

ANNALES DE L'I. H. P., SECTION B

HERVÉ RAYNAUD

Sur les graphes aléatoires

Annales de l'I. H. P., section B, tome 4, n° 4 (1968), p. 255-329

http://www.numdam.org/item?id=AIHPB_1968__4_4_255_0

© Gauthier-Villars, 1968, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section B » (<http://www.elsevier.com/locate/anihpb>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

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

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

Sur les graphes aléatoires

par

Hervé RAYNAUD

SOMMAIRE. — Ce travail se situe comme un moyen terme entre des travaux déjà classiques, ceux de A. Renyi et de P. Erdős d'une part, de J. M. Hammersley et de E. N. Gilbert d'autre part. Les graphes étudiés ici sont utilisés par les chimistes. C. Bruneau en a fait l'outil principal de sa thèse.

Définition des graphes γ .

Soit un ensemble $\{s_i\}$, $I = \{1, 2, \dots, N\}$ dont les éléments sont appelés sites.

Soit une partition E de S définie par une famille \check{E} de parties deux à deux disjointes, de réunion S , notées E_j , $j \in J$, $J = \{1, 2, \dots, r\}$, et appelées espèces.

L'espèce du site s_i , notée $E(s_i)$ est l'espèce dont il est élément.

Soit un graphe R , non multigraphe, avec ou sans boucles, de sommet E_1 , E_2, \dots, E_r .

On appelle graphe γ tout graphe pour lequel

- les sommets sont les s_i : $\forall i$, s_i compte au plus un arc incident,
- un arc ne peut avoir pour origine s_i et pour extrémité s_k que si, dans R , $E(s_i)$ précède $E(s_k)$.

Définition du graphe Γ qui se déduit d'un graphe γ .

Soit une deuxième partition M de S définie par une famille \check{M} de parties deux à deux disjointes, de réunion S , notées μ_k , $k \in K$

$$K = \{1, 2, \dots, n\}$$

et appelées monomères.

Le graphe Γ déduit d'un graphe γ donné a :

- pour sommets les monomères $\mu_1, \dots, \mu_k, \dots, \mu_n$;
- pour arcs, ou arêtes, ceux qui correspondent à des arcs ou arêtes de γ suivant la règle : toutes les fois qu'un site de μ_λ précède un site de $\mu_{\lambda'}$, on trace un arc de Γ entre μ_λ et $\mu_{\lambda'}$.

Générations aléatoires de γ .

Sur les 4 procédures étudiées de génération de γ , nous décrirons les deux principales. Pour simplifier ce résumé, supposons R non orienté.

Première procédure. — x arêtes de γ ayant été construites, on veut ajouter une arête à γ . On considère alors l'ensemble des sites qui ne sont pas extrémités d'une arête déjà construite. On forme l'ensemble des couples non ordonnés de ces sites tels que les deux sites du couple aient des espèces adjacentes dans R . On tire au hasard uniforme un couple parmi ces couples. On obtient ainsi les extrémités de l'arête supplémentaire de γ .

Pour obtenir un graphe à m arêtes, on procède ainsi m fois, à partir d'un ensemble de monomères sans arêtes.

Deuxième procédure. — x arêtes étant construites, on veut ajouter une arête à γ . On considère D , ensemble des sites qui ne sont ni origine, ni extrémité d'une quelconque des x arêtes. On considère ensuite ceux des sites de $J \in D$ tels qu'il existe au moins un autre site $d' \in D$, et telle que $E(d')$ suive $E(d)$ dans R . Pour construire la $(x + 1)^{\text{ième}}$ arête :

1° On tire un site au hasard uniforme parmi tous ceux qui ont cette propriété. Soit s_i ce site.

2° On tire un site parmi tous ceux qui restent dans D et dont l'espèce est adjacente à celle de s_i dans R .

Le chapitre III étudie les conditions dans lesquelles on pourra dire que les procédures sont équivalentes, conduisent aux mêmes graphes aléatoires, c'est-à-dire définissent les mêmes espaces de probabilité.

Dans le chapitre IV, on établit un certain nombre de résultats sur les urnes et sur les tirages avec contraintes en développant une théorie assez analogue à une théorie linguistique. On étudie en particulier les vidanges fortes, systèmes de contraintes sur le tirage qui permettent, avec une probabilité aussi voisine de 1 qu'on le veut de tirer une proportion de boules de l'urne aussi grande qu'on le veut.

On caractérise aussi une propriété de stationnarité des vidanges fortes dont les conséquences sont très importantes dans l'application de ces résultats aux graphes aléatoires définis dans les deux premiers chapitres.

Au chapitre VI, on montre comment les résultats précédents peuvent éclairer le plus simple des problèmes que se posaient les chimistes. On précise le cadre dans lequel, pour décrire les graphes aléatoires étudiés, on peut remplacer les tirages exhaustifs par des tirages avec remise qu'on sait facilement manier.

Il a enfin semblé utile de mener complètement à bien en indiquant avec précision, à l'intention des utilisateurs chimistes et physiciens, les calculs et leurs résultats dans les cas les plus simples, aucune difficulté de méthode ne devant se présenter dans l'extension de ces résultats aux autres cas possibles.

SUMMARY. — This work may be considered as midway between the classical work of A. Renyi and P. Erdős, and the work of J. M. Hammersley and E. N. Gilbert. The graphs studied have been applied to problems in chemistry: C. Bruneau used them as the principal tool for his doctoral-thesis work.

Definition of a γ -graph

Let S be a set $S = \{S_i\}_{i \in I}$, $I = \{1, 2, \dots, N\}$. The elements of this set we shall call sites.

We define a partition E of S by a family $\overset{v}{E}$ of mutually disjoint subsets E_j , $j \in J$, $J = \{1, 2, \dots, r\}$ of union S .

These subsets E_j we call species.

The species of a site S_i , is then the species of which it is an element.

Let R be a graph R , not a multigraph, with or without loops, and with vertices E_1, E_2, \dots, E_r . We then call γ -graph any graph which satisfies the following conditions:

- i)* The vertices are the S_i .
- ii)* For each vertex the sum of its degrees or demidegrees is less than or equal to one.
- iii)* An arc has initial vertex S_i and terminal vertex S_k if and only if $E(S_i)$ precedes $E(S_k)$ in R .

Definition of the Γ -graph corresponding to a γ -graph

Let a second partition M of S be defined by a family M of mutually disjoint subsets $\{\mu_k, k \in K\}$, $K = \{1, 2, \dots, n\}$ with union S . These μ_k we call monomers.

The Γ -graph corresponding to a given γ -graph is then defined as follows:

- i)* Its vertices are the monomers $\mu_1, \mu_2, \dots, \mu_n$.
- ii)* Its arcs or edges are those corresponding to the arcs or edges of γ

under the rule obtained by creating an arc joining μ_λ to $\mu_{\lambda'}$ if and only if a site of μ_λ precedes a site of $\mu_{\lambda'}$.

Random generation of γ graphs

Of the four procedures for generating γ graphs, we will only describe here the two principal ones. For simplicity in this summary we will suppose that R is not directed.

First procedure. — Having constructed x edges of γ , we wish to construct the $(x + 1)$ th edge of γ . Consider the set of all sites which are not yet the ending point of an already constructed edge, and construct the set of all those unordered pairs of these sites such that the two sites of the pair have species adjacent in R. Select one of these pairs at random. This gives us the $(x + 1)$ th edge of γ . To obtain a graph with M edges, we repeat this process M times, starting with a set of monomers without edges.

Second procedure. — As before, having constructed x edges of γ , we wish to construct an $(x + 1)$ th edge of γ . Let D be set of all sites which are neither initial vertex nor terminal vertex of one of the X edges, and consider those sites $d \in D$ such that there exists at least one other site $d' \in D$ such that $E(d')$ follows $E(d)$ in R. To construct the $(x + 1)$ th edge:

- i) Choose at random a site S_i from those which have this property.
- ii) Choose at random a site amongst those which remain in D whose species are adjacent to that of S_i in R.

In chapter III we examine the conditions under which these proceedings can be said to be equivalent, in that they produce the same random graphs. In chapter IV we establish a certain number of results on urns, and on random choice under constraints, and develop a theory analogous to a linguistic theory. We examine in particular the « vidanges fortes » systems of constraints on the random sampling which permits the choice, with a probability arbitrarily close to 1, of an arbitrarily large proportion of balls from one urn. We characterise then a fixed-point property of « vidanges » whose consequences are extremely important in the application of the results on random graphs developed in the first two chapters.

In chapter VI we show how the preceding results can clarify certain simple problems in chemistry. We define the cases in which, to describe the random graphs studied, we can replace the sampling without replacement by sampling with replacement which is easier to deal with.

It seemed useful to conclude by showing for the benefit of potential physicist and chemist users, the calculations and their results for the sim-

plest cases; there should be no difficulty in extending these results to other possible cases.

In conclusion, the results of this thesis suggests a considerable number of research problems whose chemical application could be of great interest in the present theory of macro-molecules.

INTRODUCTION

Dans l'étude des hauts polymères, les chimistes utilisent souvent les graphes aléatoires. Dans le cadre de ce modèle, ils peuvent définir axiomatiquement la « gélification » d'un composé. Suivant les auteurs, la définition axiomatique de cet état de « gel » et les contraintes imposées aux « molécules » des modèles varient; nous ne prétendons pas donner ici une solution aux problèmes physico-chimiques, mais nous souhaitons proposer aux chimistes un langage commode et aux mathématiciens un nouveau type de problèmes.

Ce travail se situe comme un moyen terme entre des travaux déjà classiques, ceux de A. Renyi et P. Erdős [12] d'une part, J. M. Hammersley [7] et E. N. Gilbert [5] d'autre part. Tandis que les premiers ont pu obtenir des résultats très forts — mais sur des graphes qui ne peuvent pour le moment servir de modèles aux chimistes — les seconds ont dû, par exemple, abandonner la recherche des méthodes de calcul exact des probabilités de percolation.

Nous espérons que les problèmes de graphes étudiés ici, plus souples que les problèmes de percolation et plus généraux que ceux posés par la famille définie par A. Renyi et P. Erdős, se révéleront à la fois accessibles et utiles. Ils ont déjà été utilisés par C. M. Bruneau, qui en a fait l'outil principal de sa thèse ([2] et [3]). C'est d'ailleurs sur sa suggestion que certaines appellations des êtres mathématiques que nous allons rencontrer ont été inspirées par le vocabulaire chimique. Nous nous en excusons d'avance auprès du lecteur mathématicien qui trouverait arbitraire une part de notre terminologie. Quant au vocabulaire habituel de la théorie des graphes, il a été emprunté à C. Berge [1]. Dans les pages qui suivent, nous allons tout d'abord définir quatre familles de graphes aléatoires, étudier certains problèmes de passage d'une famille à une autre, puis démontrer, dans un certain cadre d'équivalence asymptotique, des théorèmes de stabilité concernant leur structure. Nous appliquerons ensuite ces théorèmes à quelques graphes simples des familles en question.

I

DÉFINITIONS PRÉLIMINAIRES

I. Définition des graphes γ .

Soit un ensemble $S = \{ s_i \}_{i \in I}$, $I = \{ 1, 2, \dots, N \}$, dont les éléments sont appelés *sites*.

Soit une partition E de S définie par une famille \check{E} de parties deux à deux disjointes, de réunion S , notées E_j , $j \in J$, $J = \{ 1, 2, \dots, r \}$ et appelées *espèces*.

L'espèce du site s_i , notée $E(s_i)$, est l'espèce dont il est élément.

Soit un graphe R , non multigraphe, avec ou sans boucles, de sommets E_1, E_2, \dots, E_r . R pourra être orienté ou non.

1° Si R est orienté :

On appellera graphe γ tout graphe orienté pour lequel :

- les sommets sont s_1, s_2, \dots, s_N ;
- pour chaque sommet, la somme de ses demi-degrés est inférieure ou égale à un ;
- un arc ne peut avoir pour origine s_k , et pour extrémité s_k que si, dans R , $E(s_k)$ précède $E(s_k)$.

2° Si R est non orienté :

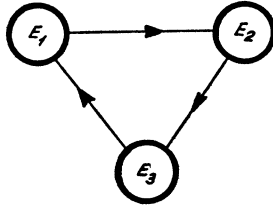
On appellera graphe γ tout graphe non orienté pour lequel :

- les sommets sont s_1, s_2, \dots, s_N ;
- chaque sommet est adjacent à un sommet au plus ;
- chaque arête joint des sommets qui sont d'espèces adjacentes dans R .

Exemple. — Éclairons notre définition d'un exemple, qui nous a été suggéré par B. Roy, que nous remercions à cette occasion.

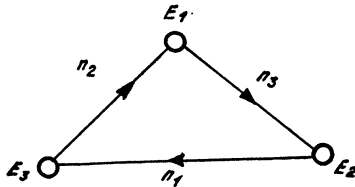
Soit un ensemble S formé de N_1 sites d'espèce E_1 ,
 N_2 sites d'espèce E_2 ,
 N_3 sites d'espèce E_3

et pour lequel le graphe R orienté est le suivant :



Les sites étant considérés comme discernables, combien existe-t-il de graphes γ distincts et comptant ?

n_3 arcs allant d'un site de l'espèce E_1 à un site de l'espèce E_2 ,
 n_1 arcs allant d'un site de l'espèce E_2 à un site de l'espèce E_3 ,
 n_2 arcs allant d'un site de l'espèce E_3 à un site de l'espèce E_1 .



Il y a :

$$\frac{N_1!}{(N_1 - n_3)!} \quad \text{arrangements de } n_3 \text{ sites parmi les } N_1 \text{ d'espèce } E_1.$$

et donc :

$$\frac{N_1!}{(N_1 - n_3)!} \frac{N_2!}{(N_2 - n_3)!} \quad \text{façons de choisir les } n_3 \text{ premiers arcs.}$$

Il reste alors, de même :

$$\frac{(N_2 - n_3)!}{(N_2 - n_3 - n_1)!} \frac{N_3!}{(N_3 - n_1)!} \quad \text{façons de choisir les } n_1 \text{ suivants,}$$

et enfin :

$$\frac{(N_3 - n_1)!}{(N_3 - n_1 - n_2)!} \frac{(N_1 - n_3)!}{(N_1 - n_3 - n_2)!} \quad \text{façons de choisir les } n_2 \text{ derniers.}$$

soit au total :

$$\frac{N_1 ! N_2 ! N_3 !}{(N_1 - n_2 - n_3) ! (N_2 - n_3 - n_1) ! (N_3 - n_1 - n_2) !}$$

graphes γ distincts.

II. Définition des graphes Γ .

Soit une deuxième partition M de S définie par une famille $\overset{\vee}{M}$ de parties deux à deux disjointes, de réunion S , notées μ_k ; $k \in K$; $K = \{1, 2, \dots, n\}$ et appelées monomères.

Remarque. — L'intérêt de cette deuxième partition provient de ce que, dans la suite, nous étudierons des graphes dans lesquels, le nombre N de sites devenant très grand, le nombre r d'espèces reste constant tandis que le nombre n des monomères est du même ordre de grandeur que le nombre de sites N .

Se donner la *composition* d'un monomère μ_k c'est se donner, pour toute espèce E_j , le nombre μ_{kj} de sites de μ_k appartenant à l'espèce E_j .

1° Si R est orienté :

Un graphe Γ déduit d'un graphe γ donné est un graphe

- dont les sommets sont les monomères $\mu_1, \dots, \mu_k, \dots, \mu_n$;
- dont les arcs sont aussi nombreux que ceux de γ . A chaque arc de γ correspond un arc et un seul de Γ par la bijection : toutes les fois qu'un site de μ_λ précède dans γ un site de $\mu_{\lambda'}$, on trace un arc de Γ entre μ_λ et $\mu_{\lambda'}$.

2° Si R est non orienté :

Un graphe Γ déduit d'un graphe γ donné est un graphe

- dont les sommets sont les monomères $\mu_1, \dots, \mu_k, \dots, \mu_n$;
- dont les arêtes sont aussi nombreuses que celles de γ . A chaque arête de γ correspond une arête et une seule de Γ par la bijection : toutes les fois qu'un site de μ_λ est adjacent dans γ à un site de $\mu_{\lambda'}$, on trace une arête de Γ entre μ_λ et $\mu_{\lambda'}$.

Remarques :

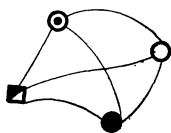
- Si γ est orienté, Γ est orienté. Si γ n'est pas orienté, Γ n'est pas orienté.
- Γ peut être multigraphe avec boucles.

Exemple de passage d'un graphe γ à un graphe Γ .

Supposons que les graphes que nous allons représenter comptent 4 espèces de sites :

- E_1 dont les sites seront représentés par des \blacksquare , au nombre de 4.
- E_2 dont les sites seront représentés par des \bullet , au nombre de 2.
- E_3 dont les sites seront représentés par des \circ , au nombre de 3.
- E_4 dont les sites seront représentés par des \odot , au nombre de 3.

Soit R non orienté :



Soit le graphe μ :

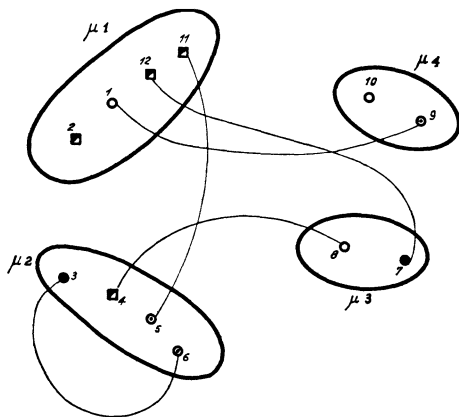


FIG. 1.

R ne comptant pas de boucles, deux sites de même espèce ne peuvent être reliés par une arête de γ .

Supposons, comme indiqué sur la figure 1, que :

- le monomère μ_1 comprend les sites s_{11}, s_{12}, s_1, s_2 ,
- le monomère μ_2 comprend les sites s_3, s_4, s_5, s_6 ,
- le monomère μ_3 comprend les sites s_7, s_8 ,
- le monomère μ_4 comprend les sites s_9, s_{10} .

On en déduit le graphe :

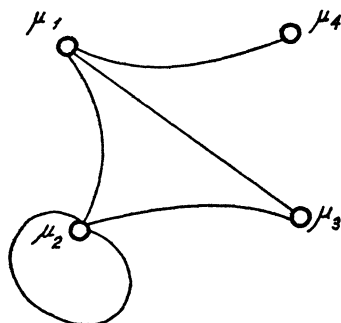


FIG. 2.

Il est évidemment possible de regrouper sur un même schéma les deux représentations :

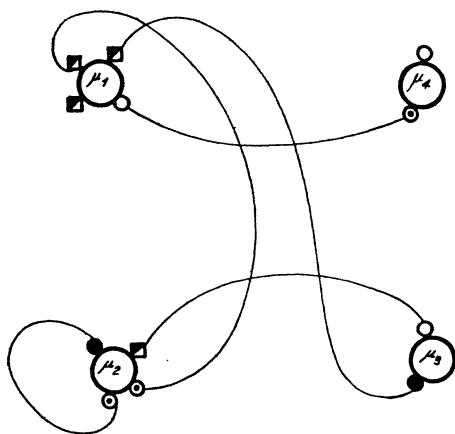


FIG. 3.

