ANNALES SCIENTIFIQUES DE L'É.N.S.

PAUL LÉVY

Systèmes markoviens et stationnaires. Cas dénombrable

Annales scientifiques de l'É.N.S. 3^e série, tome 68 (1951), p. 327-381

<http://www.numdam.org/item?id=ASENS_1951_3_68__327_0>

SYSTÈMES MARKOVIENS ET STATIONNAIRES. CAS DÉNOMBRABLE

PAR M. PAUL LÉVY.

INTRODUCTION.

1. L'origine du présent travail a été une conversation, au cours du symposium de Berkeley (août 1950), avec M. K. L. Chung, qui a attiré mon attention sur un problème non résolu, qui m'a d'abord paru facile à résoudre, et qui sera résolu par le théorème II.6.1 du présent travail. Mais la démonstration de ce théorème n'est simple que pour les systèmes sans états instantanés. Son extension aux systèmes à états instantanés oblige à distinguer plusieurs cas, et nous avons été conduit finalement à proposer une classification des processus étudiés, qui distingue sept types, dont quatre pour les systèmes sans états instantanés et trois pour les systèmes à états instantanés.

Les principaux résultats ont été résumés dans quatre Notes présentées à l'Académie des Sciences (P. Lévy [3]; *voir* la bibliographie à la fin du présent travail) (1).

Les théorèmes fondamentaux de la théorie dont il s'agit ont été exposés, dès 1936, pour le cas des chaînes, par A. Kolmogoroff [1] et [2], et depuis, par W. Doeblin [1] et [2], et, pour le cas du paramètre continu, par J. L. Doob [1] et [2] (2). Le présent travail contient quand même un grand nombre de résultats nouveaux et il nous a paru utile de présenter un exposé d'ensemble qui puisse être lu même par des lecteurs n'ayant aucune connaissance préalable du sujet. Nous évitons de renvoyer aux Mémoires de Kolmogoroff dont l'un ne contient que des énoncés sans démonstration et dont l'autre est rédigé en russe.

(1) Je voudrais m'excuser ici d'un malentendu, dû sans doute à ma difficulté à suivre une conversation en anglais, qui m'avait fait croire que la littérature du sujet était presque inexistante, et présenter d'abord comme nouveaux des résultats qui ne l'étaient qu'en partie.

(2) *Voir* aussi W. Feller [1]. Ce travail ne concerne pas uniquement le cas dénombrable, mais contient des théorèmes généraux qu'il applique en particulier à ce cas. Il ne suppose pas la stationnarité.

Dans tout ce travail, nous avons cherché à dépasser aussi vite que possible le point de vue de Bernoulli, c'est-à-dire celui de l'étude des probabilités de passage, pour arriver à l'étude des propriétés probables ou presque sûres de la fonction aléatoire $H(t)$ ou, dans le cas des chaînes, de la suite $\{H_n\}$. Certaines de ces propriétés ont déjà été considérées par A. Kolmogoroff, W. Doeblin, et J. L. Doob. Mais nous adoptons plus systématiquement le point de vue indiqué et montrons que, dans certains cas, la démonstration la plus simple d'une propriété des probabilités de passage repose sur la loi forte des grands nombres. Bien entendu, il ne s'agit pas de contester qu'il soit logique de définir d'abord les probabilités de passage et que la manière naturelle de formuler analytiquement notre problème repose sur leur introduction.

Le chapitre I est consacré aux chaînes de Markoff (cas du paramètre discontinu). C'est surtout un rappel de résultats connus de A. Kolmogoroff et W. Doeblin; il y a cependant quelques remarques nouvelles, utiles pour la suite.

Dans le chapitre II, nous commençons l'étude du cas du paramètre continu. Nous y indiquons la classification déjà mentionnée plus haut; les différents types de processus que nous introduisons se distinguent essentiellement les uns des autres par la nature de l'ensemble des points de discontinuité. La fin du chapitre est consacrée à l'étude des systèmes sans états instantanés. Nous ne faisons que résumer au paragraphe II.7 une question traitée par J. L. Doob, relative aux équations différentielles de Kolmogoroff, et donnons au paragraphe II.8 la démonstration de deux théorèmes généraux qui semblent nouveaux. Après quelques remarques sur les propriétés presque sûres de $H(t)$ et l'étude de quelques exemples, nous terminons ce chapitre par l'esquisse d'une théorie générale de ce que nous appelons les états fictifs. Elle est nécessaire pour arriver à une définition infinitésimale des processus analogue à celle qui résulte pour le cas fini des équations différentielles de Kolmogoroff; mais il n'est pas certain qu'on puisse arriver à un résultat réellement satisfaisant.

Le chapitre III est consacré à l'étude des systèmes à états instantanés. Ceux pour lesquels les probabilités de passage sont mesurables forment deux types, le cinquième et le sixième. Nous indiquons la forme la plus générale des processus de ces deux types, et montrons que les théorèmes généraux du paragraphe II.8 restent applicables. Il ne saurait en être de même pour le septième type, qui introduit des probabilités de passage non mesurables. Nous nous contentons d'en indiquer un exemple qui généralise un exemple de J. L. Doob [1] et une nouvelle généralisation dont il y a lieu de se demander si elle ne comprend pas tous les processus de ce type; ce serait un théorème analogue à celui qui a été démontré pour le sixième type. Nous n'avons pas résolu cette question.

2. NOTATIONS; NOTIONS GÉNÉRALES COMMUNES AU CAS DES CHAINES ET AU CAS CONTINU. — Le système étudié sera désigné par $\mathcal{S}_r$ s'il y a un nombre fini r d'états possibles et par $\mathcal{S}_\omega$ s'il y en a une infinité dénombrable. Les différents états seront désignés par A_h, A_k, ..., les indices étant supposés entiers. Pour des exemples particuliers, il pourra être commode de les faire varier de zéro à l'infini, ou de $-\infty$ à $+\infty$, ou même de considérer des systèmes de deux indices. Pour la théorie générale, leur nombre importe seul.

Le résultat des expériences s'exprime, dans le cas des chaînes, par une suite d'indices H_n, dans le cas continu, par une fonction $H(t)$ égale à l'instant t à l'indice de l'état réalisé à cet instant.

La loi de l'évolution ou loi du *processus stochastique*, supposé markovien et stationnaire, est bien définie par les *probabilités de passage*. Dans le cas continu nous les désignerons par

$$P_{h,k}(t) = \Pr\{H(t_2) = k / H(t_1) = h\} \qquad (t = t_2 - t_1 \geqq 0).$$

Ces fonctions doivent vérifier les conditions bien connues

$$(\mathcal{I}) \qquad P_{h,k}(0) = \delta_{h,k},$$

$$(\mathcal{P}) \qquad P_{h,k}(t) \geqq 0,$$

$$(\mathcal{T}) \qquad \sum_k P_{h,k}(t) = 1,$$

$$(\mathcal{C}) \qquad P_{h,k}(s+t) = \sum_j P_{h,j}(s)\, P_{j,k}(t) \qquad (s, t \geqq 0),$$

dont la première est inutile si l'on ne considère que les valeurs positives de t, ce qui suffit pour définir le processus ([3]). Remarquons tout de suite qu'il n'en résulte pas nécessairement que $P_{h,k}(+0)$ existe et ait la valeur $\delta_{h,k}$.

Inversement, tout système de fonctions $P_{h,k}(t)$ vérifiant ces conditions définit un processus markovien et stationnaire. La définition d'un processus est donc exactement équivalente à la définition d'un système de fonctions $P_{h,k}(t)$, définies pour $t > 0$ et vérifiant les conditions $(\mathcal{P})$, $(\mathcal{T})$ et $(\mathcal{C})$.

Pour définir une *fonction aléatoire* $H(t)$, il faut se donner une condition supplémentaire, qui sera en principe une condition initiale, relative à un instant t_0 qu'on peut supposer pris pour origine. Elle sera alors de la forme

$$\Pr\{H(0) = h\} = w_h,$$

avec $w_h \geqq 0$ et $\Sigma w_h = 1$, et les probabilités absolues seront

$$W_k(t) = \Pr\{H(t) = k\} = \sum_h w_h P_{h,k}(t).$$

([3]) Les lettres par lesquelles nous représentons ces équations sont les initiales des mots : *initial*, *positif*, *total*, *Chapman*.

Ainsi, pour un même processus, on aura une infinité de fonctions aléatoires particulières. Il faut remarquer que chacune de ces fonctions conserve le caractère markovien, mais n'est en général pas stationnaire. Remarquons aussi que, bien qu'il soit souvent commode de considérer un instant initial, il peut arriver que la fonction soit prolongeable à gauche et définie de $-\infty$ à $+\infty$; tel sera précisément le cas si elle est stationnaire.

Toutes ces remarques s'appliquent en particulier au cas des chaînes; il n'y a qu'à remplacer la variable continue t par l'entier n. Nous écrirons dans ce cas $P_{h,k}^{(n)}$ au lieu de $P_{h,k}(n)$ et $W_k^{(n)}$ au lieu de $W_k(n)$.

CHAPITRE I.

Les chaines stationnaires de Markoff.

I.1. Le point de vue de Bernoulli. — Dans le cas des chaînes, toutes les probabilités de passage sont bien définies par la seule donnée des probabilités

$$p_{h,k} = P_{h,k}^{(1)},$$

assujetties à vérifier les équations $(\mathfrak{R})$ et $(\mathfrak{C})$. Les autres $P_{h,k}^{(n)}$ s'en déduisent par l'un ou l'autre des systèmes d'équations de récurrence

$$\left.\begin{aligned} (\mathcal{R}_1)\qquad & P_{h,k}^{(n+1)} = \sum_j P_{h,j}^{(n)} p_{j,k} \\ (\mathcal{R}_2)\qquad & P_{h,k}^{(n+1)} = \sum_j p_{h,j} P_{j,k}^{(n)} \end{aligned}\right\} \quad (n = 1, 2, \ldots),$$

qui sont des cas particuliers de $(\mathcal{C})$. On vérifie sans peine que toutes les conditions $(\mathfrak{R})$, $(\mathfrak{C})$ et $(\mathcal{C})$ sont alors vérifiées pour les coefficients obtenus de cette manière. On peut même observer que cette vérification est inutile; il existe sûrement un processus markovien et stationnaire bien défini par l'itération de l'opération élémentaire définie elle-même par la donnée des $p_{h,k}$.

Pour les probabilités absolues, on a un système unique d'équations de récurrence

$$(\mathcal{R})\qquad W_k^{(n+1)} = \sum_j p_{j,k} W_j^{(n)}.$$

Dans le cas d'un système $\mathfrak{S}_r$, on obtient la solution générale de ces équations en cherchant d'abord des solutions de la forme

$$W_k^{(n)} = c_k s^n.$$

En portant cette forme dans les équations $(\mathcal{R})$, on obtient un système de r équations indépendantes de n; en annulant leur déterminant, on obtient une équation en s, de degré r. A chaque racine d'ordre ρ correspondent ρ systèmes distincts de solutions, soit tous de la forme $c_k s^n$, soit de la forme plus générale $P_k(n)s^n$, les $P_k(n)$ étant des polynomes de degrés $< \rho$. En combinant toutes ces solutions, on obtient la solution générale de $(\mathcal{R})$.

Pour chaque h fixe, les $P_{h,k}^{(n)}$ sont des solutions de $(\mathcal{R})$. On obtient ainsi r solutions linéairement indépendantes; on le vérifie immédiatement pour $n = 0$. Comme elles sont bornées, on voit que n'importe quelle solution reste finie, quand n augmente indéfiniment. Donc, pour toutes les racines de l'équation en s, on a $|s| \leq 1$ et, si $|s| = 1$, le degré des polynomes $P_k(n)$ est toujours nul. On en déduit immédiatement :

a. Les $P_{h,k}^{(n)}$, considérés comme fonctions de n, ne peuvent comprendre que trois sortes de termes : des constantes, des termes périodiques et des termes s'annulant comme des exponentielles pour n infini (c'est-à-dire que tous les $W_k^{(h)}$ seront majorés par $c\sigma^n$, où $s < \sigma < 1$, c étant une constante convenablement définie).

b. Les moyennes

$$Q_{h,k}^{(n)} = \frac{1}{n} \sum_{1}^{n} P_{h,k}^{(\nu)} \tag{I.1}$$

ont, pour n infini, des limites $P_{h,k}$ et les $P_{h,k}$ vérifient les conditions $(\mathfrak{L})$ et $(\mathfrak{E})$, écrites en supprimant l'indice n.

c. Pour que les fonctions $P_{h,k}^{(n)}$ elles-mêmes tendent vers ces limites, il faut et il suffit qu'elles n'aient pas de termes périodiques.

d. On déduit de $(\mathfrak{E})$ que 1 est racine de l'équation en s. Il y a donc au moins une solution formée de constantes, c'est-à-dire qu'au moins une des fonctions aléatoires déduites du processus est stationnaire.

Enfin un raisonnement très simple, dû à M. Fréchet ([1], p. 108), montre que :

e. Les termes périodiques ont nécessairement des périodes entières et $\leq r$. En d'autres termes, elles correspondent à des racines (autres que l'unité) d'équations binomes $s^\rho = 1$ (ρ entier ≤ 1).

Les méthodes et les résultats précédents ne sont pas tous applicables aux systèmes $\mathcal{S}_\omega$. En ce qui concerne les méthodes, observons que cela n'a plus de sens de dire que, quand on a le nombre voulu de solutions linéairement indépendantes, on a la solution générale du système d'équations $(\mathcal{R})$. La méthode de M. Fréchet pour démontrer la propriété *e* n'est absolument pas susceptible de généralisation. En ce qui concerne les résultats, il suffit d'un exemple simple pour montrer que certains d'entre eux sont en défaut. Considérons, en

effet, un processus pour lequel on ait $p_{h,k}=0$ si $k \leq h$; alors la suite des H_n est constamment croissante et, quels que soient h et k, il vient un moment où $P_{h,k}^{(n)}=0$. Donc les $P_{h,k}$ existent, mais sont tous nuls et ne vérifient pas la condition $(\mathfrak{C})$, et l'énoncé d devient complètement faux.

La manière la plus naturelle d'étudier ces questions, si l'on veut rester au point de vue de Bernoulli, est d'appliquer au système $(\mathcal{R})$ les méthodes connues de la théorie des équations intégrales; il s'agit au fond de représenter la fonction w_h de l'entier k par une combinaison linéaire de fonctions fondamentales. On arrive ainsi à l'énoncé a, sauf en ce qui concerne la rapidité de la convergence vers une limite des solutions sans termes périodiques. On en déduit immédiatement b et c, à cela près que $(\mathfrak{C})$ est remplacé par

$$(\mathfrak{C}') \qquad \sum_k P_{h,k} \leq 1.$$

Nous n'insisterons pas ici sur ces questions, désirant arriver aussi vite que possible au second point de vue, qui nous donnera entre autres une démonstration très simple de l'énoncé b, modifié comme il vient d'être dit, et nous conduira aussi, mais moins simplement, à définir toutes les possibilités concernant l'allure asymptotique des $P_{h,k}^{(n)}$.

Rappelons d'autre part que, dans A. Kolmogoroff [1] et [2], est énoncé en particulier (en termes un peu différents de ceux que nous employons) le résultat suivant :

Quels que soient h et k, il existe une certaine progression arithmétique $n=n_0+\nu\rho$ (ρ entier ≥ 1; $\nu=0, 1, 2, \ldots$) telle $P_{h,k}^{(n)}$ soit nul si n n'appartient pas à cette progression et tende vers une limite (positive ou nulle) si n, appartenant à cette progression augmente indéfiniment. La démonstration n'étant donnée que dans le second Mémoire, rédigé en russe, nous ne savons pas sur quelle méthode elle repose. En tout cas les différentes distinctions introduites dans d'autres énoncés du premier Mémoire, dont nous parlons plus loin, s'introduisent bien naturellement, comme nous le verrons, si l'on utilise les méthodes qui reposent sur la loi forte des grands nombres.

I.2. GROUPES; GROUPES FINAUX (4), SEMI-FINAUX, ET DE PASSAGE; GROUPES CYCLIQUES. — Pour faciliter le langage, nous dirons qu'un événement est *possible*, ou qu'il

(4) C'est à dessein que nous écrivons *finaux*, bien que la forme théoriquement correcte de ce pluriel soit *finals*. L'usage conduit souvent à abandonner certaines formes grammaticales irrégulières, et qui, par suite, choquent l'oreille s'il ne s'agit pas de mots assez usuels pour que l'oreille y soit habituée. Il ne semble y avoir aucune raison de sauver le pluriel *finals*.

D'autre part, nous avons hésité à employer le mot *groupe* dans un sens différent de celui qu'il a généralement en analyse. Le terme *groupement* employé par M. Fréchet présente un autre inconvénient : dans le langage courant, il signifie *répartition en groupes*. Le mot *classe* employé par certains auteurs étrangers est peut être préférable.

peut se produire, lorsque sa probabilité est positive et cela au moins pour certains choix des données qui peuvent être indéterminées.

Nous dirons que deux états A_h et A_k sont *échangeables* quand le passage de A_h à A_k est possible (c'est-à-dire qu'il existe au moins un n pour lequel $P_{h,k}^{(n)} > 0$), ainsi que le passage de A_k à A_h.

Nous dirons qu'un ensemble G d'états est un *groupe* si deux quelconques de ses états sont échangeables et si aucun état n'appartenant pas à G n'est échangeable avec ceux de G. Un groupe est défini par un seul de ses éléments; c'est la réunion de cet élément et de tous ceux qui sont échangeables avec lui.

Un groupe G est *final* si, un état du groupe étant réalisé, le système ne peut pas sortir du groupe; en d'autres termes si, pour tous les $A_h \in G$ et $A_l \notin G$, $p_{h,l} = 0$ et, par suite, $P_{h,l}^{(n)} = 0$ pour tout $n > 0$. Il peut ne comprendre qu'un seul état A_h qui alors ne peut que se succéder à lui-même, indéfiniment, (c'est-à-dire que $p_{h,h} = 1$). Un groupe non final peut aussi se réduire à un seul état A_h; alors $0 < p_{h,h} < 1$.

Désignons par γ_h la probabilité que le système, s'il est initialement dans l'état A_h d'un groupe G, reste indéfiniment dans des états de ce groupe. Les γ_h sont tous égaux à un pour un groupe final, tous < 1 pour un groupe non final. D'ailleurs, si A_h et A_k appartiennent au même groupe, on a évidemment

$$\gamma_h \geqq P_{h,k}^{(n)} \gamma_k$$

et, comme au moins un des $P_{h,k}^{(n)}$ est positif, les γ_h sont tous nuls ou tous positifs, donc, si le groupe n'est pas final, tous nuls ou tous compris entre zéro et un. Le groupe sera dit *de passage* dans le premier cas et *semi-final* dans le second (5).

Nous verrons que tous ces cas existent réellement et que, pour un groupe semi-final, la borne inférieure des α_h peut être positive ou nulle; leur borne supérieure est nécessairement un.

Un groupe peut, d'autre part, être *cyclique* ou *acyclique*. On dit qu'il est cyclique d'ordre r s'il se décompose en un certain nombre r de sous-ensembles ou *sous-groupes*, $G_1, G_2, \ldots, G_r$, tels que, si $H_n \in G_\rho$, on ait nécessairement $H_{n+1} \in G_{\rho+1}$ ($r > 1$; $\rho = 1, 2, \ldots, r$; si $\rho = r$, G_{r+1} désigne G_1); en d'autres termes, il y a entre ces sous-groupes une permutation circulaire non aléatoire, le hasard n'intervenant à chaque opération que pour le choix d'un état dans le sous-groupe imposé.

Cette distinction est indépendante de la précédente. Si l'on a défini un groupe acyclique d'une des trois catégories distinguées d'abord (final, semi-final ou de passage), on n'a qu'à remplacer chaque état A_h par r états $A_h^{(\rho)}$ ($\rho = 1$,

(5) Nous introduisons une terminologie nouvelle. Dans le cas des systèmes $\mathfrak{S}_r$, il n'y a pas de groupes semi-finaux et il n'y avait pas d'inconvénient, pour la théorie de ces systèmes, à confondre les notions de groupe de passage et de groupe non final.

2, ..., r) qui se succéderont dans l'ordre des ρ croissants avant qu'un tirage au sort ait à choisir l'état $A_k^{(1)}$ qui doit succéder à $A_h^{(r)}$. On obtient ainsi un groupe cyclique de la même catégorie que le groupe initial.

I.3. THÉORÈMES SUR LES GROUPES FINAUX ERGODIQUES. — 1° THÉORÈME I.3.1. — *Si le système se trouve initialement dans un état du groupe* G, *ou bien il est presque sûr que les états de ce groupe sont tous réalisés une infinité de fois, ou bien il est presque sûr qu'aucun ne l'est* ([6]).

Nous dirons que le groupe G est *ergodique* dans le premier cas et *non ergodique* dans le second. Le système (S_r ou S_ω) lui-même est ergodique si tous ses éléments constituent un seul groupe ergodique. Dans le premier cas, nous dirons que le groupe est *cycliquement ergodique* s'il est cyclique et *régulièrement ergodique* dans le cas contraire ([7]).

