors [cf. (1.6)] $C_X^x = A^x U^x \mathbf{1}_X^x + B^x T^x \int P_{dy}^x C_X^y$.

Si nous faisons l'hypothèse qu'une catastrophe ne peut arriver (avec la probabilité $1 - T^x$) que dans un état où le point n'est *pas* absorbé, nous devons substituer $U^x = 1$ et l'équation pour C_X^x se réduira à

$$(3.2) \quad C_X^x = A^x \mathbf{1}_X^x + B^x T^x \int P_{dy}^x C_X^y.$$

L'introduction d'une « catastrophe » $\mathcal{C}$ peut être ainsi interprétée comme suit. Ajoutons à l'ensemble E des états un nouvel élément, que nous dénoterons aussi par $\mathcal{C}$, et passons de la matrice de Markof P sur E à une matrice Q sur $E + \mathcal{C}$ définie par $Q_X^x = T^x P_X^x$, $Q_{\mathcal{C}}^x = 1 - T^x$ si $x \in E$, $X \in \sigma_E$, et $Q_X^{\mathcal{C}} = 0$, $Q_{\mathcal{C}}^{\mathcal{C}} = 1$ si $X \in \sigma_E$, l'élément $\mathcal{C}$ étant stochastiquement fermé. Par rapport à l'absorption et dans le cas général une interprétation analogue est possible.

Le C^x total (pour un U^x général) peut alors être exprimé par rapport aux C_X^x particuliers par

$$(3.3) \quad C^x = C^x(U) = \int C_{dy}^x U^y,$$

ce qu'on peut démontrer en utilisant le lemme 3 (Appendice, § 8) et en montrant que la solution de (3.1) est unique. Cette démonstration sera donnée plus bas pour $U^x = 1$ et peut être étendue au cas de (3.1) en remarquant qu'alors $|U^x| \leq 1$.

Nous considérons en particulier le cas où les probabilités d'absorption locale A^x ont seulement les valeurs 0 ou 1. Si alors A est l'ensemble de tous les x avec $A^x = 1$ et B son complément, nous aurons

$$(3.4) \quad A^x = \mathbf{1}_A^x, \quad B^x = \mathbf{1}_B^x.$$

En substituant (3.4) dans (3.2), on obtient en accord avec les définitions des matrices diagonales (§ 8)

$$(8.7) \quad (\mathbf{1}_A)_X^x \stackrel{\text{def}}{=} \mathbf{1}_{A \cap X}^x = \mathbf{1}_A^x \mathbf{1}_X^x, \quad (\mathbf{1}_B)_X^x \stackrel{\text{def}}{=} \mathbf{1}_{B \cap X}^x = \mathbf{1}_B^x \mathbf{1}_X^x,$$

$$(8.8) \quad (T)_X^x \stackrel{\text{def}}{=} T^x \mathbf{1}_X^x$$

et de la multiplication matricielle

$$(8.13) \quad (QP)_X^x \stackrel{\text{def}}{=} \int_B Q_{dy}^x P_X^y,$$

considérée d'une façon plus détaillée dans l'Appendice (§ 8), l'équation

$$C_X^x = (\mathbf{I}_A)_X^x + (\mathbf{I}_B TPC)_X^x$$

ou, en notation matricielle,

$$(3.5) \quad C = \mathbf{I}_A + \mathbf{I}_B TPC.$$

En multipliant les deux membres de (3.5) à gauche par $\mathbf{I}_A$ ou par $\mathbf{I}_B$ et en notant que $\mathbf{I}_A \mathbf{I}_B = \mathbf{I}_B \mathbf{I}_A = 0$, $\mathbf{I}_A \mathbf{I}_A = \mathbf{I}_A$, $\mathbf{I}_B \mathbf{I}_B = \mathbf{I}_B$ on obtient les identités importantes et souvent utilisées

$$(3.6) \quad \mathbf{I}_A C = \mathbf{I}_A, \quad \mathbf{I}_B C = \mathbf{I}_B TPC.$$

De (3.5) on obtient par induction

$$(3.7) \quad C = \sum_0^{N-1} (\mathbf{I}_B TP)^n \mathbf{I}_A + (\mathbf{I}_B TP)^N C.$$

Maintenant, en supposant d'abord que $\|T\| = \sup_{x \in E} T^x < 1$, on a [(8.42) de l'Appendice]

$$(3.8) \quad \|(\mathbf{I}_B TP)^N C\| \leq \| \mathbf{I}_B TP \|^N \|C\| \leq \|T\|^N,$$

puisque $\|\mathbf{I}_B\| = 1$ (à moins que $B = 0$, l'ensemble vide, cas banal que nous excluons), $\|P\| = 1$ et $\|C\| \leq 1$, C_X^x étant une distribution de probabilités. Ainsi le second terme dans le second membre de (3.7) tend vers zéro et

$$(3.9) \quad C = \sum_0^\infty (\mathbf{I}_B TP)^n \mathbf{I}_A,$$

où la série converge si $\|T\| < 1$. Nous savons cependant seulement que $\|T\| \leq 1$. Si nous remplaçons T^x dans (3.9) par θT^x , θ étant un nombre réel, la série dans le second membre de (3.9) est convergente pour $0 \leq \theta < 1$, donc une fonction analytique de θ . Comme tous ses termes sont ≥ 0 et que sa valeur, étant une probabilité, demeure ≤ 1 , donc bornée pour $0 \leq \theta < 1$, la série demeure aussi convergente pour $\theta = 1$ et sa valeur demeure ≤ 1 . Donc (3.9) est valable non seulement pour $\|T\| < 1$, mais aussi pour $\|T\| = 1$.

A la rigueur, nous pouvons également abandonner l'interprétation des T^x comme probabilités (§ 5, 10) et les remplacer par des valeurs complexes arbitraires. Alors C devient une fonctionnelle analytique (complexe) de T , définie (au moins) pour $\|T\| \leq 1$.

Pour $\|T\| < 1$, nous pouvons écrire, à cause de (8.50)

(3.10)

$$C = (I - I_B T P)^{-1} I_A.$$

En multipliant les deux membres de (3.9) à droite par I_A , on obtient l'identité

(3.11)

$$C I_A = C$$

qui implique le fait que l'absorption se produit en A seulement, c'est-à-dire que

(3.12)

$$C_X^x = C_{A \cap X}^x.$$

Pour $T^x = 1$ (c'est-à-dire si aucune catastrophe ne se produit), (3.9) implique la probabilité totale pour qu'un point, partant de x soit absorbé éventuellement quelque part dans X . Dans le cas particulier où E est l'ensemble des entiers, de sorte que $P_X^x = \sum_{y \in X} P_y^x$, où

(3.13)

$$P_y^x = \begin{cases} p & \text{si } y = x + 1, \\ 1 - p & \text{si } y = x - 1 \end{cases}$$

et $= 0$ dans tous les autres cas, et où

$$B = \text{Ens} \{ x \in E \mid 0 < x < a + b \},$$

C_y^x avec $T^x = 1$ pour $y = x - a$ et $y = x + b$ résout le problème classique de la ruine des joueurs. En effet, (3.9) est essentiellement équivalent à la solution classique d'Abraham de Moivre. Pour $T^x = T = \text{const.}$, on obtient la fonction génératrice du problème de la durée de jeu. La généralisation qui avait été utilisée dans la théorie originale de Wald et Barnard sur l'analyse séquentielle est obtenue si dans (3.13) les conditions sont remplacées par $y = x + \beta$ et $y = x - \alpha$ respectivement, α et β étant des nombres réels positifs, alors que $T^x = T$. En particulier, la fonction génératrice utilisée par Barnard correspond au cas $\alpha = 1$, $\beta = \text{entier} > 1$. Des cas plus généraux ont été étudiés par D. Blackwell et M. A. Girshick, G. Blom, M. A. Girshick et J. H. B. Kemperman. Ce dernier a considéré le processus stochastique à accroissements indépendants général dans un espace euclidien à n dimensions et a étudié en grand détail le cas où E est l'ensemble

de tous les entiers et $P_y^x = p_{y-x}$ arbitraire (mais évidemment ≥ 0 avec $\sum_{x=-\infty}^{+\infty} p_x = 1$).

Le cas où un point cheminant arrive quelque part dans un ensemble A_1 sans être passé préalablement dans un ensemble A_2 est également contenu dans notre formule générale en prenant $A = A_1 \cup A_2$, $X = A_1$. Quelques applications de l'emploi de T^x non constant seront données dans le paragraphe 6.

4. Cas de deux régions absorbantes. — Nous allons maintenant établir une relation entre la matrice collective relative à une certaine région absorbante A_1 et celle relative à une deuxième région absorbante A_2 contenue dans A_1 . On peut penser à deux espèces de particules cheminantes, celles de première espèce étant absorbées dans A_1 , celles de seconde espèce dans A_2 seulement. Nous dénoterons ces deux matrices collectives par $C_{(A_1)}$ et $C_{(A_2)}$ respectivement et montrerons que :

THÉORÈME 1. — Si $A_1 \supset A_2$, alors

(4.1)

$$C_{(A_1)} C_{(A_2)} = C_{(A_2)} = C_{(A_2)} C_{(A_1)}.$$

On remarque que le théorème s'applique si $A_1 = A_2$ et affirme alors que C est idempotent (opérateur projectif). L'équation (4.1) suggère une analogie entre les matrices $C_{(A)}$ et les opérateurs spectraux. Ces matrices cependant ne sont pas additives dans les ensembles A .

Démonstration. — La seconde égalité dans (4.1) est banale; elle fait usage seulement du fait que d'après (3.11)

$$(4.2) \quad C_{(A_2)} = C_{(A_2)} \mathbf{I}_{A_2}$$

et que d'après (3.5) et (3.6)

$$(4.3) \quad C_{(A_1)} = \mathbf{I}_{A_1} + \mathbf{I}_{B_1} C_{(A_1)}.$$

Alors la définition de la multiplication matricielle (8.37) implique pour les matrices diagonales $\mathbf{I}_X$ et $\mathbf{I}_Y$ que

$$(4.4) \quad \mathbf{I}_X \mathbf{I}_Y = \mathbf{I}_{X \cap Y},$$

par conséquent comme $A_2 \subset A_1$ et dès lors $B_1 \subset B_2$, $A_1 \cap B_1 = A_2 \cap B_2 = 0$,

$$A_1 \cup B_1 = A_2 \cup B_2 = E, \quad A_2 \cap B_1 = \emptyset, \quad A_1 \cup B_2 = E,$$

$$(4.5) \quad \mathbf{I}_{A_2} \mathbf{I}_{A_1} = \mathbf{I}_{A_1} \mathbf{I}_{A_2} = \mathbf{I}_{A_2},$$

$$(4.6) \quad \mathbf{I}_{B_1} \mathbf{I}_{B_2} = \mathbf{I}_{B_2} \mathbf{I}_{B_1} = \mathbf{I}_{B_1},$$

$$(4.7) \quad \mathbf{I}_{A_2} \mathbf{I}_{B_1} = \mathbf{I}_{B_1} \mathbf{I}_{A_2} = \emptyset,$$

$$(4.8) \quad \mathbf{I}_{A_1} \mathbf{I}_{B_2} = \mathbf{I}_{B_2} \mathbf{I}_{A_1} = \mathbf{I}_{A_1} - \mathbf{I}_{A_2} = \mathbf{I}_{B_2} - \mathbf{I}_{B_1}.$$

De (4.2), (4.3), (4.5), (4.7) on déduit immédiatement

$$C_{(A_2)} C_{(A_1)} = C_{(A_2)} \mathbf{I}_{A_2} \mathbf{I}_{A_1} + C_{(A_2)} (\mathbf{I}_{A_2} \mathbf{I}_{B_1}) C_{(A_1)} = C_{(A_2)} \mathbf{I}_{A_2} + \emptyset = C_{(A_2)}.$$

Nous donnerons deux démonstrations de la première partie de l'équation (4.1). La seconde démonstration qui est purement algébrique comme la précédente sera donnée plus tard et sous une forme légèrement généralisée de manière à obtenir en même temps un autre résultat.

La première démonstration est basée sur l'interprétation probabiliste de $C_{(A_1)}$ et $C_{(A_2)}$,

$C_{(A_2)}^x$ est la probabilité de l'événement suivant $\mathcal{A}$: un point partant de x arrive finalement dans $A_2 \cap X$ sans qu'une catastrophe soit arrivée, et y est absorbé. Maintenant considérons la région $D \stackrel{\text{def}}{=} A_1 \cap B_2$, c'est-à-dire la partie entre A_1 et A_2 . Si $\mathcal{A}$ arrive, alors le point cheminant peut avoir ou ne pas avoir passé par un point de D . Les deux cas étant exclusifs et exhaustifs $C_{(A_2)}^x$ est la somme des deux probabilités correspondantes. Dans le premier cas la probabilité est la même que si la plus grande région A_1 , avait été la région absorbante pourvu que le point ait été absorbé non seulement dans X mais encore dans $A_2 \cap X$. Donc sa probabilité est $C_{(A_1)}^x = C_{(A_1)} \mathbf{I}_{A_2}^x$. Dans le second cas il y a un point $y \in D$ où le point cheminant vient pour la première fois dans D . La probabilité pour qu'il arrive dans un petit ensemble dy est la probabilité d'être absorbé là si A_1 avait été la région absorbante, c'est-à-dire $(C_{(A_1)})_{dy}^x$. Ceci doit être multiplié par la probabilité que le point aille de y à un point de $X \cap A_2$ (toujours sans catastrophe), A_2 étant la région absorbante, c'est-à-dire par $C_{(A_2)}^y$ est intégré sur $y \in D$. Finalement nous obtenons

$$(4.9) \quad C_{(A_2)}^x = C_{(A_1)}^x \mathbf{I}_{A_2 \cap X} + \int_D C_{(A_1)}^x_{dy} C_{(A_2)}^y$$

ou en notation matricielle

$$(4.10) \quad C_{(A_2)} = C_{(A_1)} \mathbf{I}_{A_2} + C_{(A_1)} \mathbf{I}_D C_{(A_2)},$$

mais $\mathbf{I}_D = \mathbf{I}_{B_1} - \mathbf{I}_{B_2}$; $C_{(A_1)} \mathbf{I}_{B_1} = C_{(A_1)} \mathbf{I}_{A_1} \mathbf{I}_{B_1} = 0$, donc

$$C_{(A_2)} = C_{(A_1)} (\mathbf{I}_{A_2} + \mathbf{I}_{B_1} C_{(A_2)}) = C_{(A_1)} C_{(A_2)},$$

d'après une identité analogue à (4.3). Ce qui prouve le théorème.

Avant de donner la seconde démonstration, considérons un cas très particulier. Soit E l'ensemble de tous les entiers. Les intégrations peuvent alors (et d'une manière générale si E est dénombrable) être remplacées par des sommations et nous connaissons les matrices et les fonctions de seconde espèce si nous connaissons leurs valeurs pour les ensembles qui ne contiennent qu'un seul élément. Soit B_2 l'intervalle $-a < x < b$ et B_1 l'intervalle $-d < x < b$, où $-a < -d < 0 < b$ (a , b et d sont des entiers positifs). En outre, nous prenons $x = 0$ et nous prenons pour X l'ensemble réduit au seul élément $-l$ avec $-l \leq -a$ (l entier). Supposons, en outre, que le processus soit invariant de sorte que $P_y^x = P_{y+k}^{x+k} = p_{y-x}$ pour tous les entiers x , y et k et $T^x = \text{const.}$ Alors (4.9) devient, avec $y = -j$,

$$(4.11) \quad C_{(A_2)} - l = C_{(A_1)} - l + \sum_d^{a-1} j C_{(A_1)-j}^0 C_{(A_2+j)-l+j}^0$$

où

$$A_2 + j = \text{Ens} \{ x + j \mid x \in A_2 \}.$$

Avec des notations différentes, cette relation a été trouvée par J. H. B. Kemperman dans sa Thèse de doctorat à Amsterdam (p. 71).