Il est commode d'utiliser un modèle chimique pour se rappeler la notation. Les monomères sont analogues à des molécules qui porteraient des sites réactifs. Ces sites réactifs — les sites de nos graphes — sont de plusieurs espèces chimiques; certaines réagissent entre elles, certaines ne peuvent réagir et donner lieu à la réalisation d'une liaison. C'est R qui nous fournit cette information. Connaître la composition des monomères, c'est connaître la nature des monomères chimiques mis en présence. Quant aux arêtes de γ , elles représenteront les liaisons qui s'établissent entre les sites au fur

et à mesure que la réaction chimique avance, tandis que les arêtes de Γ représenteront les liaisons entre monomères.

Dans les graphes que nous allons étudier, la fonctionnalité (terme chimique) f_k , donnée par :

$$f_k = \sum_{j=1}^r \mu_{kj},$$

du monomère μ_k ne suffira pas, en général, à la détermination de ses propriétés. Cependant, ce total est intéressant car dorénavant :

- Nous ne considérons pas de monomères de fonctionnalité 1.
- Nous appellerons pré-anti-nœuds les monomères de fonctionnalité 2.
- Nous appellerons pré-nœuds les monomères de fonctionnalité strictement supérieure à 2.

II

GÉNÉRATION ALÉATOIRE DE γ ET DE Γ

Un graphe aléatoire est un espace de probabilités (Ω, \mathcal{A}, P) . L'ensemble Ω a pour éléments les réalisations possibles du graphe γ (et donc Γ) à partir de l'ensemble S , d'un graphe R , de deux partitions E et M de S .

Cet ensemble est donc fini. \mathcal{A} est l'algèbre de Boole des parties de Ω et P est une probabilité dont la valeur sera définie sur chaque élément de Ω .

Ce qui fait l'intérêt essentiel de ces graphes aléatoires c'est qu'il existe dans la nature, des phénomènes auxquels ils ressemblent. Ces phénomènes sont progressifs. Ils partent d'un ensemble de monomères entre lesquels aucun arc, aucune arête, n'existent, puis il en apparaît un certain nombre, puis un peu plus, puis plus encore, jusqu'à ce que le processus s'arrête.

Il en est ainsi dans les processus de « polyagrégation » de C. M. Bruneau ([2], [3]).

Sous le nom de procédures, nous allons maintenant décrire des méthodes d'obtention d'une probabilité P . Les procédures π_1 et π_2 peuvent décrire des processus réels, en relation avec les mécanismes chimiques. La procédure π_3 ne représente plus un processus réel. C'est une méthode abstraite pour obtenir une probabilité P , mais qui, dans certaines conditions, conduit aux mêmes espaces de probabilité que π_1 et π_2 . L'intérêt de π_3 est d'être susceptible d'un traitement probabiliste beaucoup plus commode que

π_1 et π_2 . Quant à la procédure π_4 , elle a les mérites des trois autres, mais son intérêt dans les applications nous semble pour le moment beaucoup plus hypothétique.

Dans les trois premières procédures, le nombre m d'arêtes du graphe γ est fixé. Dans la quatrième, c'est une variable aléatoire qui dépend d'un paramètre p . Le problème qui va par la suite se poser est, en termes généraux, le suivant :

A partir d'un ensemble de monomères, d'un graphe R , d'une procédure et d'un nombre d'arêtes (respectivement d'un paramètre p) donnés, quelles propriétés peut-on déduire sur le graphe aléatoire engendré ?

La description des 4 procédures se trouve légèrement modifiée suivant que R est orienté ou non orienté. Pour simplifier le langage et éviter une lecture fastidieuse, nous n'avons décrit que le cas R non orienté. Le lecteur effectuera sans peine la transposition. Par ailleurs, dans toute la suite de ce chapitre, par soucis d'abréviation, on a remplacé « tirer d'une façon exhaustive et au hasard uniforme parmi les éléments du type x qui n'ont pas encore été tirés » par « tirer un x ».

Procédure π_1 : équi-probabilité des arêtes.

C'est une procédure progressive dans laquelle on tire une première arête, puis une seconde, et ainsi de suite. C'est en gros la procédure utilisée par Renyi.

x arêtes ayant été tirées, on veut tirer une arête supplémentaire de γ .

On considère alors l'ensemble des sites qui ne sont pas extrémités d'une arête déjà tirée.

On forme l'ensemble des couples non ordonnés de ces sites tels que les deux sites du couple aient des espèces adjacentes dans R (si (s_i, s_j) est un tel couple, alors $E(s_i)$ est adjacent à $E(s_j)$ dans R). On tire au hasard uniforme un couple parmi ces couples. On obtient ainsi les extrémités de l'arête supplémentaire de γ .

On procède ainsi m fois, zéro arêtes étant tirées au départ et l'on obtient les m arêtes d'un graphe γ .

Exemple. — Reprenons le graphe γ de la page 263. Nous nous étions donnés :

1° Deux monomères de quatre sites :

— μ_1 , formé de trois sites \blacksquare et un site \circ ;

— μ_2 , formé d'un site \blacksquare , un site \bullet et deux sites \odot .

2° Deux monomères de deux sites :

- μ_3 , formé d'un site ● et d'un site ○;
- μ_4 , formé d'un site ⊙ et d'un site ○.

3° Cinq arêtes.

Supposons que ces arêtes aient été tirées suivant la procédure π_1 dans l'ordre :

$$(s_{11}, s_5), (s_{12}, s_7), (s_1, s_9), (s_3, s_6), (s_4, s_8).$$

Le tirage de la première aurait eu lieu entre 53 possibles

$$\left(\frac{1}{2} [(11 - 2)3 + (11 - 3)4 + (11 - 1)2 + (11 - 2)3] \right).$$

Le tirage de la seconde entre 37

$$\left(= \frac{1}{2} [(9 - 2)(3 + (9 - 2)3 + (9 - 1)2 + (9 - 1)2)] \right).$$

Celui de la troisième, entre 23; de la quatrième entre 13; de la cinquième entre 4.

L'évaluation des probabilités associées aux différents graphes possibles est déjà longue. Quand le nombre de sites et d'arêtes augmente, elle devient inextricable.

Remarques. — Dans le cadre d'une interprétation chimique, la procédure π_1 équivaut à l'hypothèse : quel que soit l'état d'avancement d'une réaction, la probabilité associée aux différentes liaisons chimiques qui peuvent se produire est la même pour toutes les liaisons possibles.

— On pourrait dès maintenant discuter des contraintes sur m , nombre des arêtes, qui ne peut être choisi arbitrairement. Un seul type de contrainte sur m devant nous servir par la suite, nous traiterons ce point ultérieurement.

Procédure π_2 : le graphe enfilé arête par arête.

La procédure que nous allons décrire ici est donc un autre processus dont l'aboutissement est un graphe aléatoire construit arête par arête et tel que, pour obtenir la probabilité des graphes à x arêtes on n'ait besoin que de la connaissance de la probabilité des graphes à $(x - 1)$ arêtes.

Après le tirage de la $(x - 1)^{\text{ième}}$ arête, $N = 2(x - 1)$ sites restent non tirés, à partir desquels nous formerons les arêtes suivantes de γ . Appelons D

cet ensemble de sites restants. Considérons les sites $d \in D$ tels qu'il existe au moins un site $d' \in D$ vérifiant la relation $E(d)$ adjacente à $E(d')$ dans R .

Pour tirer la $x^{\text{ième}}$ arête :

1° Je tire alors (*) un site parmi tous ceux qui ont la propriété précédente. Soit s_i ce site.

2° Je tire ensuite un second site parmi tous ceux dont l'espèce est adjacente à l'espèce de s_i dans R (Il est clair que ceci est toujours possible à cause de la condition vérifiée par s_i).

On peut se demander pour quelle raison nous avons appelé cette procédure un enfilage.

Imaginons une petite fille devant un lot de perles en groupes de faible effectif (les monomères), et possédant :

— des brins de fil trop courts pour lui permettre de relier plus de deux perles avec le même brin ;

— une règle R qui ne lui permet de joindre entre elles que certaines couleurs de perles.

Elle choisira alors une perle, puis une autre, et ainsi de suite.

On peut encore dire que les sites, dans la procédure π_2 , sont choisis un par un.

Le lecteur peut se poser dès maintenant la question de l'équivalence entre π_1 et π_2 . A partir d'un ensemble de monomères, d'un graphe R et d'un nombre d'arêtes m , π_1 et π_2 donnent, sur Ω , des probabilités P_1 et P_2 . π_1 et π_2 auront été équivalentes si et seulement si P_1 est identique à P_2 . L'étude des conditions d'équivalence entre procédures fera l'objet du chapitre suivant.

Les deux procédures que nous venons de décrire peuvent être arrêtées là. Dans la description de réactions chimiques réversibles, C. M. Bruneau a trouvé efficace de poursuivre le tirage en ajoutant aux m arêtes autant d'arêtes fictives qu'il lui était nécessaire. Voyons de plus près les notions qui lui ont été utiles.

Le tirage de m arêtes ayant été réalisé, il reste $N - 2m$ sites à tirer, dont un certain nombre pourrait donner lieu à des arêtes si l'on poursuivait le tirage suivant les procédures π_1 ou π_2 . Supposons donc que l'on continue la procédure de tirage jusqu'à ce qu'il ne reste plus de couples d'espèces adjacentes dans R , et que l'on ait pu ainsi tirer m_1 arêtes supplémentaires.

Il reste alors $N - 2m - 2m_1$ sites qui pourraient donner des couples qui seraient convenables si leurs espèces étaient adjacentes dans R . On

(*) Avec la convention de la page 266.

convient alors de poursuivre le tirage comme si R était un graphe complet avec boucles, et l'on appelle $m_2 =$ partie entière de $\frac{1}{2}(N - 2m - 2m_1)$, le nombre maximum d'arêtes ainsi ajoutées.

Par définition, nous appellerons :

arêtes réalisées les m premières;

pré-arêtes les m_1 suivantes;

et non-arêtes les m_2 dernières, qui ne joignent donc que des sites d'espèces non adjacentes dans R . Quand on parlera, sans plus préciser, des arêtes, il s'agira des $m + m_1 + m_2$ arêtes.

Tous les sites, sauf au plus un, sont alors l'origine ou l'extrémité d'une arête d'un des trois types (Pour éviter toute difficulté, on supposera en pratique toujours N pair).

Le graphe γ sera dit saturé si, après le tirage réel de m arêtes, on ne peut plus tirer de pré-arêtes ($m_1 = 0$ c'est-à-dire l'ensemble des couples de sites restant non tirés d'espèces adjacentes dans R est vide).

On appellera graphe de soutien le graphe Γ obtenu après le tirage des ($m + m_1 + m_2$) arêtes.

Remarques. — Du point de vue chimique, le rapport des nombres m_1 et m pourra servir à caractériser le degré d'avancement d'une réaction, tandis que la connaissance des m_2 non-arêtes peut permettre de prévoir le résultat d'une substitution chimique.

— En terminologie classique, on appelle branche d'un graphe un ensemble d'arêtes et de sommets, de degré 2 au plus, connexe. On appellera ici *branche* d'un graphe γ un ensemble de sites et d'arêtes réalisées tels que les sommets et arêtes correspondant de Γ forment une branche. On appellera *pré-branche* d'un graphe γ un ensemble de sites, d'arêtes réalisées, de pré-arêtes, et de non-arêtes tels que les sommets et arêtes correspondants de Γ forment une branche du graphe de soutien.

Branches et pré-branches peuvent être cycliques, pendantes, flottantes. Si N est pair, il n'y a cependant pas de pré-branches pendantes. Il ne peut exister de pré-branches flottantes non cycliques. C'est à la description des pré-branches d'une famille de graphes aléatoires que nous nous sommes intéressés dans [11].

Procédure π_3 ou enfilage pré-branche par pré-branche.

Les procédures π_1 et π_2 étaient donc, en quelque sorte, progressives : pour connaître la probabilité associée aux graphes à m arêtes il suffisait de partir du résultat pour les graphes à $(m - 1)$ arêtes, et d'ajouter, en

quelque sorte, une arête au graphe. La procédure π_3 , globale, ne pourra être appliquée que dans des conditions restrictives par rapport aux conditions qui permettraient d'appliquer π_1 et π_2 . Elle ne pourra, en particulier, être définie, que si le nombre des arêtes du graphe saturé obtenu par les premières procédures, est fixé, dès que les cardinaux des espèces sont fixés, et cela quelles que soient les valeurs de ces cardinaux. C'est donc une condition sur R que l'on cherche. Nous dirons que R possède la propriété A si et seulement si le nombre m d'arêtes du graphe saturé est déterminé par les seuls cardinaux des espèces E_i .

On a alors le théorème :

THÉORÈME 1. — R possède la propriété A si et seulement si chaque composante connexe de R (cf. fig. 4) :

- ou bien est un graphe complet avec boucles (graphe du type (1));
- ou bien est tel que ses sommets puissent être répartis en deux classes disjointes, chaque sommet à l'intérieur d'une classe étant non adjacent aux sommets de sa classe mais adjacent à tous les sommets de l'autre classe (graphe du type (2)).

Nous démontrerons ce théorème dans une formulation ensembliste qui nous semble mieux mettre en évidence sa généralité.

Sur S , ensemble des sites muni de sa partition en espèces, soit une partition définie par une famille de $(m + 1)$ sous-ensembles deux à deux disjoints, composée :

- d'un sous-ensemble dont on ne peut extraire deux sites d'espèces adjacentes dans R ;
- de m sous-ensembles de deux sites d'espèces adjacentes dans R .

Si, quels que soient les cardinaux des espèces, m ne dépend que de ceux-ci et pas de la façon dont on pourrait choisir les sous-ensembles, on dit que R a la propriété A' . La nouvelle forme équivalente du théorème sera donc :

THÉORÈME 1'. — Pour que R ait la propriété A' , il faut et suffit que chacune de ses composantes connexes soit du type (1) ou du type (2).

Démonstration. — Démontrons tout d'abord l'implication dans le sens « il suffit ».

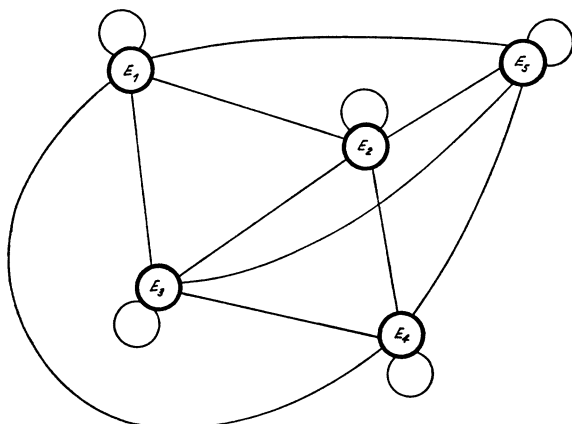
Dans tout ce qui suit $[N]$ signifie : « partie entière de N ».

1° Chaque composante connexe du type graphe complet, si la somme des

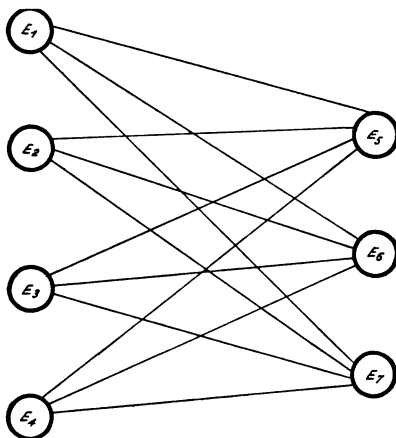
cardinaux des espèces qui la composent est N_1 , donnera $\left\lfloor \frac{N_1}{2} \right\rfloor$ couples de sites convenables.

2° Soit une composante C de l'autre type. Ses sommets sont répartis entre deux classes, C_1 et C_2 , de cardinaux respectifs (en sites), $|C_1|$ et $|C_2|$. Elle donnera nécessairement $\inf \{ |C_1|, |C_2| \}$ couples de sites convenables.

L'implication dans le sens « il faut » est moins immédiate.



Graphe du type (1).



Graphe du type (2).

FIG. 4.

LEMME. — *Tout graphe connexe fini admet un point qui n'est pas d'articulation.* Soit un graphe connexe fini. Suivant C. Berge, appelons $d(x, y)$ le plus petit nombre d'arêtes par lesquelles il faut passer pour aller de x à y , x et y désignant deux sommets quelconques du graphe. Par définition, l'écartement de x , $e(x)$, est égal à $\max_y d(x, y)$. Tout sommet d'écartement maximum est un sommet périphérique, et puisque notre graphe est fini, il existe des sommets d'écartement maximum.

Il ne reste plus qu'à montrer qu'un sommet périphérique n'est jamais point d'articulation.

Soit x_0 un sommet périphérique. Il existe au moins un x_1 tel que

$$d(x_0, x_1) = \max_{y, y'} d(y, y').$$

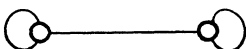
Or, si x_0 était point d'articulation, le sous-graphe obtenu en effaçant x_0 et les arêtes qui lui sont adjacentes contiendrait au moins deux composantes connexes, à l'intérieur de l'une desquelles on trouverait x_1 . En prenant x_2 , sommet d'une autre composante, on voit qu'il existerait un chemin de longueur minimum joignant x_1 à x_2 et plus long que $d(x_0, x_1)$, ce qui est impossible.


Considérons maintenant un graphe R_0 à r sommets ayant la propriété A, les cardinaux des espèces E_i étant arbitrairement choisis. Supposons R_0 connexe. Il existe donc au moins une espèce E_x qui ne soit pas point d'articulation dans R_0 . E_x est adjacente à au moins une autre espèce E_y . Supposons que l'on choisisse alors un cardinal pour E_y , supérieur au cardinal de E_x . On peut alors commencer à écrire une partition π par $|E_x|$ couples contenant chacun un site d'espèce E_x et un site d'espèce E_y . Il faut donc que le sous-graphe R'_0 à $(r - 1)$ sommets ait aussi la propriété, car l'espèce E_x étant vide la propriété doit être vraie pour les autres couples à former de la partition π .

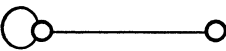
Par ailleurs, puisque E_x n'était pas point d'articulation, R'_0 est connexe. Tout graphe R_0 est connexe à r sommets jouissant de la propriété A peut s'obtenir à partir d'un des graphes R'_0 connexes à $(r - 1)$ sommets jouissant de la propriété A en ajoutant un sommet et en le reliant à certains des $(r - 1)$ autres sommets.

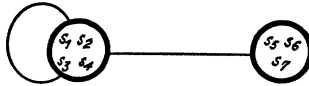
Supposons donc que R, connexe, jouisse de la propriété A et compte deux sommets.

Trois graphes non isomorphes peuvent être dessinés avec ces trois sommets :

1°  , qui est un graphe complet, convenable d'après le début de notre démonstration;

2°  , qui est le graphe le plus simple du second type convenable;

3°  , qui ne peut convenir. Il suffit par exemple que l'espèce avec boucle compte quatre sites et l'espèce sans boucle trois sites pour que la propriété ne soit plus vérifiée :



$$\pi_1 = \{ \{ (s_1, s_2), (s_3, s_4) \}, \{ s_5, s_6, s_7 \} \}$$

$$\pi_2 = \{ \{ (s_1, s_5), (s_2, s_6), (s_3, s_4) \}, \{ s_7 \} \}$$

On peut donc remarquer que tout graphe R connexe à deux sommets jouissant de la propriété A est soit du type (1), soit du type (2).

Supposons maintenant que le théorème soit vrai pour tout graphe R'_0 connexe à $(r - 1)$ sommets.