Démontrons maintenant le théorème énoncé. Soit α_h la probabilité que le système, partant de l'état A_h, le reprenne au moins une fois; elle est positive, puisque le système, en quittant l'état A_h, ne peut prendre qu'un état A_k du groupe et peut, si $k \neq h$, revenir à A_h, en une ou plusieurs opérations. La probabilité que le système reprenne au moins n fois l'état A_h est alors α_h^n. Si alors $\alpha_h = 1$, la probabilité α'_h que le système reprenne une infinité de fois l'état A_h est aussi égale à un. Si $\alpha_h < 1$, α'_h, évidemment $\leq \alpha_h^n$ (quel que soit n) est nul. Donc chaque α'_h est égal, soit à zéro, soit à un.

Il reste à montrer que $\alpha'_h = 1$ pour un état du groupe entraîne $\alpha'_k = 1$ pour tous les autres. Comme il y a une suite presque sûrement infinie de racines $N_0 = 0$, N_1, N_2, ... de $H_n = h$, dire que le passage de A_h à A_k est possible, c'est dire qu'il y a une probabilité positive $\beta_{h,k}$ qu'un intervalle $(N_{\rho-1}, N_\rho)$ contienne une racine de $H_n = k$. Cette probabilité est évidemment indépendante de ρ, puisque le processus est stationnaire, et du passé, puisqu'il est markovien. On se trouve donc dans le cas classique de Bernoulli. La circonstance indiquée (réalisation de l'état A_k pour au moins un n compris entre $N_{\rho-1}$ et N_ρ) se produit donc presque sûrement une infinité de fois ([8]), C. Q. F. D.

2° THÉORÈME I.3.2. — *Si le système se réduit à un groupe final ergodique, le rapport des fréquences de deux états quelconques* A_k *et* A_h *tend presque sûrement vers la limite* $\frac{\beta_{h,k}}{\beta_{k,h}}$.

([6]) *Cf.* A. KOLMOGOROFF [1].

([7]) Nous nous excusons d'abandonner aussi bien la terminologie de M. Fréchet que celle de A. Kolmogoroff qui appelle *classe récurrente* ce que nous appelons *groupe ergodique*. Il nous semble que l'essentiel, pour définir le caractère ergodique, n'est pas l'existence d'une limite (vraie ou au sens de Cesaro) indépendante de h, mais le fait que, presque sûrement, le système reprenne une infinité de fois n'importe lequel des états possibles. Dans le cas des espaces continus, c'est de même que, presque sûrement, la trajectoire traverse une infinité de fois n'importe quel domaine donné.

([8]) Le même raisonnement montre qu'on aboutit à une contradiction en supposant qu'un état n'appartenant pas à un groupe final puisse être réalisé une infinité de fois.

Nous pouvons évidemment faire abstraction des états autres que A_h et A_k; alors la suite des H_n se décompose en suites partielles S_1, S_2, ..., de termes consécutifs égaux à h, alternant avec des suites S'_1, S'_2, ... de termes égaux à k. Désignons respectivement par L_ν et L'_ν les longueurs des suites S_ν et S'_ν (c'est-à-dire leurs nombres de termes). En conservant les notations du 1°, on a évidemment

$$\Pr\{L_\nu = r\} = \beta_{h,k}(1-\beta_{h,k})^{r-1} \qquad (r = 1, 2, \ldots), \tag{I.3.1}$$

et, par suite, en supprimant l'indice ν,

$$E\{L\} = \frac{1}{\beta_{h,k}}, \qquad \sigma^2\{L\} = \frac{1-\beta_{h,k}}{\beta_{h,k}}, \tag{I.3.2}$$

et des formules analogues pour les L'. Toutes ces variables aléatoires sont d'ailleurs indépendantes; on est encore dans le cas de Bernoulli.

Les nombres des réalisations des événements A_h et A_k, à un instant quelconque, sont toujours de la forme

$$U_n = L_1 + L_2 + \ldots + L_\nu + \theta\, L_{\nu+1},$$
$$U'_n = L'_1 + L'_2 + \ldots + L'_\nu + \theta' L'_{\nu+1}$$

(θ, $\theta' \geq 0$ et ≤ 1). Si ν est très grand, les derniers termes sont négligeables et les nombres considérés sont des infiniment grands presque sûrement équivalents à $\nu E\{L\}$ et $\nu E\{L'\}$. Compte tenu de (I.3.2), le théorème énoncé en résulte.

3° *Remarques.* — *a.* On peut préciser le théorème 1.3.2 en appliquant les principaux théorèmes connus sur le cas de Bernoulli; loi des écarts, loi du logarithme itéré. C'est pour rendre cette application plus facile pour le lecteur que nous avons indiqué l'expression $\sigma^2\{L\}$.

Il faut remarquer que ces applications introduisent explicitement la variable ν, c'est-à-dire le nombre des séries partielles de chaque espèce. On peut aussi chercher la loi dont dépend U'_n au moment où U_n atteint une valeur donnée $\rho+1$, c'est-à-dire au moment où $n = N_\rho$. Pour cela il vaut mieux partir de la remarque que l'augmentation de U'_n quand n varie de $N_{\nu-1}$ à N_ν a pour valeur probable

$$\beta_{h,k} E\{L'\} = \frac{\beta_{h,k}}{\beta_{k,h}}.$$

Ces augmentations étant indépendantes, on en déduit une nouvelle démonstration du théorème I.3.2, et il devient facile de le préciser dans le sens indiqué. On en déduit ensuite aisément les lois dont dépendent U_n et U'_n lorsque leur somme est donnée.

Si l'on fait abstraction des états autres que A_h et A_k, cette somme n'est autre que n. Si au contraire on en tient compte, le problème n'est plus aussi simple.

Nous verrons, en effet, que le rapport $\frac{U_n}{n}$, c'est-à-dire la fréquence de l'état A_h, a une limite presque sûre, qui peut être positive ou nulle; le problème ne se présente pas du tout de la même manière dans les deux cas.

b. Si le système ne se réduit pas à un seul groupe final ergodique, il y a plusieurs cas à distinguer. Dans tous les cas, la conclusion relative à la limite du rapport $\frac{U'_n}{U_n}$ est immédiate, le seul cas non trivial ayant été traité. Il peut naturellement arriver que U_n et U'_n restent tous les deux nuls et l'on ne peut pas parler de ce rapport. Dans tous les autres cas, il a presque sûrement une limite et cette limite est en général aléatoire, la loi dont elle dépend dépendant de l'état initial. Si U_n et U'_n restent tous les deux finis, mais sans borne supérieure certaine (une telle borne ne peut d'ailleurs être que un), il peut arriver que l'ensemble des valeurs possibles pour cette limite soit l'ensemble des nombres rationnels positifs; il est facile de donner des exemples de cette circonstance. Enfin U_n et U'_n ne peuvent devenir tous les deux infinis en même temps que si A_h et A_k appartiennent à un même groupe ergodique; il ne faut pas oublier dans ce cas que la probabilité α_l que le système atteigne ce groupe dépend de l'état initial A_l et que, si $\alpha_l < 1$, le rapport considéré a une probabilité $1 - \alpha_l$ d'être de la forme $\frac{0}{0}$, et seulement une probabilité α_l de tendre vers la limite indiquée ci-dessus.

4° *Limites des fréquences.* — Plaçons-nous toujours dans le cas d'un système réduit à un groupe ergodique unique et choisissons un état A_0 comme dénominateur commun. Désignons par N_1, N_2, ..., la suite des n pour lesquels $H_n = 0$ et par c_h le nombre probable des réalisations de l'état $A_h (h \neq 0)$ entre deux réalisations consécutives de A_0.

On a évidemment, en posant $c_0 = 1$,

$$E\{N_\rho - N_{\rho-1}\} = \sum_0^\infty c_h = S,$$

et il y a deux cas à distinguer suivant la nature de la série Σc_h. Comme c_h est la fréquence relative des états A_h et A_0, tous les c_h sont multipliés par un même nombre positif et fini si l'on change l'état choisi pour dénominateur commun; cette distinction est donc indépendante de ce choix.

Si alors S est fini, la loi forte des grands nombres montre que $\frac{N_\rho}{\rho}$ tend presque sûrement vers S, quand ρ augmente indéfiniment. Donc la fréquence de l'état A_0, qui est la valeur de $\frac{\rho}{n}$ pour $N_\rho \leqq n < N_{\rho+1}$, tend presque sûrement vers $\frac{1}{S}$ et celle de n'importe quel état A_h tend presque sûrement vers $\frac{c_h}{S}$. Au contraire si S est infini, toutes ces fréquences tendent presque sûrement vers zéro.

Nous dirons que le groupe final considéré est *fortement ergodique* dans le premier cas et *faiblement ergodique* dans le second. Ces définitions nous paraissent naturelles, puisque dans le second cas, s'il est bien presque sûr que chaque état est réalisé une infinité de fois, la portée de ce fait est diminuée par la rareté de ces réalisations (9). Naturellement, dans chacun des cas, le groupe peut être cyclique ou non cyclique.

5° *Remarques.* — *a.* La question se pose tout naturellement, pour les groupes fortement ergodiques, de préciser la loi forte de la même manière qu'au 3°, *a*. Ce problème oblige à introduire une nouvelle distinction, suivant que le second moment

$$\sigma^2 = \sigma^2\{N_\rho - N_{\rho-1}\} = E\{(N_\rho - N_{\rho-1})^2\} - S^2$$

est fini ou infini. Dans le premier cas, l'écart réduit $\frac{N_\rho - S_\rho}{\sigma\sqrt{\rho}}$ est asymptotiquement une variable laplacienne réduite, et l'on peut le borner supérieurement par la loi du logarithme itéré. Il n'en est pas de même dans le second cas; il peut arriver, par exemple, que $N_\rho - S_\rho$ soit asymptotiquement de la forme $c\rho^\alpha X$, α étant compris entre $\frac{1}{2}$ et 1, et X dépendant d'une loi stable d'exposant $\frac{1}{\alpha}$. Comme il est bien connu, des circonstances très variées peuvent être réalisées dans ces conditions.

b. Il existe naturellement des processus de *types mixtes* pour lesquels le système peut, en partant d'un état initial A_0, évoluer de plusieurs manières différentes et, par suite, arriver à l'un ou l'autre de plusieurs groupes qui ne soient pas du même type. Il semble inutile d'insister sur cette extension triviale.

1.4. Exemples. — Il est bien évident, dans le cas d'un système S_r, qu'il y a nécessairement au moins un état réalisé une infinité de fois et que cet état doit appartenir à un groupe final fortement ergodique. Nous nous proposons de montrer par des exemples qu'au contraire, pour les systèmes S_ω, toutes les circonstances prévues par la théorie sont possibles.

1° Il peut n'y avoir aucun groupe final. Tel est le cas d'un système à évolution

(9) Dans le premier cas, les fréquences limites étant positives, A. Kolmogoroff appelle ce cas *cas positif*. Nous préférons ne pas abuser du mot *positif*, quand il ne dispense pas le lecteur de rechercher les définitions pour savoir ce qui est positif. D'autre part, F. G. Foster [1], réunit le type faiblement ergodique et le type non ergodique en un seul type, qu'il appelle type *dissipatif*. Ses systèmes *semi-dissipatifs* sont des systèmes *composés*, c'est-à-dire ne se réduisant pas à un seul groupe; un groupe n'est jamais semi-dissipatif. Pour l'étude des systèmes composés, il nous paraît préférable de mettre avant tout ce caractère composé en évidence.

irréversible où, par exemple, les H_n ne peuvent varier qu'en croissant (c'est-à-dire que $p_{h,k}=0$ si $k<h$). Si en plus $p_{h,h}=0$, il n'y a aucun groupe. Si, au contraire, les $p_{h,h}$ sont >0 et <1, chaque état A_h forme un groupe de passage.

On peut généraliser beaucoup cet exemple en considérant une infinité de groupes de passage $G_1, G_2, \ldots$, tels que le système ne puisse quitter un de ces groupes que pour un groupe d'indice plus élevé.

2° Considérons un système comprenant un état A_0 et d'autres états répartis en familles $F_h (h=1, 2, \ldots)$. Si le système est initialement à l'état A_0, il y aura une probabilité α_h que l'état suivant appartienne à F_h; dans ce cas, après un séjour dans F_h de longueur L_h presque sûrement finie, il reviendra à l'état A_0 (nous appelons longueur le nombre des états successifs). On peut donner à $E\{L_h\}=\mu_h$ n'importe quelle valeur positive. Il suffit de constituer F_h par un état A_h pour lequel

$$p_{h,0}=\frac{1}{\mu_h}, \qquad p_{h,h}=1-\frac{1}{\mu_h}$$

(on peut aussi définir F_h de manière à associer à une valeur donnée de μ_h des valeurs arbitrairement grandes de $\sigma_h=\sigma\{L_h\}$. Au contraire, si μ_h n'est pas entier, σ_h a une borne inférieure positive. Si μ_h est entier, on peut annuler σ_h en supposant F_h formé de μ_h états qui se succèdent dans un ordre déterminé).

On a alors $c_h=\alpha_h\mu_h$ et $S=1+\Sigma\alpha_h\mu_h$. Il est facile de rendre cette série convergente ou divergente, donc de former des processus fortement ou faiblement ergodiques. Il est facile aussi, dans ce second cas, d'en former pour lesquels les c_h tendront vers zéro ou, au contraire, augmenteront indéfiniment avec h. On peut, plus généralement, réaliser n'importe quelle suite de valeurs positives données. On aurait pu être tenté de croire qu'on peut ranger les différents états dans l'ordre des c_h décroissants. Il en est évidemment ainsi dans le cas d'un groupe fortement ergodique; on voit que cela n'est pas toujours vrai pour un groupe faiblement ergodique.

3° Considérons maintenant un système défini par

$$p_{h,0}=\alpha_h, \qquad p_{h,h+1}=1-\alpha_h \qquad (h=0, 1, \ldots; \ 0\leq\alpha_h<1).$$

Il constitue évidemment un groupe unique à moins que tous les α_h d'indices assez grands ne soient nuls. La probabilité que, en partant de $H_0=0$, on ait $H_\nu=\nu$ pour $\nu=1, 2, \ldots, h$, est

$$\beta_h=(1-\alpha_0)(1-\alpha_1)\ldots(1-\alpha_{h-1}).$$

Les β_h forment donc une suite non croissante, variant depuis un jusqu'à une limite positive ou nulle β; *à cela près, elle est absolument quelconque.*

Il y a manifestement deux cas à distinguer :

Premier cas : la série $\Sigma\alpha_h$ *est convergente, donc* $\beta>0$. — Il y a donc une proba-

bilité positive β que la suite des H_ν soit constamment croissante, c'est-à-dire que l'on ait indéfiniment $H_\nu = \nu$. La probabilité que cette série croissante soit au contraire interrompue par un retour à l'état A_0 est $1-\beta$. La probabilité qu'il y ait ainsi n retours est $(1-\beta)^n$; la probabilité qu'il y en ait une infinité est nulle. Il est, par suite, presque sûr que la suite des H_ν finit par augmenter indéfiniment, après un nombre de retours à A_0 dont la valeur probable, d'après un calcul fait plus haut [formule (I.3.2)], est $\frac{1}{\beta}$. Le système est donc non ergodique.

Deuxième cas : la série $\Sigma\alpha_h$ est divergente, donc $\beta = 0$. — Alors, la suite des H_n ne peut pas augmenter indéfiniment; il y a presque sûrement, tôt ou tard, un retour à l'état A_0, et, par suite, une infinité; le système est ergodique.

Entre deux retours à l'état A_0, un autre état A_h ne peut être réalisé qu'une fois et a une probabilité β_h de l'être. Son nombre probable de réalisations est donc $c_h = \beta_h$ et le système est fortement ergodique si la série $\Sigma\beta_h$ est convergente et faiblement ergodique si cette série est divergente. Les deux cas sont effectivement possibles.

Une variante de l'exemple précédent s'obtient en prenant

$$p_{h,h} = 1-\gamma_h, \qquad p_{h,0} = \gamma_h\alpha_h, \qquad p_{h,h+1} = \gamma_h(1-\alpha_h), \qquad \text{avec} \quad 0 < \gamma_h < 1.$$

Alors, l'état A_h pourra se répéter et le nombre probable de ces répétitions, avant le passage à un des états A_0 et A_{h+1}, est $\mu_h = \frac{1}{\gamma_h}$. Alors, $c_h = \beta_h\mu_h$, formule analogue à celle relative à l'exemple donné au 2° ci-dessus.

4° *Exemple de groupe semi-final.* — Modifions l'exemple du 3° en supposant, qu'en plus des états A_0, A_1, ..., il y ait un ou plusieurs autres états, formant une famille F, d'où le retour aux états A_h ($h = 0, 1, \ldots$) ne soit plus possible. Posons

$$\alpha_h = \alpha'_h + \alpha''_h \qquad \text{et} \qquad p_{h,F} = \alpha'_h, \qquad p_{h,0} = \alpha''_h, \qquad p_{h,h+1} = 1-\alpha_h$$

($p_{h,F}$ indiquant la probabilité du passage direct de A_h à la famille F). Supposons que α''_h soit positif pour une infinité de valeurs de h, et qu'au moins un α'_h soit positif. Alors les états A_h d'indices positifs ou nuls forment un groupe G qui n'est pas final. Si la série $\Sigma\alpha_h$ est divergente, c'est un groupe de passage, que le système doit quitter tôt ou tard pour atteindre F. Mais si la série $\Sigma\alpha_h$ est convergente, il y a, comme tout à l'heure, une probabilité positive β que H_ν augmente indéfiniment sans retour à A_0, ce qui suffit pour conclure que G est un groupe semi-final.

Si d'ailleurs $\beta' < 1-\beta$ est la probabilité d'un retour à A_0, il y a une probabilité $\beta\beta'^n$ que l'augmentation indéfinie de H_ν se produise aprés exactement n

retours à A_0 et une probabilité totale $\frac{\beta}{1-\beta'}$ qu'elle finisse par se produire. De même, si $\beta''=1-\beta-\beta'$ est la probabilité d'une fuite vers F sans retour préalable à A_0, la probabilité totale d'une telle fuite est $\frac{\beta''}{1-\beta'}$.

Le principe de Bayes permet d'obtenir autrement ces probabilités. Les deux seules circonstances finalement possibles ayant chaque fois des probabilités β et β'' de se produire sans nouveau retour à A_0, leurs probabilités totales sont respectivement

$$\frac{\beta}{\beta+\beta''}, \quad \frac{\beta''}{\beta+\beta''} \qquad (\beta+\beta''=1-\beta').$$

5° *Remarque générale sur les groupes semi-finaux.* — Soit λ_h la probabilité que le système, placé initialement dans un état A_h d'un groupe semi-final G, reste indéfiniment dans ce groupe. Par définition des groupes semi-finaux, $0<\lambda_h<1$, pour tous les états du groupe. Dans l'exemple qui précède, λ_h tend vers un pour h infini, tandis que la borne inférieure des λ_h est positive.

On peut modifier cet exemple de manière à annuler cette borne inférieure. L'indice h variera de $-\infty$ à $+\infty$, l'ensemble des états possibles comprenant le groupe G des états A_h, et la famille F; nous prendrons

$$p_{h,F}=\alpha'_h, \qquad p_{h,-h}=\alpha''_h, \qquad p_{h,h+1}=1-\alpha_h \qquad (\alpha_h=\alpha'_h+\alpha''_h)$$

et supposerons $\alpha''_h=0$ (donc $\alpha_h=\alpha'_h$) pour $h\leq 0$. S'il existe des h arbitrairement grands pour lesquels $\alpha''_h>0$, et que les α_h soient tous <1, G est bien un groupe. Si au moins un des α'_h est positif, il n'est pas final; si, de plus la série $\sum_0^\infty \alpha_h$ est convergente, l'augmentation indéfinie des H_n est possible et c'est un groupe semi-final. Cela n'empêche pas de supposer que certains des α'_h, d'indices négatifs, forment une suite infinie ayant pour limite un. Comme évidemment $\lambda_h<1-\alpha'_h$, la borne inférieure est bien zéro.

Au contraire, *leur borne supérieure est toujours un.* C'est une conséquence immédiate d'un théorème général connu (*cf.* P. Lévy [1], p. 129) : *si une propriété de la suite* $\{H_n\}$ *a une probabilité,* γ, *si* γ_n *désigne la nouvelle évaluation de cette probabilité quand on connaît tous les* H_ν *d'indices* $\leq n$, *les ensembles de suites* $\{H_n\}$ *définis, l'un par la propriété considérée, l'autre par* $\lim\gamma_n=1$, *coïncident, à un ensemble de probabilité nulle près.* Ici, la propriété considérée est que tous les H_n soient des indices d'états de G; γ est positif; γ_n est la valeur de λ_h pour $h=H_n$. Il existe donc une suite partielle de λ_h qui tend vers un et, si le système peut rester indéfiniment dans G, c'est en parcourant les états d'une telle suite.