Grâce à (4.10), on peut trouver $C_{(A_2)X}^x$ pour tout x et X si la matrice $C_{(A_1)}$ est complètement connue et $C_{(A_2)X}^x$ pour $x \in D$ seulement. Par itération comme dans le paragraphe 3, nous pouvons cependant exprimer complètement $C_{(A_2)}$ en fonction de $C_{(A_1)}$.

En effet

$$C_{(A_1)} \mathbf{I}_D C_{(A_2)} = C_{(A_1)} \mathbf{I}_{B_2} C_{(A_2)} = C_{(A_1)} \mathbf{I}_{B_2} T P C_{(A_2)},$$

d'après (3.6) et (4.10) est équivalente à

$$(4.12) \quad C_{(A_2)} = C_{(A_1)} \mathbf{I}_{A_2} + C_{(A_1)} \mathbf{I}_D T P C_{(A_2)}.$$

Au moyen du même processus itératif que celui utilisé précédemment ceci donne

$$(4.13) \quad C_{(A_2)} = C_{(A_1)} \sum_0^{\infty} (\mathbf{I}_D T P C_{(A_1)})^n \mathbf{I}_{A_2}$$

ou d'après (8.50) si $\|T\| < 1$

(4.14)

$$C_{(A_2)} = C_{(A_1)} (I - I_D T P C_{(A_1)})^{-1} I_{A_2}.$$

Dans le cas particulier où $A_1 = E$, donc $B_1 = 0$, $D = B_2$, $C_{(A_1)} = I$; ceci nous ramène à (3.10).

La solution (4.14) est particulièrement simple si D se compose seulement d'un seul point d . En notant par $I + Q$ le facteur du milieu du dernier membre de (4.14) nous avons donc

(4.15)

$$(I - I_d T P C_{(A_1)}) (I + Q) = I$$

ou

(4.16)

$$Q = I_d T P C_{(A_1)} + I_d T P C_{(A_1)} Q$$

montrant que Q_x^x s'annule à moins que $x = d$ (généralement $x \in D$) auquel cas

$$Q_x^d = T^d (P C_{(A_1)})_x^d + T^d (P C_{(A_1)})_d Q_x,$$

c'est-à-dire

(4.17)

$$Q_x^d = T^d \{ I - T^d (P C_{(A_1)})_d^d \}^{-1} (P C_{(A_1)})_x^d.$$

La solution (4.14) devient alors

(4.18)

$$C_{(A_2)} x = C_{(A_1) A_2 \cap x} + C_{(A_1) d} Q_{A_2 \cap x}^d.$$

Si E est dénombrable, (4.17) et (4.18) donnent une méthode récurrente pour calculer successivement les $C_{(A_k)}$ ($k = 1, 2, 3, \dots$) si nous prenons $B_0 = 0$ (par conséquent $C_{(A_0)} = I$) et ajoutons à chaque fois un élément à B_k .

Nous supposons, pour simplifier les notations, que les éléments $x, y, z, \dots$ de E sont les entiers $1, 2, 3, \dots$ eux-mêmes, dans l'ordre dans lequel ils sont pris dans B , de sorte que

$$B_k = \{1, \dots, k\}, \quad A_k = \{k+1, k+2, \dots\}$$

et nous noterons C_k au lieu de $C_{(A_k)}$. En outre, nous remarquerons que pour $y \geq k$,

$$(P C_{k-1})_y^k = \sum_{z=1}^{\infty} P_z^k (C_{k-1})_y^z = \sum_{z=1}^{K-1} P_z^k (C_{k-1})_y^z + P_y^k$$

puisque $(C_{k-1})_y^x = 0$ si $x \geq k$ à moins que $x = y$. La relation récursive

devient alors

$$(C_k)_y^x = 0 \quad \text{si } y \leq k$$

et

$$(4.19) \quad (C_k)_y^x = (C_{k-1})_y^x + (C_{k-1})_k^x \left\{ 1 - T^k P_k^k - T^k \sum_{z=1}^{k-1} P_z^k (C_{k-1})_k^z \right\}^{-1} T^k \\ \times \left\{ P_y^k + \sum_{z=1}^{k-1} P_z^k (C_{k-1})_y^z \right\} \quad \text{si } y \geq k+1,$$

dans laquelle nous rappelons que $(C_{k-1})_y^x$ et $(C_k)_y^x$ sont fonctions d'une infinité de variables indépendantes $T^1, T^2, T^3, \dots$ (les indices supérieurs ne sont pas des exposants!).

§. Seconde démonstration et l'identité fondamentale de Wald. — La seconde démonstration de la première équation de (4.1) peut être donnée de plusieurs manières, et consiste simplement à vérifier (4.1), (4.10) ou (4.12) au moyen de (3.5), (3.9) ou (3.10), appliqués à $C_{(A_1)}$ et $C_{(A_2)}$. Nous choisirons la forme suivante. A la place de (3.5), nous pouvons écrire $I - C = I_B - I_B T P C$ ou

$$(5.1) \quad I_B(I - TP) = (I - I_B TP)(I - C)$$

ou si $\|T\| < 1$

$$(5.2) \quad I - C = (I - I_B TP)^{-1} I_B(I - TP).$$

D'où, si nous multiplions les deux membres à droite par une matrice bornée D , nous trouvons

$$(5.3) \quad CD = D$$

si D satisfait l'identité

$$(5.4) \quad I_B(I - TP)D = 0.$$

En prenant maintenant, comme précédemment,

$$C = C_{(A_1)}, \quad B = B_1, \quad D = C_{(A_2)},$$

(5.4) est satisfaite. Car, en vertu de (3.6),

$$I_{B_1} C_{(A_2)} = I_{B_1} T P C_{(A_2)};$$

si les deux membres sont multipliés à gauche par I_{B_1} , (4.6) prouve (5.4).

Donc (5.3), c'est-à-dire la première partie de (4.1), est vraie.

Les identités (5.1), (5.2), conduisent à d'autres résultats intéres-

sants. Comme l'ensemble X dont dépend D_x^x n'intervient pas, pour autant que D n'entre dans les équations que comme dernier facteur (à droite), nous pouvons tout aussi bien le remplacer par une fonction f^x de première espèce. Alors, du point de vue formel, notre résultat précédent (§.3), (§.4) devient : f^x satisfait l'équation

$$(\S.5) \quad \int C_{dy}^x f^y = f^x$$

(c'est-à-dire f est une fonction propre de C appartenant à la valeur propre 1), si

$$\mathbf{I}_B^x f^x = \mathbf{I}_B^x T^x \int P_{dy}^x f^y,$$

c'est-à-dire si

$$(\S.6) \quad f^x = T^x(Pf)^x \quad \text{pour } x \in B.$$

D'autre part, si (§.5) est vérifiée, (§.2) montre que (§.6) vaut également. Le résultat que (§.5) est impliqué par (§.6) est valable si f est bornée, et si les séries intervenant implicitement dans C et dans le membre de droite de (§.2) sont uniformément convergentes. Pour cela il suffit que $\|\mathbf{I}_B T P \mathbf{I}_B\| < 1$, c'est-à-dire que $|T^x| < (P_B^x)^{-1}$ pour chaque x dans B , ou, plus généralement, que $\|(\mathbf{I}_B T P \mathbf{I}_B)^n\| < 1$ pour quelque n .

Dans l'application la plus importante, cependant, f n'est pas bornée, et l'associativité doit alors être garantie d'une autre manière, mettons par les conditions du lemme 4 (Appendice, § 8). Nous avons alors :

THÉOREME 2. — *S'il existe une $f_0^x \geq 0$ et un $T_0^x \geq 0$ tels que*

$$(\S.7) \quad f_0^x \geq T_0^x(Pf_0)^x \quad \text{si } x \in B$$

et que

$$(\S.8) \quad \|\mathbf{I}_B P T_0 \mathbf{I}_B\| = \sup_{x \in B} \int_B P_{dy}^x T_0^y \leq 1.$$

alors pour deux constantes non négatives quelconques θ et ϵ avec $\theta < 1$, pour tout T^x et f^x avec $|f^x| \leq \epsilon f_0^x$ et avec $|T^x| \leq \theta T_0^x$ pour tout $x \in B$ les équations (§.5) et (§.6) sont équivalentes ⁽¹⁾.

(¹) Cf. J. H. B. Kemperman, th. 2 (p. 14) pour les cas où E est l'axe réel, le processus est invariant (mais pas nécessairement stationnaire), T^x et T_0^x sont constants, et A est une demi-droite. Pour le cas général non stationnaire, cf. § 10.

Démonstration. — Nous avons, grâce à (3.9) et (5.2),

$$C = \sum_0^\infty {}^n (\mathbf{I}_B TP)^n \mathbf{I}_A,$$

$$\mathbf{I} - C = \sum_0^\infty {}^n (\mathbf{I}_B TP)^n \mathbf{I}_B (\mathbf{I} - TP),$$

pourvu que les membres de droite existent, ce que nous prouverons en premier lieu. De (5.8) on obtient par induction

$$[(\mathbf{I}_B | T | P)^n]_X^x = \theta^n [(\mathbf{I}_B T_0 P)^n]_X^x \leq \theta^n \mathbf{I}_B^x T_0^x.$$

Puisque pour $n = 1$ nous avons

$$(\mathbf{I}_B T_0 P)_X^x = \mathbf{I}_B^x T_0^x P_X^x \leq \mathbf{I}_B^x T_0^x,$$

et, en supposant que $[(\mathbf{I}_B T_0 P)^n]_X^x \leq \mathbf{I}_X^x T_0^x$, nous obtenons

$$\begin{aligned} [(\mathbf{I}_B T_0 P)^{n+1}]_X^x &= [(\mathbf{I}_B T_0 P)(\mathbf{I}_B T_0 P)^n]_X^x \leq \{(\mathbf{I}_B T_0 P)(\mathbf{I}_B T_0)\}_X^x \\ &= \int \mathbf{I}_B^x T_0^x P_{dy}^x \mathbf{I}_B^y T_0^y = \mathbf{I}_B^x T_0^x \int_B P_{dy}^x T_0^y \leq \mathbf{I}_B^x T_0^x \quad \text{d'après (5.8).} \end{aligned}$$

Alors

$$\left\{ \sum_0^\infty {}^n (\mathbf{I}_B | T | P)^n \right\}_X^x \leq \sum_0^\infty \theta^n \mathbf{I}_B^x T_0^x = (1 - \theta)^{-1} \mathbf{I}_B^x T_0^x,$$

de sorte que C et le membre de droite de (5.2) existent absolument.

Donc, pour $R \stackrel{\text{def}}{=} \sum_0^\infty {}^n (\mathbf{I}_B | T | P)^n$, R existe, ainsi que $\sum_0^\infty {}^n (\mathbf{I}_B TP)^n$.

$(R|f|)^x$ existe aussi, puisque

$$(R|f|)^x = \left\{ \sum_0^\infty {}^n (\mathbf{I}_B | T | P)^n |f| \right\}_X^x \leq c \sum_0^\infty \theta^n \{(\mathbf{I}_B T_0 P)^n f_0\}_X^x.$$

(L'ordre de la sommation et de l'intégration peut être inversé parce que R existe et est positif.)

En outre,

$$\{(\mathbf{I}_B T_0 P)^n f_0\}_X^x \leq (\mathbf{I}_B f_0)^x$$

pour tout n , ce qu'on prouve encore par induction, comme suit.

Pour $n = 1$,

$$\{(\mathbf{I}_B T_0 P) f_0\}_X^x = \mathbf{I}_B^x T_0^x (P f_0)^x \leq \mathbf{I}_B^x f_0^x \quad \text{d'après (5.7).}$$

En supposant que l'inégalité soit vraie pour n , nous avons pour $n+1$

$$\begin{aligned} \{(\mathbf{I}_B T_0 P)^{n+1} f_0\}^x &= \{[(\mathbf{I}_B T_0 P)^n (\mathbf{I}_B T_0 P)] f_0\}^x \\ &= \{(\mathbf{I}_B T_0 P)^n (\mathbf{I}_B T_0 P) f_0\}^x \quad \text{d'après lemme 4 (Appendice, § 8),} \end{aligned}$$

parce que $[(\mathbf{I}_B T_0 P) f_0]^x$ et $\{(\mathbf{I}_B T_0 P)^n (\mathbf{I}_B T_0 P)\}_X^x$ existent pour tout $x \in E$ et $X \in \sigma_E$ ce qu'on a démontré plus haut et parce que

$$\begin{aligned} \{(\mathbf{I}_B T_0 P)^n (\mathbf{I}_B T_0 P) f_0\}^x &\leq \{(\mathbf{I}_B T_0 P)^n \mathbf{I}_B f_0\}^x \leq \{(\mathbf{I}_B T_0 P)^n f_0\}^x \\ &\leq (\mathbf{I}_B f_0)^x < \infty \quad \text{d'après la supposition d'induction.} \end{aligned}$$

Donc $(R|f|)^x \leq c(1-0)^{-1} (\mathbf{I}_B f_0)^x$ et existe.

Maintenant, de l'équation (5.2) il résulte que $(Cf)^x$ existe aussi et alors, également

$$\{(1-C)f\}^x = f^x - \int C_{dy}^x f^y.$$

D'après (5.2) nous avons

$$\begin{aligned} \{(1-C)f\}^x &= \left[\sum_0^\infty (\mathbf{I}_B TP)^n \mathbf{I}_B (1-TP) \right] f^x \\ &= \left[\sum_0^\infty (\mathbf{I}_B TP)^n \mathbf{I}_B (1-TP) f \right]^x, \end{aligned}$$

selon le lemme 4 (Appendice, § 8), parce que

$$1^\circ \left[\sum_0^\infty (\mathbf{I}_B TP)^n \mathbf{I}_B (1-TP) \right]_X^x \text{ existe pour tout } X \in \sigma_E;$$

2° $(\mathbf{I}_B (1-TP) f)^x$ existe pour $x \in E$, puisque

$$(\mathbf{I}_B TP f)^x \leq (\mathbf{I}_B |T| P |f|)^x \leq c \mathbf{I}_B^x T_0^x (P f_0)^x \leq c \mathbf{I}_B^x f_0^x;$$

3° $\left\{ \sum_0^\infty (\mathbf{I}_B TP)^n \mathbf{I}_B (1-TP) f \right\}^x < \infty$ pour tout $x \in E$, car

$$\left| \left(\sum_0^\infty (\mathbf{I}_B TP)^n \mathbf{I}_B f \right)^x \right| \leq (R|f|)^x < \infty$$

et

$$\begin{aligned} \left| \left(\sum_0^\infty (\mathbf{I}_B TP)^n \mathbf{I}_B TP f \right)^x \right| &\leq (R \mathbf{I}_B |T| P |f|)^x \\ &\leq c \mathbf{I}_B^x \sum_0^\infty (\mathbf{I}_B T_0 P)^n \mathbf{I}_B T_0 P f_0^x \\ &\leq c \mathbf{I}_B^x \sum_0^\infty (\mathbf{I}_B T_0 P)^n f_0^x, \end{aligned}$$

suivant (5.7)

$$\leq c \theta (1 - \theta)^{-1} (1_B f_0)^x$$

comme on a démontré plus haut, donc $< \infty$.