En vertu d'une propriété que nous venons d'établir l'ensemble des graphes R_0 connexes à r sommets convenables, peut s'obtenir à partir de l'ensemble des graphes R'_0 par adjonction d'un sommet et des arêtes correspondantes. Montrons alors que, dans ces conditions, R_0 est soit du type (1) soit du type (2).

1° Si R'_0 était complet, alors R_0 est complet.

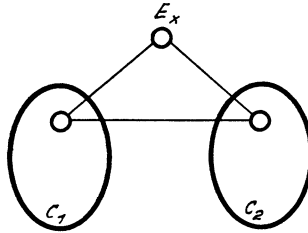
a) *Le sommet ajouté porte une boucle* : en effet, supposons que E_x , sommet ajouté, ne soit pas une boucle. Puisque R'_0 est complet, il existe une partition π composée de couples de sites d'espèces engendrant R'_0 sauf au plus un. Dès que l'effectif de l'espèce E_x est supérieur à deux, on voit qu'il existe des partitions qui conduisent à des valeurs de m différentes. E_x est donc une boucle.

b) *Le sommet ajouté est adjacent à tous les autres* : en effet, supposons qu'il existe E_{x_0} sommet non adjacent à E_x . Supposons que les cardinaux en sites de E_{x_0} et de E_x soient impairs tandis que les cardinaux des autres sommets sont tous pairs. On voit alors que, suivant la procédure choisie m peut varier d'une unité.

2° Si R'_0 était du type (2), ses sommets étaient donc répartis en deux classes disjointes que j'appelle C_1 et C_2 .

a) Si E_x est lié à un sommet d'une classe, il est disjoint de tous les sommets de l'autre classe.

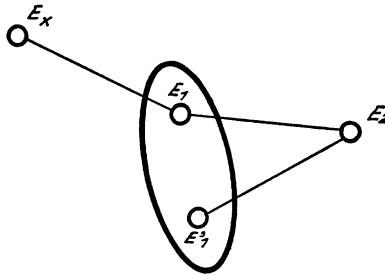
S'il était lié à l'un d'eux, on aurait effectivement dans R_0 le sous-graphe



Il suffit de se reporter à un exemple antérieur pour voir qu'on pourrait toujours choisir les effectifs des différentes espèces de façon que m varie avec le choix de π , que E_x porte ou non une boucle.

b) Si E_x est lié à un sommet d'une classe, il est lié à tous les sommets de l'autre classe.

Ici encore, si E_x était lié à un sommet d'une classe et disjoint d'un autre sommet de cette classe, on pourrait très bien se ramener à un choix de couples de sites parmi les espèces suivantes, et respectant le graphe suivant



E_x , sommet ajouté n'est pas relié à E'_1 . Si chaque espèce contient un site, la procédure qui consiste à choisir le site de E_1 et le site de E_2 pour former le premier couple empêche la formation du second, que l'on aurait obtenu en tirant tout d'abord le site de E_x et le site de E_1 , puis le site de E_2 et le site de E'_1 .

E_x est donc lié à E_1 , que E_x soit ou non une boucle.

c) Enfin montrons que E_x ne peut pas être une boucle. E_x est donc relié à tous les sommets de la classe C_1 . Supposons que E_x soit une boucle. On peut s'arranger pour que l'effectif de C_1 soit par exemple supérieur de deux

unités à l'effectif de C_2 . Soit une première partition π_1 qui forme, d'une certaine façon, le nombre maximum de couples de sites convenables près dans les espèces de C_1 et C_2 . Il reste donc deux sites dans C_1 . Si E_x comprend deux sites, on pourra alors continuer π_1 en prenant le couple formé des deux sites de E_x . On pourrait aussi obtenir π_2 en formant deux derniers couples avec chacun un site de E_x et un site de C_1 . E_x n'est donc pas une boucle.

d) Remarquons enfin que si R n'est pas connexe, puisque les cardinaux des E_i sont arbitraires, chaque composant connexe doit bien avoir la propriété.

Dans le cas où R est un graphe orienté, on aura le théorème correspondant :

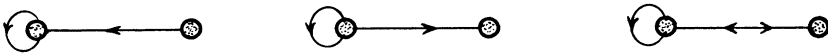
THÉORÈME 2. — Si l'on considère les ensembles maximaux (au sens de la cardinalité) de sous-ensembles disjoints de deux sites tels que l'espèce du premier précède l'espèce du second dans R , pour que le cardinal d'un tel ensemble, les cardinaux des espèces étant choisis arbitrairement, ne dépende pas de l'ensemble choisi, il faut et suffit que chaque composante de R :

- (1)' ou bien soit un graphe où tous les sommets sont des boucles et tel que chaque couple d'espèces soit joint par au moins un arc;
- (2)' ou bien soit un graphe bipartite tel que chaque sommet soit connecté à tous ceux de la classe de sommets à laquelle il n'appartient pas par au moins un arc et seulement à ceux-là.

Indiquons rapidement la démonstration qui ressemble à la précédente.

1° A l'intérieur d'une composante connexe, tous les sommets sont des boucles, ou aucun sommet n'est une boucle.

En effet, on pourrait sinon avoir comme sous-graphe de R :



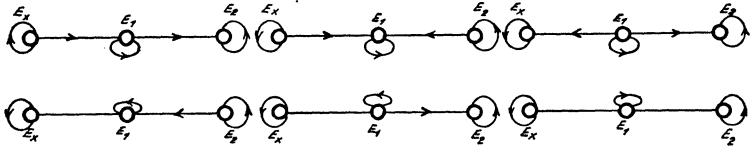
dont on vérifiera facilement qu'ils n'ont pas la propriété voulue.

D'une façon générale, tout sous-graphe de R obtenu en retirant une espèce et les arcs qui ont cette espèce pour origine et pour extrémité doit encore avoir la propriété.

2° Si tous les sommets sont des boucles, on démontre par récurrence que, puisque pour obtenir une composante connexe à r espèces on doit partir

d'une composante connexe à $(r - 1)$ espèces, le sommet E_x ajouté doit être connecté à tous les autres.

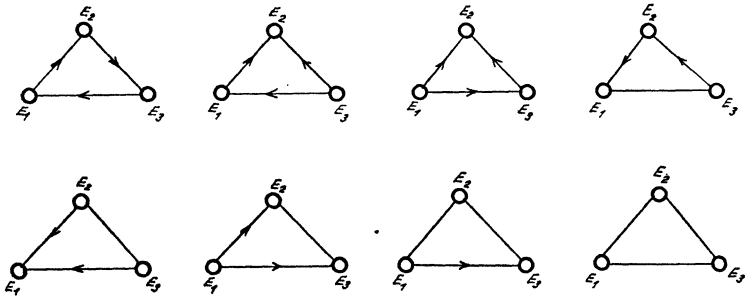
Supposons qu'il en existe un auquel il ne soit pas connecté. On pourrait donc mettre en évidence un des sous-graphes suivants :



en prenant $E_x = 3$, $E_1 = 2$ et $E_2 = 3$ on voit facilement qu'on peut écrire 3 ou 4 couples ordonnés suivant la méthode d'écriture choisie.

3° Si aucun sommet n'est une boucle dans la composante, on va encore raisonner par récurrence et chercher à ajouter un sommet à une composante connexe à $(r - 1)$ sommets :

a) E_x ne peut pas avoir un arc qui le joigne à un sommet d'une classe et un autre arc à un sommet de l'autre classe. En effet, si cela était, on pourrait mettre en évidence un des huit sous-graphes :

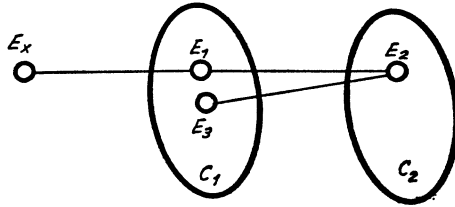


Il suffit de prendre $|E_1| = 2N$, $|E_2| = |E_3| = N$ pour imaginer une écriture des couples qui infirme la propriété. E_x ne peut donc être connecté qu'aux sommets d'une seule classe.

Supposons qu'il ne soit pas connecté à au moins un. On isolerait un sous-graphe à quatre sommets, E_x , deux sommets E_1 et E_3 dans C_1 et E_2 dans C_2 , E_3 n'étant pas connecté à E_x . En prenant :

$$|E_x| = |E_1| = |E_3| = |E_2| = N,$$

on peut toujours imaginer des procédures qui infirment la propriété.



On remarque que, dans le cas où l'on impose à R la propriété d'être symétrique, on retombe sur les graphes (1) et (2) du théorème précédent.

DÉFINITIONS :

1° On appellera *pré-nœud* tout monomère comptant trois sites au moins, et *pré-anti-nœud* tout monomère comptant exactement deux sites (pré-nœuds et pré-anti-nœuds donneront respectivement les nœuds et les anti-nœuds du graphe de soutien). Dans la figure 3, page 264, μ_1 et μ_2 , pré-nœuds, sont des nœuds. μ_4 au contraire est resté pré-anti-nœud.

2° Par extension naturelle, nous appellerons *site pré-nodal* (resp. *pré-anti-nodal*) tout site appartenant à un pré-nœud (resp. à un pré-anti-nœud).

Définition de la procédure π_3 .

Malgré l'aspect rébarbatif de la description formelle de π_3 , nous exhortons le lecteur à ne pas se laisser rebuter et à poursuivre ses efforts qui devraient être éclairés par de nombreux exemples. Pour définir π_3 , il faut supposer que :

- a) L'ensemble des monomères à un seul site est vide.
- b) L'ensemble des sites pré-nodaux a un cardinal pair.
- c) Le graphe R est du type (1) ou du type (2); par conséquent, en vertu du théorème 1 : si les monomères et m sont donnés, m_1 et m_2 sont fixés également.

1^{re} cas : si R est du type (1).

1° On tire un site pré-nodal.

2° On tire une arête puis un site : — si ce dernier site tiré est pré-nodal, on a tiré la première pré-branche, et l'on reprend, avec les sites et arêtes non tirés la procédure au 1° ;

— si ce dernier site tiré est pré-anti-nodal, on tire le second site du pré-anti-nœud, puis on continue, avec les sites et les arêtes non tirés, la procédure comme au 2°.

3° S'il reste des pré-anti-nœuds non tirés après le tirage du dernier site pré-nodal, on considère comme tiré de l'ensemble des sites restant un certain site fixé à l'avance, par exemple celui des sites restants qui porte l'indice le plus élevé. On tire ensuite une arête puis un site.

a) Si ce dernier site tiré est le second site du pré-anti-nœud dont le premier site a été « tiré » d'une façon déterministe, on a obtenu une pré-branche cyclique. On continue alors avec les sites restants comme au 3°.

b) Si ce dernier site tiré fait partie d'un autre pré-anti-nœud que celui dont on avait tiré le premier site, on tire le second site du pré-anti-nœud, on tire une arête, puis un site parmi les sites restants, que l'on considère comme le dernier site tiré. On se trouve ramené au a) du 3°.

2^e cas : Si R est du type (2) :

Ses sommets sont divisés en deux classes C_1 et C_2 , l'effectif de sites de C_1 étant supérieur à l'effectif de sites de C_2 , le nombre des non-arêtes sera positif.

1° On tire un site pré-nodal. C'est le dernier site tiré.

A. — Il appartient à C_1 . On tire une arête dans l'ensemble des arêtes :

a) c'est une non-arête. On tire alors un second site parmi les sites de C_1 . C'est le dernier site tiré ;

b) c'est une pré-arête ou une arête réalisée. On tire alors un second site parmi les sites de C_2 . C'est le dernier site tiré.

B. — Il appartient à C_2 . On tire alors une arête parmi les pré-arêtes et les arêtes réalisées, puis un site de C_1 . C'est le dernier site tiré.

2° Si le dernier site tiré est pré-anti-nodal, on tire le second site du pré-anti-nœud. C'est le dernier site tiré, et l'on poursuit le tirage comme en 1° A.

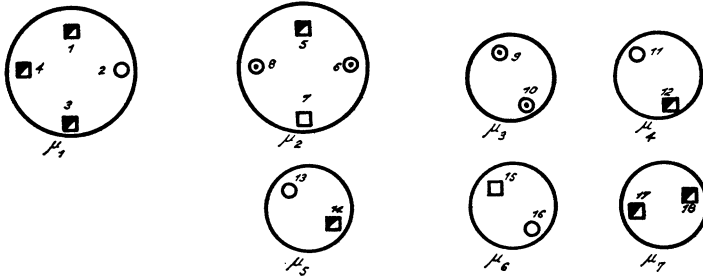
Si le dernier site tiré est pré-nodal, on a terminé une pré-branche et on recommence comme en 1°.

3° S'il ne reste plus de sites pré-nodaux, on poursuit la procédure comme lorsque R est du type (1) : On tire le premier site suivant une méthode non aléatoire, par exemple le site d'indice le plus élevé, puis on procède comme en 1°, A et 1°, B.

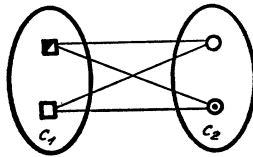
a) Si le dernier site tiré est le second site du pré-anti-nœud auquel appartenait le premier site, on a terminé une pré-branche cyclique, et l'on se trouve ramené au 3°.

b) Si le dernier site tiré n'appartient pas au pré-anti-nœud auquel appartenait le 1^{er} site, on tire le second site du pré-anti-nœud du dernier site tiré, puis on procède comme en 1°, A et 1°, B. On se trouve alors ramené à 3°, a).

Exemple. — Soit l'ensemble de monomères :



Soit R le graphe du type (2)



Card $C_1 = \text{Card } C_2 + 2$. On aura donc une non-arête et 8 arêtes réalisées ou pré-arêtes.

Supposons qu'il y ait une pré-arête.

1° Je tire un site parmi $(s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8)$. Supposons que ce soit s_4 . s_4 appartient à C_1 .

2° Je tire une arête. Supposons que ce soit la non-arête. Je tire alors un site parmi les sites non tirés de C_1 . Supposons que ce soit s_{15} . Je tire donc aussitôt s_{16} .

3° Puisque s_{16} appartient à C_2 , je tire une arête entre les arêtes non tirées. Supposons que ce soit une arête réalisée. Je tire un site non encore tiré de C_2 . Supposons que ce soit s_5 . Ma première pré-branche est tirée.

Supposons que :

la seconde pré-branche tirée soit $(s_1, \text{arête réalisée}, s_8)$;

la troisième pré-branche tirée soit $(s_3, \text{arête réalisée}, s_9, s_{10}, \text{arête réalisée}, s_7)$;

la quatrième pré-branche tirée soit $(s_2, \text{arête réalisée}, s_{17}, s_{18}, \text{arête réalisée } s_6)$.

4° Il ne reste plus que des sites situés sur des pré-anti-nœuds. On tire donc le site 14, puis une arête; supposons que ce soit la pré-arête, puis un site;

supposons que ce soit le site 13. La première pré-branche cyclique est terminée. π_4 et la dernière arête forment alors nécessairement une seconde pré-branche cyclique. Le graphe Γ obtenu est le suivant :

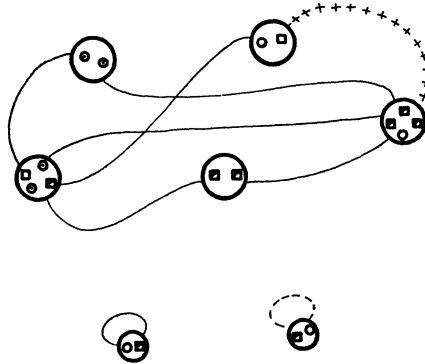


FIG. 5.

Pour se représenter simplement la procédure π_3 , le lecteur pourra reprendre l'image de la petite fille et du collier de perles en cherchant suivant quelle méthode elle pourrait pratiquement construire un graphe de soutien, qui ait de plus quelques chances de ressembler aux graphes de soutien obtenus au moyen des procédures π_1 et π_2 .

Par ailleurs, cette procédure π_3 , qui semble bien compliquée, est susceptible d'un traitement probabiliste plus simple que les procédures π_1 et π_2 , dans les questions de comportement asymptotique qui vont nous intéresser par la suite. Par conséquent, dans la zone d'équivalence des procédures, on aura avantage à « passer en procédure π_3 ».

La dernière procédure, susceptible d'un traitement plus « commode » que les trois autres ne peut être réservée qu'à des cas très particuliers, à notre connaissance. Nous ne la développerons guère par la suite, son domaine d'équivalence avec les autres procédures restant très étroit.

Procédure π_4 : le graphe de soutien rigide.

1° A partir de l'ensemble donné de monomères, d'un graphe R_1 complet avec boucles, on engendre, suivant la procédure π_1 un graphe aléatoire auxiliaire γ_1 , saturé.

2° Toutes les fois qu'une arête de γ_1 joint deux sites d'espèces non adjacentes dans R , on baptise cette arête non-arête.

3° On considère les arêtes de γ_1 qui ne sont pas non-arêtes. Pour chacune

d'elles, on effectue un tirage indépendant : avec la probabilité p l'arête sera baptisée arête réalisée, avec la probabilité $(1 - p)$ elle sera baptisée pré-arête.

Du point de vue chimique, ce mode de génération revient à réaliser dans un premier temps une certaine distribution des sites en position aléatoire, tout site ayant une chance constante de se retrouver face à face avec tout autre, puis d'effectuer la réaction sans que des remaniements de structure puissent avoir lieu au fur et à mesure de l'avancement de cette réaction, qui est mesuré par p .

Voyons maintenant en guise de conclusion à ce chapitre comment d'autres problèmes de graphes aléatoires peuvent être considérés comme relevant des procédures décrites plus haut.

1° Les graphes étudiés par A. Renyi et P. Erdős sont des graphes qu'on peut obtenir au moyen de π_1 .

On les obtient à partir de N monomères de $N - 1$ sites. Chacun de ces $N - 1$ sites est d'une espèce différente des $N - 2$ sites restants du monomère. Chacun des $N - 1$ monomères restants porte un site et un seul d'une des espèces du premier. En réitérant cette opération, on trouvera $(N - 1) + (N - 2) + \dots + 1 = \frac{N(N - 1)}{2}$ espèces de sites et les $N(N - 1)$ arêtes possibles viendront de ce que le graphe R est réduit à un ensemble de boucles :

Exemple :

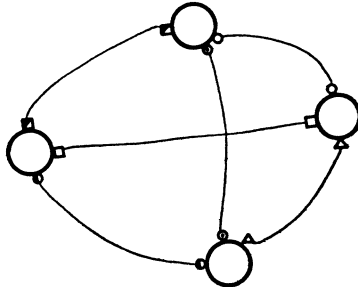


FIG. 6.

— Considérons le réseau carré plan des problèmes de percolation classiques. Ses monomères (sommets du réseau) peuvent être indexés par un ensemble

$$I = \{(i_1, i_2); i_1 \text{ et } i_2 \in \mathbb{Z}\}.$$

Tous les monomères sont composés de 4 sites.

Considérons le monomère (i_1, i_2) , et distinguons ses quatre sites numérotés :

$$(1, i_1, i_2), (2, i_1, i_2), (3, i_1, i_2) \text{ et } (4, i_1, i_2).$$

L'espèce du site $(1, i_1, i_2)$ sera exclusivement adjacente dans R à l'espèce du site $(3, i_1 - 1, i_2)$;

l'espèce du site $(2, i_1, i_2)$ à celle du site $(4, i_1, i_2 + 1)$;

l'espèce du site $(3, i_1, i_2)$ à celle du site $(1, i_1 + 1, i_2)$;

l'espèce du site $(4, i_1, i_2)$ à celle du site $(2, i_1, i_2 - 1)$.