6° *Autres exemples.* — Comme au 4°, nous allons former un groupe G non final, en partant d'un groupe final G' et en modifiant certains des $p_{h,k}$ de manière à permettre la

fuite vers une famille F d'où le retour à G soit impossible. Pour simplifier, nous ne modifierons que les $p_{1,k}$, en les multipliant par un facteur q compris entre zéro et un; la probabilité disponible $1-q$ sera celle du passage de A_1 à F. Dans ces conditions, il est bien évident que si G′ était un groupe ergodique, les retours successifs à l'état A_1 finiront par entraîner le passage à F; G est alors un groupe de passage. Dans le cas contraire, il y a une probabilité positive d'échapper au retour à A_1 et G est semi-final.

Opérons maintenant d'une manière analogue sur r groupes finaux G'_ν qui deviendront des groupes de passage ou semi-finaux $G_\nu (\nu = 1, 2, \ldots, r)$; après l'état $A_1^{(\nu)}$ distingué dans G_ν, il y aura seulement une probabilité q_ν que le système reste dans G_ν; mais supposons que, dans le cas contraire, au lieu de passer à une nouvelle famille F, il passe dans un autre des états $A_1^{(\rho)}$; désignons par $q_{\nu,\rho} (\rho \neq \nu)$, la probabilité du passage de $A_1^{(\nu)}$ à $A_1^{(\rho)}$; si l'on pose

$$q_{\nu,\nu} = q_\nu, \qquad \text{on doit avoir} \qquad \sum_\rho q_{\nu,\rho} = 1.$$

Dans ces conditions, les G_ν deviennent des sous-groupes, qui s'échangent suivant une loi bien définie par les $q_{\nu,\rho}$. Supposons que l'ensemble forme un groupe unique G. Si tous les G'_ν étaient ergodiques, il est bien évident que G le sera aussi. Mais si un ou plusieurs G'_ν étaient non ergodiques, il est possible que le système reste indéfiniment dans un des G_ν correspondants, et cette circonstance finira presque sûrement par se produire; G est aussi un groupe non ergodique. Cela était bien évident; ceux qui rendent leur proie finissent par la perdre définitivement, si tout le monde ne fait pas de même.

Si les G_ν sont en nombre infini, une nouvelle circonstance est possible. Il peut arriver que les $q_{\nu,\rho}$ définissent une évolution non ergodique, qui fasse apparaître comme possible une augmentation indéfinie de l'indice ν. Si alors les G'_ν étaient tous ergodiques, cette augmentation finira presque sûrement par se produire. Dans le cas contraire, il pourra arriver qu'un des G_ν garde sa proie; mais, si la probabilité de cette circonstance est le terme général d'une série convergente, il reste encore possible que le système quitte successivement tous les G_ν et que ν augmente indéfiniment. Ainsi G a plusieurs manières bien différentes d'être non ergodique.

I.5. Retour au point de vue de Bernoulli. — 1° *Existence des limites en moyenne des* $P_{h,k}^{(n)}$. — Si $U_{k,n}$ désigne le nombre des termes égaux à k dans la suite H_1, H_2, ..., H_n et si $Q_{h,k}^{(n)}$ est défini comme au I.1.*b*, on a évidemment

$$Q_{h,k}^{(n)} = E\left\{\frac{1}{n} U_{k,n} \middle/ H(0) = h\right\}.$$

On sait que, quand des variables aléatoires bornées ont une limite presque sûre, leur valeur probable tend vers cette limite. Or nous avons vu que la fréquence $\frac{1}{n} U_{k,n}$ a une limite presque sûre, positive ou nulle, quand A_k appartient à un groupe ergodique, et si le système est, pour n assez grand, dans les états de ce groupe. Dans tous les autres cas, cette fréquence tend évidemment vers zéro au moins presque sûrement. Donc, dans tous les cas, il y a une limite presque sûre Φ_k, pouvant être aléatoire. Par suite,

$$Q_{h,k}^{(n)} \to P_{h,k} = E\{\Phi_k / H(0) = h\}.$$

Donc : *la probabilité de passage* $P_{h,k}^{(n)}$ *a, pour n infini, une limite au sens de Cesaro et cette limite* $P_{h,k}$ *est la valeur probable de* Φ_k, *dans l'hypothèse* $H(o) = h$.

Si A_h et A_k appartiennent à un même groupe final, $P_{h,k}$ a une valeur P_k, indépendante de h, positive seulement dans le cas d'un groupe fortement ergodique. Si A_l n'appartient pas à ce groupe et que α_l soit la probabilité de l'atteindre en partant de A_l, on a évidemment $H_{l,k} = \alpha_l P_k$. Dans tous les autres cas, les limites $P_{h,k}$ sont nulles.

Nous attirons l'attention sur le raisonnement qui précède. Il peut paraître naturel de commencer par l'étude approfondie du point de vue de Bernoulli. On voit qu'il peut arriver aussi qu'un théorème, qui n'est pas du tout évident au point de vue de Bernoulli, soit un corollaire immédiat d'un théorème lui-même facile à démontrer et même, à notre avis, presque intuitif, si l'on est habitué à la loi forte des grands nombres.

2° *Limites des* $P_{h,k}^{(n)}$ *dans le cas d'un groupe acyclique.* — Il est bien évident que, si A_h et A_k appartiennent à un groupe cyclique (d'ordre $r > 1$) et fortement ergodique, $P_{h,k}^{(n)}$ ne peut être positif que si n est de la forme $n_0 + \nu r$ (ν entier, n_0 dépendant de h et k) et, pour ces valeurs, tend au moins en moyenne vers $rP_k > o$; donc dans ce cas, $P_{h,k}^{(n)}$ ne tend pas vers P_k. Si A_k appartient seul à un tel groupe, il peut arriver aussi que $P_{h,k}^{(n)}$ n'ait pas de limite. Nous nous proposons de montrer que; *dans tous les autres cas,* $P_{h,k}^{(n)}$ *tend, pour n infini, vers une limite* (évidemment égale à la limite en moyenne $P_{h,k}$).

Nous allons d'abord nous placer dans le cas où A_h et A_k appartiennent à un même groupe ergodique et acyclique, c'est-à-dire démontrer que :

Théorème I.5. — *Si* A_h *et* A_k *appartiennent à un même groupe* G, *ergodique et acyclique,* $P_{h,k}^{(n)}$ *tend, pour n infini, vers* P_k.

Ce théorème une fois démontré, il sera facile de conclure dans tous les cas.

3° *Remarques préliminaires.* — Ce théorème est connu dans le cas d'un système S_r (*voir* W. Doeblin [1], n° 5). Rappelons seulement que sa démonstration repose sur le fait que, d'après l'équation de Chapman, les $P_{h,k}^{(n+n')}$ sont des moyennes pondérées des $P_{l,k}^{(n)}$. Il en résulte une réduction indéfinie de l'intervalle de variation possible de ces probabilités pour (k fixe et n croissant), à laquelle on ne peut échapper que dans le cas cyclique.

Rappelons aussi que, pour reconnaître si l'on est dans le cas cyclique, il suffit de considérer un seul état, par exemple A_0. Si $P_{0,0}^{(n)}$ n'est positif que pour des valeurs de n ayant r pour plus grand commun diviseur, A_0 appartient à un groupe cyclique d'ordre r. Cela est évident, et d'ailleurs bien connu (*cf.* A. Kolmogoroff [1]).

4° Ne nous occupons d'abord que de l'état A_0. Supposons-le réalisé initialement, et désignons par $N_0 = 0, N_1, N_2, \ldots$, la suite, presque sûrement infinie, des racines de $H_n = 0$. Toutes les différences $X_\nu = N_\nu - N_{\nu-1}$ dépendent de la même loi. Comme elles sont indépendantes, la donnée de cette loi détermine celles des N_ν et, par suite, toutes les probabilités

$$P_{0,0}^{(n)} = \sum_{\nu=1}^{n} \Pr\{N_\nu = n\}.$$

La réciproque est vraie aussi (*cf.* A. Kolmogoroff [1]); mais nous n'aurons pas à l'utiliser.

D'après ces remarques, nous pouvons nous borner à l'étude d'un type simple de chaînes. Pourvu que n'importe quelle loi possible pour les X (compte tenu du fait que G est ergodique et acyclique) soit réalisée pour au moins une chaîne de ce type, la conclusion relative aux $P_{0.0}^{(n)}$ sera générale. Cela ne nous empêchera pas d'introduire dans nos raisonnements les autres $P_{h,k}^{(n)}$; mais la conclusion relative au cas où $h = k = 0$ pourra seule être généralisée.

Prenons, comme type simple, celui défini au I.4.3°, dans le cas ergodique, où $\beta = 0$. On a donc

$$\Pr\{X = h\} = \gamma_h = \beta_{h-1} - \beta_h \geqq 0 \qquad \left(\sum_{1}^{\infty} \gamma_h = 1\right).$$

Le groupe G étant supposé acyclique, les valeurs possibles de X, c'est-à-dire les h pour lesquels $\gamma_h = \alpha_{h-1}\beta_{h-1} > 0$, n'ont pas de diviseur commun autre que 1 ([10]). Remarquons tout de suite que, dans ces conditions, on peut déterminer un nombre m tel que les valeurs possibles inférieures à m soient aussi sans diviseur commun. Remarquons, d'autre part, que

$$(I.5.1) \qquad P_{0,k}^{(n)} = 0 \quad (n < k), \qquad P_{0,k}^{(n)} = \beta_k P_{0,0}^{(n-k)} \leqq \beta_k \quad (n \geqq k).$$

Choisissons un nombre r, supérieur à m et qui devra, en outre, être assez grand pour vérifier des conditions qui seront indiquées plus loin. Extrayons de la suite des H_n une suite partielle H'_ν, comprenant tous les termes $\leqq r$ et ceux-là seulement. Pour cette suite, il y a au plus $r+1$ valeurs possibles, formant un groupe acyclique. Elle est d'ailleurs markovienne et stationnaire. Le théorème à démontrer étant déjà connu dans ce cas, les probabilités de passage $P_{h,k}^{\prime(n)}$ relatives à cette chaîne ont pour n infini une limite $P'_{h,k} = P'_k$ et il

([10]) Cela n'empêche pas que ces valeurs possibles peuvent appartenir toutes à une progression arithmétique $h = h_0 + \nu\rho$ ($\nu = 0, 1, 2, \ldots$); seulement h_0 et ρ devront être premiers entre eux. La suite des N_ν aura alors un caractère cyclique, mais non celle des H_n.

résulte du fait que cette limite est aussi une limite presque sûre de la fréquence de l'état A_k que

$$(I.5.2) \qquad P'_k = \frac{P_k}{P_0 + P_1 + \ldots + P_r} = \frac{\beta_k}{S_r} \qquad \left(S_r = \sum_0^r \beta_h \right)$$

(puisque $P_k = \beta_k P_0$; $\beta_0 = 1$).

Désignons par $\mathcal{A}_n$ et $\mathcal{A}'_n$ les hypothèses $H_n \leqq r$ et $H_n > r$ et leurs probabilités par $1 - \eta$ et η. Dans l'hypothèse $\mathcal{A}_n$ et en supposant toujours $H_0 = 0$, la probabilité de $H_n = k \leqq r$ est

$$(I.5.3) \qquad P''^{(n)}_{0,k} = \frac{1}{1-\eta} P^{(n)}_{k,0}.$$

Dans cette hypothèse, H_n figure dans la suite des H'_ν avec un rang *aléatoire* ν et

$$P''^{(n)}_{0,k} = E\{ P'^{(\nu)}_{0,k} \}.$$

Or, on peut choisir $\bar{\nu}$ assez grand pour que $\nu \geqq \bar{\nu}$ entraîne

$$| P'^{(\nu)}_{0,k} - P'_k | < \varepsilon,$$

puis $\bar{n}$ assez grand pour que $n > \bar{n}$ entraîne

$$\Pr\{ \nu < \nu' \} < \varepsilon \quad (11).$$

On a alors

$$| P''^{(n)}_{0,k} - P'_k | < 2\varepsilon,$$

c'est-à-dire que, pour n infini, $P''^{(n)}_{0,k}$ tend vers P'_k.

Il reste à montrer que, pour r assez grand,

$$P'_k - P_k \quad \text{et} \quad \operatorname{Max} | P''^{(n)}_{0,k} - P^{(n)}_{0,k} |$$

peuvent être rendus arbitrairement petits. Il en résultera bien qu'en prenant d'abord r, puis n assez grand, $P^{(n)}_{0,k} - P_k$ (qui ne dépend pas de r) peut être rendu arbitrairement petit, donc que $P^{(n)}_{0,k}$ tend, pour n infini, vers P_k.

Pour cela, il faut distinguer deux cas.

Premier cas : $S = \sum_0^\infty \beta_h$ *est fini.* — C'est le cas fortement ergodique. D'après (I.5.1), on a

$$\Pr\{ \mathcal{A}'_n \} = \sum_{r+1}^\infty P^{(n)}_{0,k} \leqq \sum_{r+1}^\infty \beta_k ;$$

donc, pour r assez grand, on peut rendre cette probabilité inférieure, quel que

(11) Le rang $\bar{n}$ de $H'_{\bar{\nu}}$ dans la suite des H_n est en effet une variable aléatoire presque sûrement finie. Il n'y a qu'à choisir $\bar{n}$ de manière que

$$\Pr\{ n > \bar{n} \} < \varepsilon.$$

soit n, à un nombre arbitrairement petit et, d'après (I.5.3), il en est de même de $|P''^{(n)}_{0,k} - P^{(n)}_{0,k}|$. D'autre part, d'après (I.5.2), $P'_k = \frac{\beta_k}{S_r}$ tend pour r infini vers $P_k = \frac{\beta_k}{S}$, ce qui termine la démonstration.

Deuxième cas : la série $\sum \beta_h$ est divergente. — C'est le cas faiblement ergodique. Les P_k sont tous nuls et les P'_k tendent vers zéro. Donc, pour r assez grand, $P'_k \leq P'_0 < \varepsilon$ et, pour n assez grand, compte tenu de (I.5.3),

$$P^{(n)}_{0,k} \leq P''^{(n)}_{0,k} < \varepsilon,$$

de sorte que le résultat est encore vrai dans ce cas.

5° La démonstration du théorème énoncé est maintenant immédiate. Le résultat obtenu pour $P^{(n)}_{0,0}$ s'étend naturellement à $P^{(n)}_{k,k}$, qui tend vers P_k. Si N_1 désigne la première racine de $H_n = k$ et si $H_0 = h \neq k$, on a

$$P^{(n)}_{h,k} = \sum_1^n \Pr\{N_1 = \nu\} P^{(n-\nu)}_{k,k}.$$

On peut déterminer $\bar{\nu}$ et $\bar{n}$ de manière que

$$\Pr\{N_1 > \bar{\nu}\} < \varepsilon,$$
$$|P^{(n)}_{k,k} - P_k| < \varepsilon \qquad (n > \bar{n}).$$

Alors, pour $n > \bar{\nu} + \bar{n}$, on a

$$|P^{(n)}_{h,k} - P_k| < 2\varepsilon, \qquad \text{C. Q. F. D.}$$

6° Le théorème I.5 étant ainsi démontré, il n'y a aucune difficulté à conclure dans tous les autres cas.

a. Si A_k n'appartient pas à un groupe ergodique, le système partant de A_h a une probabilité α d'atteindre A_k au moins une fois. Dans ce cas, il y a presque sûrement une racine N de $H_n = k$ qui est la dernière réalisation de cet état. Pour tout n assez grand, on a

$$P\{N = n\} \leq \Pr\{N \geq n\} < \varepsilon, \qquad \text{donc} \quad P^{(n)}_{h,k} < \alpha\varepsilon \leq \varepsilon,$$

c'est-à-dire que $P^{(n)}_{h,k}$ tend vers zéro.

b. Si A_h et A_k appartiennent à un groupe cyclique d'ordre r et ergodique, la suite des $H_{n_0+\nu r}$ ($\nu = 0, 1, 2, \ldots$) constitue une suite du type acyclique. Si l'on détermine n_0 de manière que $P^{(n_0+\nu r)}_{h,k}$ soit positif (au moins pour certains ν), le théorème I.5 nous montre que $P^{(n)}_{h,k}$, sûrement nul si $n - n_0$ n'est pas multiple de r, tend, pour $n - n_0 = \nu r$ et ν infini, vers une limite qui ne peut être que rP_k (puisqu'en tout cas P_k est la limite en moyenne).

c. Si A_k appartient seul à un groupe ergodique G, α_h désignant toujours la probabilité que le système partant de A_h atteigne G, il est évident, si le groupe G n'est pas cyclique, que $P_{h,k}^{(n)}$ tend vers $H_{h,k}=\alpha_h P_k$. Si le groupe est cyclique d'ordre r, on ne sait pas à l'avance de quelle manière le système atteindra le groupe, c'est-à-dire quelle sera la valeur de n_0 ($n_0=1, 2, \ldots, r$) pour laquelle les $P_{h,k}^{(n_0+\nu r)}$ pourront être positifs. On peut évidemment répartir comme on veut la probabilité entre ces r possibilités; donc, $\varphi(n)$ étant une fonction de la variable entière n, non négative, de période r et de moyenne dans une période égale à 1, on peut définir des processus pour lesquels

$$\lim[P_{h,k}^{(n)}-\alpha_h P_k \varphi(n)]=0.$$

Naturellement, si le groupe G est faiblement ergodique, c'est-à-dire si $P_k=0$ et aussi si $\alpha_h=0$, $P_{h,k}^{(n)}$ tend vers zéro. Si, au contraire, G est cyclique d'ordre r et fortement ergodique et que, de plus, α_h soit positif, $P_{h,k}^{(n)}$ sera *en général* asymptotiquement périodique, c'est-à-dire différera infiniment peu d'une fonction $\alpha_h P_k \varphi(n)$, $\varphi(n)$ étant périodique (de période r), non négatif, en moyenne égale à un.

A cela près, cette fonction est quelconque. En effet, si le système arrive au groupe G, nous ne saurons pas toujours à l'avance comment il y arrivera. Nous pouvons supposer, par exemple, que l'état qui suit immédiatement A_h ne puisse qu'appartenir à un des r sous-groupes de G, ou à un ensemble d'états d'où il ne puisse pas revenir à G; nous pouvons répartir la probabilité comme nous le voulons entre ces $h+1$ possibilités, c'est-à-dire que, sans autre restriction que $0\leqq\alpha_h\leqq 1$ et les conditions triviales indiquées pour $\varphi(n)$, nous pouvons prendre n'importe quelle détermination de α_h et $\varphi(n)$.

On remarque qu'en particulier $\varphi(n)$ peut avoir une période sous-multiple de r, ou même être constant (c'est pour cela que, dans le cas considéré, $P_{h,k}^{(n)}$ n'est qu'*en général* asymptotiquement périodique).

7° *Résumé des principaux résultats. — a. La limite en moyenne* $P_{h,k}$ *existe toujours. Elle est de la forme* $\alpha_{h,k}P_k$, P_k *étant positif si* A_k *appartient à un groupe* G *fortement ergodique et dans ce cas seulement;* $\alpha_{h,k}$ *est la probabilité du passage de* A_h *à ce groupe.*

b. Pour qu'il y ait des oscillations asymptotiquement périodiques, il faut que le groupe G *soit cyclique et fortement ergodique; il suffit que, de plus,* A_k *appartienne aussi à ce groupe. S'il n'y a pas de telles oscillations,* $P_{h,k}^{(n)}$ *tend vers* $P_{h,k}$. *S'il y en a,* $P_{h,k}(t)$ *est la somme d'un terme périodique, à période entière et de moyenne* $P_{h,k}$, *et d'un terme tendant vers zéro.*

CHAPITRE II.

LE CAS CONTINU.
NOTIONS GÉNÉRALES ET ÉTUDE DES SYSTÈMES SANS ÉTATS INSTANTANÉS.

II.1. LES COEFFICIENTS λ_h ET μ_h. — Désignons par $\Phi_h(t)$ la probabilité que le système, s'il est initialement dans l'état A_h, y reste constamment pendant l'intervalle de temps (o, t). C'est évidemment une fonction non croissante et, de plus on a, quels que soient τ positif et n entier positif

$$\Phi_h(n\tau) = [\Phi_h(\tau)]^n.$$

Il en résulte évidemment que $\Phi_h(t)$ est de la forme

$$\Phi_h(t) = e^{-\lambda_h t} \tag{II.1.1}$$

($\lambda_h \geqq o$; les valeurs limites o et $+\infty$ ne sont pas exclues.

Nous attirons l'attention sur le fait que ce résultat simple, bien connu et dont il est inutile de rappeler les nombreuses applications, ne suppose rien d'autre que le caractère markovien et stationnaire de l'évolution. La signification du coefficient λ_h est d'ailleurs très simple : $\lambda_h\, dt$ est, à $O(dt^2)$ près, la probabilité qu'il y ait au moins un changement d'état pendant l'intervalle de temps $(t, t+dt)$, si $H(t) = h$ (12).