Si (5.6) est vraie, le troisième membre de (5.8') est égal à zéro, le premier membre aussi et (5.5) est vraie.

On démontre de la même manière le passage de (5.5) à (5.6) et d'une façon plus simple en utilisant l'équation (5.1).

Nous allons appliquer maintenant ce résultat au cas d'un processus à accroissements indépendants. Nous prenons f^x de la forme exponentielle

$$(5.9) \quad f^x = e^{\xi x}$$

et $T^x = T = \text{const.}$, ξ et T étant des nombres réels ou complexes.

Alors si [cf. (2.8)]

$$(5.10) \quad P_X^x = \int P_{d\nu} \mathbf{1}_X^{x+\nu},$$

nous aurons

$$(5.11) \quad (Pf)^x = \int P_{d\nu} f^{x+\nu} = \int P_{d\nu} e^{\xi(x+\nu)} = \varphi(\xi) f^x,$$

où

$$(5.12) \quad \varphi(\xi) \stackrel{\text{def}}{=} \int P_{d\nu} e^{\xi \nu}$$

est la fonction caractéristique de la fonction de répartition. On peut certainement changer l'ordre des intégrations si les intégrales dans (5.11) convergent absolument, c'est-à-dire si (5.12) converge absolument, c'est-à-dire si $\varphi(\text{Re} \xi)$ existe. Donc nous avons établi le :

THÉOREME 3. — *Pour chaque ξ pour lequel $\varphi(\text{Re} \xi)$ existe, $f^x = e^{\xi x}$ est une fonction propre de la matrice P d'un processus de Markof stationnaire à accroissements indépendants, satisfaisant (5.10) et appartenant à la valeur propre $\varphi(\xi)$.*

La substitution de (5.9) et (5.11) dans (5.6) montre que cette dernière équation est satisfaite si et seulement si

$$(5.13) \quad T \varphi(\xi) = 1$$

(dès que B n'est pas vide), et vaut alors pour tout $x \in E$ (et pas seulement $\in B$).

Dans le but d'appliquer le théorème 2, nous choisirons $f_0^x = e^{\xi_0 x}$,

ξ_0 étant réel et tel que Pf_0 , donc $\varphi(\xi_0)$ existe, et $T_0^x = T_0 = \text{const.}$, réel > 0 , et $\leq (\|P_B\|^B)^{-1}$. La dernière condition, donc (5.8), est toujours satisfaite si $T_0 \leq 1$. La condition (5.7) (avec le signe d'égalité) est satisfaite si

$$(5.14) \quad T_0 \varphi(\xi_0) \leq 1.$$

En outre, nous prendrons un T et un ξ satisfaisant (5.13), avec $|T| T_0^{-1} = 0 < 1$, et tels que $(\xi_0 - \text{Re} \xi)x \geq -\ln c$ pour tout x dans B . Alors, en écrivant

$$(5.15) \quad C = \sum_0^\infty T_n C_{(n)}$$

de sorte que $C_{(n)}^x$ est la probabilité d'absorption dans X après le $n^{\text{ième}}$ échelon, nous aurons

$$(5.16) \quad (f^{-1} C f)^x = e^{-\xi x} \sum_0^\infty T^n \int C_{(n)}^x dy e^{\xi y}$$

et nous trouvons que (5.13) entraîne, pour $x = 0$, le membre de gauche de (5.16) étant égal à 1,

$$(5.17) \quad \sum_0^\infty \varphi(\xi)^{-n} \int C_{(n)}^0 dy e^{\xi y} = 1,$$

L'identité (5.17) est connue comme l'identité fondamentale de Wald. Nous avons vu que c'est un cas spécial de l'identité $Cf = f$, valable pour tout f satisfaisant $f = TPf$ sur B sous les conditions du théorème 2.

Une généralisation partielle pour des processus stochastiques arbitraires est considérée dans l'Appendice (§ 10).

6. Le problème des boucles. — Nous considérons un processus stationnaire sans absorption. Comme cependant nous désirons considérer les propriétés de segments initiaux de longueur donnée n du parcours, nous devons admettre la discontinuité de processus à un moment quelconque. Nous partirons donc de l'expression (3.1) et nous substituerons

$$(6.1) \quad B^x = B \text{ const.}, \quad A^x = 1 - B,$$

où B est une variable auxiliaire. En outre, nous prendrons $U^x = T^x$ et

un point initial déterminé $x = a$. Nous obtenons

$$(6.2) \quad C^a = \sum_0^{\infty} B^n (1 - B) ((TP)^n T)^a,$$

de sorte que (non C^a , mais) $C^a(1 - B)^{-1}$ est la fonction génératrice avec B comme variable auxiliaire.

Pour certaines applications de la théorie de Markof et d'autres processus stochastiques, par exemple dans la statistique chimique des molécules à longue chaîne (*cf.*, par exemple, G. King, 1948, 1949; E. W. Montroll, 1950; Ch. M. Tchen, 1951; J. J. Hermans, M. S. Klamkin et R. Ullman, 1952 et la littérature signalée dans ces articles) il peut être d'une certaine utilité d'avoir des méthodes pour étudier l'occurrence de « boucles » dans le parcours, c'est-à-dire de retours du point mobile à une position où il s'est trouvé auparavant. Sans entrer dans ces applications et sans préjuger de sa valeur pratique ou de sa capacité à fournir des résultats non triviaux dans des cas pratiquement importants, nous allons exposer une telle méthode ici, qui en tout cas pourrait fournir une ligne d'attaque de ces problèmes et de problèmes similaires, et qui est susceptible d'être généralisée pour être adaptée à d'autres problèmes.

Évidemment, la probabilité de la coïncidence de $\underline{x}_k$ et $\underline{x}_l$ pour $k \neq l$ est nulle si les $\underline{x}_k$ ont des distributions continues.

Pour éviter des complications hors de propos, nous supposons que ces distributions sont purement discontinues, c'est-à-dire que E est un ensemble dénombrable. Dans certaines des applications mentionnées, E est un tel ensemble, à savoir un réseau de points (par exemple un réseau tétraédrique) dans l'espace. Selon l'Appendice (8.1), la matrice de transition P_x^x est alors complètement déterminée par la matrice ordinaire (infinie) $P_y^x (x \in E, y \in E)$, et les intégrales $\int \mathbf{r}_{d,x} f^x$ deviennent des sommes (en général infinies) $\Sigma F_x f^x$.

C^a est maintenant une fonction analytique d'une infinité dénombrable de variables T^x , c'est-à-dire d'un vecteur dans un espace à une infinité de dimensions. Les dérivées partielles d'une fonction des T^x par rapport à ces variables peuvent être considérées comme les composantes d'un autre vecteur dans l'espace dual : le gradient. Nous désignerons l'opé-

rateur de différentiation par ∇ , avec

$$(6.3) \quad \nabla_x \stackrel{\text{def}}{=} \frac{\partial}{\partial T^x}.$$

L'indice est écrit en bas parce que : 1° dans l'analyse vectorielle les dérivées par rapport à un vecteur contravariant constituent un vecteur covariant; 2° il sera trouvé que la généralisation de (6.3) à E non dénombrable conduit à des fonctions de seconde espèce.

$$(6.4) \quad C = \sum_0^\infty p_n T^n$$

est une fonction génératrice ordinaire (T^n étant ici la $n^{\text{ième}}$ puissance de T) des probabilités $p_n = P[\underline{n} = n]$ des valeurs prises par une variable aléatoire $\underline{n}$, les dérivées successives de C par rapport à T donnent pour $T=1$ les moments factoriels successifs

$$(6.5) \quad [C]_{T=1} = 1.$$

$$(6.6) \quad \left[\frac{dC}{dT} \right]_{T=1} = \Sigma n p_n = \mathcal{E} \underline{n},$$

$$(6.7) \quad \left[\frac{d^2 C}{dT^2} \right]_{T=1} = \Sigma n^2 p_n = \mathcal{E} \underline{n}^2,$$

où $\mathcal{E}$ est le symbole pour l'espérance mathématique, tandis que

$$(6.8) \quad x^{!k} \stackrel{\text{def}}{=} x(x-1) \dots (x-k+1)$$

désigne la puissance factorielle d'ordre k de x .

En écrivant effectivement les sommations matricielles implicites dans (6.2), nous obtenons une suite infinie de termes, dont chacun est la probabilité d'un parcours défini de longueur prédéterminée n , multipliée par $T^{x_0} \dots T^{x_n} B^n (1-B)$. Si le parcours passe k fois par la position x , l'application de ∇_x suivie de la substitution de $T=1$ donne k fois la probabilité du parcours. Donc, si $\underline{n}_x$ désigne le nombre de fois qu'un parcours passe par x , nous avons

$$(6.9) \quad [\nabla_x C^a]_{T=1} = \mathcal{E} \underline{n}_x.$$

De la même manière nous trouvons

$$(6.10) \quad [\nabla_y \nabla_x C^a]_{T=1} = \mathcal{E} \underline{n}_y \underline{n}_x - \mathcal{E} \underline{n}_y \delta_{yx} = \begin{cases} \mathcal{E} \underline{n}_x^{!2} & \text{si } y = x, \\ \mathcal{E} \underline{n}_y \underline{n}_x & \text{si } y \neq x, \end{cases}$$

.....

Il suit que

$$(6.11) \quad \sum_{x \in E} \left[\frac{1}{2} \nabla_x \nabla_x C^a \right]_{T=1} = \frac{1}{2} \mathcal{E} \sum_{x \in E} n_x^{12}.$$

Maintenant si $n_x = 0$ ou $n_x = 1$, alors $n_x^{12} = 0$; si $n_x = 2$, c'est-à-dire si x est un point double du parcours, alors $\frac{1}{2} n_x^{12} = 1$. Donc (6.11) donne l'espérance du nombre de points doubles dans un parcours, un point triple, quadruple, k -uple, étant compté comme d'habitude en géométrie algébrique, pour $3, 6, \frac{1}{2}k(k-1)$ points doubles. Donc, nous avons établi le

THEOREME 4. — Si $\underline{D}$ désigne la variable aléatoire qui sur chaque parcours est égale au nombre de points doubles ou boucles, comptés comme il est indiqué ci-dessus, nous avons

$$(6.12) \quad \mathcal{E} \underline{D} = \left[\frac{1}{2} \square C^a \right]_{T=1},$$

où $\square$ désigne l'opérateur laplacien généralisé

$$(6.13) \quad \square \stackrel{\text{def}}{=} \sum_{x \in E} (\nabla_x)^2 = \sum_{x \in E} \left(\frac{\partial}{\partial T^x} \right)^2.$$

Des relations analogues existent évidemment pour le nombre de points triples, etc.

Dans la démonstration de (6.12), nous avons utilisé seulement le fait que C^a est une série de puissances en T^x , mais *non* sa forme spéciale (6.2). Donc le résultat vaut aussi pour des processus arbitraires non stationnaires et pour les processus non markoviens. Dans le cas spécial (6.2), les membres de gauche (6.9), (6.10) et (6.11) se calculent aisément. En effet, nous avons

$$(6.14) \quad \nabla_x ((TP)^n T) = \sum_{m=0}^n (TP)^m \mathbf{1}_x (PT)^{n-m}.$$

$$(6.15) \quad \begin{aligned} \nabla_y \nabla_x ((TP)^n T) &= \sum_{k+l+m=n-1} (TP)^k \mathbf{1}_y P(TP)^l \mathbf{1}_x (PT)^m \\ &+ \sum_{k+l+m=n-1} (TP)^k \mathbf{1}_x P(TP)^l \mathbf{1}_y (PT)^m. \end{aligned}$$

Donc nous obtenons

$$(6.16) \quad \left[\frac{1}{2} \square C^a \right]_{T=1} = (1-B) \sum_{k,l,m} B^{k+l+m+1} \sum_{x \in E} (P^k)_x^a (P^{l+1})_x^x (P^m)_E^x.$$

Mais $(PT)^x = \int P_{dy}^x T^y = P_E^x = 1$ si $T^x = 1$ pour tout x . Donc

$$\sum_0^\infty B^m (P^m)_E^x = (1-B)^{-1}$$

et (6.12), (6.16) peuvent aussi s'écrire

$$(6.17) \quad \mathcal{E} \underline{D} = \sum_{x \in E} \Phi_x^a (\Phi_x^x - 1),$$

où

$$(6.18) \quad \Phi_y^x \stackrel{\text{def}}{=} \left(\sum_0^\infty B^n P^n \right)_y^x = ((1-BP)^{-1})_y^x.$$

Dans le cas spécial, où le processus stationnaire de Markoff est invariant, (6.17) sera simplifié encore plus. Car dans ce cas Φ_x^x est indépendant de x , de sorte que la sommation s'étend sur Φ_x^a seulement et donne $\Phi_E^a = (1-B)^{-1}$. Donc, pour un processus invariant, (6.17) devient

$$(6.19) \quad \mathcal{E} \underline{D} = (1-B)^{-1} (\Phi_0^a - 1)$$

Exprimé en fonction des espérances conditionnelles $\mathcal{E}_{(n)} \underline{D}$ pour n donné, le membre de gauche de (6.19) est $\sum_0^\infty B^n (1-B) \mathcal{E}_{(n)} \underline{D}$, de sorte que (6.19) montre que $\mathcal{E}_{(n)} \underline{D}$ est le coefficient de B^n dans $(1-B)^{-2} (\Phi_0^a - 1)$; par suite

$$(6.20) \quad \mathcal{E}_{(n)} \underline{D} = \sum_1^n (n-l+1) (P^l)_0^a.$$

Ici $(P^m)_0^a$ est la probabilité pour qu'un parcours de longueur m soit fermé, c'est-à-dire que sa première et sa dernière $[(m+1)^{\text{ième}}]$ positions coïncident (qu'il contienne ou non des points doubles). Comme tant de cas particuliers de théorèmes généraux, ce résultat est d'ailleurs banal : dans un parcours de longueur n une boucle peut avoir une longueur quelconque l , $1 \leq l \leq n$ (nous parlons d'une boucle de longueur 1, si

le point chevinant ne saute pas, mais reste au même état). Pour l , la boucle commence au 1^{er} , ..., $(n-l+1)^{\text{ième}}$ point; dans le cas d'un processus invariant ces $n-l+1$ cas ont des probabilités égales, valant $(P^l)_0$, c'est-à-dire la valeur du produit l -uple de convolution de p_v [cf. (2.8)] pour la valeur $v=0$.

La méthode indiquée ici peut être généralisée pour les processus stochastiques *arbitraires* et appliquée à d'autres problèmes que l'espérance du nombre moyen de boucles.