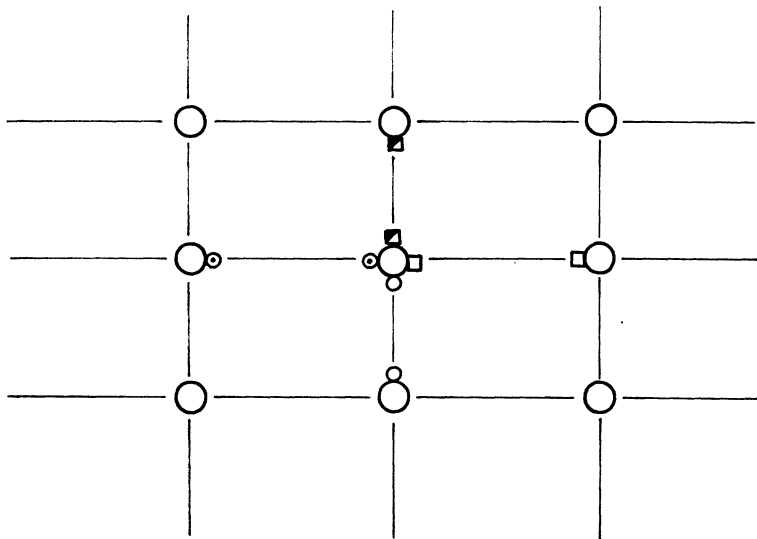


FIG. 7.

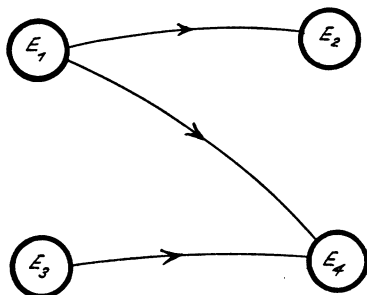
On voit que l'on exprime ainsi sans ambiguïté la contrainte du réseau. En utilisant encore ici la procédure π_1 , on voit que l'on retombe sur le mode de description le plus courant des processus de percolation. Il est clair que cet exemple s'étend immédiatement à tous les réseaux de configuration périodique.

Remarquons que si l'on se limite à un nombre fini de monomères, on tombe sur un graphe qui appartient exactement à la classe des graphes que nous avons définie et dans laquelle on suppose sites, espèces et monomères en nombres finis.

La procédure π_2 nous a semblé pouvoir fournir des modèles commodes pour les phénomènes sociologiques dans lesquels la relation entre deux personnes n'est pas symétrique, et l'initiative de cette relation réservée à l'une d'entre elles. On représentera les personnes par des monomères, et

leurs profils, relativement à la relation qu'on désire étudier, par des sites. Supposons que deux sortes de personnes existent, les forts et les faibles; qu'un fort puisse influencer un fort comme un faible et qu'un faible ne puisse influencer qu'un autre faible. Enfin supposons que les individus s'influencent au hasard à un instant donné chaque individu pouvant choisir parmi ceux qu'il peut influencer éventuellement en fonction d'une certaine capacité à être influencé.

La procédure π_2 pourrait rendre compte de ce schéma. Par exemple, chaque monomère porterait huit sites de deux espèces et il y aurait quatre espèces de sites : E_1 , E_2 , E_3 , E_4 . Un monomère pourrait être composé soit de quatre sites d'espèce E_1 et de quatre sites d'espèce E_2 , soit de quatre sites d'espèce E_3 et de quatre sites d'espèce E_4 . Le graphe R serait le suivant :



On pourrait dire que « le monomère x influence le monomère y » si et seulement si il existe un chemin de x vers y . On pourrait alors étudier par exemple la taille des groupes influencés par les mêmes monomères, la taille de l'arbre maximum dans le graphe, etc.

Ce genre d'application peut évidemment s'étendre à π_1 . On peut obtenir ainsi de nombreux modèles différents de sociogrammes.

Enfin, dans le travail originel de C. M. Bruneau, on s'intéressait essentiellement à des graphes pour lesquels les monomères, réduits à quelques types simples (sites de deux espèces; distinction entre pré-nœuds et différentes sortes de pré-anti-nœuds) étaient une infinité dénombrable.

La seule difficulté provenait alors du tirage des arêtes, dans le cas de la première procédure.

Nous verrons au cours des chapitres suivants comment cette difficulté a pu être tournée.

Nous allons établir maintenant un certain nombre de résultats probabilistes et combinatoires. Nous montrerons dans les derniers chapitres qu'on peut appliquer ces résultats aux graphes que nous venons de définir.

III

QUELQUES THÉORÈMES GÉNÉRAUX

A. — Équivalence entre deux procédures.

On dira que π_i et π_j sont équivalentes si, à partir des mêmes monomères et du même nombre d'arêtes réalisées (respectivement d'arc réalisés) quels qu'ils soient, ils conduisent au même graphe aléatoire, c'est-à-dire à la même probabilité dans (Ω, \mathcal{A}, P) , puisque Ω et \mathcal{A} sont identiques dans les deux procédures.

Dans tout ce qui suit nous avons distingué le cas R orienté du cas R non orienté et nous avons cherché à quelles conditions les procédures pouvaient être équivalentes.

ÉQUIVALENCE ENTRE π_1 ET π_2

On peut distinguer utilement les *procédures « arête par arête »* comme π_1 et π_2 des *procédures globales*, comme π_3 et π_4 . π_1 et π_2 peuvent être considérées comme des processus dont l'état final décrit l'espace de probabilités (Ω, \mathcal{A}, P) cherché, en supposant par exemple que chaque unité de temps entraîne l'addition d'une arête jusqu'à ce que l'on atteigne le nombre d'arêtes fixé à l'avance. Ajouter une arête au graphe garde alors une signification descriptive : on laisse le processus évoluer une unité de temps supplémentaire. Il n'en n'est pas de même pour π_3 et π_4 . On ne peut associer à π_4 de processus, puisque la construction du graphe aléatoire se fait en deux temps. Quant au processus associé à π_3 et dans lequel, à chaque unité de temps on ajouterait une arête, une non-arête ou une anti-arête aux pré-branches déjà décrites, on voit que seul son état final compte, et que si l'on cherche à obtenir un graphe à $x + 1$ arêtes réalisées, après avoir obtenu le graphe à x arêtes réalisées, il faudra reprendre au départ la procédure π_3 .

Dans le cas de procédures « arête par arête », on pourra parler d'*équivalence arête par arête*.

Deux procédures arête par arête π et π' , conduisent, après le tirage de x arêtes aux probabilités P_x et P'_x . On dit de π et π' sont équivalentes arête par arête si, pour toute réalisation à x arêtes ω_x du graphe, P_{x+1} conditionné par ω_x est égal à P'_{x+1} conditionné par ω_x .

Il est clair que deux procédures équivalentes arête par arête sont équivalentes, tandis que la réciproque est, en général, fautive.

I. *R est non orienté.*

Supposons que $m - 1$ arêtes aient été tirées et que les deux procédures π_1 et π_2 aient donné le même graphe.

Tirons maintenant au hasard la $m^{\text{ième}}$ arête.

Appelons N' le nombre des sites non encore tirés, et dont les espèces ont au moins une espèce adjacente dans R non vide de sites; pour toute espèce E_i ayant au moins une espèce adjacente dans R encore non vide, soient :

N'_i le nombre des sites d'espèce E_i n'ayant pas encore été tirés;

N'_i^* le nombre des sites n'ayant pas encore été tirés et dont les espèces sont adjacentes de E_i dans R , ce nombre étant diminué d'une unité au cas où E_i serait sommet d'une boucle;

M' le nombre total des arêtes distinctes entre lesquelles on tirera la $m^{\text{ième}}$ suivant π_1 .

LEMME 2. — Pour que les deux tirages soient équivalents, il faut alors que $\frac{1}{N'_i} + \frac{1}{N'_j}$ soit indépendant de i et j , et égal au rapport $\frac{N'}{M'}$. Dans le cas d'une boucle, on aura $\frac{2}{N'_i - 1} = \frac{N'}{M'}$.

En effet, tirer l'arête $s_u s_v$, si les deux tirages sont équivalents, doit avoir la même probabilité avec les deux procédures. Supposons que $E(s_u) = E_i$, $E(s_v) = E_j$. Avec la procédure π_1 , la probabilité de tirage de $s_u s_v$ est $\frac{1}{M'}$, et avec π_2 , $\frac{1}{N'} \frac{1}{N'_i} + \frac{1}{N'} \frac{1}{N'_j}$. D'où le résultat.

On a alors le

THÉORÈME 3. — Pour que π_1 et π_2 soient équivalentes arête par arête il faut et il suffit

1° que R soit complet avec boucle (c'est-à-dire du type (1)) ou :

2° que R soit un graphe bipartite sans boucle tel que chaque sommet d'une classe, disjoint des sommets de sa classe, soit joint à tous les autres (c'est-à-dire du type (2)).

Démonstration. — On convient de ne pas envisager le cas où l'ensemble des sommets de R contiendrait des espèces isolées non boucles.

Nous allons commencer par montrer le théorème dans le sens « il faut ».

1° Si R n'est pas réduit à une boucle, il ne peut contenir de boucle isolée.

Soit en effet une telle boucle; supposons qu'elle ait lieu sur l'espèce E_i et que le $m^{\text{ième}}$ tirage donne le couple de sites (s_{i_1}, s_{i_2}) appartenant à cette espèce.

Suivant la procédure π_1 , la probabilité associée à ce tirage était $\frac{1}{M'}$ tandis qu'il était $\frac{2}{(N'_i - 1)N'}$ suivant la procédure π_2 .

Si R n'est pas réduit à une boucle, c'est qu'en dehors de la boucle E_i , il contient soit une autre boucle, soit une paire d'espèces adjacentes non vides de sites (Lorsque $m = 1$ au moins). Lors du tirage de l'arête suivante N'_i aura diminué de deux unités, mais pas les autres N'_j , $j \neq i$. N' aura également décré de deux unités, et l'on aura eu, pour au moins un couple j, k convenable

$$\frac{2}{N'(N'_i - 1)} = \frac{1}{N'} \left(\frac{1}{N'_j} + \frac{1}{N'_k} \right) = \frac{1}{M'}$$

$$\frac{2}{(N' - 2)(N'_i - 3)} = \frac{1}{(N' - 2)} \left(\frac{1}{N'_j} + \frac{1}{N'_k} \right) = \frac{1}{M' - 2N + 3}$$

en effet, le « M' » du $(m + 1)^{\text{ième}}$ tirage sera égal à :

$$M' - \left[\frac{N(N - 1)}{2} - \frac{(N - 2)(N - 3)}{2} \right] = M' - 2N + 3$$

On peut donc écrire :

$$\frac{2}{N'_i - 1} = \frac{1}{N'_j} + \frac{1}{N'_k} = \frac{2}{N'_i - 3}$$

ce qui est impossible.

2° Si une boucle est liée à un sommet, celui-ci est forcément une boucle.

Soit E_i la boucle et E_j le sommet lié à la boucle. Il faut alors que la probabilité au $m^{\text{ième}}$ tirage de tirer l'arête s_{i_1}, s_{i_2} par exemple (s_{i_1} et s_{i_2} particularisés tels que $E(s_{i_1}) = E(s_{i_2}) = E_i$) soit égale à la probabilité de tirer l'arête (s_{j_1}, s_{j_1}) (s_{j_1} particularisé tel que $E(s_{j_1}) = E_j$).

Soit :

$$\frac{1}{N'} \frac{2}{N'_i} = \frac{1}{N'} \left(\frac{1}{N'_i} + \frac{1}{N'_j} \right)$$

qui entraîne $N'_i = N'_j$, qui ne peut être vérifié au long du tirage que si les

sommets adjacents à E_i sont les mêmes que les sommets adjacents à E_j . E_j est donc une boucle.

3° Toute composante connexe de R qui contient une boucle est le seul sous-graphe connexe complet de R qui contienne toutes les boucles de R .

En effet, si E_j est une espèce adjacente à la boucle E_i , non seulement c'est une boucle, mais encore les sommets adjacents de E_i doivent être les mêmes que les sommets adjacents de E_j . Si E_i n'a pas d'autres sommets adjacents que lui-même et E_j , E_i et E_j forment à eux deux un graphe complet connexe.

On raisonne ensuite de proche en proche, et, puisque R est fini, on a le résultat.

4° R est connexe.

a) R réduit à une boucle est connexe.

b) Si R contenait une boucle et n'était pas réduit à une boucle, il pourrait être réduit à un graphe complet.

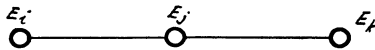
c) S'il n'était pas réduit à un graphe complet, c'est qu'il contiendrait deux composantes connexes au moins dont une au moins ne compterait pas de boucle et au moins deux sommets.

Tirons donc une arête et supposons qu'elle soit justement tirée entre les deux espèces dont nous venons de nous assurer l'existence. En écrivant l'égalité du lemme à l'intérieur d'une autre composante connexe, on voit que $\frac{N'}{M'}$ reste invariant. Cependant, pour les deux sommets qui nous intéressent, l'écriture du lemme oblige $\frac{N'}{M'}$ à croître. Le graphe est donc connexe.

5° Si R est un graphe sans boucle (c'est-à-dire n'est pas un graphe complet) il est du type (2).

Si R n'a que deux sommets, la propriété est vérifiée.

Si R compte trois sommets ou plus, trois espèces au moins E_i, E_j, E_k sont les sommets d'un sous-graphe de R ayant l'allure suivante



Mieux, pour toute espèce on peut trouver un sous-graphe de R de la forme précédente et qui comprenne parmi ses sommets l'espèce envisagée. Puisque la probabilité pour tirer une des arêtes (s_i, s_j) (telle que $E(s_i) = E_i, E(s_j) = E_j$) doit être la même que la probabilité de tirer une arête (s_j, s_k)

(telle que $E(s_k) = E_k$), il faut donc que, tout au long du tirage, l'on ait :

$$\frac{N'}{M'} = \frac{1}{N_i^*} + \frac{1}{N_j^*} = \frac{1}{N_j^*} + \frac{1}{N_k^*}$$

$$\downarrow$$

$$N_i^* = N_k^*$$

Pour que cette égalité soit vraie tout au long de n'importe lequel des tirages possibles, il faut et il suffit que les espèces adjacentes de E_i soient les mêmes que celles de E_k .

Dans ces conditions, E_i ne saurait être adjacent à E_k , car alors, pour la même raison, E_k serait une boucle, ce qui est contradictoire avec l'hypothèse. On possède donc un algorithme pour séparer les deux classes de sommets répondant à la propriété du théorème : on choisit une espèce au hasard. On appelle sa classe C_1 . Puisque R est connexe et fini, en procédant de proche en proche on déterminera successivement la classe de tous les sommets.

Montrons maintenant l'implication dans le sens « il suffit ».

Cette démonstration est immédiate en considérant, d'une part le cas où R est complet et d'autre part le cas où R est un graphe bipartite sans boucle.

Les probabilités de tirage de chaque arête, quel que soit le nombre d'arêtes déjà tirées sont les mêmes dans les deux procédures.

Il s'ensuit deux corollaires immédiats.

COROLLAIRE 1. — Si R est connexe, symétrique, et tel qu'on puisse procéder au tirage du graphe pré-branche par pré-branche, alors π_1 et π_2 sont strictement équivalentes pour R .

COROLLAIRE 2. — Si π_1 est strictement équivalente à π_2 pour R , indépendamment du nombre d'arêtes fixé, le nombre des non-arêtes du graphe est fixé dès que l'on connaît les cardinaux des espèces, et l'on peut procéder au tirage du graphe pré-branche par pré-branche.

Le cas où R est orienté, bien que comportant des raisonnements similaires, va nous conduire à des résultats forts différents. Ceci provient de ce qu'une boucle E qui donnait $\frac{|E|(|E|-1)}{2}$ couples en donne $|E|(|E|-1)$,

que deux espèces E_1 et E_2 adjacentes qui donnaient $\frac{|E_1||E_2|}{2}$ couples en donnent soit $|E_1||E_2|$ soit $2|E_1||E_2|$.

La démonstration étant très voisine de la précédente, nous ne ferons que l'indiquer rapidement.

II. *R est orienté.*

Appelons ici de même, après le tirage du $m^{\text{ième}}$ arc :

— N' le nombre de sites non encore tirés, et dont les espèces sont suivies dans R par au moins une espèce non vide de sites.

Et, pour toute espèce E_i ayant au moins une espèce adjacente dans R encore non vide :

— N'_i le nombre des sites d'espèce E_i n'ayant pas encore été tirés;

— N_i^* le nombre des sites n'ayant pas encore été tirés dont les espèces suivent E_i dans R , diminué d'une unité, au cas où E_i serait sommet d'une boucle;

— M' le nombre total des arcs distincts entre lesquels on tirera la $(m+1)^{\text{ième}}$ suivant π_2 .

Le lemme prend alors la forme suivante :

LEMME 3. — Si les deux procédures sont équivalentes, \forall_i , tout au long du tirage, $\frac{1}{N_i^*} = \frac{N'}{M'}$, quantité indépendante de i .

On en déduit immédiatement le théorème correspondant.

THÉORÈME 4. — Pour que π_1 et π_2 soient équivalentes, il faut et il suffit que les sommets de R soient répartis en 3 classes (dont deux éventuellement vides) :

— Une classe dont tous les sommets sont de demi-degré intérieur nul, chaque membre de la classe étant ascendant de tous les sommets des deux autres classes.

— Une classe dont tous les sommets sont de demi-degré extérieur nul, chaque membre de la classe étant descendant de tous les sommets des deux autres classes.

— Une classe dont les sommets ont un demi-degré extérieur et un demi-degré intérieur différents de zéro, et qui engendrent de plus entre eux un sous-graphe complet avec boucles.

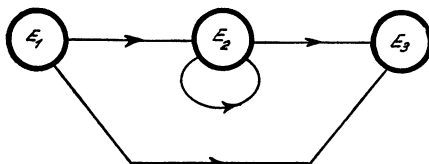
Commençons ici encore par démontrer l'implication dans le sens « il faut ».

1° Puisque N_i^* est indépendant de i , il faut que tous les sommets de demi-degré extérieur différent de zéro aient les mêmes descendants, sinon le $m^{\text{ième}}$ tirage pourrait changer N_i^* pour un certain i et laisser inchangé N_j^* pour un certain j , ce qui est impossible.

2° En conséquence tout sommet qui a un ascendant et un descendant est une boucle. Plus précisément, si E_1, E_2, E_3 sont trois espèces de R telles que l'on ait un sous-graphe de R identique à :



Alors l'ensemble des arcs qui relient entre eux E_1, E_2 et E_3 contient forcément le sous-graphe :



3° Pour la même raison, les boucles ayant un demi-degré extérieur non nul ont les mêmes descendants, c'est-à-dire forment entre elles un graphe complet avec boucles; les sommets sans boucle de demi-degré extérieur différent de zéro sont disjoints.

4° Il ne reste plus que les sommets qui ont un demi-degré extérieur nul. Descendants d'au moins un autre sommet du graphe ils sont descendants de tous, et puisqu'ils ont un demi-degré extérieur nul, ils sont nécessairement disjoints.