Nous désignerons par T le temps au bout duquel le système quitte l'état initial A_h. La loi dont dépend la variable auxiliaire $U = \lambda_h T$ ne dépend d'aucun paramètre. C'est la première loi de Laplace. On sait que $E\{U\} = 1$. Donc la durée probable de chaque séjour du système dans l'état A_h est

$$E\{T\} = \mu_h = \frac{1}{\lambda_h}. \tag{II.1.2}$$

Nous dirons que ce coefficient μ_h est la *vie probable* de A_h.

Si $\lambda_h = o$ (donc $\mu_h = \infty$), l'état A_h est un *état final*. Si le système arrive à cet état, son évolution est terminée.

Si $\lambda_h = \infty$ (donc $\mu_h = o$), A_h est un *état instantané*; si λ_h est fini, l'état A_h sera dit *stable* (sa *stabilité* est mesurée par μ_h).

---

(12) Ces coefficients, avec d'autres notations, ont été considérés par J. L. Doob [1], qui les définit par la formule $P'_{h,h}(+o) = -\lambda_h$. Nous pensons qu'il y a intérêt à les introduire directement. Mais, dans la suite, nous supposerons connue la définition du processus par les fonctions $P_{h,k}(t)$, qui a été rappelée dans l'introduction et qui donne sans doute la seule base possible, en tout cas la plus naturelle, pour l'étude générale des processus markoviens et stationnaires.

Nous appellerons *variable du type* U, ou simplement désignerons par U_1, U_2, ..., des variables dépendant de la première loi de Laplace. Elles joueront naturellement dans la suite un rôle important.

II.2. ÉTATS INSTANTANÉS ET ÉTATS FICTIFS. — Les états instantanés jouent un rôle essentiel pour l'étude de certains processus. Ainsi il peut arriver que $P_{h,k}(t)$ ait une valeur P_k indépendante aussi bien de h que de t. Alors les valeurs successives de $H(t)$ résulteront d'expériences absolument indépendantes les unes des autres et, s'il y a plusieurs états effectivement possibles, il ne peut pas arriver que toutes les expériences faites pendant un intervalle fini, si petit soit-il, donnent toutes le même résultat. Donc, $H(t)$ ne peut pas rester constant pendant un tel intervalle; tous les états sont instantanés.

Il arrive d'autres fois que l'introduction d'états instantanés semble résulter d'une mauvaise définition du processus et qu'il vaille mieux les éliminer. Ainsi, supposons que les états A_h se succèdent nécessairement dans l'ordre des h croissants, que $\mu_1 = 0$, μ_0 et μ_2 étant positifs. Si le système est initialement dans l'état A_0, l'état A_1 aura une existence instantanée à l'instant $T = \mu_0 U$. D'ailleurs, la probabilité que T ait une valeur donnée t est nulle; donc $P_{0,1}(t) = 0$, et naturellement, tous les autres $P_{h,1}(t)$ sont nuls aussi, pour tout t positif. Donc, non seulement l'état A_1, à cause de son caractère instantané, est inobservable, mais il ne joue aucun rôle dans la définition du processus par les fonctions $P_{h,k}(t)$. Il vaut mieux l'éliminer et considérer que le système passe directement de l'état A_0 à l'état A_2.

D'une manière générale, nous dirons qu'un état A_l est un *état fictif* s'il n'intervient pas dans la définition du processus par les fonctions $P_{h,k}(t)$, c'est-à-dire si $P_{h,l}(t) = 0$ quels que soient $t > 0$ et l. Si l'on veut éliminer les états fictifs, on a évidemment le droit de le faire ([13]). Mais nous verrons que, dans des cas moins simples que celui qui vient d'être indiqué, il peut être utile d'introduire des états fictifs, nécessairement instantanés, pour définir le fonctionnement du processus. Naturellement un état fictif ne peut être qu'instantané. En effet, pour un état stable, on a

$$P_{h,h}(t) \geqq e^{-\lambda_h t} > 0,$$

de sorte qu'un tel état ne peut pas être fictif. Par contre, l'exemple indiqué au début du présent paragraphe montre qu'un état instantané peut n'être pas fictif.

([13]) S'il y en a au plus une infinité dénombrable, leur réunion conserve le caractère d'état fictif. S'il y en a une infinité non dénombrable, ou bien leur réunion F conserve le caractère fictif, c'est-à-dire que $P_{h,F}(t) = 0$ quels que soient $t > 0$ et h; ou bien elle le perd, mais en même temps le système étudié cesse d'être un système S_ω à une infinité dénombrable au plus d'états réellement possibles. Ce qui est dit dans le texte est donc toujours exact pour les systèmes S_ω.

II.3. Systèmes sans états instantanés. La durée d'une évolution donnée pour un tel système. — 1° Pour un système sans état instantané, chaque état A_h est, à l'intérieur d'un intervalle donné $(0, t)$, presque sûrement réalisé sur un ou plusieurs intervalles $(i_h, i'_h, \ldots)$, ou n'y est pas réalisé. Désignons par $\mathcal{E}'_t$ la réunion de tous ces intervalles ouverts, et par $\mathcal{E}_t$ l'ensemble complémentaire, c'est-à-dire l'ensemble des points où $H(t)$ est discontinu. Il est nécessairement mesurable. Montrons que : *sa mesure* $\mathbf{m}$ *est presque sûrement nulle*. En effet, dans le cas contraire, elle aurait une valeur probable positive μ. Or $\frac{\mu}{t}$, qui est la probabilité qu'un point θ choisi au hasard entre zéro et t appartenant à $\mathcal{E}_t$, peut se calculer en choisissant θ avant la détermination de $H(t)$, ce qui donne

$$\mu = \int_0^t \Pr\{ t \in \mathcal{E}_t \}\, dt = 0.$$

Donc $\mathbf{m}$ est presque sûrement nul ([14]), C. Q. F. D.

Donc : *t est exactement la somme des temps passés par le système dans les différents états.*

2° Proposons-nous maintenant d'étudier le temps T qu'il faut au système pour parcourir une succession donnée d'états. Leur ordre est évidemment indifférent. Il nous suffit de connaître, pour chaque indice h, le nombre de réalisations n_h de l'état A_h. Le temps total passé dans cet état est alors de la forme $\mu_h S_h$, S_h désignant la somme de n_h variables indépendantes du type U. Elle dépend d'une loi bien connue de Pearson. Il nous suffira de rappeler que

$$\Pr\{ S_h < s \} = \int_0^s \frac{e^{-u} u^{n_h - 1}}{(n_h - 1)!}\, du \qquad (s > 0,\ n_h > 0)$$

et que

$$E\{ S_h \} = \sigma^2\{ S_h \} = n_h.$$

On a ensuite évidemment

$$T = \Sigma \mu_h S_h,$$

tous les S_h étant indépendants les uns des autres.

3° Théorème II.3.1. — T *est presque sûrement fini ou infini en même temps que sa valeur probable* $a = \sum n_h \mu_h$.

Ce théorème est évident dans le cas où l'un au moins des n_h est infini (on peut d'ailleurs le vérifier en appliquant à la somme S_h le raisonnement qui va

([14]) Nous supposons bien entendu que le théorème de Fubini soit applicable. Nous n'insisterons pas ici sur ce point; nous reviendrons sur la légitimité de cette application dans un cas analogue (II.6.2°).

être appliqué à la somme T). Il est évident aussi dans le cas où a est fini (si la variable aléatoire positive T n'était pas presque sûrement finie, sa valeur probable serait infinie).

Supposons donc a infini, mais tous les n_h finis et supposons d'abord tous les μ_h au plus égaux à un. Posons

$$T_h=\sum_1^h \mu_k S_k, \qquad a_h=E\{T_h\}, \qquad b_h=\sigma\{T_h\}.$$

On a

$$b_h^2=\sum_1^h n_k\mu_k^2\leq\sum_1^h n_k\mu_k=a_h,$$

d'où, puisque a_h augmente indéfiniment avec h, $b_h=o(a_h)$. Dans ces conditions, quelque petit que soit ε et quelque grand que soit s, l'inégalité de Tchebycheff montre que, pour h assez grand,

$$Pr\{T_h\leq s\}<\varepsilon,$$

et, comme T_h ne peut que croître avec h, il en résulte bien que sa limite T est presque sûrement infinie.

Il reste à montrer que le résultat subsiste si certains des μ_h sont <1. Remplaçons-les par 1, ce qui ne peut que diminuer T, T_h, a et a_h. Mais a reste infini; en effet, ou bien on n'a changé qu'un nombre fini de termes $n_h\mu_h$, et cela ne peut pas changer la nature de la somme, ou bien on en a changé une infinité; alors les termes réduits sont ≥ 1 et la somme est encore infinie. Le résultat obtenu dans le cas précédent montre alors que, malgré la diminution de certains termes, T est presque sûrement infini (c'est-à-dire que T_h augmente indéfiniment avec h). Cela est vrai *a fortiori* sans cette modification.

4° *Remarque.* — Il est utile, pour la suite, d'attirer l'attention sur le cas particulier du théorème précédent mentionné au début de la démonstration : *Un même état* A_h, *non instantané, ne peut pas être réalisé une infinité de fois en un temps fini.*

5° *L'image moyenne d'une succession donnée d'états.* — Nous appellerons ainsi la figure obtenue en remplaçant chacun des intervalles $i_h^{(\nu)}$ où $H(t)=h$ par un intervalle $\bar{i}^{(\nu)}$ ayant pour longueur la longueur probable μ_h de $i_h^{(\nu)}$. L'ordre de ces intervalles ne sera pas changé et il n'y aura toujours aucun vide entre eux; en d'autres termes, chacun de ces intervalles aura pour extrémité de gauche le point dont l'abscisse est la longueur totale des intervalles placés à sa gauche. En d'autres termes encore, l'ensemble complémentaire $\bar{\mathcal{E}}$ sera de mesure nulle.

Il y a naturellement une correspondance biunivoque (aléatoire) entre l'axe des t et son image; on peut la définir avec précision en supposant que, quand t décrit $i_h^{(\nu)}$, $\bar{t}$ soit une fonction linéaire croissante de t et décrive $\bar{i}_h^{(\nu)}$. Alors $\bar{t}$ est une fonction continue $\varphi(t)$ de t, toujours croissante, et qui varie avec t de zéro à l'infini. La relation $\bar{t}=\varphi(t)$ fait correspondre $\bar{\mathcal{E}}$ à l'ensemble $\mathcal{E}$ des points de discontinuité de $H(t)$.

6° Théorème II.3.2. — *Dans le cas où $\Sigma n_h \mu_h$ est fini, T dépend d'une loi absolument continue, à densité de probabilité positive de zéro à l'infini.*

Nous supposons bien entendu un au moins des n_h positifs; nous ne restreignons rien en supposant que ce soit n_0. Alors $T=\mu_0 U+T'$, T' étant non négatif et indépendant du premier terme (15). Ce premier terme dépendant d'une loi absolument continue à densité de probabilité positive dans $(0,\infty)$, il en est de même de la somme dans l'intervalle $(\mathrm{Min}\,T',\infty)$. Il reste donc à montrer que T (et *a fortiori* $T'\leqq t$) peut être inférieur à un nombre τ arbitrairement petit.

Choisissons d'abord un nombre h assez grand pour que

$$\sum_{h+1}^{\infty} n_k\mu_k = a-a_h < \frac{\tau}{4},$$

et, par suite,

$$\Pr\left\{(T-T_h)\leqq\frac{\tau}{2}\right\}>\frac{1}{2} \qquad \left(T_h=\sum_{1}^{h}\mu_k S_k\right).$$

D'autre part, pour que $T_h<\frac{\tau}{2}$, il suffit que chacune des variables U qui interviennent dans cette somme soit $<\tau/2\sum_0^h n_k\mu_k$. Le nombre de ces variables indépendantes étant $\sum_0^h n_k<\infty$, la probabilité α de cette circonstance est positive et il vient

$$\Pr\{T<\tau\}>\frac{\alpha}{2}>0, \qquad \text{C. Q. F. D.}$$

II.4. Classification des processus. 1° *Remarque préliminaire.* — Plaçons-nous toujours dans le cas des processus sans états instantanés et observons d'abord que l'ensemble $\bar{\mathcal{E}}$ défini au II.3.5° est de mesure nulle et fermé (puisqu'il est le complément d'une réunion d'intervalles ouverts). Montrons que : *à cela près il est quelconque.*

(15) Au premier terme, on pourrait écrire S_0 au lieu de U. En écrivant U, nous évitons d'utiliser les propriétés de la loi de Pearson rappelées plus haut.

En effet, si l'on se donne un tel ensemble $\bar{\mathcal{E}}$ (soit de $-\infty$ à $+\infty$, soit de o à $+\infty$), son complément est la réunion d'au plus une infinité dénombrable d'intervalles ouverts; on peut donc les dénombrer $(\bar{\imath}_1, \bar{\imath}_2, \ldots)$ et faire correspondre à chaque $\bar{\imath}_h$ un état A_h du système. On aura ainsi évidemment l'image moyenne d'un certain processus, dans lequel l'ordre de succession des états est défini, la rapidité de l'évolution étant seule aléatoire.

Naturellement, si plusieurs des intervalles $\bar{\imath}_h$ sont égaux, on peut les faire correspondre à un même état A_h, qui sera ainsi réalisé plusieurs fois. Dans le cas contraire, tous les états introduits d'abord sont nécessairement distincts.

2° *Classification des fonctions* $H(t)$ *possibles.* — Cette classification repose sur celle des ensembles $\mathcal{E}$, qui sont de quatre types possibles.

Premier type, ou *type fini.* — Dans n'importe quel intervalle fini, $\mathcal{E}$ ne comprend qu'un nombre fini de points. Ces points sont alors des sauts de $H(t)$. Réciproquement, si $H(t)$ n'a pas d'autres discontinuités que des sauts, $H(t)$ est du type fini. Naturellement, s'il n'y a pas d'état final, le nombre total des sauts (de zéro à l'infini) est infini.

Deuxième type, ou *type transfini.* — $\mathcal{E}$ est un ensemble bien ordonné, avec au moins un point d'accumulation à distance finie. On peut alors utiliser la numération transfinie pour numéroter, soit les points de discontinuité T_ν, soit les valeurs successives H_ν de $H(t)$. L'instant initial sera $T_0 = t_0$ et la valeur initiale $H(t_0 + o) = H_0$, et l'on aura toujours $H(T_\nu + o) = H_\nu$. Les indices ν utilisés forment ce que A. Denjoy appelle une *tranche commençante* $(1, \bar{\nu})$ de l'ensemble des nombres transfinis de la première classe. Le nombre $\bar{\nu}$ est évidemment un nombre quelconque de cette classe [puisqu'on peut toujours faire l'image de la tranche $(1, \bar{\nu})$ sur l'axe des t]. Mais, bien entendu, si $\bar{\nu}$ est fini ou si $\bar{\nu} = \omega$, on est dans le cas fini; il faut que $\bar{\nu} > \omega$ pour qu'on soit dans le cas transfini.

Observons, en outre, que $\bar{\nu} = \rho + 1$ implique l'existence d'un état final A_ρ, pour lequel $\lambda_\rho = o$.

Troisième type. — $\mathcal{E}$ est dénombrable, mais n'est pas bien ordonné.

On remarque que le premier type est caractérisé par le fait que $H(t - o)$ et $H(t + o)$ existent partout. L'ensemble des deux premiers est caractérisé par le fait que $H(t + o)$ existe partout. Le troisième type est donc caractérisé par l'apparition des points où $H(t + o)$ n'existe pas.

Il y a évidemment en tout, à ce point de vue, quatre espèces de points de discontinuité, qui se distinguent par l'existence ou la non-existence de $H(t - o)$ et $H(t + o)$. Toutes les quatre sont possibles pour le troisième type. On pourrait évidemment le diviser en plusieurs subdivisions, suivant celles de ces espèces qui sont effectivement réalisées. On

pourrait aussi mettre en évidence le nombre des ensembles dérivés de $\mathcal{E}$ qui ne sont pas vides (ce qui introduirait une classification transfinie, pour ce type comme pour le précédent). Cela ne nous a pas paru nécessaire.

Quatrième type. — $\mathcal{E}$ est toujours de mesure nulle, mais n'est pas dénombrable.

Il résulte du 1° que ces quatre types sont effectivement possibles. Nous en donnerons d'ailleurs des exemples. On remarque que cette classification repose sur des caractères qui ne sont pas modifiés si l'on remplace $\mathcal{E}$ par son image moyenne $\overline{\mathcal{E}}$. On peut donc parler indifféremment du type de $\mathcal{E}$, de $\overline{\mathcal{E}}$ ou de $H(t)$.

3° *Cas des systèmes comportant des états instantanés.* — Ces systèmes seront étudiés au Chapitre III; mentionnons tout de suite qu'ils conduisent à introduire deux types nouveaux; pour les distinguer, nous introduirons l'ensemble E_h des racines de $H(t) = h$.

Cinquième type. — Tous les E_h sont mesurables et il existe au moins un état instantané A_h réalisé sur un ensemble E_h de mesure positive [dire que A_h est instantané revient à dire que E_h ne contient aucun intervalle].

Sixième type. — Il existe au moins un E_h non mesurable (donc au moins deux); disons tout de suite que la réunion des E_h non mesurables peut remplir des intervalles, ou au contraire former un ensemble discontinu de mesure positive).

Naturellement, la définition de $\overline{\mathcal{E}}$ ne s'applique plus ici, ou du moins devrait être convenablement précisée.

4° *Classification des processus eux-mêmes.* — La succession des états possibles, que nous avons jusqu'ici supposée connue, est naturellement le plus souvent aléatoire, et le type même de la fonction $H(t)$ qui sera réalisée peut dépendre à la fois du hasard et du choix de l'état initial. La complexité des six types qui viennent d'être définis étant croissante et la distinction entre le cas où les fonctions $P_{h,k}(t)$ sont toutes mesurables et celui où certaines de ces fonctions ne le sont pas ayant aussi une grande importance, nous sommes conduits aux définitions suivantes :

Un processus est du $n^{\text{ième}}$ *type* ($n \leq 6$) *si toutes les fonctions* $P_{h,k}(t)$ *sont mesurables, si la fonction* $H(t)$ *peut être du* $n^{\text{ième}}$ *type, et si elle ne peut pas être d'un type de rang* $> n$. *Il est du septième type si une au moins des fonctions* $P_{h,k}(t)$ *n'est pas mesurable* (il y en aura alors au moins quatre qui ne le sont pas).

II.5. Les coefficients $p_{h,k}$. — 1° *Définition.* — Il est entendu qu'il n'est plus question que des systèmes sans états instantanés; tous les λ_h sont donc finis.

Considérons l'instant aléatoire T_1, où le système quitte l'état initial A_h. Il

peut à ce moment prendre un quelconque des états *autres que* A_h. Désignons par $p_{h,k}$ la probabilité que ce soit l'état A_k. Ces coefficients $p_{h,k}$ vérifient évidemment les conditions

$$p_{h,k} \geqq 0, \qquad p_{h,h} = 0, \qquad p_h = \sum_k p_{h,k} \leqq 1,$$

$p'_h = 1 - p_h$ étant la probabilité de la non-existence de $H(T_1 + 0)$. On voit donc que :

La condition $p_h = 1$ (quel que soit h) *exprime que* $H(t+0)$ *existe en tout point où* $H(t-0)$ *existe. Elle est nécessaire pour que le processus considéré appartienne à un des deux premiers types. Elle est vérifiée, en outre, pour certains processus du troisième type* [ceux pour lesquels $H(t+0)$ peut ne pas exister, mais seulement en des points où $H(t-0)$ n'existe pas non plus].

2° *Conséquences immédiates de la définition.* — Supposons d'abord $p_h = 1$ (quel que soit h). A l'état initial succéderont successivement des états d'indices $H_1, H_2, \ldots$, qui forment une chaîne de Markoff, du type étudié au chapitre I, avec seulement la condition restrictive $p_{h,h} = 0$. Ainsi, *la donnée des* $p_{h,k}$ *définit la loi dont dépend la succession des états réalisés, tant qu'il n'y a eu qu'un nombre fini de changements d'états. Les coefficients* λ_h *antérieurement définis déterminent la loi dont dépend la rapidité de l'évolution* ([16]). Donc : *si l'on connaît à la fois les* $p_{h,k}$ *et les* λ_h *et si* $p_h = 1$ (quel que soit h), *la loi dont dépend l'évolution est complètement définie jusqu'à l'instant* T_ω (si en particulier T_ω est infini, elle est définie depuis l'instant initial jusqu'à l'infini).

A l'instant T_ω, la loi de l'évolution ultérieure n'est plus définie. Si nous voulons définir un processus du type transfini, $H(T_\omega + 0)$ devra exister; mais c'est une variable aléatoire pour laquelle nous pouvons nous donner arbitrairement une nouvelle loi de probabilité, définie par des formules de la forme

$$p_{\omega,k} = \Pr\{H(T_\omega + 0) = k\}, \qquad \text{avec} \qquad p_{\omega,k} \geqq 0, \qquad \sum_k p_{\omega,k} = 1.$$

Mais on n'a pas toujours ainsi toutes les solutions du problème. Il peut arriver que les $p_{\omega,k}$ dépendent de la succession des états qui ont précédé l'instant T_ω. En d'autres termes, on ne pourrait pas considérer qu'il existe un état fictif réalisé à l'instant T_ω, mais il pourra exister plusieurs états fictifs qu'il faudra distinguer, parce que la suite de l'évolution dépendra de celui qui est réalisé. Nous reviendrons sur cette question, Retenons pour le moment que : *après l'instant* T_ω, *la loi dont dépend la suite de l'évolution n'est pas complètement définie par la donnée des coefficients* $p_{h,k}$ *et* λ_h.