Pour une fonctionnelle arbitraire Φ d'une fonction T^x , nous posons (cf. Hadamard, 1910; Fréchet, 1912, 1914; van Dantzig, 1935)

$$(6.21) \quad \nabla_x \Phi(T) = \left(\frac{\partial}{\partial T} \right)_X \Phi(T) \stackrel{\text{def}}{=} \lim_{\varepsilon \rightarrow 0} \frac{\Phi(T + \varepsilon \mathbf{1}_X) - \Phi(T)}{\varepsilon}$$

si cette limite existe. Sous certaines conditions de régularité, qui sont, par exemple, satisfaites dans le cas où Φ est analytique, cette quantité est une fonction complètement additive de l'ensemble X . La définition duale pour une fonctionnelle Ψ d'une fonction V_X de seconde espèce serait

$$(6.22) \quad \left(\frac{\partial}{\partial V} \right)^x \Psi(V) = \lim_{\varepsilon \rightarrow 0} \frac{\Psi(V + \varepsilon \mathbf{1}_X) - \Psi(V)}{\varepsilon}$$

qui fournit pour Ψ suffisamment régulière une fonction de première espèce. Pour une matrice P , nous pouvons poser

$$(6.23) \quad \left(\frac{\partial}{\partial P} \right)_A^a \Psi(P) = \lim_{\varepsilon \rightarrow 0} \frac{\Psi(P + \varepsilon \mathbf{1}_A \mathbf{1}_X^a) - \Psi(P)}{\varepsilon},$$

où $(\mathbf{1}_A \mathbf{1}_X^a)_X^x = \mathbf{1}_A^x \mathbf{1}_X^a$ est à distinguer de

$$\mathbf{1}^a \mathbf{1}_A = \int \mathbf{1}_{dx}^a \mathbf{1}_X^x = \mathbf{1}_A^a.$$

Dans cet article nous aurons affaire seulement à (6.21). En appliquant l'opérateur ∇_x au processus stochastique général

$$(6.24) \quad C = (1-B) \sum_0^\infty B^n \iint \dots \int P_{d.c_0 \dots d.x_{n-1}}^{(n)} T^{x_0} \dots T^{x_{n-1}}$$

à partir de quoi le cas spécial (6.2) s'obtient de nouveau par la particularisation (2.5) avec $P_{(n)} = P$ indépendant de n et $P_{(0)X} = \mathbf{1}_X^a$, nous obtenons

$$(6.25) \quad (\nabla_X C)_{T=1} = \mathcal{E} n_X,$$

où la variable aléatoire $\underline{n}_X$ pour X donné quelconque est le nombre de fois que le processus passe par une position dans X . Pour $X = E$, $\underline{n}_E = \underline{n}$ est le nombre de pas effectués jusqu'à ce que le processus soit arrêté.

De la même manière nous trouvons, comme généralisation de (6.10),

$$(6.26) \quad (\nabla_Y \nabla_X C)_{T=1} = \mathcal{E}(\underline{n}_X \underline{n}_Y - \underline{n}_{X \cap Y}).$$

Dans le cas de distributions continues des sauts, le problème des boucles devient banal : leur probabilité est alors nulle. Nous pouvons cependant considérer des problèmes apparentés, par exemple la question de savoir combien de fois un parcours repasse dans des positions situées à moins d'une distance donnée de l'une de ses positions précédentes, sans prétendre que ceci puisse être utilisé pour la solution du problème généralisé des boucles (« quasi-boucles »), problème lui-même auquel il semble assez difficile de donner une forme précise.

Nous supposons ici que E est un espace euclidien ou plus généralement un espace métrique, dans lequel deux positions x et y ont une distance, désignée par r^{xy} . Plus généralement nous pouvons prendre une fonction quelconque des deux positions f^{xy} (mis à part les conditions d'intégrabilité), et former l'opérateur

$$(6.27) \quad \square_{(f)} \stackrel{\text{def}}{=} \iint f^{xy} \nabla_{dx} \nabla_{dy}$$

qui devient dans le cas d'un ensemble fini E le laplacien généralisé

$$(6.28) \quad \square_{(f)} = \sum_i \sum_j f^{ij} \frac{\partial}{\partial T^i} \frac{\partial}{\partial T^j},$$

où les f^{ij} sont des *constantes* (indépendantes des T^i). Sous des conditions suffisantes de régularité, l'ordre des dérivations peut être permuté, de sorte que (6.27) s'évanouit alors identiquement pour un f antisymétrique $f^{xy} = -f^{yx}$. Par suite nous nous limiterons à f symétrique $f^{xy} = f^{yx}$.

Pourvu que la permutation des intégrations soit permise, (6.26) donne

$$(6.29) \quad \left(\frac{1}{2} \square_{(f)} C \right)_{T=1} = \frac{1}{2} \mathcal{E} \iint \underline{n}_{dx} \underline{n}_{dy} f^{xy} - \frac{1}{2} \mathcal{E} \int \underline{n}_{dx} f^{xx}.$$

Ici les intégrales après le second et le premier signe espérance deviennent respectivement pour un parcours de longueur $\underline{n} - 1$, passant

respectivement $\underline{n}_1, \dots, \underline{n}_r$ fois par r différents points $\underline{x}_1, \dots, \underline{x}_r$

(de sorte que $\underline{n}_X = \sum_{i=1}^r \underline{n}_i \mathbf{1}_X^{x_i}$),

$$(6.30) \quad \int \underline{n}_{dx} f^{xx} = \sum_{i=1}^r \underline{n}_i f^{x_i x_i}$$

$$(6.31) \quad \iint \underline{n}_{dx} \underline{n}_{dy} f^{xy} = \sum_{i=1}^r \sum_{j=1}^r \underline{n}_i \underline{n}_j f^{x_i x_j}$$

Par suite, le membre de droite de (6.29) est l'espérance de

$$(6.32) \quad \underline{S}_{(f)} \stackrel{\text{def}}{=} \frac{1}{2} \sum_{i \neq j} \underline{n}_i \underline{n}_j f^{x_i x_j} + \sum_i \frac{1}{2} \underline{n}_i (\underline{n}_i - 1) f^{x_i x_i}.$$

Cette quantité est la somme des $f^{x_i x_j}$ ($i, j = 1, 2, \dots, n$) sur toutes les paires $i \neq j$, plus la somme des $f^{x_i x_i}$ sur tous les points multiples, comptés comme des points doubles multiples comme plus haut.

Le second terme s'évanouit identiquement : 1° si $f^{xx} = 0$ pour tout x , par exemple si f^{xy} est la distance r^{xy} de x et y ; 2° si $\underline{n}_i$ prend seulement les valeurs 0 et 1 (sauf pour une probabilité nulle).

Dans le dernier cas, qui arrive toujours si les probabilités de transition ont des distributions continues, nous pouvons omettre les $\underline{x}_i$ avec $\underline{n}_i = 0$, de sorte que $\underline{n}_i = 1$ pour tout i et (6.32) devient

$$(6.33) \quad \underline{S}_{(f)} = \frac{1}{2} \sum_{i \neq j} f^{x_i x_j},$$

où les $\underline{x}_i$ sont les $\underline{n}$ ($= r$) différents points par lesquels passe le parcours. Donc dans ce cas l'opérateur (6.29) donne l'espérance de la somme des valeurs de f pour toutes les paires formées avec les différents points du parcours.

Si en particulier f^{xy} est la distance r^{xy} de x et y , c'est $\frac{1}{2} \underline{n} (\underline{n} - 1)$ fois la distance moyenne, prise sur le parcours, de deux positions par lesquelles passe le point mobile.

Si, au lieu de $f^{xy} = r^{xy}$, nous prenons pour f la fonction caractéristique d'une relation $R(x, y)$ entre deux positions (par exemple $R(x, y) \stackrel{\text{def}}{=} (r^{xy} \leq a)$, où a est un nombre non négatif), c'est-à-dire

$$(6.34) \quad f^{xy} = \begin{cases} 1 & \text{si } R(x, y) \text{ est satisfait,} \\ 0 & \text{dans le cas contraire,} \end{cases}$$

alors (6.33) devient la moitié du nombre de paires ordonnées des différentes positions dans un parcours pour lesquelles $R(x, y)$ est satisfait (par exemple qui sont à une distance $\leq a$). Dans le cas d'une relation symétrique c'est le nombre de paires non ordonnées.

Dans le cas markovien stationnaire, nous pouvons encore calculer aisément le membre de gauche de (6.29). Comme précédemment nous obtenons pour un f^{xy} symétrique

$$\frac{1}{2} \square_{(f)} C = (1 - B) \sum_{k, l, m} B^{k+l+m} \iint \langle (TP)^k \rangle_{dx} \langle P(TP)^{l-1} \rangle_{dy}^x f^{xy} \langle (PT)^m \rangle_y.$$

Pour $T^x = 1$, ceci devient, puisque $\langle PT \rangle^y$ est alors aussi égal à 1,

$$(6.35) \quad \frac{1}{2} (\square_{(f)} C)_{T=1} = (1 - B) \sum_{k, l, m} B^{k+l+m} \iint \langle P^k \rangle_{dx} \langle P^l \rangle_{dy}^x f^{xy}.$$

Dans le cas d'un processus invariant donné par (2.8), ceci se simplifie encore davantage, si $f^{x,y} = f^{x-y} = f^{y-x}$.

Nous avons alors

$$\iint \langle P^k \rangle_{dx} \langle P^l \rangle_{dy}^x f^{x,y} = \int P_{dz}^{(k)} \int P_{dv}^{(l)} \int \mathbf{I}_{dx}^z \int \mathbf{I}_{dy}^{x+v} f^{y-x} = \int P_{dv}^{(l)} f^v,$$

($P_z^{(l)}$ étant le $l^{\text{ème}}$ produit de convolution de p_z), puisque $\int \mathbf{I}_{dx}^z = 1$ et $\int P_{dz}^{(h)} = 1$. Donc le coefficient de $(1 - B) B^n$ devient, puisque l doit être ≥ 1 ,

$$(6.36) \quad \mathcal{E}_{(n)} \frac{1}{2} \sum_{i \neq j} f^{x_i x_j} = \sum_{l=1}^n l (n - l + 1) \int P_{dv}^{(l)} f^v.$$

Comme auparavant, cette relation peut être aisément établie d'une façon élémentaire.

Si f^v a une transformée de Fourier $\chi(t)$

$$(6.37) \quad f^v = \int e^{ivt} \chi(t) dt,$$

(6.36) devient à cause de $\int P_{dv}^{(l)} e^{ivt} = (\varphi(t))^l$, où $\varphi(t) \stackrel{\text{def}}{=} \int P_{dv} e^{ivt}$,

$$(6.38) \quad \mathcal{E}_{(n)} \frac{1}{2} \sum_{i \neq j} f^{x_i x_j} = \sum_{l=1}^n l (n - l + 1) \int \varphi(t)^l \chi(t) dt \\ = \int \frac{n - (n+1)\varphi(t) + \varphi(t)^{n+1}}{(1 - \varphi(t))^2} \varphi(t) \chi(t) dt.$$

Pour p et f donnés, donc φ et χ , (6.38) donne l'espérance cherchée sous la forme d'une intégrale.

APPENDICE.

7. Fonctions de première et de seconde espèce. — D'après Kolmogoroff, un champ de probabilité sur un ensemble E est donné au moyen d'une fonction complètement additive d'ensemble sur E , fonction définie et non négative pour tous les ensembles appartenant à un σ -corps σ_E de sous-ensembles de E et prenant la valeur 1 sur E .

Plus généralement, au lieu d'un σ -corps, considérons un δ -corps (nous nous conformons à la terminologie employée par Hahn-Rosenthal), c'est-à-dire un système δ_E de sous-ensembles de E qui a les propriétés suivantes :

1° δ_E est un corps [c'est-à-dire que s'il contient deux ensembles X et Y il contient leur union (réunion) $X \cup Y$, leur différence (leurs différences) que nous pouvons désigner par $X - Y$ et $Y - X$ respectivement, par conséquent aussi leur intersection (partie commune) $X \cap Y$];

2° δ_E contient l'intersection $\bigcap_1^\infty X_n$ de toute suite $\{X_n\}$ d'ensembles X_n qu'il contient; de plus on suppose pour simplifier que

3° $E \in \sigma_E$, où $\sigma_E \stackrel{\text{def}}{=} \text{le plus petit } \sigma\text{-corps contenant } \delta_E$, c'est-à-dire que E est l'union d'une suite d'ensembles $\{E_n\}$ appartenant à δ_E .

Il en résulte qu'en remplaçant E_n par $\bigcup_1^n E_k$ que nous pouvons supposer sans restriction que

$$(7.1) \quad E_{n+1} \supset E_n, \quad E_n \in \delta_E \text{ pour tout } n.$$

De plus, 1° et 2° impliquent que l'union $X = \bigcup_1^\infty X_n$ d'une suite d'ensembles $\{X_n\}$ appartenant à δ_E appartient aussi à δ_E si (et seulement si) tous les X_n sont contenus dans un $Y \in \delta_E$. En effet, dans ce cas, $X \subset Y$ et

$$X = Y - \bigcap_1^\infty (Y - X_n).$$

Par conséquent δ_E est un σ -corps si et seulement si $E \in \delta_E$.

L'emploi de δ -corps au lieu de σ -corps présente l'avantage d'éviter l'emploi des « valeurs » de la fonction $\pm \infty$ (cf., par exemple, Hahn-Rosenthal, Halmos, etc.).

Généralisant la terminologie introduite précédemment (D. van Dantzig, 1935), nous appellerons *fonction de deuxième espèce*, toute fonction réelle *définie* et *complètement additive* sur δ_E . Nous dénoterons par F_X la valeur que F prend sur l'ensemble X . Si nous appelons une suite d'ensembles $X_n \in \delta_E$ qui sont mutuellement disjoints et qui ont X pour union, une *partition* de X et si nous dénotons ceci par $\{X_n\} \in D(X)$, par conséquent

$$(7.2) \quad \{X_n\} \in D(X) \stackrel{\text{def}}{=} \bigcup X_n = X \quad \text{et} \quad X_m \cap X_n = \emptyset \quad \text{si} \quad m \neq n$$

($\emptyset$ étant l'ensemble vide), nous aurons alors

$$(7.3) \quad \{X_n\} \in D(X) \rightarrow \sum_n F_{X_n} = F_X.$$

La fonction valeur absolue de F qui est elle-même une fonction de seconde espèce est dénotée par $|F|$ et définie par

$$(7.4) \quad |F|_X \stackrel{\text{def}}{=} \sup_{\{X_n\} \in D(X)} \sum_n |F_{X_n}|.$$

D'autre part, nous définirons une *fonction de première espèce* comme une fonction f définie pour tout $x \in E$ et *bornée et δ -mesurable* sur tout $X \in \delta_E$. Par conséquent, si nous dénotons par f^x [au lieu de la notation habituelle $f(x)$] la valeur que f prend sur x et par $\text{Ens}\{x \in E \mid A(x)\}$ l'ensemble de tous les $x \in E$ pour lesquels la proposition $A(x)$ est valable nous aurons :

$$(7.5) \quad 1^\circ \quad X \in \delta_E \rightarrow \|f\|^X < \infty,$$

où

$$(7.6) \quad \|f\|^X \stackrel{\text{def}}{=} \sup_{x \in X} \|f^x\|$$

et

$$(7.7) \quad 2^\circ \quad \text{Ens}\{x \in X \mid f^x \in C\} \in \delta_E$$

pour tout C réel et $X \in \delta_E$.

Les fonctions d'ensemble $\|f\|^X$ ne sont pas additives, par conséquent ce ne sont pas des fonctions de deuxième espèce.