Exemple :

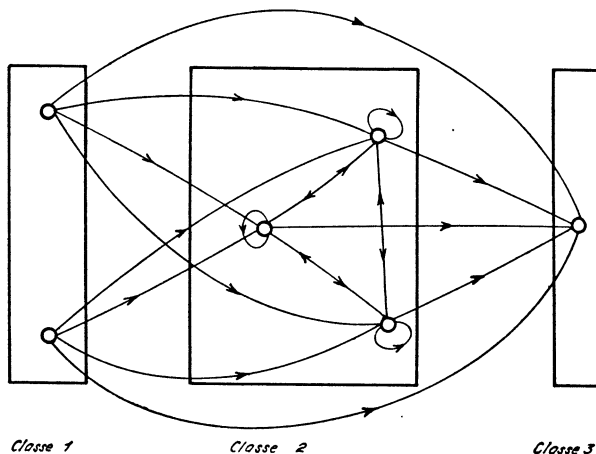


FIG. 8.

L'implication dans le sens « il suffit » est alors immédiate, puisque, à chaque nouvelle arête tirée, les procédures π_1 et π_2 confèrent la même probabilité.

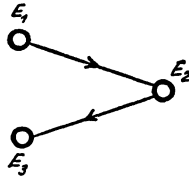
Remarque. — On voit que dans le cas R orienté, l'équivalence entre π_1 et π_2 n'entraîne la possibilité d'effectuer π_3 que si :

1° R est complet avec boucles;

2° R est bipartite, sans boucle. Ses sommets sont répartis entre deux classes à l'intérieur desquelles les sommets sont disjoints, tandis que tous les sommets de C_1 sont reliés à tous les sommets de C_2 par des arcs qui sont tous de C_1 vers C_2 .

On s'en persuadera sur le contre-exemple suivant :

R :



$$s_1 \in E_1, \quad s_2 \in E_2.$$

La probabilité de tirer (s_1, s_2) est : $\frac{1}{(|E_1| + |E_2|) + (|E_2| + |E_3|)}$.

Elle devrait aussi être égale à :

$$\frac{|E_1|}{|E_1| + |E_2| + |E_3|} \cdot \frac{1}{|E_2|}$$

Ce qui entraîne :

$$|E_1| + |E_2| + |E_3| = |E_1| (|E_1| + |E_3|)$$

qui ne peut être vérifié en général, bien que π_3 puisse être décrite.

ÉQUIVALENCE ENTRE π_1 , π_2 ET π_3

THÉORÈME 5. — Dans les cas où π_1 et π_2 sont équivalentes, si le nombre des sites pré-nodaux est pair, π_3 est équivalente à π_1 et π_2 et conduit à une distribution de probabilité P uniforme.

Démonstration. — Il est clair que la procédure π_3 permet d'obtenir tous les graphes obtenus au moyen des procédures π_1 ou π_2 . Envisageons en effet

un graphe quelconque de Ω . On peut toujours le parcourir pré-branche par pré-branche. Une fois toutes les pré-branches décrites restent les boucles qu'on ne peut parcourir que d'une seule façon.

I. *Supposons tout d'abord que γ soit saturé.*

1° *R est un graphe complet avec boucles.* — Soit N le nombre total de sites, $2v$ le nombre total de sites pré-nodaux.

Soit ω une réalisation du graphe. Celle-ci est entièrement déterminée par les arêtes de Ω , c'est-à-dire par la connaissance de $\frac{N}{2}$ couples de sites disjoints. Si on utilise la procédure π_2 , on voit qu'une permutation Θ des N premiers entiers est une écriture d'un tel ensemble de couples (¹). Mais toute permutation déduite de Θ en permutant les $\frac{N}{2}$ couples de nombres consécutifs convient, ainsi que celle obtenue en transposant deux nombres formant un couple.

D'où la probabilité de ω :

$$\frac{\left(\frac{N}{2}\right)! 2^{\frac{N}{2}}}{N!}$$

dans les procédures π_1 et π_2 .

Dans la procédure π_3 , considérons le tirage jusqu'au dernier site pré-nodal, et supposons qu'il reste $2p$ sites pré-anti-nodaux. La probabilité associée à cette première partie de tirage s'écrit :

$$\frac{1}{N-1} \cdot \frac{1}{N-3} \cdots \frac{1}{2p+1} \frac{1}{v} \cdot \frac{1}{v-2} \cdots \frac{1}{4} \frac{1}{2}.$$

Quant à la probabilité associée aux pré-branches cycliques, elle s'écrit :

$$\frac{1}{2p-1} \frac{1}{2p-3} \cdots \frac{1}{5} \frac{1}{3}$$

Le lecteur s'en convaincra facilement sur un exemple en se reportant à la définition de la procédure. On voit que la probabilité associée à un tirage complet ne dépend que de N et de v .

Or, pour obtenir une écriture quelconque du graphe par la procédure π_3 , on peut utiliser un quelconque des tirages déduits du précédent, par le

(¹) Chaque couple étant formé d'un nombre de la permutation de rang impair et du nombre de rang immédiatement supérieur (par exemple).

produit de permutations entre les pré-branches et par le renversement de certaines pré-branches (le 1^{er} site de l'une étant le dernier site de l'autre).

$2^v \binom{v}{2}!$ tirages obtenus au moyen de π_3 donnent donc le même graphe.

Chaque graphe a donc la probabilité :

$$\frac{1}{N-1} \frac{1}{N-3} \cdots \frac{1}{5} \frac{1}{3} \frac{1}{(v-2) \dots 4 \cdot 2} \frac{2^v \binom{v}{2}!}{N!} = \frac{\left(\frac{N}{2}\right)! 2^{\frac{N}{2}}}{N!}.$$

Si R est un graphe complet avec boucles, le théorème est donc démontré, si la procédure est poussée jusqu'au bout, si γ est saturé.

2^o R est un graphe du type (2) et les nombres des sites appartenant aux deux classes C_1 et C_2 sont les mêmes. — Il n'y a pas de non-arête. La probabilité associée à un graphe par π_2 est facile à écrire. Envisageons donc un certain tirage τ des couples suivant π_2 ; il a pour probabilité :

$$\frac{1}{N} \binom{1}{\frac{N}{2}} \frac{1}{(N-2)} \frac{1}{\left(\frac{N}{2}-1\right)} \frac{1}{(N-4)} \frac{1}{\left(\frac{N}{2}-2\right)} \cdots \frac{1}{2} = \frac{1}{\left(\frac{N}{2}\right)!} \frac{1}{N(N-2) \dots 2}$$

Parmi les tirages effectués suivant π_2 , combien donnent le même graphe de soutien que τ ?

1^o Tous ceux qui s'en déduisent par une transposition des deux sites à l'intérieur d'un couple;

2^o tous ceux qui s'en déduisent par une permutation des couples.

D'où la formule qui donne la probabilité de tirer suivant π_2 le graphe défini par τ :

$$\frac{\left(\frac{N}{2}\right)! 2^{\frac{N}{2}}}{N(N-2) \dots 2 \left(\frac{N}{2}\right)!} = \frac{1}{\left(\frac{N}{2}\right)!}$$

Ω est donc muni par π_2 d'une probabilité uniforme.

Et suivant π_3 ?

Le tirage des sites pré-nodaux introduit un facteur $\frac{1}{v(v-2) \dots 4 \cdot 2}$.

Quant aux autres sites, il est clair que, après le tirage d'un site de rang quelconque, si l'on ne compte pas ce dernier site tiré, le nombre des sites tirés de la classe C_1 est égal au nombre des sites tirés de la classe C_2 . Comme le

site que l'on va tirer est de la classe à laquelle n'appartient pas le dernier site tiré, la probabilité associée à un tirage suivant π_3 est :

$$\frac{1}{\left(\frac{N}{2}\right)!} \frac{1}{v(v-2) \dots 4 \cdot 2}$$

Pour les mêmes raisons que dans le cas précédent, $2^{\frac{v}{2}} \left(\frac{v}{2}\right)!$ tirages conviennent, et le graphe a bien la probabilité :

$$\frac{1}{\left(\frac{N}{2}\right)!}$$

3° R est un graphe du type (2) mais les nombres des sites appartenant aux classes C_1 et C_2 diffèrent. — Si ces nombres de sites diffèrent, c'est nécessairement d'un nombre pair. En effet, puisque les sites pré-nodaux sont en nombre pair, et que les sites pré-anti-nodaux sont aussi (et par nécessité) en nombre pair, le nombre total des sites est pair. Puisque la somme des cardinaux des deux classes est paire, leur différence est paire (Pour s'en convaincre, considérer le système en nombres entiers $x+y=2p$, $x-y=2q+1$, $2y+2q+1=2p$, impossible).

Nous allons ici encore montrer que l'espace Ω est muni, par les trois procédures, d'une probabilité uniforme en ayant recours à un artifice. On définit ainsi des procédures π'_1 , π'_2 et π'_3 qui s'obtiennent à partir de π_1 , π_2 et π_3 en ajoutant au système des monomères des pré-anti-nœuds fictifs, exclusivement composés de sites d'une espèce fictive, mais appartenant à celle des deux classes dont le cardinal est le plus petit. Le nombre de sites ainsi ajoutés doit rétablir l'égalité entre les cardinaux des deux classes. Sur ce nouvel ensemble de monomères, on procède alors au tirage suivant les procédures π_1 , π_2 ou π_3 . Un graphe de soutien étant ainsi construit, on supprime alors les pré-anti-nœuds fictifs et les arêtes qui y parvenaient en les remplaçant par des non-arêtes. On procède enfin au tirage sans remise des pré-arêtes.

Il est clair que les espaces Ω de définition sont les mêmes. Pour s'en convaincre on pourra toujours considérer un graphe réalisé et imaginer une réalisation de π'_1 , π'_2 , π'_3 ou π_1 , π_2 , π_3 qui y conduise.

Nous allons montrer maintenant que, d'une part π'_1 , π'_2 , π'_3 sont équivalentes entre elles, et qu'elles sont équivalentes à π_1 , π_2 , π_3 . C'est immédiat :

en vertu de l'équivalence entre π_1, π_2, π_3 si les effectifs des deux classes de sites sont les mêmes, π'_1, π'_2 et π'_3 conduisent aux mêmes graphes avec sites fictifs avec la même probabilité. Comme un certain nombre de ces graphes donne le même graphe avec non-arêtes, et que ce nombre est indépendant de la procédure suivie et ne dépend pas même du graphe, π'_1, π'_2 et π'_3 sont équivalentes.

1° L'équivalence entre π_3 et π'_3 est immédiate. Supposons que C_1 soit la classe de sites excédentaire. A tous les tirages effectués suivant π'_3 et qui donnent un tirage τ de π_3 correspond bien une probabilité indépendante de τ .

2° Puisque π_1 est équivalente à π_2 (et π'_1 équivalente à π'_2), il suffit de montrer, par exemple, que π_1 est équivalente à π'_1 . Ici encore, on obtiendra le résultat en montrant que π_1 munit Ω d'une probabilité uniforme. En effet, soit N_1 le cardinal de C_1 , N_2 le cardinal de C_2 et $N'_2 = N_1 - N_2$.

La probabilité associée à un certain tirage τ suivant π_1 est indépendante de τ , et égale à :

$$\frac{2}{N_1 N_2} \frac{2}{(N_1 - 1)(N_2 - 1)} \cdots \frac{2}{(N'_2 + 1)1} \frac{2}{N'_2(N'_2 - 1)} \cdots \frac{2}{2 \cdot 1}$$

Soit :

$$\frac{2^{\binom{N_2 + N'_2}{2}}}{N_2! N_1!}$$

Puisque $N_2! \binom{N'_2}{2}!$ tirages donnent le même graphe, π_1 munit bien Ω d'une probabilité uniforme.

On voit que, dans le cadre de ce raisonnement, tous les sites ayant été distingués, les seuls effectifs d'espèces qui entrent en compte sont ceux des classes, comme il fallait s'y attendre.

II. Si γ n'est pas saturé et compte x arêtes réalisées.

Dans le cas de π_1 et π_2 , la probabilité associée à une réalisation γ_x du graphe est la probabilité pour que les x premières arêtes tirées soient justement les x arêtes de γ_x tandis que les suivantes pourraient être arbitraires. Mais puisque π_1 est équivalente à π_2 , la probabilité pour que les x premières arêtes tirées soient les x arêtes fixées à l'avance est $x!$ fois égale à la pro-

babilité pour que la suite des x premières arêtes tirées soit exactement tirée dans un ordre fixé à l'avance. Si R est du type (1), cette probabilité est :

$$2^x \frac{1}{N} \frac{1}{N-1} \cdots \frac{1}{N-2x+1}$$

Si R est du type (2), cette probabilité est :

$$2^x \frac{1}{N_1} \frac{1}{N_2} \frac{1}{N_1-1} \frac{1}{N_2-1} \cdots \frac{1}{N_1-x+1} \frac{1}{N_2-x+1}$$

Dans les deux cas, elle ne dépend que de x et pas de la suite particulière d'arêtes choisies.

Dans le cas de π_3 , on pourrait commencer par procéder au tirage du graphe de soutien, sans qu'aucune arête soit réalisée, puis tirer sans remise, parmi les pré-arêtes, la proportion fixée à l'avance d'arêtes réalisées, puisque, d'après la description même de π_3 , les deux tirages sont indépendants.

Il revient évidemment au même d'effectuer le tirage du graphe de soutien en supposant toutes les arêtes réalisées, puis de tirer sans remise les pré-arêtes, en nombre exactement complémentaire. Pour tous les graphes de soutien différents les résultats finaux seront, évidemment, équiprobables. Dès lors, la probabilité pour qu'un graphe non saturé soit formé de x arêtes fixées à l'avance sera égale au produit de la probabilité pour que le graphe de soutien contienne ces x arêtes et que les x arêtes réalisées soient justement celles-là. Il est clair que cette probabilité ne dépend que de x et pas des arêtes choisies.

B. — L'équivalence en pré-branches.

Nous utiliserons, plus ou moins explicitement par la suite (cf. chapitre V), une équivalence entre procédures plus faibles que l'équivalence que nous venons d'étudier.

Relation d'équivalence entre pré-branches.

Deux pré-branches sont équivalentes, ou sont « les mêmes », si et seulement si il existe une application de l'ensemble des éléments de la première dans l'ensemble des éléments de la seconde :

1° qui respecte l'ordre de succession des sites et des arêtes quand on chemine d'un site pré-nodal vers l'autre;

2° et telle que les sites et les arêtes en correspondance soient de même espèce et de même type.

Modulo cette relation d'équivalence, l'ensemble des pré-branches distinctes a donc la puissance du dénombrable. L'ensemble des pré-branches d'un graphe aléatoire (Ω, \mathcal{A}, P) est donc isomorphe à un ensemble de suites dont les termes, $x_1, x_2, \dots, x_i, \dots$, sont des nombres entiers presque tous nuls, ensemble muni d'une probabilité θ déduite de P . Puisque le nombre des monomères du graphe est fixé, le nombre des pré-branches et donc la somme des x_i sont fixés.

Dans le cas où le nombre des monomères tend vers l'infini d'une certaine façon, on utilisera les suites :

$$\left(\frac{x_i}{\sum x_i} \right)$$

et on dira que deux procédures sont asymptotiquement équivalentes en pré-branches, si les deux suites $\left(\frac{x_i}{\sum x_i} \right)$ obtenues convergent en probabilité vers la même suite (k_i) quand le nombre des monomères tend vers l'infini d'une certaine façon.

Remarques. — Il est clair que deux procédures strictement équivalentes sont équivalentes en pré-branches. Si de plus, quand le nombre des monomères tend vers l'infini d'une certaine façon, une des deux procédures conduit à une suite $\left(\frac{x_i}{\sum x_i} \right)$ qui converge en probabilité, alors les deux procédures seront asymptotiquement équivalentes en pré-branches.

Ce n'est évidemment que dans le cadre d'une telle équivalence, que parler de l'équivalence entre π_4 ⁽¹⁾ et les autres procédures peut avoir un sens (Dans π_4 , la réalisation des arêtes dépend d'un tirage avec remise).

Si, en particulier, R est un graphe complet avec boucles, nous verrons au chapitre suivant que π_3 est équivalente en pré-branches à π_4 .

Enfin, dans l'idée de ne pas alourdir exagérément cet article, nous ne nous sommes pas intéressés ici au cas où R est orienté, qui serait d'ailleurs inutile pour la suite.

Il est clair que, parmi les pré-branches, les chaînes cycliques jouent un rôle particulier. On a le :

LEMME 4. — Dans le cadre de l'équivalence stricte π_1, π_2, π_3 , et pour un nombre de sites fixé, le graphe pour lequel l'espérance du nombre de boucles formées sera le plus grand, sera engendré par un système de monomères

(1) Dont on trouvera la définition page 280.

exclusivement formés d'un site d'espèce A et d'un site d'espèce B, et un graphe R du type (2).

En effet :

Les boucles ne se forment qu'à partir des pré-anti-nœuds qui restent. Or, à partir de la procédure π_3 , et d'un système exclusivement formé de pré-anti-nœuds, chaque arête fait disparaître deux sites, en fermant ou non une boucle.

Supposons tout d'abord R complet avec boucles.

A chaque tirage d'une nouvelle arête, si le nombre de sites restant est $N - 1$, la probabilité de fermer la boucle en cours est $\frac{1}{N - 1}$. L'espérance du nombre de boucles est donc :

$$\frac{1}{N - 1} + \frac{1}{N - 3} + \dots + \frac{1}{5} + \frac{1}{3} + 1$$

Supposons ensuite R du type (2) et un système dont les espèces appartiennent aux classes A et B. Si les extrémités à refermer sont de même espèce, la probabilité de refermer la boucle est zéro. Si ce sont deux extrémités d'espèces différentes, c'est $\frac{1}{N}$ (s'il reste $N - 1$ sites). L'espérance est $\frac{1}{2}$

donc bornée par :

$$2\left(\frac{1}{N} + \frac{1}{N - 2} + \dots + \frac{1}{4} + \frac{1}{2} + 1\right)$$

Or, en tant qu'espérance ce nombre n'est atteint que par le système exclusivement composé de pré-anti-nœuds formés d'un site dont l'espèce est dans la classe A et d'un site dont l'espèce est dans la classe B.

Enfin, supposons que l'effectif total de la classe A soit supérieur à l'effectif total de la classe B. On utilisera alors la procédure π'_1 dont on sait qu'elle est équivalente arête par arête à π_1 . L'espérance du nombre de boucles obtenues — dans le graphe de soutien — va pouvoir encore être bornée. Dans le cas le plus favorable, en effet, si N_A et N_B désignent les effectifs totaux respectifs des classes A et B, pour maximiser l'espérance du nombre de boucles, on peut supposer des $\frac{N_A - N_B}{2}$ pré-anti-nœuds fictifs, qu'ils commencent par s'accoupler à $\frac{N_A - N_B}{2}$ pré-anti-nœuds A - A. Le

tirage se poursuit alors comme précédemment, mais avec un $\frac{N}{2}$ égal à :

$$N_B + \frac{N_A - N_B}{2} = \frac{N_A + N_B}{2}$$

C'est-à-dire avec $N = N_A + N_B$. On retrouve la formule :

$$2\left(\frac{1}{N} + \frac{1}{N-2} + \dots + \frac{1}{2} + 1\right)$$

d'où le :

THÉORÈME 6. — Quand le nombre des monomères tend vers l'infini, si le nombre des sites pré-nodaux tend vers l'infini avec le même ordre de grandeur, le rapport du nombre de cycles du graphe de soutien au nombre de pré-branches converge en probabilité vers zéro.

Si N_0 est le nombre de pré-anti-nœuds et N_1 le nombre de pré-branches, l'espérance du rapport que nous venons de définir reste bornée par :

$$\frac{\text{Log } N_0 + C + \varepsilon\left(\frac{1}{N_0}\right)}{N_1}$$

où C désigne la constante d'Euler.