([16]) Remarquons tout de suite que, si certains des λ_h sont nuls, ils correspondent à des états finaux et, pour ces valeurs de h, la donnée des $p_{h,k}$ est devenue inutile.

La conclusion de ces remarques est le théorème suivant :

THÉORÈME II.5. — 1° *Un processus du type fini est entièrement défini par la donnée des coefficients* λ_h *et* $p_{h,k}$; 2° *pour un processus du type transfini, ces coefficients suffisent à déterminer la loi de l'évolution jusqu'à l'instant* T_ω ([17]).

Nous verrons qu'il peut arriver aussi qu'un processus du troisième type soit complètement défini par la donnée des coefficients λ_h et $p_{h,k}$, La propriété des processus du type fini que nous venons d'établir n'est donc pas caractéristique.

3° *Remarques relatives au cas fini.* — Pour qu'un proeessus soit de type fini :

a. Il suffit que les λ_h soient bornés supérieurement, donc les μ_h inférieurement. En effet, dans ces conditions, une suite infinie d'états ne peut pas être parcourue en un temps fini.

b. Il faut que $p'_h = 0$ (quel que soit h).

Le rapprochement de ces deux conditions montre que les conditions triviales imposées séparément aux λ_h et aux $p_{h,k}$ ne suffisent pas pour qu'il existe un processus formé avec ces coefficients. Si tous les λ_h sont bornés par un nombre fini λ, et si un au moins des p'_h est positif, un tel processus ne peut pas exister.

Il ne semble pas exister de condition nécessaire et suffisante *simple* pour reconnaître s'il existe un processus du type fini correspondant aux coefficients donnés. La marche à suivre est la suivante : il faut que $p_h = 1$, pour tous les h; les H_n forment alors une chaîne de Markoff bien définie par les $p_{h,k}$ et pour laquelle il est facile de calculer les $p^{(n)}_{h,k}$. Chacune des chaînes ainsi possibles a une durée probable $T = \Sigma N_k \mu_k$ (N_k étant le nombre de réalisations de A_k). Il s'agit de savoir si T est presque sûrement infini.

On remarque qu'une condition nécessaire est que $E\{T\}$ soit infini, c'est-à-dire que, quel que soit h, on ait

$$\sum_k \mu_k \sum_n \Pr\{H_n = k / H_0 = h\} = \infty .$$

Elle n'est pas suffisante.

Si T peut être fini, des données complémentaires permettent, comme nous l'avons dit, de former une infinité de processus du type transfini correspondant aux coefficients donnés. Il peut aussi en exister qui appartiennent à d'autres types; nous reviendrons plus loin sur ce cas.

II.6. LES COEFFICIENTS $p'_{h,k}$. — *Leurs relations avec les coefficients* λ_h et $p_{h,k}$. — 1° Proposons-nous d'évaluer approximativement $P_{h,k}(t)$ pour t très petit. Nous montrerons dans un instant que l'erreur commise en négligeant la possibilité qu'il y ait plus d'un changement pendant l'intervalle de temps t es $o(dt)$.

([17]) *Cf.* W. Feller [1] et J. L. Doob [2].

Admettons-le provisoirement. Il en résulte évidemment que

$$P_{h,h}(t) = e^{-\lambda_h t} + o(t) = 1 - \lambda_h t + o(t),$$

et, si $k \neq h$,

$$P_{h,k}(t) = (1 - e^{-\lambda_h t})\, p_{h,k} + o(t) = \lambda_h p_{h,k} t + o(t).$$

Par suite, on a $P_{h,k}(+0) = \delta_{h,k}$, les dérivées $P'_{h,h}(0)$ et $P'_{h,k}(0)$ existent et ont les valeurs

$$\text{(II.6.1)} \qquad \begin{cases} p'_{h,h} = P'_{h,h}(0) = -\lambda_h, \\ p'_{h,k} = P'_{h,k}(0) = \lambda_h p_{h,k} \qquad (k \neq h). \end{cases}$$

Ces équations montrent l'équivalence de la donnée des $p'_{h,k}$ et de celle de l'ensemble des coefficients λ_h et $p_{h,k}$. Elles sont en effet résolues par rapport aux premiers. Inversement, si l'on se donne les $p'_{h,k}$, on en déduit λ_h, puis les $p_{h,k}$, du moins pour les valeurs de h n'annulant pas λ_h. Mais, si $\lambda_h = 0$, l'état A_h est final, et il n'est pas nécessaire de lui associer des coefficients $p_{h,k}$. Donc la donnée des $p'_{h,k}$ est, au point de vue de la définition du processus, exactement équivalente à celle des λ_h et des $p_{h,k}$. On voit aussi que les conditions triviales qu'il faut imposer aux $p'_{h,k}$ sont

$$(\mathfrak{T}') \qquad p'_{h,h} \leq 0, \qquad p'_{h,k} \geq 0 \qquad (k \neq h),$$

$$(\mathfrak{S}') \qquad \sum_{k \neq h} p'_{h,k} \leq |p'_{h,h}|,$$

cette dernière condition devenant une égalité, soit si l'état A_h est final, soit si, à l'instant T_1 où le système le quitte, $H(T_\omega + 0)$ est presque sûrement bien défini.

2° Démontrons maintenant le résultat admis au début du 1°. Il ne s'agit pas, bien entendu, de démontrer que la probabilité qu'il y ait plus d'un changement pendant l'intervalle de temps $(0, t)$ est $o(t)$; cela est faux si $p'_h > 0$, puisque cette probabilité est au moins $(1 - e^{-\lambda_h t}) p'_h$, donc $\lambda_h p'_h t - O(t^2)$. Il s'agit (que k soit $= h$ ou $\neq h$) de la probabilité d'un changement, autre que le premier, et *intéressant l'état* A_k, de manière à modifier la probabilité de sa réalisation à l'instant t. Si T_1 est l'instant du premier changement et si $T_1 < t$, il s'agit de montrer que la probabilité d'un second changement intéressant l'état A_k est très petite si t, et, *a fortiori*, $t - T_1$, sont très petits. La probabilité de $T_1 < t$ étant $\lambda_h t - O(t^2) = O(t)$, la probabilité que ces deux circonstances soient réalisées successivement sera bien $o(t)$.

D'abord, si l'état A_k est réalisé après un premier changement, la probabilité d'un second changement est majorée par $\mu_k t$. Dans ce cas le résultat énoncé est exact.

D'autre part, si l'état A_k n'est pas réalisé après le premier changement, il s'agit de majorer la probabilité qu'un chemin *conduisant de* A_h à A_k puisse être

parcouru en un temps t. Pour démontrer que cette probabilité tend vers zéro avec t, il suffit d'observer que le temps qui s'écoule, depuis l'instant T_1 où le système quitte l'état A_h pour prendre un état inconnu, mais autre que A_k, jusqu'à la première réalisation de A_k est une variable aléatoire, finie ou infinie, mais presque sûrement positive. La probabilité qu'elle soit $< t$ tend donc vers zéro avec t. C. Q. F. D.

II.7. Les équations de Kolmogoroff. — 1° Dans le cas qui nous occupe, les équations différentielles formées par A. Kolmogoroff [1] à propos de problèmes plus généraux, s'écrivent aisément, au moins si l'on se contente d'un calcul formel. D'une part, en dérivant par rapport à s les équations $(\mathcal{C})$ du paragraphe 2 de l'introduction, puis faisant $s = 0$, il vient

$$(\mathcal{K}_1) \qquad P'_{h,k}(t) = \sum_j p'_{h,j} P_{j,k}(t).$$

Si l'on opère de la même manière après avoir échangé t et s, on trouve

$$(\mathcal{K}_2) \qquad P'_{h,k}(t) = \sum_j p'_{j,k} P_{h,j}(t).$$

Il n'y a pas lieu d'être surpris de trouver deux expressions différentes des mêmes dérivées. Considérons les fonctions $P_{h,k}(t)$ comme les éléments d'un tableau à double entrée dans lequel h indique la ligne et k la colonne. Dans le système $(\mathcal{K}_1)$, on peut fixer k ; on a ainsi un système d'équations vérifié par les fonctions d'une même colonne. Le système $(\mathcal{K}_2)$ correspond au contraire à un groupement par lignes (18).

Dans le cas d'un système $\mathcal{S}_r$ sans états instantanés, les équations $(\mathcal{K}_1)$ et $(\mathcal{K}_2)$ sont toujours exactes et, jointes aux conditions initiales $P_{h,k}(+0) = \delta_{h,k}$, un quelconque de ces systèmes détermine parfaitement les $P_{h,k}(t)$. Dans le cas qui nous intéresse (systèmes $\mathcal{S}_\omega$ sans états instantanés), bien que nous ayons établi l'existence des $p'_{h,k}$ et la validité des conditions initiales $P_{h,k}(+0) = \delta_{h,k}$, la légitimité des dérivations effectuées n'est pas évidente. Cette question a été discutée par J. L. Doob [2]. Nous nous contenterons d'énoncer ses résultats : *l'existence des* $P'_{h,k}(t)$ *et la validité de* $(\mathcal{K}_1)$ *caractérisent les processus des deux premiers types; l'existence des* $P'_{h,k}(t)$ *et la validité de* $\mathcal{K}_2$ *caractérisent les processus pour lesquels* $H(t-0)$ *est presque sûrement bien défini sur tout l'axe des* t.

2° On peut être surpris, les systèmes $(\mathcal{K}_1)$ et $(\mathcal{K}_2)$ ayant l'aspect canonique,

(18) Rappelons que les rôles respectifs de ces deux systèmes apparaissent plus nettement encore si l'on étudie le problème général des processus markoviens (non nécessairement stationnaires). Alors, au lieu de $P_{h,k}(t-s)$, il faut écrire $P_{h,k}(s, t)$ (s, instant initial; t instant final). Les deux systèmes d'équations donnent alors respectivement les dérivées par rapport à s et par rapport à t.

que la solution du problème de Cauchy ne soit pas toujours bien déterminée. Cela prouve simplement que les théorèmes de Cauchy ne peuvent pas s'étendre sans précaution aux systèmes d'ordre infini. Il y a d'ailleurs à ce point de vue une différence importante entre les deux systèmes d'équations $(\mathcal{K}_1)$ et $(\mathcal{K}_2)$, provenant de ce que

$$\sum_j |p'_{h,j}| = \lambda_h(1+p_h) \leqq 2\lambda_h.$$

Donc, dans chaque équation de $(\mathcal{K}_1)$, la somme des coefficients est finie. Si, de plus, les λ_h ont une borne supérieure finie λ (ce qui implique que le processus étudié soit du type fini, donc $p'_h = 1$), la méthode de Cauchy s'applique sans difficulté et montre que toutes les fonctions $P_{h,k}(t)$ sont représentables par des séries entières majorées par la série

$$e^{2\lambda t} = \nu \sum_0^\infty \frac{(2\lambda t)^n}{n!}.$$

3° Supposons toujours $\lambda_h \leqq \lambda < \infty$ et désignons par $P^{(n)}_{h,k}(t)$ la probabilité que, si $H(o) = h$, l'état A_k soit réalisé à l'instant t après n sauts. Dans le cas fini, on a toujours

$$P_{h,k}(t) = \sum_0^\infty P^{(n)}_{h,k}(t). \tag{II.7.1}$$

De plus, si $\lambda_h \leqq \lambda < \infty$ et si le passage de A_h à A_k en n sauts est possible, il est presque évident que $P^{(n)}_{h,k}(t)$ est positif de zéro à l'infini et a, pour t infiniment petit, une valeur principale de la forme $c_{h,k}t^n$. Dans le cas contraire, $P^{(n)}_{h,k}(t)$ est constamment nul (sauf pour $t = o$, si $n = o$, $h = k$).

Par suite, *la formule* (II.7.1) *représente, pour t infiniment petit, un développement asymptotique du premier membre, dans lequel le* $(n+1)^{ième}$ *terme est* $c_{h,k}t^n + o(t^n)$. On remarque, en outre, que chaque fonction $P_{h,k}(t)$ est, ou bien constamment nulle, ou bien constamment positive. Nous ne faisons ici que mentionner ce théorème, sur lequel nous reviendrons au paragraphe II.8.

4° Si les λ_j ne sont pas bornés, les résultats précédents ne sont en général plus exacts. On peut toutefois faire une remarque sur la fonction

$$S^{(n)}_h(t) = \sum_k P^{(n)}_{h,k}(t)$$

qui est la probabilité que, si $H(o) = h$, il y ait exactement n sauts avant l'instant t. Un raisonnement analogue à celui fait plus haut (II.4.2°) à propos

du second saut montre que, dans le cas fini, s'il y a n sauts, la probabilité du $(n+1)^{\text{ième}}$ reste très petite pour t très petit, c'est-à-dire que

$$S_h^{(n+1)}(t) = o[S_h^{(n)}(t)] \qquad (t \to o).$$

Par suite, la formule

$$1 - e^{-\lambda_h t} = \sum_1^\infty S_h^{(n)}(t)$$

donne un développement asymptotique du premier membre, pour t infiniment petit. Ce résultat subsiste dans le cas transfini, à cela près qu'au second membre la série doit être transfiniment prolongée. Il ne subsiste pas, bien entendu, s'il peut exister des points (non connus à l'avance), pour lesquels $H(t+o)$ n'existe pas.

Il peut arriver que le résultat, vrai pour la somme $S_h^{(n)}(t)$, le soit pour chaque terme (c'est-à-dire que le développement (II.5.1) de $P_{h,k}(t)$ ait le caractère asymptotique déjà signalé dans le cas où $\lambda_h \leqq \lambda < \infty$. Mais, comme nous l'avons dit, il n'en est pas toujours ainsi. En tout cas, dans un groupe cyclique, le développement considéré ne peut avoir un caractère asymptotique qu'en faisant abstraction des termes nuls. Il est facile de définir d'autres cas où ce caractère asymptotique n'existe pas. Il serait peut-être plus intéressant, à défaut de condition nécessaire et suffisante, de définir des cas aussi étendus que possibles où le développement considéré est sûrement un développement asymptotique.

II.8. Deux théorèmes sur les fonctions $P_{h,k}(t)$. — 1° Lemme II.8.1. — *Si le système peut passer de l'état A_h à l'état A_k, l'instant $T_1^{(k)}$ où il atteint A_k pour la première fois est une variable aléatoire absolument continue, à densité de probabilité positive de zéro à l'infini.*

L'état A_k ne pouvant pas être réalisé une infinité de fois en un temps fini, on peut numéroter les instants $T_\nu^{(k)}$ ($\nu = 1, 2, \ldots,$) de ses réalisations successives. Nous n'excluons pas le cas où $k = h$; alors $T_1^{(k)}$ sera l'instant où le système revient pour la première fois à l'état A_h après l'avoir quitté. S'il n'y a pas ν réalisations de A_k, nous considérerons que $T_\nu^{(k)}$ est infini.

On a évidemment

$$T_1^{(k)} = \Sigma \mu_j S_j,$$

N_j désignant le nombre des réalisations de l'état A_j qui précèdent la première réalisation de A_k (donc $N_k = o$ si $k \neq h$ et 1 si $k = h$), et S_j étant la somme de N_j variables aléatoires indépendantes du type U défini au paragraphe II.1.

Le théorème est tout à fait analogue au théorème II.3.2; mais ici les N_j sont aléatoires.

Tout d'abord, l'état initial étant A_h, $T_1^{(k)}$ est la somme d'un premier terme $\mu_h U$ et d'une somme partielle non négative et indépendante de ce premier terme. Donc cette variable aléatoire dépend d'une loi absolument continue et à densité de probabilité positive depuis un certain minimum m jusqu'à l'infini. Il s'agit de montrer que $m = 0$, c'est-à-dire que, quel que soit $\tau > 0$, on a

$$\varphi(\tau) = \Pr\{T_1^{(k)} < \tau\} > 0.$$

Désignons par $\psi_\tau(N)$ la probabilité conditionnelle de la même inégalité quand on connaît les N_j. Par hypothèse, le système peut passer de A_h à $\mathring{A}_k$, c'est-à-dire qu'il y a une probabilité positive α que, partant de A_h, il prenne un chemin conduisant à A_k et dont la durée de parcours probable $\Sigma N_h \mu_h$ soit finie. Pour chacun de ces chemins, d'après le théorème II.3.2, $\psi_\tau(N) > 0$. Donc

$$\varphi(\tau) = E\{\psi_\tau(N)\} > 0, \qquad \text{C. Q. F. D.}$$

2° On peut faire une objection au raisonnement qui précède : l'expression $\psi_\tau(N)$ est-elle bien une variable aléatoire ? En d'autres termes, sa valeur probable n'est-elle pas l'intégrale d'une fonction non mesurable, c'est-à-dire une expression dépourvue de sens ?

Ce serait un progrès important, pour la théorie générale des processus stochastiques, de pouvoir arriver, à l'aide d'un théorème suffisamment général, à éviter d'avoir à examiner cette objection dans chaque cas particulier. Il semble que pour nos systèmes $\mathcal{S}_\omega$ sans états stationnaires, on ne risque pas de se trouver en présence de fonctions non mesurables. Plus généralement, dans la théorie des probabilités dénombrables, une fonction non aléatoire, déduite sans ambiguïté des données, donc *nommée*, est mesurable si l'on n'a pas introduit dans les données des fonctions non mesurables. Aussi ne reviendrons-nous pas dans la suite sur les objections de ce genre. Mais dans le cas qui nous occupe nous allons indiquer brièvement un raisonnement qui évite cette objection.

D'abord, en faisant abstraction des états d'indices $l > r$, on obtient un système $\mathcal{S}_r$ pour lequel il n'y a pas de difficulté. On en déduit, si h et k sont $\leq r$, que $\sum_0^r N_j \mu_j$ est une variable aléatoire; donc les séries infinies $\Sigma N_j \mu_j$ et $\Sigma \mu_j S_j$, dans les cas de probabilité α où toutes les deux sont convergentes, représentent aussi des variables aléatoires. On peut alors déterminer l de manière que

$$\Pr\left\{\sum_{l+1}^{\infty} \mu_j S_j > \frac{\tau}{2}\right\} < \frac{\alpha}{4},$$

puis c de manière que,

$$\Pr\{N_1 \mu_1 + \ldots + N_l \mu_l > c\} < \frac{\alpha}{4}.$$

On aura alors

$$\Pr\left\{N_1 \mu_1 + \ldots + N_l \mu_l \leq c,\ \sum_{l+1}^{\infty} \mu_j S_j \leq \frac{\tau}{2}\right\} > \frac{\alpha}{2}.$$

D'autre part, les termes $U_{j,\nu}$ qui interviennent dans les sommes S_1, S_2, ..., S_l sont indépendants de N_1, ..., N_l, S_{l+1}, S_{l+2} ...; il y a donc une probabilité positive β qu'ils soient tous inférieurs à $\frac{\tau}{2c}$, donc que

$$\mu_1 S_1 + \ldots + \mu_l S_l < \frac{\tau}{2c}(N_1\mu_1 + \ldots + N_l\mu_l).$$

Il vient donc finalement

$$\Pr\left\{\sum_1^\infty \mu_j S_j < \tau\right\} > \frac{\alpha\beta}{2} > 0, \qquad \text{C. Q. F. D.}$$

3° Posons $\varphi(\tau) = \alpha F(\tau)$, de sorte que $F(t)$ est la fonction de répartition conditionnelle de $T_1^{(k)}$, relative au cas où $H(0) = h$ et où cette variable aléatoire existe (19). Désignons de même par β la probabilité de l'existence de $T_1^{(k)}$ si $H(0) = k$, et, par $G(t)$ la fonction de répartition de cette variable, si elle existe.

Plaçons-nous alors dans le cas où $H(t) = h$ et désignons par $T_1^{(k)}$, $T_2^{(k)}$, ..., les extrémités de gauche des intervalles correspondant aux réalisations successives de A_k. Toutes les différences $T_n^{(k)} - T_{n-1}^{(k)}$, sont indépendantes, et dépendent de la même loi, définie par $\beta G(t)$, de sorte que la probabilité de l'existence de $T_n^{(k)}$ est $\alpha\beta^{n-1}$ et, si cette variable existe, sa fonction de répartition est le produit de composition de *Volterra-Stieltjes*

$$F_n(t) = F(t) \star [G(t)]^{\star(n-1)} \quad (20).$$

Sous la même condition, la probabilité qu'une valeur donnée de t appartienne au $n^{\text{ième}}$ intervalle de réalisation de A_k est alors

$$\int_0^t e^{-\lambda_k(t-u)}\, dF_n(u) = e^{-\lambda_k t} \star F(t) \star [G(t)]^{\star(n-1)}.$$

Si alors A_k est réalisé à l'instant t, t appartient à un des intervalles de réalisation, dont le rang n ne peut être que fini. On en déduit que, pour $k \neq h$

$$P_{h,k}(t) = \sum_1^\infty \alpha\beta^{n-1} \int_0^t e^{-\lambda_k(t-u)}\, dF_n(u), \tag{II.8.1}$$

c'est-à-dire, symboliquement

$$P_{h,k}(t) = e^{-\lambda_k t} \star \alpha F(t) \star [1 - \beta G(t)]^{\star(-1)}, \tag{II.8.1'}$$

---

(19) Nous préférons mettre en évidence cette fonction $F(t)$ qui varie de zéro à un. Mais les formules seraient un peu plus simples en conservant la notation $\varphi(t)$.