On remarque que $|F|_X$ et $\|f\|^X$ sont finies si $X \in \delta_E$ et peuvent être considérées comme normes de F et f sur X si X est fixé. En particulier si $|F|_X$ et $\|f\|^X$ sont bornées sur δ_E alors aussi F_X et f^X sont bornées et l'on peut étendre leur définition et celle de F_X à tout X appartenant à un σ -corps sur E à savoir au plus petit σ -corps σ_E sur E qui contient δ_E . Quand nous rencontrerons des fonctions bornées (en particulier des probabilités), nous supposerons que leur définition a été étendue de cette manière. Dans ce cas $|F|_E$ et $\|f\|^E$ sont aussi finies et ont les propriétés ordinaires des normes de F et f sur E et nous omettrons souvent les suffixes E .

Plus généralement on peut admettre que les valeurs de F et f sont des nombres complexes et même dans certains cas sont pris dans des espaces de Banach duaux (arbitraires)

L'intégrale de f par rapport à F existe sur tout sous-ensemble $X \in \delta_E$ et sera dénotée par

$$\int_X F_{dx} f^x$$

ou plus brièvement par $(Ff)_X$. C'est, d'après les théorèmes bien connus, une fonction de seconde espèce satisfaisant à l'inégalité

$$(7.8) \quad |(Ff)|_X \leq |F|_X \|f\|^X$$

pour tout $X \in \delta_E$. De même, d'après des théorèmes connus, nous avons

$$(7.9) \quad |(Ff) - (Gg)|_X \leq |F - G|_X \|f\|^X + |G|_X \|f - g\|^X,$$

inégalité qui est souvent appliquée. En particulier, si nous désignons par $\text{var}_X f$ la variation de f sur X , c'est-à-dire

$$(7.10) \quad \text{var}_X f \stackrel{\text{def}}{=} \sup_{x \in X} f^x - \inf_{x \in X} f^x$$

et pour toute partition $\{X_n\}$ de X

$$(7.11) \quad \text{var}_{\{X_n\}} f \stackrel{\text{def}}{=} \sup_n \text{var}_{X_n} f,$$

nous avons si $\{X_n\} \in D(X)$, $x_n \in X_n$ pour tout n

$$(7.12) \quad |(Ff)_X - \sum_n F_{X_n} f^{x_n}| \leq |F|_X \text{var}_{\{X_n\}} f.$$

Ceci résulte de (7.9) avec $G = F$ et $g^x = f^{x_n}$ si $x \in X_n$.

8. Matrices. — Soient E et E' deux ensembles arbitraires sur chacun desquels des δ -corps δ_E et $\delta_{E'}$ de sous-ensembles sont donnés satisfaisant les conditions du paragraphe 7.

Nous considérons maintenant une fonction P [que nous appellerons une « matrice généralisée » ou, plus explicitement, une matrice (E, E') ,] déterminant un nombre réel (occasionnellement un nombre complexe) P_X^Y comme fonction de 1° un élément $Y \in E'$ et 2° un sous-ensemble $X \in \delta_E$ sujet aux conditions :

1° pour un $X = A \in \delta_E$ fixé, P_A^Y détermine une fonction de première espèce sur E' (fonction désignée par P_A) ;

2° pour $Y = b \in E'$ fixé P_X^b détermine une fonction (désignée par P^b) de deuxième espèce sur E .

Si, en particulier, E est dénombrable (fini ou énumérable) et si δ_E contient tout ensemble se réduisant à un seul élément, on a

$$(8.1) \quad P_X^Y = \sum_{x \in X} P_x^Y$$

et si E' est aussi dénombrable P_X^Y est déterminé par la matrice ordinaire rectangulaire finie ou infinie P_x^Y .

La norme pour Y fixé est donnée par

$$(8.2) \quad |P^Y|_X = \sup_{\{X_n\} \in D(X)} \sum |P_{X_n}^Y|$$

qui est elle-même une matrice. Sa norme pour X fixé est donnée par

$$(8.3) \quad \|P\|_X^Y = \sup_{Y \in E'} |P^Y|_X.$$

Ce n'est plus une matrice dans le sens donné ci-dessus, Y étant un ensemble, et non plus un élément de E' et $\|P\|_X^Y$ n'étant d'ailleurs pas complètement additive sur X . La matrice P est dite *bornée* si $\|P\|_X^Y$ est bornée pour $X, Y \in \delta_E$. Dans ce cas on peut définir P_X^Y pour tout $X \in \sigma_E$ et

$$(8.4) \quad \sup_{X, Y} \|P\|_X^Y = \|P\|_E^{E'} < \infty.$$

Au lieu du symbole $\|P\|_E^{E'}$, nous utiliserons aussi le symbole $\|P\|$. En particulier, les ensembles E et E' peuvent coïncider. Dans ce cas, nous

supposerons que les corps δ_E et $\delta_{E'}$ coïncident aussi. Nous avons ici comme cas particulier la matrice unité désignée par $\mathbf{I}$ (iota) et définie par

$$(8.5) \quad \mathbf{I}_X^x \stackrel{\text{def}}{=} \begin{cases} 1 & \text{si } x \in X, \\ 0 & \text{s'il n'en est pas ainsi.} \end{cases}$$

pour un $X = A$ fixé $\mathbf{I}_A$ est la fonction caractéristique de l'ensemble A ; pour un α fixé, $\mathbf{I}_\alpha$ peut être appelée la « fonction caractéristique de seconde espèce d'un point α »; elle prend la valeur 1 sur tout ensemble contenant α , et 0 sur tout autre ensemble. Évidemment $\mathbf{I}$ est bornée avec $\|\mathbf{I}\| = 1$, et

$$(8.6) \quad \mathbf{I}_{X_1}^x \mathbf{I}_{X_2}^x = \mathbf{I}_{X_1 \cap X_2}^x.$$

Plus généralement à tout sous-ensemble $A \in \delta_E$ correspond une matrice aussi dénotée par $\mathbf{I}_A$ et définie par

$$(8.7) \quad (\mathbf{I}_A)_X^x \stackrel{\text{def}}{=} \mathbf{I}_{A \cap X}^x.$$

Dans le cas d'un $E = E'$ fini, elle correspond, d'après (8.1), à une matrice ayant la valeur 1 pour tous les éléments de la diagonale principale dont les suffixes appartiennent à A , et dont tous les autres éléments sont nuls. Évidemment $\mathbf{I}_A$ est aussi bornée et $\|\mathbf{I}_A\| = 1$ à moins que A ne soit vide.

A toute fonction f de première espèce correspond une matrice dénotée par $f\mathbf{I}$ et définie par

$$(8.8) \quad (f\mathbf{I})_X^x \stackrel{\text{def}}{=} f^x \mathbf{I}_X^x = \begin{cases} f^x & \text{si } x \in X, \\ 0 & \text{s'il n'en est pas ainsi.} \end{cases}$$

Dans le cas d'un $E = E'$ dénombrable, la matrice qui correspond à $f\mathbf{I}$ grâce à (8.1) est la matrice diagonale dont l'élément de la diagonale principale qui correspond à x a la valeur f^x . Évidemment $f\mathbf{I}$ est bornée si et seulement si f l'est et l'on a

$$\|f\mathbf{I}\| = |f|^E.$$

Revenons maintenant au cas général où E et E' peuvent être différents. Soit P une matrice comme auparavant et f une fonction de première espèce. Pour tout $X \in \delta_E$ et pour tout $Y \in E'$ fixé, écrivons $Pf\mathbf{I}$ au lieu de $P(f\mathbf{I})$; alors

$$(8.9) \quad (Pf\mathbf{I})_X^Y = \int_X P_{\alpha x} f^x$$

existe et est complètement additive sur X pour $X \in \delta_E$ et est mesurable et bornée en y sur tout $Y \in \delta_{E'}$; d'ailleurs

$$(8.10) \quad \|Pf\|_X^Y \leq \|P\|_X^Y \|f\|_X^X,$$

de sorte que Pf est une matrice. En particulier,

$$(Pf)^y = (Pf\mathbf{1})_E^y$$

si elle existe pour tout y , est une fonction mesurable sur E' et une fonction de première espèce sur E' quand $(Pf)^y$ est bornée sur tous les $y \in \delta_{E'}$, quand P et f sont bornées.

D'autre part, soit F une fonction de deuxième espèce sur E' . Alors

LEMME 1. — Pour tout $X \in \delta_{E'}$, $Y \in \delta_E$,

$$(FP)_{Y,X} \stackrel{\text{def}}{=} \int_Y F_{dy} P_X^y$$

existe et est une fonction de deuxième espèce sur δ_E et sur $\delta_{E'}$.

Démonstration. — La complète additivité par rapport à Y sur δ_E , pour X constant résulte de celle pour $\int_Y F_{dy} f^y$. Pour Y constant, $(FP)_{Y,X}$ est simplement additive sur X . Pour l'additivité complète il est dès lors suffisant de prouver que $\lim_{n \rightarrow \infty} (FP)_{Y,X_n} = 0$ si X_n est une suite décroissante d'ensembles $X_n \in \delta_E$ dont l'intersection est vide.

Choisissons un $\varepsilon > 0$ et posons

$$(8.11) \quad B_n = \text{Ens } \{y \in Y \mid |P^y|_{X_n} \geq \varepsilon\},$$

à cause de la mesurabilité de $P_{X_n}^y$, donc de $|P|_{X_n}^y$ pour tout $X_n, B_n \in \delta_{E'}$. Alors, comme $|P^y|$ est complètement additive et ≥ 0 pour tout $y \in E'$ et comme $X_{n+1} \subset X_n$, nous avons aussi $B_{n+1} \subset B_n$. De plus, $\bigcap_1^\infty B_n = 0$, puisque, s'il existait un $y \in \bigcap_1^\infty B_n$, alors pour cet y $|P^y|_{X_n} \geq \varepsilon$ pour tout n , ce qui est contraire à la complète additivité de $|P^y|$ puisque $\bigcap_1^\infty X_n = 0$.

De même, pour tout n suffisamment grand $|F|_{B_n} \leq \varepsilon$. Alors si C_n est

le complément de B_n par rapport à Y : $C_n = Y - B_n$, alors

$$(FP)_{Y, X_n} = (FP)_{B_n, X_n} + (FP)_{C_n, X_n}.$$

Mais

$$|(FP)_{B_n, X_n}| \leq \int_{B_n} |F|_{dy} |P^y|_{X_n} \leq |F|_{B_n} \|P\|_{X_n}^{B_n} \leq \varepsilon C_1,$$

avec $C_1 = \|P\|_{X_n}^{B_n}$. Et

$$|(FP)_{C_n, X_n}| \leq \int_{C_n} |F|_{dy} |P^y|_{X_n} \leq |F|_{C_n} \|P\|_{X_n}^{C_n} \leq C_2 \varepsilon,$$

avec $C_2 = |F|_Y$, comme $C_n \subset Y$ et $\|P\|_{X_n}^{C_n} = \sup_{y \in C_n} |P^y|_{X_n} \leq \varepsilon$, d'après la définition de B_n et de C_n . Par conséquent, $|(FP)_{Y, X_n}| \leq (C_1 + C_2) \varepsilon$ pour tout n suffisamment grand, $\varepsilon > 0$ étant donné, c'est-à-dire

$$\lim_{n \rightarrow \infty} (FP)_{Y, X_n} = 0. \quad \text{C. Q. F. D.}$$

On en déduit aisément que pour tout X, Y

$$(8.12) \quad |(FP)_{Y, X}| \leq |F|_Y \|P\|_X^Y.$$

Si, en particulier, F et P sont bornées, FP l'est aussi.

Si F et P sont non bornées et si FP existe, alors FP est complètement additive, pourvu que FP existe absolument, c'est-à-dire

$$\int_{E'} |F|_{dy} |P^y|_X < \infty.$$

COROLLAIRE. — Soient E, E', E'' des ensembles arbitraires avec les δ -corps correspondants $\delta_E, \delta_{E'}, \delta_{E''}$; supposons que x, y, z et X, Y, Z appartiennent à E, E', E'' et $\delta_E, \delta_{E'}, \delta_{E''}$ respectivement. Si P et Q sont des matrices (E, E') et (E', E'') respectivement, alors $\int_Y Q_{dy}^z P_X^y$ existe pour tout $X \in \delta_E, Y \in \delta_{E'}, Z \in \delta_{E''}$ et détermine pour Y fixé une matrice (E, E'') .

Si Q et P sont bornées, on peut définir le produit généralisé de ces matrices

$$(8.13) \quad (QP)_X^z \stackrel{\text{def}}{=} \int_{E'} Q_{dy}^z P_X^y.$$

On peut étendre cette définition au cas où Q et P sont non bornées,

pourvu qu'elles satisfassent la condition suivante

$$(8.14) \quad \int_{E'} |Q^z|_{dy} |P^y|_x \text{ existe pour tous les } z \in E'' \text{ et } X \in \delta_E$$

et est bornée sur tout $z \in \delta_{E''}$ pour chaque $X \in \delta_E$.

Si cette condition est satisfaite, QP aussi est une matrice.

Des inégalités (8.10) et (8.12), nous déduisons que

LEMME 2. — Une matrice bornée P transforme les fonctions bornées de première espèce sur E , f en de telles fonctions sur E' et les fonctions bornées de seconde espèce sur $\delta_{E'}$ en de telles fonctions sur δ_E ; on a

$$(8.15) \quad \|Pf\| \leq \|P\| \cdot \|f\|,$$

$$(8.16) \quad \|FP\| \leq \|F\| \cdot \|P\|$$

(où les indices E et E' ont été omis).

LEMME 3. — (Voir aussi Herbert Robbins, 1948). — Si f et F sont des fonctions de première et de seconde espèce sur E et $\delta_{E'}$ respectivement et si P est une matrice, alors, si $Z \in \delta_E$ et $Y \in \delta_{E'}$,

$$(8.17) \quad \int_X \left\{ \int_Y F_{dy} P_{dx}^y \right\} f^x = \int_Y F_{dy} \int_X P_{dx}^y f^x,$$

c'est-à-dire que l'ordre des intégrations peut être inversé.

Démonstration. — Si nous posons pour des $U \in \delta_E$, $V \in \delta_{E'}$ arbitraires

$$(8.18) \quad F'_V \stackrel{\text{def}}{=} F_{Y \cap V}, \quad f'^x \stackrel{\text{def}}{=} f^x \mathbf{1}_X^x, \quad P'_{\mathcal{U}} \stackrel{\text{def}}{=} \mathbf{1}_Y^y P_{X \cap U}^y,$$

alors $F'P'$ et f' sont bornées et (8.17) est équivalent à

$$(8.19) \quad \int_E \left\{ \int_{E'} F'_{dy} P'_{dx}^y \right\} f'^x = \int_{E'} F'_{dy} \int_E P'_{dx}^y f'^x,$$

par conséquent il est suffisant de prouver le théorème pour des fonctions et matrices bornées, les intégrations étant étendues sur la totalité des ensembles E et E' (que nous omettrons dès lors). Nous omettrons les primes et écrirons (8.19) sous la forme abrégée

$$(8.20) \quad (FP)f = F(Pf).$$

Posons alors

$$(8.21) \quad Pf \stackrel{\text{def}}{=} g, \quad \text{c'est-à-dire} \quad g^y \stackrel{\text{def}}{=} \int_E P_{dx}^y f^x,$$

$$(8.22) \quad FP \stackrel{\text{def}}{=} G, \quad \text{c'est-à-dire} \quad G \stackrel{\text{def}}{=} \int_{E'} F_{dy} P_{X'}^y.$$

Nous avons à montrer que

$$(8.23) \quad Gf = Fg, \quad \text{c'est-à-dire} \quad \int_E G_{dx} f^x = \int_{E'} F_{dy} g^y.$$

Soit $\varepsilon > 0$ donné. Comme f est borné on peut trouver une partition finie $\{X_n\} \in D(E)$, où $n \leq N$ et $\text{var}\{x_n\} f \leq \varepsilon$.