La démonstration est immédiate : le rapport considéré est une variable aléatoire strictement positive dont l'espérance converge vers zéro. Elle converge donc vers zéro en probabilité.

IV

LANGAGE EXHAUSTIF ET MARKOV ASSOCIÉ

DÉFINITIONS

Urne U.

Soit une urne U contenant un ensemble S de N boules $\{s_1, s_2, \dots, s_N\}$.

Soit une partition E de S définie par une famille E de parties deux à deux disjointes, de réunion S, notées $E_j, j \in J, J = \{1, 2, \dots, r\}$, et appelées espèces.

L'espèce de la boule s_i , notée $E(s_i)$ est l'espèce dont elle est élément.

Processus associé.

A cette urne, nous associons un processus discret, qui évolue entre r états que nous noterons — sans confusion possible — également E_1, E_2, \dots, E_r . Ce processus sera muni d'une grammaire g , application de l'ensemble des suites d'éléments de \check{E} dans l'ensemble des parties de \check{E} . $\varphi(n)$ désignera la suite des n premiers états du processus associé. On dira qu'une grammaire g est d'ordre m si et seulement si $\forall n$, l'image de $\varphi(n)$ par g dépend au plus des m dernières espèces écrites de $\varphi(n)$. Le processus associé sera dit d'ordre m également.

Probabilités de transition.

Les transitions ont lieu à chaque unité de temps.

Supposons que le système ait évolué jusqu'à l'instant n . Son évolution est décrite par une suite $\varphi(n)$ de n éléments de \check{E} . Soit $\{E_{j_1}, E_{j_2}, \dots, E_{j_p}\}$ l'image de $\varphi(n)$ par g . Alors, parmi les $N_{j_1}, N_{j_2}, \dots, N_{j_p}$ boules d'espèces respectives E_{j_1}, \dots, E_{j_p} restant dans l'urne, je tire *sans remise* une boule au hasard uniforme. Elle est d'espèce E_j . J'écris E_j à la suite de $\varphi(n)$ et j'ai obtenu $\varphi(n+1)$. Cette procédure définit bien de façon unique les probabilités de transition du processus.

Markov associé.

En général, le processus que nous venons de décrire a des probabilités de transition qui ne sont pas indépendantes du passé, jusqu'à l'instant ns les nombres $N_{j_1}, N_{j_2}, \dots, N_{j_p}$ dépendent de tous les éléments de $\varphi(n)$.

Envisageons une grammaire g d'ordre m fini, et un processus d'ordre m associé mais dans lequel les probabilités de transition seraient obtenues au moyen d'un tirage *avec remise*, la procédure étant en tout point identique par ailleurs. Il est clair qu'on aura obtenu ainsi un système de Markov, homogène, stationnaire d'ordre m . C'est le Markov associé.

Espèces morcelantes.

Ce seront des espèces telles que, lorsqu'elles sont écrites à droite de $\varphi(n)$, $g(\varphi(n)) = g(\phi)$ où $g(\phi)$ représente par définition l'ensemble d'espèces parmi lesquelles on tire la boule initiale. Quand on laisse le processus évoluer sans arrêter son évolution, celle-ci finit par s'arrêter au bout d'un temps fini, au plus égal à la somme des cardinaux des espèces. On dit alors que le processus a écrit une phrase ϕ , ϕ désignant la suite des espèces représentant les états pris par le processus tout au long de l'évolution.

Un mot de la phrase ϕ est alors une sous-suite d'espèces consécutives de ϕ débutant soit par la première espèce de ϕ , soit par la première espèce suivant

une espèce morcelante, et se terminant par la première espèce morcelante rencontrée parmi les espèces écrites à droite du début du mot.

Remarque. — Dans la procédure π_3 précédente, la procédure de tirage conduit à une urne U contenant $N + N_1 + N_2 + N_3$ boules :

- N sites répartis entre deux classes d'espèces;
- N_1 arêtes réalisées;
- N_2 pré-arêtes;
- N_3 non-arêtes.

La grammaire est une grammaire d'ordre 2 que nous définirons précisément par la suite.

Les espèces morcelantes sont les sites pré-nodaux. Un mot est une pré-branche. Dans le cadre de l'équivalence π_1, π_2, π_3 , une phrase ne comprend pas les boucles de la fin, car il faudrait changer la grammaire.

Langage exhaustif.

Posons $N_j = |E_j|$ et $N = \sum_j N_j$. On dira que « la taille de l'urne U

tend vers l'infini » quand on considérera une suite d'urnes munies de la même grammaire g , contenant des boules d'espèces semblables, et telles que les N_j tendent vers l'infini tandis que les rapports $\frac{N_j}{N}$ tendent vers des limites finies k_j strictement comprises entre zéro et un.

Un langage exhaustif est défini au moyen d'un triplet $(\overset{\vee}{E}, g, (k_j))$ où $\overset{\vee}{E}$ désigne un ensemble $\{E_1, \dots, E_r\}$;

g désigne une grammaire définie sur l'ensemble des suites d'éléments de $\overset{\vee}{E}$;

k_j est une suite de r nombres réels positifs (strictement) et de somme égale à l'unité.

Étudier un langage exhaustif ce sera étudier les différents mots écrits (et leurs fréquences d'apparition) quand le processus associé à l'urne U munie de la partition E définie par la famille $\overset{\vee}{E}$ de parties et grammaire g se déroule à partir de l'origine des temps, et quand la taille de l'urne tend vers l'infini.

Langage bernoullien associé.

Au moyen du même triplet $(\overset{\vee}{E}, g, k_j)$, étudier le langage bernoullien associé, ce sera étudier les différents mots écrits (et leurs fréquences d'apparition) par le Markov associé, c'est-à-dire quand les tirages sont effectués avec remise, et quand la taille de l'urne U tend vers l'infini.

A. — La grammaire libérale.

La grammaire libérale est celle qui n'interdit aucune transition : l'image par g de $\varphi(n)$ est \check{E} , tout entier, pour tout $\varphi(n)$.

Le processus associé n'est qu'un tirage sans remise d'une urne contenant plusieurs espèces de boules. Le Markov associé est un tirage avec remise de la même urne.

L'évolution du processus pourra toujours avoir lieu jusqu'à l'instant N .

D'intéressants résultats ont été établis par A. M. Mood [9] en utilisant des expressions asymptotiques de la loi hypergéométrique. Exprimons-les dans notre terminologie :

Soit un langage $(\check{E}, g_0, (k_j))$ où

$$\check{E} = \{ E_1, \dots, E_r \};$$

g_0 = la grammaire libérale ;

(k_j) = une suite de nombres réels quelconques.

Dans la phrase écrite par le processus associé, quand le processus venant d'un état E_i arrive dans un état $E_j \neq E_i$, il y reste m unités de temps, puis quitte l'état E_j pour l'état $E_k \neq E_j$; on dit que le processus a écrit une séquence de longueur m et d'espèce E_j .

THÉORÈME DE MOOD N° 1. — Soit t_{im} le nombre de séquences de longueur m et d'espèce E_i écrites dans une phrase et :

$$t_i = \sum_{m=1}^{N_i} t_{im}$$

Quand la taille de l'urne tend vers l'infini, le n -uplet de composantes :

$$x_i \frac{t_i - Nk_i(1 - k_i)}{\sqrt{N}}$$

suit une loi asymptotiquement normale, les variances et co-variances étant données par les formules :

$$\text{var } ij = k_i^2 k_j^2 \qquad \text{var } ii = k_i^2 (1 - k_i)^2$$

Si $r = 2$, le nombre :

$$x_0 = \frac{t - Nk_1 k_2}{\sqrt{Nk_1 k_2}}$$

où t est le nombre total des séquences d'une quelconque des deux espèces, suit asymptotiquement une loi normale de moyenne nulle et de variance unité.

THÉORÈME DE MOOD N° 2. — On obtiendrait la même convergence en loi en procédant à l'écriture d'une phrase de N espèces du Markov associé (suivant un résultat classique par ailleurs pour les Markov).

THÉORÈME DE MOOD N° 3. — τ désigne maintenant le nombre total de séquences $\sum_i t_i$. Dans le cas du processus associé comme dans celui du Markov associé, la variable

$$z = \frac{\tau - N(1 - k_i^2)}{N}$$

converge en loi vers une variable normale de variance

$$\sigma^2 = \Sigma k_i^2 = 2\Sigma k_i^3 + (\Sigma k_i^2)^2.$$

De ces théorèmes on déduit le

THÉORÈME 6. — Toujours dans le cadre du langage $(\check{E}, g_0, (k_j))$, pour le processus associé, comme pour le Markov associé, quand la taille de l'urne U tend vers l'infini, le rapport $\frac{t_i}{\tau}$ converge en probabilité vers une limite égale à :

$$\frac{k_i(1 - k_i)}{1 - \sum_i k_i^2}.$$

Démonstration. — Avec les notations antérieures, pour le processus associé comme pour le processus de Markov associé, on peut écrire :

$$t_i = Nk_i(1 - k_i) + x_i\sqrt{N}$$

$$\tau = N(1 - \Sigma k_i^2) + z\sqrt{N}$$

$$\frac{t_i}{\tau} = \frac{k_i(1 - k_i) + \frac{x_i}{\sqrt{N}}}{1 - \Sigma k_i^2 + \frac{z}{\sqrt{N}}}$$

$\frac{x_i}{\sqrt{N}}$ et $\frac{z}{\sqrt{N}}$ sont des variables aléatoires qui convergent en probabilité vers zéro.

Puisque $\sum k_i = 1 \forall i$, et $k_i > 0 \forall i$, $\sum k_i^2 < 1$.

Dans les deux cas, $\frac{t_i}{\tau}$ converge donc en probabilité vers $\frac{k_i(1 - k_i)}{1 - \sum k_i^2}$.

C'est ce théorème qui nous a donné l'idée des « généralisations » qui vont suivre.

Supposons que l'espèce E_i soit morcelante. Nous appellerons mot de type « i » tout mot se terminant par E_i , et donc ne contenant pas E_i dans ses espèces antérieures.

Soit m_i un mot fixé de type « i ». On a le

THÉORÈME 7. — Quand la taille de l'urne U tend vers l'infini le rapport du nombre de mots m_i au nombre total de mots écrits dans une phrase converge en probabilité vers une quantité qui ne dépend que des k_j , et qui est la même pour le processus associé et pour le Markov associé (On effectuera le calcul, très facile, de cette limite, sur le Markov associé).

Démonstration. — Commençons par éluder une difficulté : si au cours d'une évolution le processus associé reste dans l'état E_i pendant plusieurs instants consécutifs il aura écrit autant de mots composés de la seule espèce E_i .

Puisque l'urne U contient N_i boules d'espèce E_i , chaque phrase φ contiendra N_i mots pouvant être le mot m_i . Soit alors l'ensemble des événements notés A_n : le $n^{\text{ième}}$ mot écrit par le processus au cours de son évolution est le mot m_i (n varie de 1 à N_i).

Soit I_{A_n} une fonction qui prend la valeur 1 sur A_n et la valeur 0 ailleurs. Nous étudions donc le rapport

$$X_{N_i} = \frac{1}{N_i} \sum_{n=1}^{N_i} I_{A_n}.$$

Puisque l'évolution du système est décrite par une permutation quelconque des N boules, avec une probabilité égale à $\frac{1}{N!}$, on en déduit que

$$\frac{1}{N_i} \sum_{n=1}^{N_i} P(A_n) = P(A_1)$$

$$\frac{2}{N_i(N_i - 1)} \sum_{1 \leq j < k \leq N_i} P(A_j \cap A_k) = P(A_1 \cap A_2)$$

puisque en effet, le nombre de permutations des boules qui réalisent A_k , V_k est le même que le nombre de permutations des boules qui réalisent A_1 .

Or : $E[X_{N_i}] = P(A_1)$, E désignant l'espérance mathématique.

$$\begin{aligned} E[(X_{N_i} - EX_{N_i})^2] &= E[(X_{N_i})^2] - (E[X_{N_i}])^2 \\ &= \frac{1}{N_i^2} E \left[\left(\sum_{\mu} I_{A_{\mu}} \right) \left(\sum_{\nu} I_{A_{\nu}} \right) \right] - (P(A_1))^2 \\ &= \frac{1}{N_i^2} E \left[\sum_{\mu \neq \nu} I_{(A_{\mu} \cap A_{\nu})} + \sum_{\mu} I_{A_{\mu}} \right] - (P(A_1))^2 \\ &= \frac{N_i - 1}{N_i} P(A_1 \cap A_2) + \frac{P(A_1)}{N_i} - (P(A_1))^2 \\ &= P(A_1 \cap A_2) - (P(A_1))^2 + \frac{P(A_1) - P(A_1 \cap A_2)}{N_i} \\ &= [P(A_1)][P(A_2/A_1) - P(A_1)] + \frac{P(A_1) - P(A_2 \cap A_2)}{N_i} \end{aligned}$$

Quand la taille de l'urne tend vers l'infini, $P(A_2/A_1) \rightarrow P(A_2) = P(A_1)$ et la variance de X_{N_i} tend vers zéro. X_{N_i} converge en probabilité vers son espérance.

Dans le Markov associé, la théorie des événements récurrents montre classiquement que $X_{N_i} \xrightarrow{P} P(A_1)$. La limite de $P(A_1)$, qui ne dépend elle-même que des k_i sera facile à calculer.

B. — Quelques propriétés des grammaires d'ordre m .

Dans tout ce qui suit, nous supposons toujours que les grammaires étudiées sont d'ordre fini.

DÉFINITIONS

Soit ν le nombre d'instantants pendant lesquels le processus associé à un langage $(E, g, (k_i))$ évolue. Soit N le cardinal de l'ensemble des boules de l'urne U associée.

Vidange.

g est une vidange pour (k_i) si et seulement si $\nu \xrightarrow{P, S} \infty$ quand la taille de l'urne tend vers l'infini. g est une vidange forte si $\frac{\nu}{N} \xrightarrow{P, S} 1$ quand la taille de l'urne tend vers l'infini.

Remarque. — Une vidange forte pourrait aussi s'appeler une grammaire sans blocage.

Graphe représentatif de la grammaire g.

Au Markov associé, on fait correspondre classiquement un graphe orienté. Les sommets du graphe sont les états dans lequel peut se trouver le système (éventuellement désignés par des séquences d'au plus m espèces) et les arcs du graphe représentent les transitions permises.

PROPOSITION. — *La condition nécessaire et suffisante pour qu'un graphe représentatif d'une grammaire soit le graphe représentatif d'une vidange, est qu'il n'ait pas de sommet de demi-degré extérieur nul.*

Démonstration. — La condition suffisante est triviale. Démontrons la condition nécessaire. Supposons qu'il existe dans le graphe un sommet s_0 de demi-degré extérieur nul. Il existerait donc également dans le graphe représentatif de la grammaire un chemin fini, partant d'un sommet correspondant à un état initial, pour parvenir, après être passé par moins de :

$$r \frac{r^m - 1}{r - 1}$$

sommets, au sommet s_0 .

Quand la taille de l'urne U tend vers l'infini, la probabilité pour que le système emprunte un tel chemin au début de son évolution converge vers la probabilité du même événement dans la chaîne de Markov associée. Elle sera donc, pour U assez grande, strictement supérieure à une certaine quantité positive ε . Le processus associé n'évoluera alors pendant une durée supérieure à :

$$r \frac{r^m - 1}{r - 1}$$

unités de temps qu'avec une probabilité strictement inférieure à $1 - \varepsilon$. La grammaire g ne serait donc pas une vidange.

PROPOSITION. — *Seule la grammaire libérale permet de prolonger presque sûrement jusqu'à la dernière boule le tirage d'une urne associée, quelle que soit sa taille et quels que soient les k_i .*

En d'autres termes, c'est la seule vidange forte quels que soient les k_i .

a) Supposons au contraire que la grammaire g interdise un certain mot μ de m lettres. Sans la condition $\text{Card } E_i \neq 0, \forall i$, il suffirait de dire : soit E_k

la première lettre interdite du mot μ . g ne peut donc pas vider une urne qui ne contient que des boules d'espèce E_k et quelques boules formant le début du mot μ .

b) Dans le cadre de la condition $\text{Card } E_i \neq 0$, supposons que g interdise une certaine transition (qui conduirait donc à un mot interdit μ). En écrivant à gauche de μ toutes les espèces, puis à gauche de la séquence obtenue encore toutes les espèces, et ainsi de suite, on finira par écrire une séquence contenant toutes les espèces. Sinon, il existerait une probabilité positive pour ne jamais vider une urne telle que $\text{Card } E_i \neq 0$.

La séquence ainsi trouvée contiendra donc n_{i_0} espèces E_{i_0} . On construit alors une urne qui contient n_{i_0} boules d'espèces E_{i_0} plus une boule dont l'espèce est celle de la transition interdite. Cette urne, avec une probabilité positive, ne pourra être vidée jusqu'au bout.

Produit direct de langages.

Soit un n -uplet $(a_1, \dots, a_l, \dots, a_n)$, où les a_l sont des nombres entiers positifs.

Soient n langages :

$$L_1 = (\check{E}_1, g_1, (k_{i1})) \dots L_l = (\check{E}_l, g_l, (k_{il})) \dots L_n = (\check{E}_n, g_n, (k_{in})).$$

On appellera produit direct L de ces langages un langage (\check{E}, G, K_{i_l}) tel que :

$$1^\circ \check{E} = \bigcup_l \check{E}_l, K_{i_l} = \sum_l a_l k_{il}$$

2° G est telle que, dans φ

— a_1 espèces de \check{E}_1 sont suivies par a_2 espèces de \check{E}_2, \dots, a_n espèces de \check{E}_n par a_1 espèces de \check{E}_1 , et ainsi de suite.

— \check{V}_l , la sous-suite de φ exclusivement composée d'espèces de \check{E}_l respecte la grammaire g_l et est indépendante des autres sous-suites de φ .

Nous verrons dans la suite sur des exemples que ces conditions sont réalisables.

PROPOSITION. — *Un produit direct de langages munis d'une vidange est un langage muni d'une vidange. Un produit direct de langages munis d'une vidange forte est un langage muni d'une vidange forte. En particulier le produit direct d'un langage muni d'une vidange forte et d'un langage muni de la grammaire libérale est un langage muni d'une vidange forte.*

1° Puisque les K_i sont différents de zéro, la première partie de la proposition est évidente.

2° Soit $L_i = (\overset{\vee}{E}_i, g_i, (k_i))$ un des facteurs du produit direct.

Quand la taille de l'urne du produit tend vers l'infini la taille de l'urne qui ne contiendrait que les espèces de $\overset{\vee}{E}_i$ tendrait vers l'infini. Si N désigne le cardinal de l'ensemble des boules de l'urne produit et N_i le cardinal de l'ensemble des seules espèces de E_i , on aurait même :

$$\frac{N_i}{a_i N} \xrightarrow{N \uparrow \infty} 1$$

$$\sum_i a_i$$

$$\left(\text{puisque } \sum_{i_i} \frac{N_i}{K_{i_i} N} \xrightarrow{N \uparrow \infty} 1 \right).$$

Or, $\frac{v_i}{N} \xrightarrow{N_i \uparrow \infty} 1$, v_i désignant le nombre d'instant pendant lesquels le processus associé au langage produit évolue en passant par les espèces de $\overset{\vee}{E}_i$.
Puisque :

$$v = \sum_i v_i, \quad \frac{v}{N} = \frac{\sum_i v_i}{N} \xrightarrow{p} \frac{\sum N_i}{N} \rightarrow \sum_i \frac{a_i}{\sum a_i} = 1.$$

On trouve assez facilement des conditions nécessaires pour qu'une grammaire soit une vidange forte pour un certain n -uple (k_1, \dots, k_n) . On remarque par exemple que seule la structure des groupes finals du Markov associé compte pour que la grammaire soit une vidange forte. Si ces groupes finals sont cycliques, il existera soit un n -uple pour lequel la grammaire sera une vidange forte, soit aucun. Si les groupes finals ne sont pas cycliques nous n'avons pas su trouver de caractérisation satisfaisante.