(20) Nous définisons ce produit par la formule

$$F(t) \star G(t) = \int_0^t F(t-u)\, dG(u) = \int_0^t G(t-u)\, dF(u),$$

et nous utiliserons aussi la notation inspirée par celles de Volterra

$$[1 - \beta G(t)]^{\star(-1)} \equiv 1 + \sum \beta^n [G(t)]^{\star n}.$$

tandis que, si $h = k$, il faut tenir compte de l'intervalle initial, et il vient

$$\text{(II.8.2)} \qquad \begin{cases} P_{k,k}(t) = e^{-\lambda_k t} + \sum_1^\infty \beta^n \int_0^t e^{-\lambda_k (t-u)}\, d[G(u)]^{*n} \\ \qquad = e^{-\lambda_k t} + e^{-\lambda_k t} * \beta G(t) * [1 - \beta G(t)]^{*(-1)}. \end{cases}$$

Il nous a paru utile d'indiquer ces formules exactes. Pour la suite, il suffit de retenir que ces fonctions sont au moins égales à leurs premiers termes. Donc

$$\text{(II.8.3)} \qquad P_{h,h}(t) \geqq e^{-\lambda_h t},$$

et, si $k \neq h$,

$$\text{(II.8.4)} \qquad P_{h,k}(t) \geqq \alpha \int_0^t e^{-\lambda_k (t-u)}\, dF(u) \geqq \alpha\, e^{-\lambda_k t} F(t).$$

4° Théorème II.8.1. — *Chacune des fonctions* $P_{h,k}(t)$ *est, ou bien nulle de zéro* (inclus) *à l'infini, ou bien positive de zéro* (exclu, sauf si $h = k$) *à l'infini.*

Si en effet $k \neq h$ et $\alpha = 0$, le passage de A_h à A_k est impossible; $P_{h,k}(t)$ est toujours nul.

Si $\alpha > 0$, il résulte du lemme II.8.1 que $F(t)$ est positif pour tout $t > 0$ et, par suite de (II.8.4) que $P_{h,k}(t)$ est positif. $P_{h,h}(t)$ l'est aussi, d'après (II.8.3)

Nous verrons au chapitre III que ce théorème reste vrai pour le cinquième type et pour le sixième. Pour l'instant il n'est établi qne pour les systèmes sans états instantanés.

5° Théorème II.8.2. — *Même pour les systèmes à états instantanés, pourvu que les fonctions* $P_{h,k}(t)$ *soient mesurables, elles tendent, pour t infini, vers des limites* $P_{h,k}$.

Choisissons arbitrairement un nombre positif τ. La suite des $H(n\tau)$ est une chaîne stationnaire de Markoff. Il résulte alors du paragraphe I.5.7° que, pour n entier,

$$\text{(II.8.5)} \qquad P_{h,k}(n\tau) = P^{(1)}_{h,k}(n\tau) + P^{(2)}_{h,k}(n\tau),$$

le premier terme, considéré comme fonction de $t = n\tau$, étant une fonction périodique de période multiple de τ et le second tendant vers zéro pour n infini.

Utilisons maintenant un théorème de J. L. Doob [1], d'après lequel, si les fonctions $P_{h,k}(t)$ sont toutes mesurables, elles sont continues pour tout $t > 0$. Il résulte d'ailleurs de son raisonnement que la continuité est uniforme dans tout intervalle (t_0, ∞) $(t_0 > 0)$([21]).

([21]) Il en résulte aussi que son résultat s'applique à l'ensemble des fonctions correspondant à un

Utilisons, d'autre part, une idée de N. Kryloff et N. Bogoliouboff [1], en appliquant la formule de décomposition (II.8.5) pour deux valeurs τ' et τ'' de τ dont le rapport soit irrationnel; nous écrirons alors $P'^{(i)}$ et $P''^{(i)}$ $(i=1,2)$ au lieu de $P^{(i)}$. Nous pouvons choisir deux suites $\{t'_n\}$ et $\{t''_n\}$ de nombres indéfiniment croissants, respectivement multiples de τ' et τ'' et tels que $t'_n - t''_n$ tende vers zéro. Il en est alors de même de

$$P_{h,k}(t'_n) - P_{h,k}(t''_n), \quad P'^{(2)}_{h,k}(t'_n), \quad P''^{(2)}_{h,k}(t''_n),$$

et par suite de

$$P'^{(1)}_{h,k}(t'_n) - P''^{(1)}_{h,k}(t''_n),$$

ce qui n'est manifestement possible que si ces deux fonctions sont constantes et égales.

Il résulte alors de la formule (II.8.5) que, quel que soit τ, $P_{h,k}(n\tau)$ tend vers une constante $P_{h,k}$ indépendante de τ. La fonction $P_{h,k}(t)$ étant uniformément continue, tend vers la même constante quand t augmente indéfiniment d'une manière quelconque, C. Q. F. D.

II.9. — Propriétés presque sûres de $H(t)$. — On remarquera que ces propriétés vont être obtenues indépendamment de celles des fonctions $P_{h,k}(t)$ qui viennent d'être établies. Le point de départ est la notion bien claire de la probabilité $\alpha = \alpha_{h,k}$ que le système, partant de A_h, atteigne A_k en un temps fini $T_1^{(k)}$. Il n'est même pas nécessaire de savoir que $\alpha > 0$ entraîne $P_{h,k}(t) > 0$ dès que t dépasse la borne inférieure des valeurs possibles de $T_1^{(k)}$ [ce qui d'ailleurs se déduit bien simplement de (II.8.4)].

Cette probabilité étant définie, la notion de groupe et celles de groupe final, semi-final ou de passage se définissent exactement comme dans le cas des chaînes. La démonstration du théorème I.3.1 et la définition des groupes ergodiques subsistent aussi sans changement. Il en est de même du théorème I.3.2 de ses extensions (loi des écarts, loi du logarithme itéré), du retour aux probabilités de passage (§ I.4. 1°) qui montre simplement l'existence d'une limite en moyenne (ici, pour t infini), et de la distinction entre les groupes faiblement et fortement ergodiques, à cela près qu'il faut prendre comme point de départ, non le nombre probable c_h des réalisations d'un état A_h entre deux réalisations consécutives d'un état A_0 pris comme dénominateur commun, mais le temps probable $c_h\mu_h$ passé dans cet état. Le système sera fortement ergodique si la série $\Sigma c_h\mu_h$ est convergente et faiblement ergodique si elle est divergente. La série Σc_h elle-même ne peut d'ailleurs être convergente que dans le cas fini.

Plaçons-nous dans ce cas. La série Σc_h peut aussi bien être convergente que

---

même indice final, même si les autres ne sont pas mesurables. Rappelons d'ailleurs que, si λ_h est fini, la continuité de $P_{h,k}(t)$ est évidente. On a en effet

$$|P_{h,k}(t+\varepsilon) - P_{h,k}(t)| \leq 1 - e^{-\lambda_h \varepsilon}.$$

Si λ_k est fini, on déduit aisément le même résultat des remarques du II.6.2°.

divergente et, dans chacun de ces cas, on peut réaliser le cas voulu pour $\Sigma c_h \mu_h$ en choisissant convenablement les μ_h. Prenons par exemple $c_h = \frac{1}{h^2}$, $\mu_h = h$. La chaîne des indices H_n successivement réalisés sera fortement ergodique, et $\Pr\{H_n = k\}$ aura, pour n infini, si le groupe n'est pas cyclique, une limite positive indépendante de l'état initial. Mais le système restera plus longtemps dans les états d'indices élevés, et il en résultera qu'il deviendra faiblement ergodique et que $P_{h,k}(t)$ tend vers zéro pour t infini. Le contraire a lieu si l'on prend par exemple $c_h = \mu_h = \frac{1}{h}$ ([22]).

Considérons maintenant un processus qui ne soit pas du type fini et supposons, pour simplifier, le système réduit à un groupe ergodique unique. Le nombre probable total des changements d'états qui séparent deux réalisations consécutives de A_0 est alors nécessairement infini (puisque, dans des cas de probabilité positive, il y a au moins un point de discontinuité non isolé). La série Σc_h est donc toujours divergente. Mais la série $\Sigma c_h \mu_h$ est convergente ou divergente, suivant les cas. On peut aisément former des exemples de ces deux possibilités.

Attirons maintenant l'attention sur une circonstance spéciale relative au cas transfini : *l'extension de la notion de groupe liée au prolongement transfini de la chaîne des* H_n. Un exemple fera bien comprendre ce phénomène.

Supposons que les états A_h soient répartis en groupes de passage G_ν nécessairement parcourus dans l'ordre des ν croissants, le premier élément de chaque groupe étant imposé et la durée probable m_ν de la traversée de chaque G_ν étant finie. Si Σm_ν est fini, ν augmente indéfiniment en un temps fini; supposons qu'à ce moment il y ait une probabilité α_ν que le système revienne au premier état de G_ν, avec $\alpha_0 > 0$ et $\Sigma \alpha_\nu = 1$. Cette possibilité de retour à G_0 fait que maintenant tous les G_ν vont former un groupe unique, fortement ergodique.

D'une manière générale, il peut arriver que certains passages d'un état à un autre, qui ne sont pas possibles s'il n'y a qu'un nombre fini d'intermédiaires, le deviennent par le prolongement transfini de la chaîne des H_n et la définition des groupes se trouve modifiée en conséquence. Dans l'exemple précédent, et toutes les fois que, après un point d'accumulation de sauts, on ne peut que choisir le nouvel état (H_ω ou $H_{2\omega}$, . . ,) parmi les états antérieurs, et parcourir à nouveau la même suite d'états, toutes les différences $T_{n\omega} - T_{(n-1)\omega}$ auront la même valeur probable, et il est presque sûr que $T_{n\omega}$ augmente indéfiniment avec n. L'énumération transfinie des indices sera alors bornée à ω^2. Pour dépasser ω^2, il faut de nouveaux états toujours renouvelés; si l'on considère la

([22]) Nous hésitons à proposer de nouvelles définitions. Peut-être pourrait-on conserver pour les chaînes la terminologie du chapitre I, et, pour les systèmes *faiblement ergodiques dans le temps*, adopter l'expression de *systèmes dissipatifs* proposée par F. G. Foster [1], et déjà mentionnée plus haut (note du § I.3.4°).

suite des états fictifs réalisés aux instants $T_{n\omega}$, et dont chacun détermine la loi dont dépend le choix de $H(T_{n\omega}+o)$, certaines répétitions ne sont pas exclues; mais il faut que la fuite vers un nouvel infini soit possible; autrement dit, que l'ensemble de ces états fictifs ne constitue pas un groupe ergodique. Sous des conditions analogues les indices utilisés pour l'énumération des états successifs du système pourront atteindre et dépasser tous les nombres transfinis précédant le nombre transfini $\bar{\nu}$ qu'on aura choisi comme limite, qui peut être choisi arbitrairement dans la seconde classe. Il est en tout cas essentiel qu'aucun état ne soit réalisé une infinité de fois avant que ce nombre $\bar{\nu}$ soit atteint. Il dépendra alors du choix des μ_h que tous les indices précédant $\bar{\nu}$ soient atteints en un temps presque sûrement fini.

Remarquons que, quel que grand que soit $\bar{\nu}$, cela n'empêche pas que tous ces états peuvent former un groupe unique. Ce sera le cas, par exemple, si tous les états d'indices $\nu < \bar{\nu}$ se succèdent dans l'ordre des indices croissants, à cela près que chaque état A_ν donne une probabilité très faible α_ν au retour à A_0. Si la somme de tous les α_ν est finie, l'ensemble forme un groupe non ergodique; après un nombre plus ou moins grand de retours à A_0, la fuite vers $\bar{\nu}$ finira par se produire.

II.10. Exemples. — Les trois premiers exemples n'ont pas d'autre but que de rappeler des exemples concrets de types de processus bien connus depuis le Mémoire de J. L. Doob [2].

1° $$p_{h,h+1}=1-\alpha_h, \qquad p_{h,0}=\alpha_h, \qquad \Sigma\alpha_h<\infty, \qquad \Sigma\mu_h<\infty.$$

Alors le système atteint en un temps fini un état fictif A_ω, après lequel on définit par des coefficients $p_{\omega,h}$ un nouvel état initial. Si tous les α_h sont nuls (ce qui, si de plus $p_{\omega,0}=1$, donne l'exemple le plus simple de processus du type transfini), on a

$$E\{T_{2\omega}-T_\omega\}=\sum_0^\infty \mu_h \sum_0^h p_{\omega,k}.$$

2° $$p_{h+1,h}=1-\alpha_h, \qquad \Sigma\alpha_h<\infty, \qquad \Sigma\mu_h<\infty,$$

tous les autres $p_{h,k}$ sont nuls. A l'instant où le système quitte l'état A_h, il y a alors une probabilité α_h que $H(t+o)$ ne soit pas défini. Cela n'est possible que d'une seule manière : le passage à l'état fictif A_ω, à partir duquel les autres états seront parcourus dans l'ordre des indices décroissants, jusqu'au retour suivant à A_ω.

Remarques. — Le fait qu'il n'y ait qu'un chemin pour revenir de l'infini, donc qu'une manière d'échapper à ce que $H(t+o)$ ait une valeur déterminée, fait que le processus est bien défini par les $p_{h,k}$ et les λ_h. Il est facile de donner d'autres exemples de cette circonstance. En dehors des exemples de cette nature, ou du type fini, la donnée des λ_h et des $p_{h,k}$ ne suffit jamais à déterminer un processus.

Si $\Sigma\mu_h=\infty$, et au moins un $\alpha_h>0$, les données sont incompatibles. La seconde condition implique la possibilité du passage à un état fictif A_ω, d'où le système ne pourrait plus revenir en un temps fini. Il devrait être prévu comme état possible.

Si l'on modifie l'exemple précédent, en prenant par exemple $p_{h+2,h}=1-\alpha_h<1$, tous les

autres $p_{h,k}$ nuls, il y aurait deux chemins pour revenir de l'infini; il n'y a pas d'incompatibilité. Au contraire, une donnée supplémentaire est nécessaire, pour définir les conditions du choix entre ces deux chemins.

3° $h = 0, \pm 1, \pm 2, \ldots$; $p_{h,h+1} = 1$, tous les autres $p_{h,k}$ nuls; $\Sigma \mu_h < \infty$. $H(t)$ devient infini en un temps fini et passe instantanément de $+\infty$ à $-\infty$. A cet instant, on a un nouveau type de discontinuité; ni $H(t-0)$, ni $H(t+0)$, n'existent. On peut modifier cet exemple et rendre possible, soit le passage de $+\infty$ à une valeur finie, soit celui d'une valeur finie à $-\infty$. On peut ainsi réaliser toutes les combinaisons théoriquement prévisibles entre les divers types de singularités possibles pour $H(t)$.

4° *Exemple de processus du quatrième type.* — Établissons une correspondance biunivoque $r_h = f(h)$ entre les états A_h et les nombres rationnels r_h; $\mu_h = \varphi(r_h)$, $\varphi(r+1) = \varphi(r)$; $\varphi(r) = 0$ pour r irrationnel. Aux $r_h \geqq r$ et $< r+1$ correspondent des μ_h de somme finie S. Enfin, les A_h se succèdent dans l'ordre des r_h croissants, la rapidité de l'évolution étant seule aléatoire. Alors $f[H(t)]$ est une fonction continue, non décroissante; à chaque r_h correspond un arrêt de durée probable μ_h; aucun temps d'arrêt n'est réservé à l'ensemble des nombres irrationnels. Ils correspondent à des points de discontinuité, pour chacun desquels ni $H(t-0)$, ni $H(t+0)$ n'existent, et qui forment, sur l'axe des t, un ensemble de mesure nulle.

Il y a dans cet exemple une sorte de périodicité moyenne. Le temps probable pour que $f[H(t)]$ augmente d'une unité est toujours S, et les écarts relatifs à prévoir sont d'autant plus faibles que le plus grand des $\frac{\mu_h}{S}$ est plus petit.

On pourrait ralentir le processus en réservant du temps aux valeurs irrationnelles, qui pourraient par exemple être parcourues avec une vitesse constante donnée. Le processus serait encore markovien et stationnaire. Mais on ne pourrait plus considérer comme fictifs les états correspondant aux valeurs irrationnelles de r; il ne s'agirait plus d'un système $\mathcal{S}_\omega$.

5° *Autre exemple.* — Cet exemple différera essentiellement du précédent, parce qu'un même état fictif A_ω sera réalisé une infinité non dénombrable de fois. Il est défini par

$$p_{h,h-1} = 1 - \alpha_h > 0, \qquad p_{h,\omega} = \alpha_h, \qquad p_{0,\omega} = 1, \qquad \Sigma \alpha_h = \infty \qquad (h = 1, 2, \ldots),$$

et de plus

$$(\text{II}.10.1) \qquad S = \sum_0^\infty \mu_h \varphi(h) < \infty \qquad \left[\varphi(h) = \frac{1}{(1-\alpha_1)(1-\alpha_2)\ldots(1-\alpha_h)} \right],$$

[donc $\varphi(0) = 1$]. La divergence de $\Sigma \alpha_h$ semble exclure la possibilité pour le système de quitter l'état A_ω. Nous allons voir que, grâce à la dernière condition, il n'en est rien.

Faisant d'abord abstraction de toute considération métrique, plaçons sur l'axe des t des intervalles qui indiqueront l'ordre de succession des différents états, l'intervalle $I_h^{(\nu)}$ correspondant à la $\nu^{\text{ième}}$ réalisation de A_h après un intervalle

initial $I_0^{(0)}$. Nous pouvons placer d'abord les $I_0^{(\nu)}$, puis, par des interpolations successives, placer les $I_1^{(\nu)}$, puis les $I_2^{(\nu)}$, et ainsi de suite. Pour chaque indice h, nous commencerons par placer $I_h^{(1)}$ entre $I_0^{(0)}$ et $I_{h-1}^{(1)}$; puis nous effectuerons un tirage au sort pour savoir si cet intervalle $I_h^{(1)}$ doit être suivi par une réalisation de A_{h-1} ou par un retour à A_ω; dans ce dernier cas, $I_h^{(2)}$ sera placé entre $I_h^{(1)}$ et $I_{h-1}^{(1)}$; dans le premier cas, il sera placé à droite de $I_{h-1}^{(1)}$ et éventuellement de tous les autres intervalles $(I_{h-2}^{(1)}, \ldots)$ qui précéderaient $I_{h-1}^{(2)}$, mais à gauche de ce dernier intervalle. Dans les deux cas, on procédera à un nouveau tirage au sort pour savoir si $I_h^{(2)}$ est immédiatement suivi d'un $I_{h-1}^{(2)}$, ou s'il faut placer d'abord $I_h^{(3)}$; et ainsi de suite. Si pour chacun de ces tirages au sort on donne respectivement à ces deux éventualités les probabilités $1-\alpha_h$ et α_h, et qu'on recommence pour toutes les valeurs de h rangées dans l'ordre naturel, il est bien clair qu'au point de vue de l'ordre de succession, on aura réalisé le schéma défini par les $p_{h,k}$. D'ailleurs, puisque $\Sigma\alpha_h = \infty$, donc $\varphi(\infty) = \infty$, deux intervalles $I_h^{(\nu)}$ et $I^{(\nu')}$ finiront toujours par être séparés par une infinité d'intervalles, de sorte que, si on ne laisse aucun vide entre les intervalles placés, l'axe des t sera constitué par des intervalles dans chacun desquels $H(t)$ ne peut que décroître, l'ensemble complémentaire E_ω étant un ensemble parfait discontinu.

Sans changer l'ordre ainsi obtenu, nous pouvons réaliser l'image moyenne de l'évolution qu'il définit en donnant à chaque intervalle $I_h^{(\nu)}$ la longueur μ_h, et à E_ω une mesure nulle, c'est-à-dire que l'origine de chaque $I_h^{(\nu)}$ aura exactement pour abscisse la longueur totale des intervalles analogues qui le précèdent. Il faut bien entendu que cette longueur totale soit finie. Cela résulte de ce que, entre les origines de $I_0^{(\nu)}$ et $I_0^{(\nu-1)}$, par exemple, il y a, outre le premier de ces intervalles, N_1 intervalles $I_1^{(\nu)}$ N_2 intervalles $I_2^{(\nu)}$, et ainsi de suite. Or, $E\{N_h\} = \varphi(h)$ (la vérification par récurrence est immédiate). La distance qui sépare les deux points considérés a donc pour valeur probable l'expression (II.10.1), que nous avons supposé finie. Elle est donc elle-même finie.