Divisons, par exemple, l'intervalle fini $(-|f|^E, +|f|^E)$ en N sous-intervalles disjoints I_n et prenons $X_n \stackrel{\text{def}}{=} \text{Ens}\{x \in E \mid f^x \in I_n\}$. Alors en choisissant les $x_n \in X_n$ arbitrairement on a, d'après (7.12),

$$(8.24) \quad \left| \int G_{dx} f^x - \sum_n G_{X_n} f^{x_n} \right| \leq |G| \varepsilon.$$

Comme chacune des N fonctions P_{X_n} est bornée, nous pouvons trouver une partition finie de E' (obtenue, par exemple, en prenant l'intersection des partitions appartenant séparément aux P_{X_n}) $\{Y_m\} \in D(E')$, $m \leq M$ avec $\text{var}\{y_m\} P_{X_n} \leq \varepsilon N^{-1}$ pour tout $n \leq N$. Par conséquent, si nous choisissons $y_m \in Y_m$ arbitrairement, nous avons de nouveau d'après (7.12)

$$(8.25) \quad \left| G_{X_n} - \sum_m F_{Y_m} P_{X_n}^{y_m} \right| \leq |F| \varepsilon N^{-1}$$

et, d'après (8.24) et

$$(8.26) \quad \left| \int G_{dx} f^x - \sum_n \sum_m F_{Y_m} P_{X_n}^{y_m} f^{x_n} \right| \leq C_1 \varepsilon,$$

avec

$$C_1 \stackrel{\text{def}}{=} |F| (\|P\| + \|f\|).$$

D'autre part,

$$(8.27) \quad \left| \int P_{dx}^y f^x - \sum_n P_{X_n}^y f^{x_n} \right| \leq \|P\| \varepsilon$$

pour tout $y \in E'$, en particulier pour $y = y_m$.

Par conséquent,

$$(8.28) \quad \left| \int F_{dy} g^y - \sum^n \left(\int F_{dy} P_{X_n}^y \right) f^{x_n} \right| \leq \|F\| \|P\| \varepsilon.$$

Mais

$$(8.29) \quad \left| F_{dy} P_{X_n}^y - \sum^m F_{Y_m} P_{X_n}^{Y_m} \right| \leq \|F\| \varepsilon N^{-1}$$

puisque $\text{var}_{\{Y_m\}} P_{X_n} \leq \varepsilon N^{-1}$. Donc, comme $n \leq N$, (8.28) et (8.29) donnent

$$(8.30) \quad \left| \int F_{dy} g^y - \sum^n \sum^m F_{Y_m} P_{X_n}^{Y_m} f^{x_n} \right| \leq C_1 \varepsilon,$$

de (8.26) et (8.30) nous déduisons que

$$(8.31) \quad \left| \int G_{dx} f^x - \int F_{dy} g^y \right| \leq 2 C_1 \varepsilon.$$

Par conséquent, comme cette inégalité reste valable pour tout $\varepsilon > 0$, on a (8.23) en tenant compte de (8.19) et (8.17).

C. Q. F. D.

Cherchons une condition suffisante pour que (8.17) soit valide sur la totalité des ensembles E et E' quand F , P et f peuvent être non bornées et prouvons dès lors en faisant un raisonnement bien connu que

LEMME 4. — Si

$$(8.32) \quad H_X \stackrel{\text{def}}{=} \int_{E'} |F|_{dy} |P^y|_X,$$

$$(8.33) \quad h^y \stackrel{\text{def}}{=} \int_E P^y |f^x|_{dx}$$

existent pour tout X et y respectivement et si, de plus, une des deux inégalités suivantes :

$$(8.34) \quad \int_E H_{dx} |f^x| < \infty,$$

$$(8.35) \quad \int_{E'} |F|_{dy} h^y < \infty$$

est valide, alors la seconde de ces deux inégalités l'est aussi et

$$(8.36) \quad \int_E \left\{ \int_{E'} F_{dy} P_{dx}^y \right\} f^x = \int_{E'} F_{dy} \int_E P_{dx}^y f^x.$$

Démonstration. — Il est suffisant de prouver le théorème pour des F , P et f non négatives auquel cas $H = G$ et $h = g$. D'après le lemme 3, (8.17) est valable pour $X = E_k$, $Y = E_l$ avec k et l arbitraires [cf. (7.1)]. Dès lors, si (8.34) est valable, le second membre de (8.36) est

$$\int_E G_{dx} f^x \geq \int_{E_k} G_{dx} f^x \geq \int_{E_k} \left\{ \int_{E_l} F_{dy} P_{dx}^y \right\} f^x = \int_{E_l} F_{dy} \int_{E_k} P_{dx}^y f^x.$$

Comme la dernière expression est non décroissante si $k \rightarrow \infty$, $l \rightarrow \infty$, et bornée elle a une limite qui est le second membre de (8.36), de sorte que son existence a été prouvée, c'est-à-dire (8.35) et que

$$\int_E G_{dx} f^x \geq \int_{E'} F_{dy} g^y.$$

De même on montre que

$$\int_{E'} F_{dy} g^y \geq \int_E G_{dx} f^x$$

à cause de (8.36).

Démonstration analogue dans le second cas, c'est-à-dire si l'on suppose (8.35).

LEMME 5 (voir aussi R. G. Cooke, 1950, p. 29). — Si E_1, E_2, E_3, E_4 sont des ensembles arbitraires auxquels correspondent les δ -corps $\delta_{E_1}, \delta_{E_2}, \delta_{E_3}, \delta_{E_4}$, si t, x, y, z et T, X, Y, Z désignent des éléments arbitraires et des sous-ensembles de E_1, E_2, E_3, E_4 et $\delta_{E_1}, \delta_{E_2}, \delta_{E_3}, \delta_{E_4}$ respectivement et si M, P, Q sont des matrices $(E_1, E_2), (E_2, E_3)$ et (E_3, E_4) respectivement, alors

$$(8.37) \quad \int_X \left\{ \int_Y \varphi_{dy}^z P_{dx}^y \right\} M_T^x = \int_Y \varphi_{dy}^z P_{dx}^y M_T^x.$$

Si, en particulier, M, P et Q sont bornées ou si

$$(8.38) \quad \begin{cases} \int_E |P^y|_{dx} |M^x|_T = K_T^y, & \int_{E'} |\varphi^z|_{dy} |P^y|_X = L_X^z, \\ \int_X |L_{dx}^z| |M^x|_T < \infty & \text{ou} \quad \int_{E'} |\varphi^z|_{dy} K_T^y < \infty, \end{cases}$$

alors (8.37) est valable pour $X = E_2$ et $Y = E_3$; c'est-à-dire on peut étendre l'intégration sur la totalité des ensembles E_2 et E_3 .

Démonstration. — Elle résulte immédiatement des lemmes 3 et 4

avec, pour z et T constant,

$$F_Y = Q_Y^z \quad \text{et} \quad f^x = M_T^x.$$

Si les produits QP et PM sont aussi des matrices, par exemple, si Q , P et M sont bornées ou si QP et PM satisfont la condition (8.14) alors le produit de matrices est associatif

$$(8.39) \quad (QP)M = Q(PM),$$

c'est-à-dire

$$(8.40) \quad \int_{E_1} (QP)_{dx}^z M_T^x = \int_{E_1} Q_{dy}^z (PM)_T^y.$$

De ce lemme 5, on déduit que le calcul matriciel ordinaire peut être appliqué en employant l'intégration de Lebesgue-Stieltjes-Radon au lieu de la sommation pourvu que toutes les matrices et fonctions des deux espèces dont on s'occupe soient bornées ou plus généralement si elles satisfont (8.32), (8.33), (8.34) et (8.38) et si les produits de matrices satisfont (8.14).

Pour plus de simplicité nous omettrons désormais (sauf indication spéciale) le second cas, c'est-à-dire que nous nous restreindrons au cas de matrices et de fonctions des deux espèces qui sont bornées, et que nous supposons de plus que les ensembles E , E' , E_1 , E_2 , ..., coïncident tous. Dans ce cas les suites croissantes E_k peuvent aussi être omises puisque nous pourrions prendre $E_k = E$ pour tout k .

Si donc, P est une matrice carrée bornée, c'est-à-dire si $E = E'$ et

$$(8.41) \quad \|P\| = \|P\|_E^E = \sup_{x \in E} \sup_{\{x_n\} \in D(E)} \sum_n |P_{x_n}^x| < \infty,$$

alors on peut former des puissances de P et des polynômes et tant que les coefficients sont constants (scalaires), ces polynômes sont commutables. De (8.15), (8.16), (8.13) (8.41) on déduit aisément que

$$(8.42) \quad \|PQ\| \leq \|P\| \cdot \|Q\|.$$

Notons pour usage ultérieur les identités suivantes :

$$(8.43) \quad F\mathbf{I} = F, \quad \text{c'est-à-dire} \quad \int F_{dx} \mathbf{I}_X^x = F_X,$$

$$(8.44) \quad \mathbf{I}f = f, \quad \text{c'est-à-dire} \quad \int \mathbf{I}_{dx}^x f^x = f^x,$$

$$(8.45) \quad \mathbf{I}P = P\mathbf{I} = P, \quad \text{c'est-à-dire} \quad \int \mathbf{I}_{dy}^x P_X^y = \int P_{dy}^x \mathbf{I}_X^y = P_X^x.$$

De plus, remarquons que deux matrices diagonales sont commutables

$$(8.46) \quad \int (f\mathbf{I})_{dy}^x (g\mathbf{I})_X^y = \int (g\mathbf{I})_{dy}^x (f\mathbf{I})_X^y = f^x g^x \mathbf{I}_X^x.$$

Nous ne pouvons pas toutefois remplacer le symbole $f\mathbf{I}$ pour la matrice diagonale $f^x \mathbf{I}_X^x$ par $\mathbf{I}f$ qui désigne la fonction de première espèce (8.44). Si $f\mathbf{I}$ est suivi par un autre facteur (matricé ou fonction de première espèce), nous pouvons omettre $\mathbf{I}$, ainsi $f\mathbf{I}P = fP$ avec $(fP)_X^x = f^x P_X^x$ et, de même

$$(fg)^x = (f\mathbf{I}g)^x = f^x g^x.$$

Les intégrales sur les sous-ensembles de E peuvent être exprimées au moyen de matrices unité « partielles » $\mathbf{I}_A$ [cf. (8.7)]

$$(8.47) \quad \int_X F_{dx} f^x = \int_E F_{dy} \int_E \mathbf{I}_{dx \cap X}^y f^x = F\mathbf{I}_X f.$$

Les matrices unitaires partielles servent aussi à tronquer des fonctions

$$(8.48) \quad (\mathbf{I}_A f)^x = \mathbf{I}_A^x f^x = \begin{cases} f^x & \text{si } x \in A, \\ 0 & \text{si l'en est pas ainsi;} \end{cases}$$

$$(8.49) \quad (F\mathbf{I}_A)_X = F_{A \cap X} = \begin{cases} F_X & \text{si } X \subset A, \\ 0 & \text{si } X \subset E - A \end{cases}$$

et d'une manière analogue $\mathbf{I}_A P \mathbf{I}_B$.

Une condition suffisante pour la convergence de $\sum_1^\infty (P_n)_X^x$ (où les P_n sont des matrices bornées) pour tout x et X est la convergence de $\sum_1^\infty \|P_n\|$.

Il en résulte, en particulier, que si $\|P\| < 1$, alors d'après (8.42) $\|P^n\| < \|P\|^n < 1$, P^n désignant la $n^{\text{ième}}$ puissance de la matrice, $\sum_0^\infty P^n$ converge et est la seule inverse (à droite et à gauche) de $\mathbf{I} - P$

$$(8.50) \quad (\mathbf{I} - P)^{-1} = \sum_0^\infty P^n \quad \text{si } \|P\| < 1.$$

Plus généralement, si la matrice M peut être mise sous la forme

$$(8.51) \quad M = R + Q,$$

où R^{-1} existe (nous utilisons cette expression seulement si $XR = RX = I$ a seulement une solution $X = R^{-1}$), tandis que $\|QR^{-1}\| < 1$ ou $\|R^{-1}Q\| < 1$, alors M^{-1} existe et

$$(8.52) \quad M^{-1} = R^{-1} \sum_0^{\infty} (-QR^{-1})^n = \sum_0^{\infty} (-R^{-1}Q)^n R^{-1}.$$

La démonstration du fait que : les deux séries dans (8.52) sont identiques, convergent si seulement une des conditions d'inégalité est satisfaite, satisfont aux équations pour les inverses, et est leur seule solution, est banale.

Un cas particulier ⁽⁸⁾, bien connu se rencontre si $R = fI$ est une matrice diagonale ; (8.51) et (8.52) deviennent alors

$$(8.53) \quad M = fI + Q,$$

$$(8.54) \quad M^{-1} = \sum_0^{\infty} (-f^{-1}Q)^n f^{-1}I$$

pourvu que $f^x \neq 0$ pour tout x et, par exemple $\|f^{-1}Q\| < 1$, c'est-à-dire

$$(8.55) \quad \|Q^x\|_E = \sup_{\{X_n\} \in D(E)} \sum |Q_{X_n}^x| < |f^x|$$

pour tout x .

9. Processus de Markof invariants par rapport à un groupe transitif.

— Comme généralisation des processus invariants étudiés au paragraphe 2, nous considérons un processus de Markof sur un ensemble E tel qu'il existe un groupe G de transformations $(1,1)$ τ de E dans lui-même qui est transitif sur E . Sur E un σ -corps σ_E , invariant par rapport à toutes les transformations de G est défini ; c'est-à-dire $\tau X \in \sigma_E$ si $X \in \sigma_E$ pour tous les $\tau \in G$, où

$$(9.1) \quad \tau X \stackrel{\text{def}}{=} \text{Ens} \{ \tau x \mid x \in X \}.$$

(Un cas particulier important de ceci est celui où E est une sphère dans un espace à un nombre quelconque de dimensions, G étant le groupe de toutes les rotations de E dans lui-même.) Nous considérons

⁽⁸⁾ Voir aussi R. G. Cooke (1950, p. 31) (2.4, II) et pour le cas des matrices finites Olga Taussky (1949), et œuvres de date plus reculée indiquées dans cet article.

alors une matrice de transition P (par conséquent $P_X^x \geq 0$ pour $x \in E$, $X \in \sigma_E$ et $P_X^x = 1$) qui est invariante par rapport à toutes transformations de G .

$$(9.2) \quad P_X^x = P_{\tau X}^{\tau x}$$

pour tous les $\tau \in G$.