Heureusement, dans les conditions de l'équivalence entre π_1, π_2, π_3 , il sera très facile de voir que les grammaires dont on a besoin sont des vidanges fortes. Nous voudrions cependant, à titre d'exemple, citer deux problèmes dont la solution pourrait être utile à d'autres problèmes de graphes aléatoires.

1° La grammaire considérée est d'ordre 2. Les espèces, au nombre de deux sont a et b . Le graphe représentatif de la grammaire est le suivant :

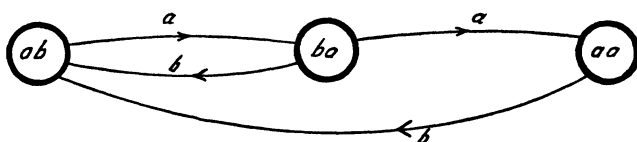


FIG. 9.

Quelles sont les conditions nécessaires et suffisantes sur k_b et k_s pour que cette grammaire soit une vidange forte ?

2° La grammaire considérée est d'ordre 2. Les espèces sont au nombre de deux, a et b . On tire les boules de l'urne deux par deux. Si on tire une boule d'espèce a au moins sur les deux, on les garde hors de l'urne. Sinon, on les remet dans l'urne. A quelle condition la grammaire ainsi définie est-elle une vidange forte ?

Des simulations que nous avons effectuées rendent plausible la conjecture suivante :

Quel que soit r , pour chaque groupe final du Markov associé, les conditions asymptotiques de passages par les différentes transitions doivent être les mêmes. Elles doivent de plus être égales au r -uple des (k_i) .

Une vidange sera morcelante s'il existe des états du Markov associé, atteints au bout d'un temps $p.s.$ fini, et tels que, si e désigne un tel état, $g(e) = g(\phi)$: toute suite d'états compris entre deux tels états sera un mot suivant les conventions de la page 300.

Une vidange morcelante sera stationnaire si et seulement si, pour tout mot m_i :

$$P(A_n^i) = P(A_1^i) \quad \text{et} \quad P(A_{n_1}^i \cap A_{n_2}^i) = P(A_1^i \cap A_2^i)$$

$P(A_n^i)$ (resp. $P(A_{n_1}^i)$) désignant la probabilité pour que le $n^{\text{ième}}$ (resp. $n_1^{\text{ième}}$) mot soit m_i ; n , n_1 et n_2 étant inférieurs au rang du dernier mot écrit.

Nous pouvons maintenant écrire le théorème fondamental dont nous avons besoin pour la suite de ce travail. Il apparaît comme une généralisation aux grammaires morcelantes stationnaires d'ordre m du théorème démontré pour la grammaire libérale. Le mode de démonstration sera identique.

THÉORÈME FONDAMENTAL 8. — Soit un langage $(\check{E}, g, (k_j))$ où

$$\check{E} = \{ E_1, E_2, \dots, E_r \};$$

g = une vidange morcelante stationnaire d'ordre m ;

(k_j) = une suite de r nombres réels strictement positifs et de somme égale à l'unité.

Soit un mot m_i .

Dans ces conditions, quand la taille de l'urne U tend vers l'infini, le rapport du nombre de mots m_i au nombre total de mots écrits dans une phrase converge en probabilité vers une quantité qui ne dépend que des k_i , et qui est la même pour le processus associé et pour la chaîne de Markov associée.

Démonstration. — Supposons que n mots écrits par le processus associé à une urne U et posons :

$$X_n = \frac{1}{n} \sum_{\alpha=1}^n I_{A_\alpha}$$

où I_{A_α} désigne une fonction égale à 1 si le $\alpha^{\text{ième}}$ mot écrit est le mot m_i , égale à zéro dans le cas contraire.

Un calcul en tout point identique à celui du théorème 6 nous conduit aux égalités :

$$\varepsilon[X_n] = P(A_1)$$

$$\varepsilon[(X_n - \varepsilon[X_n])^2] = P(A_1 \cap A_2) - P(A_1)^2 + \frac{P(A_1) - P(A_1 \cap A_2)}{n}$$

où ε désigne l'espérance mathématique, et où A_α désigne ici encore l'événement : le $\alpha^{\text{ième}}$ mot écrit par le processus associé est le mot m_i . Puisque g est une vidange, la durée d'évolution du processus peut être rendue aussi longue qu'on le veut. Puisque par ailleurs écrire un mot ne dure qu'un temps fini *p. s.*, n tend vers l'infini presque sûrement. Et comme il est évident, quand la taille de l'urne tend vers l'infini, que $P(A_1 \cap A_2) - P(A_1)^2$ tend vers zéro, on a le résultat annoncé.

PROPOSITION. — *Le produit direct d'un langage muni d'une vidange morcelante stationnaire d'ordre fini et d'un langage muni d'une autre vidange morcelante stationnaire d'ordre fini, possède également ces propriétés. Ceci reste vrai quand on effectue le produit d'un tel langage et d'un langage muni de la grammaire libérale.*

PROPOSITION. — *Dans un langage exhaustif tel que le temps passé par le Markov associé dans l'État E_i soit asymptotiquement proportionnel à k_i , toute vidange morcelante stationnaire est une vidange forte.*

1° Le Markov associé à ce langage muni d'une vidange morcelante stationnaire possède un seul groupe final qui comprend tous les états permis.

Puisqu'au bout d'un temps fini le processus retourne au départ, le Markov

associé également, et le graphe représentatif de la vidange est fortement connexe. Le seul groupe final est donc composé de l'ensemble des états du Markov associé.

2° Supposons que cette vidange morcelante stationnaire ne soit pas forte. Lorsque la phrase φ s'arrête et que le processus est bloqué dans son évolution, c'est qu'une certaine espèce de boules de l'urne U est vide. Dire que la vidange n'est pas forte c'est dire, si v désigne l'instant où l'évolution du processus s'arrête, que $\exists \eta > 0$, $\exists \alpha > 0$ tels que, pour tout N_0 il existe au moins un N supérieur à N_0 pour lequel on ait :

$$\Pr \left\{ \left| \frac{v}{N} - 1 \right| > \eta \right\} > \alpha$$

Lorsque l'événement $\left| \frac{v}{N} - 1 \right| > \eta$ est réalisé, il reste dans l'urne un nombre de boules supérieur à $[N\eta]$, et la probabilité pour que, pour tout N_0 il existe au moins un N ayant cette propriété est bornée inférieurement par une quantité strictement positive fixée.

Or, en vertu du théorème 8 les proportions des différents mots tirés jusqu'à convergent en probabilité vers les limites des mêmes proportions dans le Markov associé. La même convergence en probabilité est donc *a fortiori* vraie pour les proportions de boules tirées hors de l'urne. On peut donc choisir N_0 pour que les proportions de boules tirées dans φ s'approchent autant qu'on le veut, en probabilité, des proportions du début du tirage. D'où la contradiction : avec une probabilité bornée inférieurement il resterait dans l'urne un nombre de boules supérieur à $[N\eta]$ et dans laquelle les proportions des différentes boules convergeraient en probabilité vers les proportions de départ. Pour N_0 assez grand, la probabilité pour qu'aucune espèce de boules ne soit vide tendrait vers un, et le processus aurait une probabilité strictement positive de poursuivre son évolution.

V

APPLICATION AUX GRAPHS ALÉATOIRES DÉFINIS DANS LES DEUX PREMIERS CHAPITRES

Il nous a semblé souhaitable, en effet, de conclure ce travail par l'exemple d'une application typique, et de donner des détails de calcul suffisamment précis pour les rendre disponibles aux chimistes et permettre facilement leur vérification pour ceux qui souhaiteraient s'en servir.

Deux définitions sont encore indispensables.

1° On dira que deux monomères μ_j et $\mu_{j'}$, sont semblables, ou que ce sont les mêmes monomères si et seulement si, pour tout k , $\mu_{jk} = \mu_{j'k}$ où μ_{jk} (resp. $\mu_{j'k}$) est égal au nombre de sites d'espèce E_k du monomère μ_j (resp. $\mu_{j'}$).

2° On dira que la taille d'un graphe aléatoire tend vers l'infini quand :

a) Le nombre des monomères tend vers l'infini, tandis que les proportions relatives des différents monomères du graphe tendent vers des limites finies.

b) Le nombre des arêtes tend vers l'infini, et son rapport au nombre de sites tend vers une limite finie.

Remarque. — Dans les conditions de l'équivalence π_1, π_2, π_3 , il s'ensuit que les rapports :

$$\frac{\text{nombre de pré-arêtes}}{\text{nombre d'arêtes}}$$

et :

$$\frac{\text{nombre de non-arêtes}}{\text{nombre d'arêtes}}$$

tendent vers des limites finies.

Dans ces conditions :

THÉORÈME 9. — Soit un graphe aléatoire dont la taille peut tendre vers l'infini, les limites finies que nous venons de décrire étant fixées. Si π_1, π_2 et π_3 sont équivalentes et R non orienté :

1° Les proportions des différentes pré-branches du graphe convergent en probabilité vers des limites finies quand la taille du graphe tend vers l'infini.

2° Il existe un langage L où

— L'ensemble des espèces \check{E} est construit à partir de la réunion de l'ensemble des sites et de l'ensemble des arêtes (réalisées, pré-arêtes, non-arêtes).

— La suite (k_i) est déduite des limites finies précitées.

— La grammaire est une vidange forte, morcelante, stationnaire d'ordre 2 et telle que :

a) Il existe une application canonique entre mots de ce langage et pré-branches non cycliques du graphe.

b) Les proportions limites des différentes pré-branches du graphe sont égales aux proportions limites des différents mots.

Démonstration.

I. *R est complet avec boucles.*

La démonstration sera très allégée en supposant qu'il existe une seule espèce de sites, et que R est réduit à une boucle, l'extension au cas de plusieurs espèces ne présentant plus alors que des difficultés d'écriture.

A. Commençons par construire L.

a) \check{E} distinguera 4 espèces, pour lesquelles le calcul des k_i est immédiat :

- 1° l'espèce des sites pré-nodaux, E_1 ;
- 2° l'espèce des sites pré-anti-nodaux, E_2 ;
- 3° l'espèce des pré-arêtes, E_3 ;
- 4° l'espèce des arêtes réalisées, E_4 .

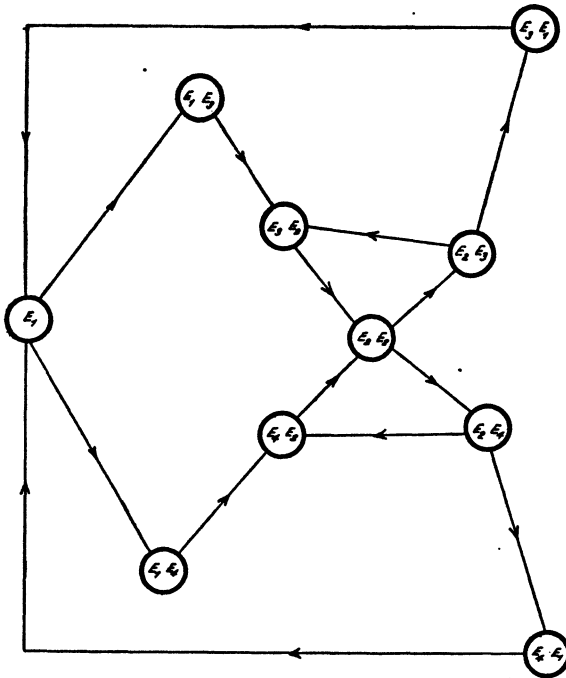


FIG. 10.

b) Pour définir la grammaire g , on peut dessiner son graphe représentatif en remarquant que les états E_3 , E_1 ou E_4 , E_1 sont morcelants.

B. Entre les mots construits suivant cette grammaire et les pré-branches non cycliques du graphe, la correspondance est bi-univoque puisque les

transitions permises sont justement celles définies pour π_3 tant que l'espèce des sites pré-nodaux non encore tirés est non vide, les différents sites à l'intérieur d'une espèce étant indiscernables.

II. R contient deux classes de sommets (C_1) et (C_2).

Ici encore, pour des raisons d'allègement de l'écriture, on supposera que chaque classe ne contient qu'un sommet.

Construisons L .

a) \check{E} distinguera 10 espèces, pour lesquelles le calcul des k_i est immédiat :

1° E_1 = sites pré-nodaux d'espèce appartenant à (C_1);

2° E_2 = sites pré-nodaux d'espèce appartenant à (C_2);

3° E_3 = sites pré-anti-nodaux d'espèce (C_1) appartenant à des anti-nœuds composés d'un site (C_2) et d'un site (C_1);

4° E_4 = sites pré-anti-nodaux d'espèce (C_2) appartenant à des anti-nœuds composés d'un site (C_2) et d'un site (C_1);

5° E_5 = sites pré-anti-nodaux d'espèce (C_1) appartenant à des anti-nœuds composés d'un site (C_1) et d'un site (C_1);

6° E_6 = sites pré-anti-nodaux (fictifs) d'espèce (C_2) appartenant à des anti-nœuds (fictifs) composés d'un site (C_2) et d'un site (C_2);

7° E_7 = sites pré-anti-nodaux d'espèce (C_2) appartenant à des anti-nœuds composés d'un site (C_2) et d'un site (C_2);

8° E_8 = arêtes réalisées;

9° E_9 = pré-arêtes;

10° E_{10} = arêtes fictives.

b) Le graphe représentatif de la grammaire est beaucoup plus compliqué que le précédent (cf. fig. 11).

Le lecteur minutieux identifiera sans peine cette grammaire à la règle de tirage de π'_3 . Ici encore, une proportion fixe de phrases peut être mise en correspondance avec les pré-branches non cycliques du graphe dont le tirage aura été obtenu au moyen de la procédure π'_3 , les différents sites à l'intérieur d'une espèce ayant été considérés comme indiscernables dans notre langage.

III. Dans les deux cas, on vérifie facilement que l'on a obtenu des vidanges morcelantes d'ordre deux, puisque tous les sommets des graphes représentatifs des grammaires sont de demi-degré extérieur supérieur ou égal à un; les états morcelants étant (E_{10} , E_1), (E_9 , E_1), (E_8 , E_1), (E_9 , E_2), (E_8 , E_2), dans le deuxième cas, (E_3 , E_1) et (E_4 , E_1) dans le premier, ils sont atteints au bout d'un temps presque sûrement fini.

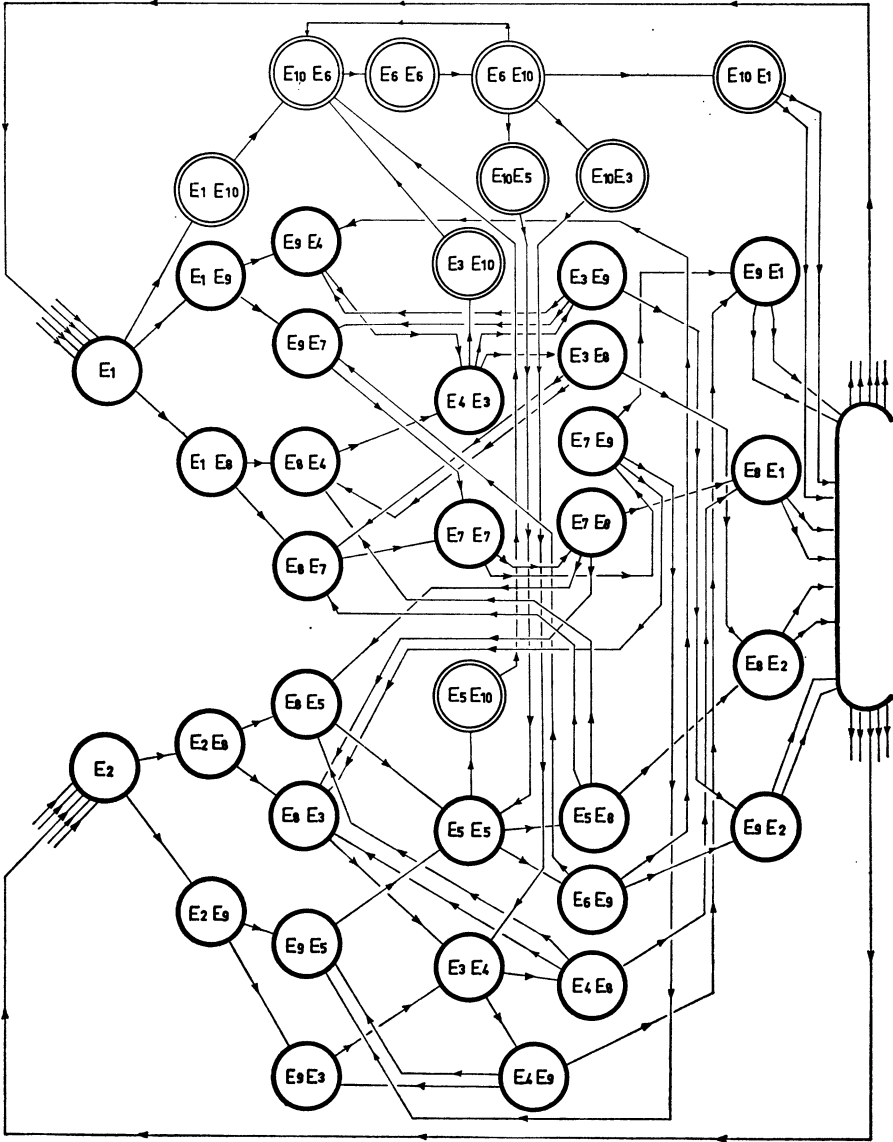


FIG. 11.

Il reste à montrer que ces vidanges morcelantes sont stationnaires, c'est-à-dire que pour tout mot m , et quels que soient n , n_1 et n_2 distincts, désignant les rangs de mots effectivement écrits au moyen des procédures π_3 et π_3 ,

$$1^\circ P(A_1) = P(A_n), \forall n ;$$

$$2^\circ P(A_{n_1} \cap A_{n_2}) = P(A_1 \cap A_2), \forall n_1 \forall n_2 ;$$

Remarquons tout d'abord que les rangs n , n_1 et n_2 ne peuvent poser de problèmes. Puisqu'on se trouve dans le cadre de l'équivalence entre π_1 et π_3 , le nombre des mots écrits au moyen de la procédure est déterminé à l'avance.

Entre les écritures de φ qui réalisent A_1 et celles qui réalisent A_n , on constate facilement l'existence d'une bijection. Il suffit par exemple de mettre en correspondance deux écritures pour lesquelles :

1° Le premier mot de la première est m ;

2° Le $i^{\text{ème}}$ mot de la première est le $(i - 1)^{\text{ème}}$ mot de la seconde, pour $i = 2, 3, \dots, n$;

3° Le $n^{\text{ème}}$ mot de la seconde est m ;

4° A partir du $(n + 1)^{\text{ème}}$ mot, les écritures sont identiques.