Naturellement, ce que nous venons de définir n'est que l'image moyenne de l'évolution. Il reste à faire les expériences qui détermineront les durées $\mu_h U_h^k$ des séjours représentés sur cette image moyenne par les intervalles $I_h^{(\nu)}$. Mais la conclusion qui précède subsiste pour le mouvement réel (d'après le théorème II.3.1). Nous avons donc bien défini un processus pour lequel l'état A_ω est réalisé pour les points d'un ensemble E_ω parfait discontinu, et de mesure nulle.

6° *Remarques.* — Considérant toujours l'exemple précédent, nous pouvons définir sur E_ω un paramètre τ continu et constamment croissant. Si en effet N'_h désigne le nombre des $I_h^{(\nu)}$ compris entre deux points donnés t_1 et t_2 de l'axe des temps, on déduit aisément de la loi forte des grands nombres que $\frac{N'_h}{\varphi(h)}$ a presque sûrement, pour h infini, une limite positive; c'est cette limite qui sera

la variation $\tau_2 - \tau_1$ de τ entre ces deux points. La durée de parcours de ces N'_h intervalles étant d'ailleurs pour h infini un infiniment petit presque sûrement équivalent à $N'_h \mu_h$, donc à $(\tau_2 - \tau_1)\mu_h \varphi(h)$, on peut dire que cette variation de τ mesure le temps consacré, à l'intérieur de l'intervalle (t_1, t_2), aux états d'indices élevés.

Cette définition de τ ayant un caractère stationnaire, il est bien clair que, si l'on prend τ comme variable, le processus reste markovien et stationnaire; $t = X(\tau)$ est alors une fonction aléatoire constamment croissante, ne croissant que par sauts, dépendant d'un processus additif et stationnaire. C'est un type de processus bien connu. On sait qu'il peut être défini par une formule du type

$$\log E\{e^{izX(\tau)}\} = \tau \int_0^\infty (e^{izu} - 1)\, dF(u), \tag{II.10.2}$$

la fonction $F(u)$ étant non décroissante, et infinie négative pour $u = 0$.

Une formule analogue, mais avec une autre fonction $F(u)$, s'applique aussi à l'image moyenne du mouvement.

Remarquons d'autre part que si, au lieu de $t = X(\tau)$, on prend $t = c\tau + X(\tau)$, l'ensemble E_ω sera maintenant de mesure positive; A_ω ne sera plus un état fictif et le processus devra être rattaché au cinquième type.

Ces remarques sont susceptibles de généralisation. Toutes les fois qu'il existe un état instantané A_ω susceptible d'être réalisé sur un ensemble parfait discontinu, on peut définir sur cet ensemble un paramètre τ, et le temps $X(\tau)$ consacré aux états autres que A_ω est une fonction aléatoire du type (II.10.2). Posant alors $t = c\tau + X(\tau)$, on obtient une famille de processus associés, dont un seul (obtenu pour $c = 0$) est du quatrième type; les autres sont du cinquième type.

Nous reviendrons plus loin, à propos des processus du cinquième type, sur les processus du quatrième type ayant au plus une infinité dénombrable d'états fictifs. Par contre si, comme dans l'exemple du 4° ci-dessus, il y a une infinité non dénombrable d'états fictifs, réalisés dans leur ensemble sur un ensemble parfait discontinu E, et si chacun d'eux est réalisé au plus une infinité dénombrable de fois, E doit être de mesure nulle, et le processus ne peut pas être associé à des processus du cinquième type. On peut encore définir sur E un paramètre τ, et étudier $t = X(\tau)$ en fonction de τ. Mais le caractère stationnaire des accroissements de $X(\tau)$ disparaît et, si dans l'exemple rappelé ci-dessus le caractère additif se conserve, il n'en sera pas toujours ainsi, car on ne saura pas toujours à l'avance quel état fictif associer à une valeur donnée de τ.

II.11. Esquisse d'une théorie générale des états fictifs. — Les remarques qui suivent ne sont qu'une esquisse, au cours de laquelle nous ferons quelques hypothèses qui semblent raisonnables, sans démontrer (ce dont nous pouvons seulement dire que cela nous semble exact) qu'elles résultent nécessairement

de celles sur lesquelles repose la définition des processus markoviens. Tout d'abord, nous admettrons qu'on peut raisonner sur un état fictif comme sur un état réellement possible. En d'autres termes si, à un certain instant t, $H(t-o)$[ou $H(t+o)$] n'a pas une valeur bien définie, nous considérerons que le système est à l'instant $t-o$ (ou $t+o$) dans un état fictif A_η (ou A_ζ), et qu'on peut parler des fonctions $P_{\eta,k}(t)$ et $P_{\zeta,k}(t)$ aussi bien que des fonctions $P_{h,k}(t)$. Alors η est une fonction non aléatoire du passé immédiat; sa connaissance résume tout ce qui reste du passé. Une expérience faite à l'instant t lui-même définit le passage de η à ζ, et la connaissance de ζ définit les conditions initiales de l'évolution ultérieure. Ainsi, il y a une transition en trois étapes (passages du passé à η, de η à ζ, de ζ à l'avenir) et, en précisant les lois de ces passages, on peut à première vue espérer compléter les renseignements déduits des $p_{h,k}$ et des λ_h et obtenir dans tous les cas une définition infinitésimale du processus. *A priori* cela n'est pas absurde, puisque, quelque petit que soit $\tau > o$, la connaissance de tous les $P_{h,k}(t)$ pour $o < t < \tau$ suffit à définir le processus. Mais on n'arrive pas à des résultats aussi simples et précis que dans le cas fini, où la donnée des $p_{h,k}$ et des λ_h est suffisante.

Occupons-nous d'abord de la première étape : le passage du passé à la valeur η de $H(t-o)$. En raison du caractère markovien du processus, il suffit de connaître, par exemple, la suite des $H(t-\frac{1}{n})$; la connaissance de chacun de ces nombres rend inutile la connaissance de ceux qui les précèdent. On peut grouper les suites possibles en classes, deux suites appartenant à une même classe si et seulement si elles coïncident à partir d'un certain moment; η a nécessairement la même valeur pour toutes les suites d'une même classe.

Mais ce n'est pas une fonction quelconque des classes ainsi définies. Nous admettrons qu'elle a nécessairement le caractère de fonction P-mesurable (la lettre P indiquant que la mesure est la probabilité de la réalisation des suites considérées). Des propriétés connues des fonctions mesurables résultent alors quelques conséquences importantes qu'un exemple fera bien comprendre.

Supposons que chaque état A_h ait deux formes possibles A'_h et A''_h, de même durée probable μ_h; $\Sigma\mu_h < \infty$ et $p_{h,h+1} = 1$, de sorte que $h = H(t)$ devient infini en un temps fini; quand le système passe de A_h à A_{h+1}, l'indice supérieur a une probabilité α_h de varier. Si alors $\Sigma\alpha_h < \infty$, cet indice a presque sûrement une valeur finale déterminée et l'on peut distinguer deux états fictifs A'_ω et A''_ω. Si $\Sigma\alpha_h(1-\alpha_h) < \infty$, on est aisément ramené au cas précédent. Mais si $\Sigma\alpha_h(1-\alpha_h) = \infty$, une fonction mesurable de la suite des indices ne peut être qu'une constante, et il n'est plus possible de distinguer deux états fictifs A'_ω et A''_ω. Cela est d'ailleurs lié au fait que les indices associés à deux valeurs de h suffisamment éloignées l'une de l'autre sont presque indépendants. Donc, quelque grand que soit h, l'indice supérieur associé à A_h finira par être oublié. Il ne reste alors aucun souvenir du passé permettant, à l'instant où $H(t)$ devient infini, de

distinguer deux états fictifs. Bien entendu, ce dernier raisonnement ne devient rigoureux qu'en vertu de l'hypothèse de mesurabilité indiquée ci-dessus.

D'une manière générale, on déduit de cette hypothèse que la durée accidentellement plus ou moins grande du séjour dans ses états successifs ne peut avoir aucune influence. On peut alors faire abstraction des valeurs effectives de t et ne tenir compte que du *temps moyen* u nécessaire à l'évolution considérée. Si alors on pose $H(t) = K(u)$, η apparaîtra comme une fonction de la suite des $K\left(u - \frac{1}{n}\right)$ (ou d'autres suites analogues, $\frac{1}{n}$ pouvant être remplacé par une autre fonction de n qui tende vers zéro). On est ainsi conduit à considérer l'espace des A_η comme une *fermeture à droite* de l'espace des A_h, dans une topologie où un état A_k serait considéré comme *voisin de* A_h *et à sa droite* si le système peut passer de A_h à A_k en un *temps moyen* très petit, donc en un temps réel très petit en probabilité. De même, l'espace des A_ζ serait une fermeture à gauche.

L'élimination du temps réel pour la définition de A_η n'épuise pas les conséquences du caractère mesurable de η. Il faut que η soit une fonction P-mesurable de la suite des $K\left(u - \frac{1}{n}\right)$. La détermination de cette suite comporte en général un certain nombre de choix, donc de bifurcations. Si ces choix sont rendus définitifs par l'exclusion de toute communication entre les différents chemins possibles, ou si du moins ces communications sont assez peu probables pour que le choix relatif à chaque $K\left(u - \frac{1}{n}\right)$ ait une influence ne tendant pas vers zéro avec $\frac{1}{\nu}$ sur les probabilités relatives à $K\left(u - \frac{1}{\nu}\right)$ ([23]), il peut y avoir plusieurs états fictifs; il peut même y en avoir une infinité non dénombrable. Si, au contraire, cette influence tend vers zéro, il ne peut y avoir qu'un état fictif.

Des remarques analogues, avec naturellement quelques différences dues au fait que les processus considérés ne sont en général pas réversibles, s'appliquent à l'ensemble des A_ζ. Il s'agit ensuite de définir, dans cet espace, une loi de probabilité qui dépende de η, la seule restriction étant l'exclusion des fonctions non mesurables. Tout cela ne semble pas permettre une définition constructive générale et simple, comparable à celle relative au cas fini. Faisons seulement quelques remarques.

2° Donnons d'abord un nouvel exemple de cas où l'ensemble des A_η n'est

---

([23]) Nous entendons par là que, F désignant un ensemble quelconque d'états A_h,

$$\operatorname*{Max}_{h,h',F}\left|P\left\{K\left(u-\frac{1}{\nu}\right)\in F/K\left(u-\frac{1}{n}\right)=h\right\}-P\left\{K\left(u-\frac{1}{\nu}\right)\in F/K\left(u-\frac{1}{n}\right)=h'\right\}\right|$$

ne tend pas vers zéro.

pas dénombrable; mais au lieu que ces états soient tous réalisés dans un ordre déterminé, il n'y aura que des réalisations isolées dont chacune comporte un choix ([24]).

Considérons à cet effet un ensemble d'états $A_{h,r}$ (h entier, r rationnel); $h = H(t)$ suivra le processus simple défini par $p_{h,h+1} = 1$, $\Sigma \mu_h < \infty$. A chaque variation de h, r_h augmentera d'une quantité rationnelle ρ_h; les ρ_h seront indépendants, et la série $\Sigma \rho_h$ presque sûrement convergente, sa somme dépendant d'une loi continue (on peut, par exemple, prendre $\rho_h = \pm \frac{1}{h}$; les valeurs possibles de la somme varient alors de $-\infty$ à $+\infty$). Il y a alors une infinité non dénombrable d'états fictifs $A_\eta = A_{\omega,R}$, R étant un nombre réel quelconque.

On peut faire une hypothèse analogue pour le retour de l'infini. Le système reviendrait de $A_{-\omega,R'}$ (R' étant choisi d'après une loi fonction de R; la plus simple est sans doute $R' = R$; dans ce cas il n'y a pas lieu de distinguer A_η et A_ζ). Mais on peut aussi supposer qu'à chaque h négatif ne corresponde qu'un état A_{-h}, le second indice n'intervenant que si $h \geq 0$. Alors il n'y a qu'un A_ζ et, bien qu'il y ait une infinité de A_η possibles, il devient inutile de les distinguer. On peut aussi supprimer les indices h négatifs et se donner, pour le choix de l'état $A_{0,r}$ qui succédera à $A_{\omega,R}$, une loi qui dépende de R d'une manière quelconque. On voit ainsi qu'un processus de type transfini peut comporter l'existence d'une infinité non dénombrable d'états fictifs A_η, et la nécessité d'introduire pour sa définition des coefficients $p_{R,h,r}$ qui dépendent du paramètre continu R.

3° Dans le cas transfini, on peut numéroter les points de discontinuité et parler de $H(T_\omega - 0)$, $H(T_{2\omega} - 0)$ ou $H(T_{\omega^2} - 0)$, par exemple, comme de variables aléatoires ayant chacune une certaine loi de probabilité dans l'espace des A_η. La même remarque s'applique si, comme dans l'exemple précédent, on peut numéroter les points de discontinuité autres que des sauts. Mais dans d'autres cas, et notamment dans tous ceux du quatrième type, cette numérotation n'est pas possible. On ne peut donc qu'introduire *a priori*, comme nous l'avons fait, la notion d'un état fictif défini par la succession de ceux qui le précèdent ou par une suite partielle extraite de cette succession.

4° Terminons ces remarques en indiquant la relation évidente entre les états A_η et les fonctions $P_{h,k}(t)$. Compte tenu toujours du caractère mesurable de η et de $\mu_k > 0$, on a

$$P_{\eta,k}(t) = \lim_{n \to \infty} P_{h_n,k}(t),$$

([24]) Remarquons d'ailleurs qu'on peut modifier l'exemple II.10.4° en introduisant des sauts dans la variation de r; ils peuvent être vers la droite ou vers la gauche et leur probabilité peut être répartie comme on veut entre les r_h rationnels et les valeurs irrationnelles de r.

$\{h_n\}$ désignant les indices d'une suite d'états dont A_η est la limite à droite (au sens de la topologie définie au 1°) et

$$P_{\eta,k} = \lim_{t \to 0} P_{\eta,k}(t).$$

Si l'on pose alors $p_\eta = \Sigma p_{\eta,k}$, on peut caractériser le cas transfini par l'ensemble des conditions

$$p_h = 1, \qquad p_\eta = 1.$$

pour tous les indices d'états réels ou fictifs. Ces formules montrent bien que cela a un sens de parler des probabilités évaluées à l'instant où un état fictif est réalisé, et qu'il y avait intérêt à se rendre compte des difficultés que présente l'étude directe des états fictifs et la définition des processus par leur intermédiaire.

CHAPITRE III.

Systèmes a états instantanés.

III.1. Un lemme de la théorie des ensembles (lemme de Baire). — *Si la réunion d'au plus une infinité dénombrable d'ensembles* E_n *recouvre entièrement un ensemble compact* $\mathcal{E}$, *on peut trouver au moins un* E_n *et un ensemble ouvert* $\mathcal{O}_n$ *intérieur à* $\mathcal{E}$, *tel que la fermeture* $\overline{E}_n$ *de* E_n *recouvre* $\mathcal{O}_n$.

Si, en effet, $\overline{E}_1$ ne recouvre pas $\mathcal{E}$, on peut trouver dans $\mathcal{E}$ un ensemble ouvert E'_1 dont la fermeture soit extérieure à $\overline{E}_1$. Si E'_1 n'est pas recouvert par $\overline{E}_2$, on peut y trouver un ensemble ouvert E'_2 dont la fermeture soit extérieure à $\overline{E}_2$; et ainsi de suite. Tous les E'_n ont au moins un point commun, qui n'appartient ainsi à aucun $\overline{E}_n$, donc à aucun E_n. Cela est en contradiction avec l'hypothèse que tout point de $\mathcal{E}$ soit recouvert par au moins un E_n. Cette hypothèse implique donc que, pour au moins un n, E'_{n-1} soit recouvert par $\overline{E}_n$ et puisse être pris pour $\mathcal{O}_n$, C. Q. F. D.

Corollaire. — *La réunion des ensembles tels que* $\mathcal{O}_n$ *est partout dense dans* $\mathcal{E}$ (cela résulte évidemment de ce que le lemme s'applique à tout ensemble compact intérieure à $\mathcal{E}$).

Dans la suite, ce lemme ne sera appliqué qu'au cas où $\mathcal{E}$ est un segment de l'axe des t.

III. 2. Processus du cinquième type. — 1° D'après la définition de ce type les fonctions $P_{h,k}(t)$ sont mesurables et il existe au moins un état instantané A_h qui

peut être réalisé sur un ensemble E_h de mesure positive. Supposons d'abord qu'il n'y en ait qu'un, soit A_0. Soit τ la mesure de la partie de E intérieure à $(0, t)$. On sait que, sauf sur un ensemble de mesure nulle, la dérivée $\frac{d\tau}{dt}$ existe et est égale à un sur E_0 et est nulle en dehors de E_0. En ne considérant que la dérivée à droite, cela veut dire que, sauf pour un ensemble e_0 de mesure nulle, un point t choisi sur E_0 a la propriété suivante : la mesure de la portion de E_0 intérieure $(t, t+u)$ est, quand u tend vers zéro, de la forme $u - o(u)$.

Or la valeur probable de cette mesure est

$$\int_0^u P_{0,0}(t)\, dt.$$

Donc $P_{0,0}(t)$ tend en moyenne vers 1 quand t tend vers zéro. D'après un théorème de J. L. Doob [1] déjà rappelé plus haut, si les fonctions $P_{h,k}(t)$ sont mesurables, elles sont continues pour tout t positif. Si elles le sont aussi pour $t = 0$, il résulte de la remarque précédente que $P_{0,0}(+0) = 1$. Il y a lieu de se demander si cette conclusion n'est pas toujours exacte pour le cinquième type.

2° Montrons maintenant que : E_0 *est un ensemble parfait discontinu et* τ *croît constamment sur cet ensemble.*

Observons à cet effet que, si E_0 est connu jusqu'à un certain point t (par exemple celui où τ atteint pour la première fois une valeur donnée), la probabilité que ce point appartienne à e_0 est nulle. De même, la probabilité qu'un intervalle donné comprenne des points de E_0 et que tous ces points appartiennent à e_0, est nulle. Il est, par suite, aussi presque impossible que, parmi l'ensemble des intervalles à extrémités rationnelles, il y en ait un seul qui ait cette propriété. Donc la portion de E_0 intérieure à un intervalle (ouvert) quelconque, est, soit vide, soit de mesure positive (puisqu'il en est ainsi s'il y a un point de E_0 n'appartenant pas à e_0). En d'autres termes encore, τ est constamment croissant sur E_0.

La fermeture $\overline{E}_0$ de E_0 est donc un ensemble parfait discontinu, et E_0 ne peut différer de $\overline{E}_0$ que par un ensemble de mesure nulle. Rien n'empêche sans doute d'imaginer un état instantané A'_0 qui serait réalisé sur un ensemble de mesure nulle, E'_0, sous-ensemble de $\overline{E}_0$. Mais ce serait un état fictif, qui n'aurait aucune chance d'être réalisé pour une valeur donnée de t. Il n'y a aucune raison de l'introduire. Il n'est donc pas nécessaire de distinguer E_0 et $\overline{E}_0$, ce qui achève la démonstration du résultat énoncé.

3° Désignons par $\theta = t - \tau = X(\tau)$ la longueur totale des intervalles extérieurs à E_0 compris entre les points 0 et τ de cet ensemble. Dans ce qui suit, nous pouvons indifféremment raisonner sur θ, ou bien sur sa valeur moyenne $\overline{\theta} = \overline{X}(\tau)$ (durée probable de la succession d'états dont θ est la durée réelle;

elle reste aléatoire puisqu'on ne sait pas à l'avance quels seront ces états). Ces fonctions sont manifestement additives, à accroissements stationnaires [c'est-à-dire que la loi dont dépend $X(\tau + c) - X(\tau)$ ne dépend que de c], et constamment croissantes. Comme nous l'avons dit plus haut, on a alors

$$\log E\{e^{izX(\tau)}\} = \tau \int_0^\infty (e^{izu} - 1)\, dN(u),$$

$N(u)$ désignant une fonction non décroissante, telle que

$$N(0) = -\infty, \qquad \int_0^\infty \frac{u}{1+u}\, dN(u) < \infty.$$

Comme elle n'est définie qu'à une constante près, on peut supposer $N(\infty) = 0$; $c\,|N(u)|$ est alors le nombre probable des sauts supérieurs à u sur un segment de longueur c de l'axe des τ; c'est-à-dire des intervalles extérieurs à E_0 et de longueurs $> u$ entre deux points τ et $\tau + c$ de E_0. Leur nombre réel est évidemment une variable de Poisson; sa loi est bien définie par cette valeur probable.

Les mêmes résultats s'appliquent évidemment à $\bar{0} = \bar{X}(t)$, $N(u)$ devant seulement être remplacé par une autre fonction $\bar{N}(u)$. Il y a d'ailleurs une différence importante entre les deux représentations analogues ainsi obtenues. La durée réelle du parcours d'une succession d'états étant une variable aléatoire continue (et cela est vrai même si cette succession est elle-même aléatoire), la fonction $N(u)$ est continue. La fonction $\bar{N}(u)$, au contraire, peut être continue ou discontinue.