Nous choisissons arbitrairement un point $a \in E$ et définissons H comme le sous-groupe comprenant tous les $\tau \in G$ laissant a invariant

$$(9.3) \quad H \stackrel{\text{def}}{=} \text{Ens} \{ \tau \in G \mid \tau a = a \}.$$

En posant pour tout $\tau \in G$

$$(9.4) \quad F_\tau \stackrel{\text{def}}{=} H\tau H = \text{Ens} \{ \tau_1 \tau \tau_2 \mid \tau_1 \in H, \tau_2 \in H \}.$$

Les ensembles F_τ sont mutuellement disjoints et ont G comme réunion. Évidemment

$$\tau \in F_\tau, \quad \tau \in H \Rightarrow F_\tau = H, \quad F_\tau \cap F_{\tau'} \neq 0 \Rightarrow \tau' \in F_\tau \Rightarrow \tau \in F_{\tau'} = F_\tau.$$

En utilisant l'axiome du choix, nous choisissons arbitrairement dans chaque F_τ un point représentatif unique γ_τ ; soit Δ l'ensemble de tous ces points représentatifs [par exemple dans le cas des rotations d'une sphère dans R_3 de centre 0 et de rayon 1, où a est le point (0, 0, 1) et η_1 et η_2 sont des rotations autour de l'axe des z d'angles respectifs φ et ψ , alors que γ est une rotation autour de l'axe des y d'angle θ , φ , ψ et θ étant les angles d'Euler; Δ est l'ensemble de tous les θ avec, par exemple $0 \leq \theta \leq \pi$].

Pour tout $x \in E$, il existe un $\tau \in G$ avec $\tau a = x$; donc un $\gamma \in \Delta$ et η , $\eta' \in H$, avec $\tau = \eta\gamma\eta'$; donc $x = \eta\gamma\eta'a = \eta\gamma a$, puisque $\eta'a = a$, $\eta' \in H$. Si nous avons aussi $x = \eta_1\gamma_1 a$, alors

$$\eta'' = \gamma^{-1}\eta^{-1}\eta_1\gamma_1 \in H, \quad \text{d'où} \quad \gamma_1 = \eta_1^{-1}\eta\gamma\eta'' \in H\gamma H, \quad \text{d'où} \quad \gamma_1 = \gamma$$

puisque les points représentatifs sont uniques, de sorte que γ est déterminé d'une façon unique par $x \in H\gamma a$, c'est-à-dire par $x = \eta\gamma a$ (ce n'est pas le cas en général pour η). Alors (9.2) entraîne, en posant $p_Y \stackrel{\text{def}}{=} P_Y^a$,

$$(9.5) \quad P_X^x = P_X^{\eta\gamma a} = P_Y, \quad \text{avec} \quad Y = \gamma^{-1}\eta^{-1}X$$

p_Y est aussi complètement additif et ≥ 0 avec $p_E = 1$ et aussi [à cause de (9.2)]

$$(9.6) \quad p_Y = p_{\eta Y}$$

pour tout $\eta \in H$ et $Y \in \sigma_E$.

Pour tout $\Gamma \subset \Delta$ tel que $H\Gamma a \in \sigma_E$ (ces Γ forment un σ -corps σ_Δ sur Δ) on pose

$$(9.7) \quad q_\Gamma \stackrel{\text{def}}{=} p_{H\Gamma a}$$

de sorte que q_Γ est complètement additif, et ≥ 0 avec $q_\Delta = 1$.

Alors pour un Y donné et un $\Gamma \in \sigma_\Delta$ variable, $p_{Y \cap H\Gamma a}$ est complètement additif dans Γ , ≥ 0 , et $\leq q_\Gamma$ et peut, par conséquent, être écrit sous la forme

$$(9.8) \quad p_{Y \cap H\Gamma a} = \int_\Gamma q_{a\gamma} l_Y^\gamma,$$

où l_Y^γ peut être choisi ≥ 0 , ≤ 1 et s'annulant sauf pour $\gamma^a \in HY$. Pour tout $Y \in \sigma_E$ fixé la fonction l_Y^γ est définie sur Δ excepté pour un ensemble de mesure q nulle et mesurable σ_Δ .

Nous introduisons maintenant :

Hypothèse A : Les quantités l_Y^γ peuvent être redéfinies de telle façon qu'elles satisfassent (9.8) et qu'elles soient définies pour tous les $\gamma \in \Delta$ et tous les $Y \in \sigma_E$ comme des fonctions à valeurs uniques de leurs arguments, pour Y fixé mesurables σ_Δ en γ et pour γ fixé complètement additives en Y .

Selon Doob (1948) une condition suffisante pour l'hypothèse A est : E est un ensemble de Borel dans un espace euclidien et σ_E consiste en tous les sous-ensembles de Borel de E . Les l_Y^γ peuvent être interprétés comme des probabilités conditionnelles, à savoir comme la probabilité pour que $x \in Y$ sous la condition $x \in H\gamma a$.

Hypothèse B : H est un groupe topologique compact.

Selon A. Haar (1933) (cf. aussi J. von Neumann, 1934, 1936; L. H. Loomis, 1945), l'hypothèse B entraîne l'existence d'une mesure de Haar, à savoir d'une fonction d'ensemble invariante à gauche, non négative, complètement additive, définie sur le σ -corps σ_H engendré par tous les sous-ensembles ouverts de H , en outre bornée et déterminée d'une façon unique à un facteur de proportionnalité près.

Les hypothèses A et B sont simultanément satisfaites, par exemple si E est une sphère euclidienne à un nombre quelconque de dimensions et G le groupe de toutes les rotations de E dans elle-même.

Hypothèse C : 1° Pour tous les $\zeta \in \Delta$ et $\gamma \in \sigma_E$ constants, $\tilde{l}_{\eta\gamma}^\zeta$ est une fonction de $\eta \in H$ mesurable.

2° Si $\Gamma \in \sigma_\Delta$ et $K \in \sigma_E$, alors $K\Gamma a \in \sigma_E$.

Par ces hypothèses C_1 et C_2 le σ -corps σ_E (donc aussi σ_Δ) est lié avec la topologie sur H .

Si μ_K est la mesure de Haar normalisée sur H , de sorte que $\mu_H = 1$, C_1 entraîne l'existence de

$$\tilde{l}_\gamma^\zeta = \int \mu_{d\eta} \tilde{l}_{\eta\gamma}^\zeta,$$

pour tous les ζ et γ . A cause de (9.6),

$$\sum \mu_{K_i} \tilde{l}_{\eta_i\gamma}^\zeta$$

est une solution de (9.8) de même que $\tilde{l}_\gamma^\zeta$, d'où aussi $\tilde{l}_\gamma^\zeta$. Donc, $\tilde{l}_\gamma^\zeta$ satisfaisant (9.8) et hypothèse A, nous pouvons prendre $\tilde{l}_\gamma^\zeta$ au lieu de $\tilde{l}_{\eta\gamma}^\zeta$ et omettre le trait. Alors nous obtenons

$$(9.9) \quad \tilde{l}_{\eta\gamma}^\zeta = \tilde{l}_\gamma^\zeta \quad \text{pour tout } \eta \in H, \quad j \in \Delta \quad \text{et } \gamma \in \sigma_E.$$

En prenant dans (9.9)

$$\gamma = K\Gamma a, \quad \text{avec } K \in \sigma_H, \quad \Gamma \in \sigma_\Delta, \quad j \in \Gamma,$$

$\tilde{l}_{K\Gamma a}^\zeta$ est, pour K variable, complètement additif, ≥ 0 , borné, invariant par rapport à tous les $\eta \in H$. Donc il a aussi les propriétés d'une mesure de Haar. Donc m_Γ^ζ étant le facteur de proportionnalité

$$(9.10) \quad \tilde{l}_{K\Gamma a}^\zeta = m_\Gamma^\zeta \mu_K = m_\Gamma^\zeta \mathbf{1}_\Gamma^\zeta \mu_K$$

puisque $m_\Gamma^\zeta = 0$ sauf $\zeta \in \Gamma$. De plus $\tilde{l}_{H\Gamma a}^\zeta = \mathbf{1}_\Gamma^\zeta$, d'où $m^\zeta = 1$.

L'extension de $\tilde{l}_{K\Gamma a}^\zeta$ à $\tilde{l}_\gamma^\zeta$ avec γ arbitraire donne

$$(9.11) \quad \tilde{l}_\gamma^\zeta = \int_H \mu_{d\eta} \mathbf{1}_\gamma^{\eta\zeta a}.$$

En substituant dans (9.8) avec $\Gamma = \Delta$, on obtient

$$(9.12) \quad p_\gamma = \int_\Delta q_{d\zeta} \int_H \mu_{d\eta} \mathbf{1}_\gamma^{\eta\zeta a},$$

et finalement par (9.5)

$$(9.13) \quad P_X^x = \int_{\Delta} q_{d\tau} \int_H \mu_{d\eta} \mathbf{I}_X^{\tau_x \eta^x a},$$

où τ_x représente l'un quelconque des $\tau \in G$ avec $\tau_x a = x$. D'un autre côté, il est clair que (9.13) satisfait à nos conditions.

Nous avons donc démontré :

THÉOREME 5. — *Si G est un groupe transitif de transformations de E dans lui-même, si H est le sous-groupe de G laissant invariant un point donné $a \in E$, μ_X la mesure de Haar normalisée sur H et Δ un système de représentations des classes bilatérales modulo H (9.4) de H dans G , alors que q_{Γ} est donné par (9.7) et si les hypothèses A, B, C sont satisfaites, alors toute matrice bornée (E, E) qui est invariante sous G , peut être représentée sous la forme (9.13).*

Si, en particulier, G est simplement transitif sur E , c'est-à-dire si H se compose seulement de l'élément unité, $\Delta = G$ et (9.7) et (9.13) simplifient en $q_{\Gamma} = p_{\Gamma a}$ et

$$(9.14) \quad P_X^x = \int_G q_{d\gamma} \mathbf{I}_X^{\tau_x \gamma^x a}.$$

Si, plus particulièrement, E et G sont identiques, nous pouvons poser $a = 1$, $\tau_x = x$ et nous obtenons $q_{\Gamma} = p_{\Gamma}$ et

$$(9.15) \quad P_X^x = \int_E p_{dy} \mathbf{I}_X^{x, y}.$$

Si G est commutatif et écrit additivement, ceci conduit à (2.8).

10. L'identité fondamentale de Wald pour des processus stochastiques arbitraires. — Exactement le même argument qui conduit à l'identité fondamentale de Wald (5.17) s'applique si E est un espace euclidien à r dimensions, x et y sont des vecteurs réels et ξ un vecteur complexe à r dimensions, ξx est le produit scalaire. Il s'applique également dans le cas plus général où E est un groupe abélien additif et $\xi = \xi_1 + i\xi_2$, où ξ_1 et ξ_2 sont des homomorphismes sur l'ensemble des nombres réels, c'est-à-dire pour tous les x et y dans E , $\xi_i x$ est un nombre réel tel que

$$\xi_i(x + y) = \xi_i x + \xi_i y \quad (i = 1, 2).$$

Nous pouvons résumer notre résultat dans le

THÉOREME 6. — *Si E est un groupe abélien (additif), si ξ_0, ξ_1 et ξ_2 sont des homomorphismes dans le groupe additif des nombres réels, si $\xi = \xi_1 + i\xi_2$ et $\varphi(\xi_0) < \infty$ et $\inf_{x \in B} (\xi_0 - \xi_1) x > -\infty$, l'identité*

$$(3.17) \quad \sum_1^{\infty} \varphi(\xi_0)^{-n} \int C_{(n)}^x dy e^{\xi y} = 1$$

est satisfaite

Une généralisation partielle de ce théorème à des processus stochastiques arbitraires, déterminés par les probabilités de passage $P_{(n)}^{x_0, \dots, x_{n-1}, x}$ satisfaisant à (2.1)-(2.4) est obtenue comme suit :

Nous faisons l'hypothèse supplémentaire que, lorsque le point cheminant arrive à l'état x_{n-1} ($n \geq 1$) après avoir passé si $n \geq 2$ par $x_0, \dots, x_{n-2}$, il existe une probabilité $A_{(n)}^{x_0, \dots, x_{n-1}}$ pour que ce point soit absorbé, donc une probabilité

$$B_{(n)}^{x_0, \dots, x_{n-1}} = 1 - A_{(n)}^{x_0, \dots, x_{n-1}}$$

pour qu'il ne soit pas absorbé.

En outre, nous introduisons les quantités auxiliaires $U_{(n)}^{x_0, \dots, x_{n-1}}$ et $T_{(n)}^{x_0, \dots, x_{n-1}}$ interprétées comme les probabilités conditionnelles pour que la catastrophe $\mathcal{C}$ n'arrive pas sous la condition que le point cheminant a passé par $x_0, \dots, x_{n-1}$ et a été (pour U) ou n'a pas été (pour T) absorbé en x_{n-1} . Finalement nous définissons $C_{(n)}^{x_0, \dots, x_{n-1}}$ comme la probabilité conditionnelle totale pour que $\mathcal{C}$ n'arrive pas sous la condition que le point a passé par $x_0, \dots, x_{n-1}$.

Pour calculer $C_{(n)}^{x_0, \dots, x_{n-1}}$ pour $n \geq 1$, nous remarquons qu'il y a deux cas complémentaires : ou bien le point est absorbé en x_{n-1} (probabilité $A_{(n)}^{x_0, \dots, x_{n-1}}$), auquel cas non- $\mathcal{C}$ a la probabilité $U_{(n)}^{x_0, \dots, x_{n-1}}$, ou bien il n'est pas absorbé (probabilité $B_{(n)}^{x_0, \dots, x_{n-1}}$) et $\mathcal{C}$ n'arrive pas (probabilité $T_{(n)}^{x_0, \dots, x_{n-1}}$); alors il saute dans un « petit » ensemble dy (probabilité $P_{(n)}^{x_0, \dots, x_{n-1}, dy}$) et il vit heureux jusqu'à la fin des siècles (c'est-à-dire $\mathcal{C}$ n'arrive plus plus tard, probabilité $C_{(n+1)}^{x_0, \dots, x_{n-1}, y}$). Donc nous avons

$$(40.1) \quad C_{(n)}^{x_0, \dots, x_{n-1}} = A_{(n)}^{x_0, \dots, x_{n-1}} U_{(n)}^{x_0, \dots, x_{n-1}} + B_{(n)}^{x_0, \dots, x_{n-1}} T_{(n)}^{x_0, \dots, x_{n-1}} \int P_{(n)}^{x_0, \dots, x_{n-1}} dy C_{(n+1)}^{x_0, \dots, x_{n-1}, y}.$$

Pour $n = 0$ nous avons une équation similaire avec la seule différence que les indices supérieurs excepté y font défaut.

Pour simplifier les formules, nous omettrons les indices inférieurs (n) et remplacerons la suite d'indices supérieurs $x_0, \dots, x_{n-1}$ par un seul symbole π (parcours). (10.1) devient

$$(10.2) \quad C^\pi = A^\pi U^\pi + B^\pi T^\pi \int P_{dy}^\pi C^{\pi y}.$$

Par substitutions successives, nous obtenons un développement formel pour C^π , π étant donné (en particulier aussi si π est « vide »)

$$(10.3) \quad \begin{aligned} C^\pi &= A^\pi U^\pi + B^\pi T^\pi \int P_{dy_1}^\pi A^{\pi y_1} U^{\pi y_1} + B^\pi T^\pi \\ &\quad \times \int P_{dy_1}^\pi B^{\pi y_1} T^{\pi y_1} \int P_{dy_1}^{\pi y_1} A^{\pi y_1 y_2} U^{\pi y_1 y_2} + \dots \\ &= \sum_{k=0}^{\infty} B^\pi T^\pi \int P_{dy_1}^\pi B^{\pi y_1} T^{\pi y_1} \int \dots \int P^{\pi y_1 \dots y_{k-1}}_{dy_k} A^{\pi y_1 \dots y_k} U^{\pi y_1 \dots y_k}. \end{aligned}$$

Pour que ce développement soit convergent et soit une solution de (10.2), il est nécessaire et suffisant que

$$(10.4) \quad R_k^\pi = B^\pi T^\pi \int P_{dy_1}^\pi B^{\pi y_1} T^{\pi y_1} \int \dots \int P^{\pi y_1 \dots y_{k-1}}_{dy_k} C^{\pi y_1 \dots y_k}$$

tende vers zéro lorsque $h \rightarrow \infty$.