Ainsi, considérons l'exemple du II.10.5°, modifié d'après une remarque du 6°; en prenant $t = \tau + X(\tau)$, on a bien un processus du cinquième type. Dans ce cas les sauts possibles de $\bar{X}(\tau)$, donc les discontinuités de $\bar{N}(u)$, sont les nombres $u_h = \sum_h^\infty \mu_k (h = 0, 1, \ldots)$ et, dans chaque intervalle, (u_h, u_{h+1}) $\bar{N}(u)$ est constant. Si l'on modifie les $p_{h,k}$, en prenant par exemple $p_{h,h-1} = p_{h,h-2} = \frac{1-\alpha_h}{2}$, (pour $h > 1$), et les autres sans changement, les sommes ci-dessus seront remplacées par des sommes partielles ayant une infinité non dénombrable de valeurs possibles; la fonction $\bar{N}(u)$ sera alors continue, mais sa nature peut encore dépendre du choix des μ_h. Si, par exemple, $\mu_h = \frac{1}{(h+1)^2}$, elle est absolument continue, si, au contraire, μ_h tend vers zéro rapidement $\left(\text{par exemple } \mu_h = q_h \text{ avec } q < \frac{1}{2}\right)$, elle ne varie que sur un ensemble parfait discontinu, de mesure nulle).

4° D'après la loi forte des grands nombres, le nombre des sauts de $X(\tau)$ [ou

de $\overline{X}(\tau)]$ supérieurs à un nombre très petit η (aussi bien que la somme des sauts inférieurs à η) est presque sûrement asymptotiquement proportionnel à la variation de τ, et, comme nous l'avons vu à propos d'un exemple (II.10.6°), on peut, en connaissant seulement la suite des sauts de $X(\tau)$, retrouver τ, à un facteur constant près. En prenant alors $t = c\tau + X(\tau)$ $(c \geq 0)$, on a une infinité de processus du cinquième type et un seul du quatrième type (pour $c = 0$).

On peut ainsi considérer qu'on a obtenu une définition constructive pour le processus le plus général du quatrième ou du cinquième type, n'ayant qu'un seul état instantané A_0 (donc l'ensemble des discontinuités est la réunion de E_0 et d'un ensemble dénombrable). En effet, dans chacun des intervalles n'appartenant pas à E_0, le processus se réduit à un processus d'un type antérieurement étudié.

Il faudrait d'ailleurs, pour que le problème d'une définition constructive apparaisse comme définitivement résolu, résoudre deux problèmes accessoires, que nous ne faisons que mentionner ici :

a. Définir pour $N(u)$ et pour $\overline{N}(u)$ le champ des fonctions possibles;

b. Le choix d'une de ces fonctions restreignant l'ensemble des processus possibles, même si l'on avait obtenu pour les trois premiers types une définition constructive maniable, il resterait à choisir parmi eux ceux qui peuvent être associés à une fonction $N(u)$ [ou $\overline{N}(u)$] donnés. En outre, pour la construction effective de $H(t)$ dans chacun des intervalles extérieur à E_0, on se trouvera en présence d'un problème de probabilité conditionnelle, puisqu'on connaît à l'avance la longueur réelle de cet intervalle [ou la variation correspondante de $\overline{X}(\tau)$] ([25]).

5° Supposons maintenant qu'il y ait plusieurs états instantanés, ou même une infinité, nécessairement dénombrable. Dans ces conditions, d'après le lemme de Baire (paragraphe III.1), il est impossible que la réunion de l'ensemble de mesure nulle qui correspond aux états fictifs et des ensembles parfaits discontinus E_η qui correspondent à ces états instantanés remplisse un intervalle. Il faut donc qu'en plus des états instantanés A_η il y ait des états stables A_h, et que l'ensemble des E_h soit partout dense sur l'axe des t. Il est même évident qu'il faut une infinité d'états stables (un même état ne peut, en effet, être réalisé qu'un nombre fini de fois en un temps fini. Or il doit y avoir une infinité d'intervalles correspondant chacun à un état stable).

Dans chacun des intervalles extérieurs aux E_η, on se trouvera dans les mêmes conditions que s'il n'y a qu'un état instantané. D'autre part, en ce qui concerne

([25]) Ce dernier problème semble plus simple que le précédent. Il l'est au moins dans le cas où, comme dans l'exemple du paragraphe II.10.5°, une valeur donnée de la variation de $\overline{X}(\tau)$ dans un de ces intervalles ne peut être réalisée que d'une seule manière.

l'ordre de succession des A_η, on n'a qu'à appliquer exactement ce qui a été dit pour les processus des trois premiers types. Ainsi, s'il n'y a que deux ensembles E_η, on peut, comme nous l'avons expliqué, choisir sur ces ensembles des paramètres τ et τ' (qui pourront être leurs mesures réelles si ces mesures ne sont pas nulles), et la probabilité du passage d'un des états considérés à l'autre pendant un intervalle correspondant à une variation $d\tau$ ou $d\tau'$ sera de la forme $\lambda\, d\tau$ ou $\lambda'\, d\tau'$, λ ne pouvant ici être infini. De même, s'il y a plus de deux états instantanés, le choix des paramètres τ permet de ramener l'étude de leur ordre de succession à celle d'un processus sans états instantanés.

Naturellement, comme nous l'avons déjà observé, l'étude des processus du quatrième type à une infinité dénombrable d'états fictifs ne rentre pas dans le cadre de la construction qui vient d'être définie.

III. 3. Processus du sixième type. — 1° Ils sont caractérisés par le fait que, les fonctions $P_{h,k}(t)$ étant encore mesurables, un au moins des E_h peut ne pas l'être. La réunion des ensembles disjoints E_h étant mesurable, il y en a alors au deux moins qui ne le sont pas.

Nous avons déjà rappelé l'exemple trivial $P_{h,k}(t) = P_k$ (indépendant de h et t). Il y a dans ce cas indépendance complète des différentes valeurs de $H(t)$, donc absence complète de toute continuité. Il est en particulier presque sûr que $H(t)$ n'est pas mesurable.

Attirons à ce sujet l'attention sur le fait suivant : une probabilité nulle est l'inverse d'une puissance. Si l'on choisit un nombre X au hasard entre zéro et un (la répartition étant uniforme), la probabilité de trouver exactement une valeur donnée est nulle; c'est ici l'inverse de la puissance du continu. Si l'on répète cette expérience de manière à considérer une fonction aléatoire $X(t)$ dont toutes les déterminations sont indépendantes, on ne peut pas dire que la probabilité que $X(t)$ prenne exactement au moins une fois une valeur donnée est encore nulle; on ne peut pas non plus dire qu'elle soit positive. C'est une forme indéterminée $\frac{\infty}{\infty}$ à laquelle il est absolument possible d'attribuer une signification ([26]).

Introduisons maintenant un second paramètre u; tous les $X(t, u)$ étant indépendants, demandons-nous quelle est la probabilité pour que, pour au moins un u, $X(t, u)$ se réduise à une fonction donnée de t. C'est encore une forme indéterminée du type $\frac{\infty}{\infty}$; mais ici nous avons au numérateur la puissance du continu et au dénominateur une puissance supérieure et il semble qu'on puisse considérer cette probabilité comme nulle.

Si l'on admet ce genre de raisonnement, on peut montrer qu'il est encore plus impossible de trouver pour $H(t)$ une fonction continue (ou même mesurable) que de trouver exactement un nombre donné par le choix au hasard de ses décimales.

2° L'exemple trivial du 1° se généralise aisément de la manière suivante : on

([26]) Les contradictions auxquelles on arrive en cherchant à lui donner une signification sont identiques à celles dont M. Émile Borel rend l'axiome de Zermelo responsable. Pour nous, elles viennent de ce qu'on cherche à mesurer des ensembles non mesurables.

considère un système d'états possibles dépendant de deux indices h et l; pour chaque valeur de h, l a un certain nombre n_h (fini ou infini) de valeurs possibles; un au moins des n_h est > 1. La variation du premier indice dépend d'un processus d'un des cinq premiers types; pour chaque h, la variation de l dépend d'un processus du type défini au 1° ci-dessus.

Les probabilités de passage de $(h, l_0$ à $k, l)$ sont alors de la forme

$$P_{h,k}(t)\,P'_{k,l}$$

(l_0 n'intervenant pas). On remarque que, si $h = k$ et $n_h > 1$, elle tend, pour $t = 0$, vers la limite $P'_{k,l}$, nécessairement > 0 et < 1).

Considérons un des h pour lesquels $n_h > 1$. Si A_h (en faisant abstraction du second indice), est un état stable, les $E_{h,l}$ ne sont pas mesurables, et leur réunion E_h est une réunion d'intervalles. Si A_h est un état instantané (la variation de h dépendant donc d'un processus du cinquième type), les ensembles non mesurables $E_{h,l}$ forment par leur réunion un ensemble parfait discontinu de mesure positive. L'ensemble des points de discontinuité peut donc comprendre des intervalles, mais n'en comprend pas nécessairement.

3° Nous allons maintenant démontrer que :

THÉORÈME III.3. — *Il n'y a pas d'autres processus du sixième type que ceux formés au* 2°.

LEMME. — *S'il y a un état* A_0 *instantané et presque sûrement réalisé pour les points d'un ensemble* E_0 *partout dense sur l'axe des t, le processus est du type trivial* $P_{h,k}(t) = P_k$ (il est entendu que nous supposons toujours les fonctions $P_{h,k}(t)$ mesurables, donc continues).

Pour simplifier la démonstration, nous admettrons l'axiome du choix sous la forme suivante : l'ensemble E_0 étant partout dense, nous pouvons, dans n'importe quel petit intervalle $(t_0 - a, t_0)$, choisir un point t' de E_0 et la connaissance de la relation $H(t') = 0$ ne modifie pas les probabilités relatives à $H(t)$ pour $t > t'$ ([27]).

L'existence de t' étant presque sûre, on a, pour $t > \varepsilon$, $t_0 = t - \varepsilon$ et $a < t_0$,

$$P_{h,k}(t) = P_{0,k}(t - t').$$

Or la fonction $P_{0,k}$ est continue. On a donc

$$P_{h,k}(t) = \lim_{a \to 0} P_{0,k}(t - t') = P_{0,k}(t - t_0) = P_{0,k}(\varepsilon).$$

Comme ε est arbitrairement petit, $P_{h,k}(t)$ est bien indépendant de h et de t.

([27]) Pour rendre ce raisonnement indépendant de l'axiome de Zermelo, il suffirait de démontrer que la suite des points $t_0 - \frac{a}{n}$ $(n = 1, 2, \ldots)$ contient presque sûrement au moins un point de E_0. C'est exact, mais non évident si l'on n'admet pas l'axiome de Zermelo.

Pour démontrer le théorème III.3, considérons un état A_0 de la nature qui caractérise le sixième type, c'est-à-dire que, au moins dans des cas de probabilité positive, l'ensemble E_0 est discontinu et non mesurable. Sa fermeture $\bar{E}_0$ est mesurable; soit τ la mesure de la partie de $\bar{E}_0$ intérieure à $(0, t)$. Nous pouvons prendre τ comme variable pour la succession des états $A_{0,l}$ qui se succèdent sur $\bar{E}_0$. Le processus qui définit cette succession est évidemment markovien et stationnaire, et l'on vérifie aussi aisément que les probabilités de passage sont mesurables, donc continues, en même temps que celles relatives au processus initial. On peut alors appliquer le lemme, qui nous montre qu'en chaque point de $\bar{E}_0$ l'indice l résulte d'un tirage au sort régi par une loi indépendante aussi bien de τ que des résultats des tirages antérieurs.

Ce résultat s'appliquant à tous les ensembles $\bar{E}_h$ analogues à $\bar{E}_0$, le théorème énoncé est établi.

Remarque. — Si, les fonctions $P_{h,k}(t)$ étant mesurables, l'ensemble des points de discontinuité peut remplir un ou plusieurs intervalles, il résulte immédiatement du lemme III.1 que le processus est du sixième type. Nous avions d'ailleurs déduit de ce lemme que cette circonstance est exclue pour le cinquième.

III.4. Processus du septième type. — Ce dernier type est caractérisé par le fait qu'au moins une des fonctions $P_{h,k}(t)$ n'est pas mesurable. Il y a alors au moins deux valeurs de k telles que, quel que soit h, $P_{h,k}(t)$ soit, ou bien non mesurable, ou bien identiquement nul; donc au moins quatre de ces fonctions ne sont pas mesurables.

J. L. Doob [1] a montré par un exemple l'existence de ce type. Il faut naturellement utiliser l'axiome de Zermelo, sans lequel on n'a jamais réussi à prouver l'existence de fonctions non mesurables. Nous généraliserons son exemple de la manière suivante :

Considérons deux ensembles F et F′ de nombres réels, qui vérifient les conditions suivantes :

a. Tout t réel peut être mis d'une manière et d'une seule sous la forme $u + u'$, où $u \in F$, $u' \in F'$;

b. Si $u_1 \in F$, $u_2 \in F$, toutes les combinaisons linéaires $u = r_1 u_1 + r_2 u_2$ à coefficients rationnels (positifs, nuls, ou négatifs) appartiennent à F;

b′. Propriété analogue pour F′;

c. F est dénombrable (et non vide, donc, compte tenu de *b*, partout dense sur l'axe des t; F′ est évidemment non dénombrable et partout dense).

En prenant pour F n'importe quel ensemble qui vérifie les conditions *b* et *c*,

on déduit aisément de l'axiome de Zermelo qu'il existe des ensembles F′ tels que toutes les conditions indiquées soient vérifiées.

Établissons une correspondance biunivoque entre les éléments de F et les états A_h; l'indice h devient une fonction de u, que nous supposons lui-même fonction bien déterminée de t; donc h est une fonction bien déterminée $f(t)$ de t. Prenant maintenant $H(t)=f(t-a)$ (a, constante inconnue), nous obtenons évidemment un processus markovien [puisque la fonction $H(t)$ est complètement déterminée par la connaissance de sa valeur en un seul point] et stationnaire. Les probabilités de passage sont discontinues, donc non mesurables (d'après le théorème de Doob rappelé au paragraphe III.2, 1°).

On remarque que, dans cet exemple, le hasard n'intervient pas, ou n'intervient que par le choix de l'état initial.

Nous appellerons *type de Doob* (généralisé) le type qui vient d'être défini et nous appellerons processus du *septième type pur* tous les processus de ce type pour lesquels les fonctions $P_{h,k}(t)$ n'ont pas d'autres valeurs que zéro ou un; le hasard ne peut intervenir que par le choix de l'état initial. Tout processus de Doob est de ce septième type pur. Nous ne savons pas si la réciproque est vraie.

On peut évidemment former des processus de type mixte par des combinaisons d'indices analogues à celles indiquées à propos du sixième type. On peut introduire ici trois indices successifs, h, k, l. Le premier indice h varierait d'après un processus d'un des cinq premiers types. Pour certaines valeurs de h (et en tout cas au moins pour une), k aurait une infinité de valeurs possibles et varierait (en fonction du paramètre τ relatif à l'ensemble E_h considéré) d'après un processus du septième type pur. Enfin, pour chacun des ensembles $E_{h,k}$ correspondant à des valeurs données des deux premiers indices, on peut introduire un troisième indice, ayant plusieurs valeurs possibles ou même une infinité dénombrable et qui, en chaque point de $E_{h,k}$, serait choisi d'après une loi fonction de h et k et indépendante aussi bien de t que des choix antérieurs.

Une question qui se pose naturellement est de savoir si par des combinaisons de ce genre on obtient le processus le plus général du septième type. Nous ne l'avons pas résolue.

III.5. Retour aux théorèmes généraux. — 1° *Continuité des fonctions* $P_{h,k}(t)$. — D'après les définitions et le théorème de Doob rappelé au paragraphe III.2. 2°, ces fonctions sont continues pour les six premiers types, sauf peut-être au point $t=0$. En ce point, pour les quatre premiers types, $P_{h,k}(+0)$ existe et a la valeur $\delta_{h,k}$. Pour le cinquième type, nous avons seulement démontré l'existence d'une limite en moyenne, encore égale à $\delta_{h,k}$; la question de savoir si $P_{h,k}(+0)$ existe (et est alors égal à $\delta_{h,k}$) n'est pas résolue. Pour le sixième type, il y a de même limite en moyenne, mais, pour la fonction relative au passage de A_{h,l_0} à $A_{k,l}$, cette limite est de la forme $\delta_{h,k}P'_{k,l}$, donc, si $h=k$ et qu'à cette valeur de k soient associées au moins deux valeurs de l, comprise entre zéro et

un. Il y a limite véritable si la variation de h dépend d'un processus d'un des quatre premiers types (le cas du cinquième restant en suspens).

Naturellement, pour un processus du septième type, celles des fonctions $P_{h,k}(t)$ qui ne sont pas mesurables ne sont pas continues.

2° *Extension du théorème* II.8.1. — Nous l'avons démontré pour les quatre premiers types. L'apparition d'un état instantané A_0 dans le cinquième type ne change rien d'essentiel. Si cet état intervient dans les successions d'états qui peuvent conduire de A_h à A_k, le temps τ passé dans cet état est une variable aléatoire du type μU (U ayant toujours la signification indiquée au paragraphe II.1) et peut être arbitrairement petit; il en est de même de la durée totale de chacune des successions d'états considérés. S'il y a plusieurs états instantanés, cela ne change rien non plus. Le théorème considéré reste vrai.

Il le reste aussi pour le sixième type, puisqu'on le déduit du précédent en multipliant certaines des probabilités de passage par des facteurs constants.

Il est manifestement faux pour le septième type, au moins pour les exemples de ce type considérés au paragraphe III.4 et pour celles des probabilités de passage qui ne sont pas mesurables.

3° Nous savons déjà que le théorème II.8.2 est vrai pour les six premiers types. Il est évidemment en défaut, pour le septième, au moins pour certaines des fonctions $P_{h,k}(t)$.

BIBLIOGRAPHIE.

Pour le cas des systèmes S_r (cas fini), nous renvoyons le lecteur à la bibliographie indiquée dans le livre de M. Fréchet [1]. Même pour le cas dénombrable, la liste qui suit n'est pas complète; nous ne mentionnons que les travaux qui nous ont paru le plus en rapport avec la présente étude.

Doeblin (W.) :

[1] *Exposé de la théorie des chaînes simples constantes de Markoff à un nombre fini d'états* (*Thèse*, Paris, 1937).

[2] *Éléments d'une théorie générale des chaînes simples constantes de Markoff* (*Ann. de l'Éc. Norm. Sup.*, 57, 1940, p. 61-101).

Doob (J. L.) :

[1] *Topics in the theory of Markoff chains* (*Trans. Amer. Math. Soc.*, 52, 1942, p. 37-64).

[2] *Markoff chains. Denumerable case* (*Trans. Amer. Math. Soc.*, 58, 1945, p. 455-473).

FELLER (W.) :

[1] *On the integro differential equations of purely discontinuous Markoff processes* (*Trans. Amer. Math. Soc.*, 48, 1940, p. 488-515 et erratum, 58, 1945, p. 474).

FORTET (R.) :

[1] *Sur l'itération des substitutions algébriques linéaires à une infinité de variables et ses applications à la théorie des probabilités en chaîne* (*Thèse,* Paris, 1939; *voir* not., p. 21-43).

FOSTER (F. G.) :

[1] *Markoff chains with an enumerable number of states and a class of cascade processes* (*Proc. Cambridge phil. Soc.*, 47, 1951, p. 77-85).

FRÉCHET (M.) :

[1] *Recherches théoriques modernes sur le Calcul des probabilités* (t. 1, fasc. III du traité publié par E. Borel et divers collaborateurs; en deux livres; *voir* not. le second livre, qui est en cours de réimpression).

KOLMOGOROFF (A.) :

[1] *Anfangsgründe der Theorie der Markoffschen Ketten mit unendlich vielen möglichen Zuständen* (*Rec. Math. de Moscou,* 1, n° 43, 1936, p. 607-610).

[2] *Chaînes de Markoff avec une infinité dénombrable d'états possibles* (en russe; *Bull. de l'Univ. d'État de Moscou.* Sect. A, 1, 1937, fasc. III).

KRYLOFF (N.) et BOGOLIOUBOFF (N.) :

[1] *Sur les propriétés ergodiques de l'équation de Smoluchovsky* (*Bull. Soc. Math. Fr.*, 64, 1936, p. 49-56).

LÉVY (P.) :

[1] *Théorie de l'addition des variables aléatoires* (Paris, Gauthier-Villars, 1937).

[2] *Processus à la fois stationnaires et markoviens pour les systèmes ayant une infinité non dénombrable d'états possibles* (*Congrès intern. des mathématiciens.* Cambridge, Mass., 1950).

[3] Quatre Notes présentées à l'Académie des Sciences (*C. R. Acad. Sc.*, 231, p. 467 et 1268; 232, p. 1400 et 1803).

YOSIDA (K.) and KAKUTANI (S.) :

[1] *Markoff process with an enumerable infinite number of possible states* (*Japan. J. Math.*, 16, 1939, p. 47-55).