En vue de généraliser le théorème 2, nous considérons des solutions du système d'équations (10.1) dans les inconnues $C_{(n)}^{x_0, \dots, x_{n-1}}$ où les U et les T ne sont pas nécessairement des probabilités (bien que les A , les B et les P le restent) mais peuvent être des suites de fonctions (non nécessairement bornées) arbitraires réelles ou complexes.

THÉOREME 7. — *Si : A. pour tous les n et tous les $\pi = (x_0, \dots, x_{n-1}) \in E^n$ il existe des nombres réels U_0^π , T_0^π tels que :*

- A. 1. $T_0^\pi \geq 0$ pour tout n et tout $\pi \in E^n$;
- A. 2. $U_0^\pi \geq 0$ pour tout n et tout $\pi \in E^n$;
- A. 3. $T_0^\pi \int P_{dy}^\pi U_0^{\pi y} \leq U_0^\pi$ pour tout n et tous les $\pi \in E^n$ avec $B^\pi \neq 0$.

Si : B. les fonctions T^π , U^π sont définies pour tout n et tout $\pi \in E^n$ si 0 et c sont des nombres réels non négatifs et si :

B. 1. $0 < 1$;

B. 2. $|T^\pi| \leq \theta T_0^\pi$ pour tout n et tout $\pi \in E^n$;

B. 3. $|U^\pi| \leq c U_0^\pi$ pour tout n et tout $\pi \in E^n$,

alors les équations (10.2) ont des solutions $C = C(U)$, données par (10.3), qui satisfont à l'équation

$$(10.5) \quad C^\pi(U) = U^\pi$$

pour tout n et tout π , si et seulement si

$$(10.6) \quad U^\pi = T^\pi \int P_{dy}^\pi U^{\pi, y}$$

pour tout n et tout π avec $B^\pi \neq 0$.

Démonstration. — En écrivant

$$(10.7) \quad Q_X^\pi = B^\pi T^\pi P_X^\pi, \quad Q_0^\pi = B^\pi T_0^\pi P_X^\pi,$$

A. 3 devient

$$(10.8) \quad \int Q_{dy}^\pi U_0^{\pi, y} \leq B^\pi U_0^\pi,$$

d'où découle l'existence du membre gauche. Les termes de la somme (10.3) avec T_0, U_0 au lieu de T, U sont ≥ 0 . Le $(k+1)^{\text{ième}}$ terme de cette somme est $\leq$ la même expression où le facteur $A^{\pi, y_1 \dots y_k}$ (qui est ≤ 1) est omis, d'où, par (10.8) avec $\pi, y_1 \dots y_{k-1}$ au lieu de π et y_k au lieu de y $\leq$ le $k^{\text{ième}}$ terme avec $A^{\pi, y_1 \dots y_{k-1}}$ remplacé par $B^{\pi, y_1 \dots y_{k-1}}$. La somme du $k^{\text{ième}}$ et du $(k+1)^{\text{ième}}$ terme est $\leq$ le $k^{\text{ième}}$ terme où le facteur $A^{\pi, y_1 \dots y_{k-1}}$ est omis. En continuant de cette sorte nous trouvons que la somme des premiers $k+1$ termes est $\leq U_0^\pi$ de sorte que la série est bornée, donc convergente. En désignant sa somme par C_0^π nous avons alors $0 \leq C_0^\pi \leq U_0^\pi$ et aussi $0 \leq R_{0,k}^\pi \leq U_0^\pi$ où $R_{0,k}^\pi$ est l'expression (10.4), avec T et C remplacés par T_0 et C_0 . De B. 1-B. 3, on conclut alors que les intégrales dans (10.3) aussi bien que la somme de la série existent absolument et que

$$R_k^\pi \leq \theta^k R_{0,k}^\pi \leq \theta^k U_0^\pi, \quad \text{d'où} \quad \lim_{k \rightarrow \infty} R_k^\pi = 0$$

de sorte que (10.3) satisfait à (10.2). Mais si une solution de (10.2)

satisfait à (10.5), (10.6) est valable; c'est un résultat banal. Si, d'autre part, (10.6) est valable en même temps (10.3), nous avons

$$\begin{aligned}
 (10.9) \quad U^\pi C^\pi &= \lim_{n \rightarrow \infty} \left[B^\pi U^\pi - \sum_1^n \int Q_{dy_1}^\pi \int \dots \int Q^{\pi y_1 \dots y_{k-1}}_{dy_k} A^{\pi y_1 \dots y_k} U^{\pi y_1 \dots y_k} \right] \\
 &= \lim_{n \rightarrow \infty} \left[\sum_1^{n-1} \int Q_{dy_1}^\pi \int \dots \int Q^{\pi y_1 \dots y_{k-1}}_{dy_k} \right. \\
 &\quad \times \left\{ B^{\pi y_1 \dots y_k} U^{\pi y_1 \dots y_k} - \int Q^{\pi y_1 \dots y_k}_{dy_{k+1}} U^{\pi y_1 \dots y_{k+1}} \right\} \\
 &\quad \left. + \int Q_{dy_1}^\pi \int \dots \int Q^{\pi y_1 \dots y_{n-1}}_{dy_n} B^{\pi y_1 \dots y_n} U^{\pi y_1 \dots y_n} \right].
 \end{aligned}$$

Les expressions entre accolades s'annulent à cause de (10.6), (10.7) (avec $\pi y_1 \dots y_{k-1}$ au lieu de π et y_k au lieu de y). Donc elles restent nulles après des intégrations répétées et la première somme entre crochets dans le deuxième membre de (10.9) s'annule. Le dernier terme entre crochets est absolument $\leq \theta^n U_0^\pi$, donc il tend vers zéro, ce qui démontre (10.5) et le théorème.

Dans la démonstration du théorème 7, le lemme 4 (Appendice, § 8), n'a pas été utilisé. On peut donc s'attendre à ce que le théorème 7 n'implique pas complètement le théorème 2. En effet, dans le cas particulier du théorème 2, nous avons

$$(10.10) \quad \begin{cases} P_{(n)}^{x_0, \dots, x_{n-1}} = P_X^{x_{n-1}}, & A_{(n)}^{x_0, \dots, x_{n-1}} = A^{x_{n-1}}, & B_{(n)}^{x_0, \dots, x_{n-1}} = B^{x_{n-1}}, \\ U_{(n)}^{x_0, \dots, x_{n-1}} = U^{x_{n-1}}, & T_{(n)}^{x_0, \dots, x_{n-1}} = T^{x_{n-1}}. \end{cases}$$

Les conditions du théorème 2, excepté (§.8) sont satisfaites, mais le théorème 7 devient

$$(10.11) \quad C^x(U) = U^x$$

si et seulement si

$$(10.12) \quad U^x = T^x \int P_{dy}^x U^y$$

de sorte qu'il reste à prouver que

$$(10.13) \quad C^x(U) = \int C_{dy}^x U^y,$$

où C_X^x est donné par (3.2) si (§.8) est satisfaite

En utilisant le lemme 4 (Appendice, § 8), on obtient (10.13), ce qui est une généralisation directe de (3.3) pour tous les U^x satisfaisant à B.3. Ainsi nous voyons que l'ensemble du théorème 7 et du lemme 4 implique le théorème 2. Il semble que pour une généralisation directe du théorème 2 une généralisation du lemme 5 avec $f^{x,x}$ au lieu de f^x soit nécessaire ⁽⁹⁾.

BIBLIOGRAPHIE.

N. ARLEY :

1943. *On the theory of stochastic processes and their application to the theory of cosmic radiation* (Thèse, Université de Copenhague).

G. A. BARNARD

1946. *Sequential test in industrial statistics* (J. Roy. Stat. Soc., Suppl. 8, p. 1-21).

M. S. BARTLETT

1944. *Negative probability* (Proc. Camb. Phil. Soc., t. 41, p. 71-73).

M. S. BARTLETT et D. G. KENDALL :

1951. *On the use of the characteristic functional in the analysis of some stochastic processes occurring in physics and biology* (Proc. Camb. Phil. Soc., t. 47, p. 65-76) ⁽¹⁰⁾.

G. BIRKHOFF

1948. *Lattice theory* (2^e édition) (Amer. Math. Soc., Colloquium Publications, vol. 25).

D. BLACKWELL et M. A. GIRSHICK :

1946. *On functions of sequences of independent chance vectors with applications to the problem of random walk in k dimensions* (Ann. Math. Stats., t. 17, p. 310-317).

A. BLANC-LAPIERRE et R. FORTET

1953. *Théorie des fonctions aléatoires*, Masson, Paris ⁽¹⁰⁾.

G. BLOM :

1949. *A generalization of Wald's fundamental identity* (Ann. Math. Stat., t. 20, p. 439-444).

L. LE CAM :

1947. *Un instrument d'étude des fonctions aléatoires : la fonctionnelle caractéristique* (C. R. Acad. Sc., t. 224, p. 701-711).

R. G. COOKE

1950. *Infinite matrices and sequence spaces*, Mac Millan and Co, Londres.

D. VAN DANTZIG :

1935. *La notion de dérivée d'une fonctionnelle* (C. R. Acad. Sc., t. 201, p. 1008-1010).

⁽⁹⁾ Cette généralisation fut obtenue ensuite dans un travail, fait en collaboration avec C. L. Scheffer (1954).

⁽¹⁰⁾ Paru après la version originelle du présent Mémoire.

1936. *Ricci-calculus and functional analysis* (*Proc. Kon. Ned. Ak. Wetensch.*, t. 39, p. 785-794).
1941. *Mathematische en empiristische grondslagen der waarschijnlijkheidsrekening* (*Ned. Tijdschrift voor Natuurkunde*, t. 8, p. 70-93).
1947. *Mathematische Statistiek* (Hoofdstuk 2, Collectieve kenmerken, voortbrengende en karakteristieke functies, Mathematisch Centrum, Amsterdam, miméographé).
1949. *Sur la méthode des fonctions génératrices* (*Colloques internationaux du Centre national de la Recherche scientifique*, t. 13, p. 29-45).
- D. VAN DANTZIG et C. L. SCHEFFER :
1954. *On arbitrary hereditary time-discrete stochastic processes, considered as stationary Markov chains, and the corresponding general form of Wald's fundamental identity* (*Proc. Kon. Ned. Ak. Wetensch.*, t. 57 (10)).
- W. DOEBLIN :
1940. *Éléments d'une théorie générale des chaînes simples constantes de Markoff* (*Ann. Éc. Norm. Sup.*, (3), t. 57, p. 61-111).
- J. L. DOOB :
1938. *Stochastic processes with an integral-valued parameter* (*Trans. Amer. Math. Soc.*, t. 44, p. 87-151).
1945. *Markoff chains, denumerable case* (*Trans. Amer. Math. Soc.*, t. 53, p. 458-473).
1948. *Asymptotic properties of Markoff transition probabilities* (*Trans. Amer. Math. Soc.*, t. 63, p. 393-421).
1953. *Stochastic Processes*, J. Wiley and Sons, New-York, (10).
- W. FELLER :
1949. *On the theory of stochastic processes, with particular reference to applications* (*Proc. Berkeley Symposium on Math. Stat. and Prob.*, t. 1, p. 403-432).
1950. *An introduction to probability theory and its applications*, vol. I, J. Wiley and Sons, New-York.
- R. FORTET :
- 1935-1938. *Sur les probabilités en chaînes* (*C. R. Acad. Sc.*, t. 201, 1935; t. 202, 1936 et t. 204, 1938).
- M. FRÉCHET :
1913. *Sur la notion de différentielle dans le calcul des fonctionnelles* (*C. R. Congrès Soc. savantes*, 1912, p. 45-58).
1914. *Sur la notion de différentielle d'une fonction de ligne* (*Trans. Amer. Math. Soc.*, t. 15, p. 135-161).
1950. *Généralités sur les probabilités. Éléments aléatoires. Traité du calcul des probabilités et de ses applications* [t. I (*Les principes de la théorie des probabilités*), fasc. III (*Recherches théoriques modernes sur le calcul des probabilités*), livre I, Gauthier-Villars, Paris].
- M. A. GIRSHICK :
1946. *Contributions to the theory of sequential analysis*, part II (*Ann. Math. Stat.*, t. 17, p. 282-291).
- A. HAAR :
1933. *Der Massbegriff in der Theorie der kontinuierlichen Gruppen* [*Ann. Math.*, (2), t. 34, p. 147-169].
- J. HADAMARD :
1910. *Leçons sur le calcul des variations* t. I, A. Hermann, Paris.

H. HAHN et A. ROSENTHAL :

1948. *Set Functions*, Albuquerque, Univ. of New Mexico Press.

P. R. HALMOS :

1950. *Measure theory*, D. van Nostrand Company.

J. J. HERMANS, M. S. KLAMKIN et R. ULLMAN :

1952. *The excluded volume of polymer chains* (*J. Chem. Phys.*, t. 20, p. 1360-1368)⁽¹⁰⁾.

J. H. B. KEMPERMAN :

1950. *The general one-dimensional random walk with absorbing barriers, with applications to sequential analysis* (Thèse de doctorat, Amsterdam ; Excelsiors Foto-Offset, La Haye).

G. W. KING :

1948. *Dependence of physical properties of high polymers on molecular structure; statistical investigations with the aid of punched card methods* (Office of Naval Research, Technical Report n° 1).

1949. *New methods in the statistical mechanics of high polymers* (Office of Naval Research, Technical Report n° 2).

P. LÉVY :

1948. *Processus stochastiques et mouvement brownien*, Gauthier-Villars, Paris.

H. LOOMIS :

1945. *Abstract congruence and the uniqueness of Haar measure* (*Ann. Math.*, (2), t. 46, p. 348-355).

E. W. MONTROLL :

1950. *Markoff chains and excluded volume effect in polymer chains* (*J. Chem. Phys.*, t. 18, p. 734-743).

J. VON NEUMANN :

1928. *Zur Theorie der Gesellschaftsspiele* (*Math. Ann.*, t. 100, p. 295-320).

1934. *Zur Haarschen Mass und topologischen Gruppen* (*Comp. Math.*, t. 1, p. 106-114).

1936. *The uniqueness of Haar measure* (*Math. Sbornik*, t. 1, p. 721-734).

J. VON NEUMANN et O. MORGENSTERN :

1944. *Theory of games and economic behavior*, Princeton University Press, 2^e éd., 1947, Princeton, New-Jersey.

OLGA TAUSSKY :

1949. *A recurring theorem on determinants* (*Amer. Math. Monthly*, t. 56, p. 672-676).

CH. M. TCHEN :

1951. *Stochastic processes and dispersion of configurations of linked events* (*J. Res. National Bureau of Standard*, t. 46, p. 480-488).

A. WALD :

1945. *Sequential analysis*, J. Wiley and Sons, New-York.

1950. *Statistical decision functions*, J. Wiley and Sons New-York.

W. WASOW :

1951. *On the duration of random walks* (*Ann. Math. Stat.*, t. 22, p. 199-216).