Il en est de même pour les écritures de φ qui réalisent A_1 et A_2 d'une part, A_{n_1} et A_{n_2} d'autre part. Puisque les probabilités de réalisation des écritures sont uniformes, et puisque autant d'écritures différentes réalisent l'un ou l'autre événement, les deux événements ont la même probabilité. Ces vidanges morcelantes sont donc stationnaires.

Nous savons que, dans ce cas, quand la taille du graphe, et donc la taille de l'urne du langage L tend vers l'infini, les proportions des différents mots écrits dans la procédure π_3 , et donc les proportions des différentes pré-branches du graphe, convergent en probabilité vers les proportions limites que donnerait le Markov associé au langage L . Puisque les probabilités associées aux différentes transitions sont des produits et des rapports de nombres réels finis et différents de zéro, les proportions limites des différentes pré-branches sont donc finies.

Le théorème est donc démontré dans le cas où R est complet et les sites du graphe sont d'une seule espèce, où R n'est pas complet et la classe d'espèces (C_1) comme la classe d'espèces (C_2) ne contiennent qu'un élément.

Dans le premier cas comme dans le second, on remarquera qu'en ne distinguant pas les espèces du premier ni celles du second dès qu'elles appartiennent à la même classe, on ne faisait que multiplier par une constante fixe le nombre de façons d'écrire les graphes. Mais nous savons que la propriété essentielle d'équi-probabilité des réalisations est toujours vraie. Un raisonnement identique pourra donc s'appliquer.

COROLLAIRE. — *Le rapport du nombre de pré-anti-nœuds appartenant à des cycles élémentaires sans sites pré-nodaux au nombre total de sites converge en probabilité vers zéro quand la taille des graphes tend vers l'infini.*

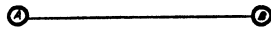
En effet, π_3 (resp. π_3) vidange morcelante stationnaire satisfaisant aux conditions de la proposition est donc une vidange forte. Le nombre des sites et des arêtes n'appartenant pas à des pré-branches est, en probabilité, un infiniment petit par rapport à N.

Remarque. — Si R est un graphe complet avec boucles, on peut encore remarquer que π_3 est le produit direct d'un langage muni d'une vidange morcelante stationnaire d'ordre deux et d'un langage muni d'une grammaire libérale (et à deux espèces : les arêtes réalisées et les pré-arêtes).

Dans le cadre de l'équivalence π_1, π_2, π_3 , il nous reste donc à calculer, en fonction des effectifs des différents types de monomères, les proportions asymptotiques des différentes pré-branches. C. M. Bruneau avait conjecturé le résultat et effectué les calculs complets dans les cas suivants (2) :

1° Les monomères contiennent une seule espèce de sites. R est réduit à une boucle;

2° Les monomères contiennent des sites de deux espèces, A et B. R est le graphe le plus simple du type (2) :



a) les pré-anti-nœuds sont tous composés d'un site d'espèce A et d'un site d'espèce B;

b) les pré-anti-nœuds sont tous composés de deux sites d'espèce A ou de deux sites d'espèce B.

Nous avons effectué le calcul dans le cas qui manquait :

c) les pré-anti-nœuds peuvent avoir les trois compositions.

Le chimiste, en pratique, n'a guère besoin que de ces résultats. Que R soit du type (1) ou du type (2), il serait en effet très facile d'étendre les résultats aux cas où les espèces seraient plus nombreuses. Nous allons donc traiter maintenant du cas 2°, c) qui redonnera, en particulier, les proportions des formules obtenues par C. M. Bruneau pour les cas 2°, a) et 2°, b).

Pour prendre une terminologie voisine de celle de C. M. Bruneau, nous appellerons E_1 et E_2 les deux espèces de sites. r désignera la limite finie du rapport

$$\frac{\text{card. } E_1}{\text{card. } E_2},$$

et sera appelé rapport stœchiométrique. On dira qu'un ensemble de monomères est stœchiométrique si et seulement si $r = 1$.

Pour fixer les idées, dans le cas non stœchiométrique, nous supposerons $r > 1$.

α_1 désignera le rapport limite du nombre de sites pré-nodaux d'espèce E_1 au nombre total de sites d'espèce E_1 .

α_2 désignera le rapport limite du nombre de sites pré-nodaux d'espèce E_2 au nombre total de sites d'espèce E_2 .

Les pré-nœuds auront une composition en sites quelconque. p_1 et p_2 désigneront les proportions limites respectives de sites d'espèce E_1 et d'espèce E_2 extrémités d'une arête réalisée. On aura nécessairement $rp_1 = p_2$.

On appellera β_1 (resp. β_2) le rapport limite du nombre de sites d'espèce E_1 (resp. E_2) situés sur des pré-anti-nœuds composés d'un site de l'espèce E_1 et d'un site de l'espèce E_2 , au nombre total de sites d'espèce E_1 (resp. E_2).

On appellera γ_1 (resp. γ_2) le rapport limite du nombre de sites d'espèce E_1 (resp. E_2) situés sur des pré-anti-nœuds composés de sites de même espèce, au nombre total de sites d'espèce E_1 (resp. E_2).

On appellera pré-branche (i, j) une pré-branche qui commence par un site pré-nodal d'espèce E_1 et se termine par un site pré-nodal d'espèce E_j .

ÉTUDE DES PROPORTIONS LIMITES POUR LES DIFFÉRENTES PRÉ-BRANCHES DANS LE CAS D'UN ENSEMBLE DE MONOMÈRES STOECHIMÉTRIQUE

Dans un tel système, on a évidemment $\beta_1 = \beta_2$. On posera $\beta_1 = \beta_2 = \gamma$. De même $p_1 = p_2 = p$.

Pré-branches (1, 1) pour lesquelles toutes les arêtes sont réalisées.

I. *Pré-branches de ce type contenant un seul pré-anti-nœud.* — (La proportion limite de pré-branches (1, 1) pour lesquelles toutes les arêtes sont réalisées et ne comportant pas de pré-anti-nœud, est évidemment nulle) :

$$p^2 \frac{\alpha_1^2}{\alpha + \alpha_2} \gamma_2$$

II. *Pré-branches (1, 1) comportant 2 pré-anti-nœuds.*

$$p^3 \frac{\alpha_1^2}{\alpha_1 + \alpha_2} (\gamma\gamma_2 + \gamma_2\gamma)$$

III. *Pré-branches (1, 1) comportant 3 pré-anti-nœuds.*

$$p^4 \frac{\alpha_1^2}{\alpha_1 + \alpha_2} (\gamma\gamma\gamma_2 + \gamma\gamma_2\gamma + \gamma_2\gamma\gamma + \gamma_2\gamma_1\gamma_2)$$

IV. *Pré-branches (1, 1) comportant 2n pré-anti-nœuds.*

$$\begin{aligned} p^{2n+1} \frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma_2 [A_{2n-1}^2 \gamma^{2n-1} + A_{2n-3}^4 \gamma^{2n-3} \gamma_1 \gamma_2 + \dots \\ + A_{2(n-i)-1}^{2i+2} \gamma^{2(n-i)-1} (\gamma_1 \gamma_2)^i \dots A_1^{2n} \gamma (\gamma_1 \gamma_2)^{n-1}] \\ = p^{2n+1} \frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma_2 \sum_{i=0}^{n-1} A_{2(n-i)-1}^{2i+2} \gamma^{2(n-i)-1} (\gamma_1 \gamma_2)^i \end{aligned}$$

où A_p^q désigne le nombre de façons de placer p objets indiscernables dans q cases discernables.

Ici $A_{2(n-i)-1}^{2i+2}$ désigne le nombre de façons de placer les $2(n-i) - 1$ pré-anti-nœuds γ le long de la chaîne formée par les i pré-anti-nœuds γ_1 et les $(i + 1)$ pré-anti-nœuds γ_2 successifs, qui déterminent entre eux $2i + 2$ cases distinctes.

D'après un résultat classique de théorie combinatoire (1)

$$A_{2(n-i)-1}^{2i+2} = \binom{2(n-i)-1+2i+2-1}{2i+2-1} = \binom{2n}{2i+1}$$

V. *Pré-branches (1, 1) comportant 2n + 1 pré-anti-nœuds.*

$$\begin{aligned} p^{2n+2} \frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma_2 [A_{2n}^2 \gamma^{2n} + A_{2n-2}^4 \gamma^{2n-2} \gamma_1 \gamma_2 + \dots \\ + A_{2(n-i)}^{2(i+1)} \gamma^{2(n-i)} (\gamma_1 \gamma_2)^i + \dots + A_0^{2n} (\gamma_1 \gamma_2)^n] \end{aligned}$$

avec $A_0^{2n} = 1$ par convention.

$$p^{2n+2} \frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma_2 \sum_{i=0}^n A_{2(n-i)}^{2(i+1)} \gamma^{2(n-i)} (\gamma_1 \gamma_2)^i$$

avec $A_{2(n-i)}^{2(i+1)} = \binom{2n+1}{2i+1}$

Pour obtenir la proportion de pré-branches (1, 1) comportant un nombre quelconque de pré-anti-nœuds on pourra utiliser une certaine façon d'écrire la somme des probabilités que nous venons de calculer en intervertissant les

(1) Cf. par exemple [4], vol. I-II-4, p. 36.

sommations, ce qui est possible parce que toutes les séries sommées sont absolument convergentes :

La proportion limite étudiée peut s'écrire

$$\frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma^2 \left\{ \begin{array}{l} p^2 \\ + p^3 2\gamma \\ + p^4 [\binom{3}{2} \gamma^2 + \binom{3}{0} \gamma_1 \gamma_2] \\ + p^5 [\binom{4}{3} \gamma^3 + \binom{4}{1} \gamma \gamma_1 \gamma_2] \\ + p^6 [\binom{5}{4} \gamma^4 + \binom{5}{2} \gamma^2 \gamma_1 \gamma_2 + \binom{5}{0} (\gamma_1 \gamma_2)^2] \\ + \dots \end{array} \right.$$

On peut effectuer séparément la sommation des colonnes (Dans la mesure où γ est < 1 , ce qui est vérifié dans notre système).

Étudions par exemple la somme des termes de la première colonne; après avoir remarqué que :

$$\binom{k+1}{k} = \frac{(k+1)!}{k!1!} = k+1,$$

on voit en mettant p^2 en facteur, qu'il reste la série

$$1 + 2\gamma p + 3(\gamma p)^2 + 4(p\gamma)^3 + \dots (n+1)(\gamma p)^n + \dots$$

dont la somme est :

$$\frac{1}{(1 - \gamma p)^2}$$

De même, pour la deuxième colonne, après avoir remarqué que :

$$\binom{k+3}{k} = (k+3)(k+2)(k+1)$$

il est facile d'en déduire, en mettant p^4 en facteur, que la série a pour somme :

$$\frac{1}{(1 - p\gamma)^4}$$

On démontrerait évidemment par récurrence que la somme des termes de la $n^{\text{ième}}$ colonne est égale à :

$$\frac{p^{2n}}{(1 - p\gamma)^{2n}} (\gamma_1 \gamma_2)^{n-1}$$

Il nous reste donc à trouver la somme de la série

$$\frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma_2 \frac{p^2}{(1 - p\gamma)^2} \left[\sum_{n=0}^{\infty} \frac{p^{2n}}{(1 - p\gamma)^{2n}} (\gamma_1 \gamma_2)^n \right]$$

Il est facile de montrer que cette série est absolument convergente : en effet, la définition de α_1, γ_1 et γ d'une part, de α_2, γ_2 et γ d'autre part, entraîne les égalités :

$$\alpha_1 + \gamma_1 + \gamma = 1 \Rightarrow \gamma_1 = 1 - \alpha_1 - \gamma$$

$$\alpha_2 + \gamma_2 + \gamma = 1 \Rightarrow \gamma_2 = 1 - \alpha_2 - \gamma$$

et comme $\alpha_1, \alpha_2, \gamma_1, \gamma_2$ et γ sont compris entre 0 et 1 dans les conditions d'application des théorèmes, on en tire les inégalités :

$$\gamma_1 < 1 - \gamma$$

$$\gamma_2 < 1 - \gamma$$

d'où

$$\gamma_1 \gamma_2 < (1 - \gamma)^2, \quad p^2 \gamma_1 \gamma_2 < (p - p\gamma)^2 < (1 - p\gamma)^2.$$

La raison de la série géométrique étudiée est donc strictement inférieure à 1 et la série absolument convergente. On peut donc légitimement calculer sa somme par la méthode usuelle.

Finalement, la porportion cherchée de pré-branches (1, 1) est égale à :

$$\frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma_2 \frac{p^2}{(1 - p\gamma)^2} \frac{1}{1 - \frac{p^2 \gamma_1 \gamma_2}{(1 - p\gamma)^2}} = \frac{\alpha_1^2 \gamma_2}{\alpha_1 + \alpha_2} \frac{p^2}{(1 - p\gamma)^2 - p^2 \gamma_1 \gamma_2}$$

La proportion des pré-branches (2, 2) comptant un nombre quelconque de pré-anti-nœuds s'en déduit immédiatement en changeant α_1 en α_2 et γ_2 en γ_1 . Elle est donc égale à :

$$\frac{\alpha_2^2 \gamma_1}{\alpha_1 + \alpha_2} \frac{p^2}{(1 - p\gamma)^2 - p^2 \gamma_1 \gamma_2}$$

Nous allons maintenant nous intéresser aux pré-branches commençant par un site d'espèce E_1 et finissant par un site d'espèce E_2 . En utilisant les procédés de calcul déjà indiqués pour les branches (1, 1) on obtient :

I. *Pré-branches (1, 2) comportant 0 pré-anti-nœud.*

$$\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p$$

II. *Pré-branches (1, 2) comportant 1 pré-anti-nœud.*

$$\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p^2 \gamma$$

III. *Pré-branches (1, 2) comportant 2 pré-anti-nœuds.*

$$\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p^3 (\gamma \gamma + \gamma_2 \gamma_1)$$

IV. *Pré-branches (1, 2) comportant 3 pré-anti-nœuds.*

$$\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p^4 (\gamma \gamma \gamma + \gamma \gamma_2 \gamma_1 + \gamma_2 \gamma \gamma_1 + \gamma_2 \gamma_1 \gamma)$$

V. *Pré-branches (1, 2) comportant 2n pré-anti-nœuds.*

$$\begin{aligned} & \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p^{2n+1} [\gamma^{2n} + A_{2n-2} \gamma^{2n-2} \gamma_1 \gamma_2 + \dots + A_{2(n-i)}^{2i+1} \gamma^{2(n-i)} (\gamma_1 \gamma_2)^i + \dots \\ & \qquad \qquad \qquad \dots + (\gamma_1 \gamma_2)^n] \\ & = \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p^{2n+1} \sum_{i=0}^n A_{2(n-i)}^{2i+1} \gamma^{2(n-i)} (\gamma_1 \gamma_2)^i \end{aligned}$$

avec $A_0^{2n+1} = 1$ par convention et $A_{2(n-i)}^{2i+1} = \binom{2n}{2i}$

VI. *Pré-branches (1, 2) comportant 2n + 1 pré-anti-nœuds.*

$$\begin{aligned} & \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p^{2n+2} [\gamma^{2n} + A_{2n-1}^3 \gamma^{2n-2} \gamma_1 \gamma_2 + \dots + A_{2(n-i)+1}^{2i+1} \gamma^{2(n-i)} (\gamma_1 \gamma_2)^i \dots \\ & \qquad \qquad \qquad \dots + A_1^{2n+1} (\gamma_1 \gamma_2)^n] \\ & = \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p^{2n+2} \sum_{i=0}^n A_{2(n-i)+1}^{2i+1} \gamma^{2(n-i)} (\gamma_1 \gamma_2)^i \end{aligned}$$

avec $A_{2(n-i)+1}^{2i+1} = \binom{2n+1}{2i}$

Pour obtenir la proportion des pré-branches (1, 2) sans pré-arêtes et ayant

un nombre quelconque de pré-anti-nœuds, on peut par exemple employer la sommation suivant les colonnes du triangle suivant :

$$\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p \left\{ \begin{array}{l} 1 \\ + p\gamma \\ + p^2[\gamma^2 + \gamma_1\gamma_2] \\ + p^3[\gamma^3 + \binom{3}{1}\gamma(\gamma_1\gamma_2)] \\ + p^4[\gamma^4 + \binom{4}{2}\gamma^2(\gamma_1\gamma_2) + (\gamma_1\gamma_2)^2] \\ + p^5[\gamma^5 + \binom{5}{3}\gamma^3(\gamma_1\gamma_2) + \binom{5}{1}(\gamma_1\gamma_2)^2] \\ + p^6[\gamma^6 + \binom{6}{4}\gamma^4(\gamma_1\gamma_2) + \binom{6}{2}^2(\gamma_1\gamma_2)^2 + (\gamma_1\gamma_2)^3] \\ + \dots \end{array} \right.$$

La somme des termes de la première colonne est égale à $\frac{1}{1 - p\gamma}$.

La somme des termes de la seconde colonne est égale à :

$$\frac{\gamma_1\gamma_2}{2!} \frac{2!}{(1 - p\gamma)^3}$$

La somme des termes de la troisième colonne est égale à :

$$\frac{(\gamma_1\gamma_2)^2}{4!} \frac{4!}{(1 - p\gamma)^5}$$

Il faut finalement calculer, si elle est absolument convergente, la somme de la série :

$$\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} \left[\frac{p}{1 - p\gamma} + \gamma_1\gamma_2 \frac{p^3}{(1 - p\gamma)^3} + (\gamma_1\gamma_2)^2 \frac{p^5}{(1 - p\gamma)^5} + \dots \right. \\ \left. \dots + (\gamma_1\gamma_2)^i \frac{p^{2i+1}}{(1 - p\gamma)^{2i+1}} + \dots \right]$$

soit :

$$\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} \frac{p}{1 - p\gamma} \sum_0^{\infty} (\gamma_1\gamma_2)^i \frac{p^{2i}}{(1 - p\gamma)^{2i}}$$

Nous avons déjà vérifié plus haut la convergence absolue de cette série dont la somme est :

$$\frac{1}{1 - \frac{p^2 \gamma_1 \gamma_2}{(1 - p\gamma)^2}}$$

La proportion limite cherchée s'écrit donc :

$$\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} \frac{p(1 - p\gamma)}{(1 - p\gamma)^2 - p^2(\gamma_1 \gamma_2)}$$

La proportion limite des pré-branches sans pré-arêtes est donc :

$$\frac{1}{(\alpha_1 + \alpha_2)[(1 - p\gamma)^2 - p^2\gamma_1\gamma_2]} [p^2(\alpha_1^2\gamma_2 + \alpha_2^2\gamma_1) + 2p(1 - p\gamma)\alpha_1\alpha_2]$$

On vérifie que, si $p = 1$, cette proportion est bien égale à 1. Elle est en effet littéralement égale à :

$$\frac{2(1 - \gamma)\alpha_1\alpha_2 + \gamma_2\alpha_1^2 + \gamma_1\alpha_2^2}{(\alpha_1 + \alpha_2)[(1 - \gamma)^2 - \gamma_1\gamma_2]}$$

pour faire apparaître $\alpha_1 + \alpha_2$ au numérateur, on écrit que :

$$\gamma_2 = (1 - \gamma) - \alpha_2 \quad \text{et} \quad \gamma_1 = (1 - \gamma) - \alpha_1.$$

Il vient alors :

$$\frac{(1 - \gamma)(\alpha_1 + \alpha_2)^2 - (\alpha_1 + \alpha_2)\alpha_1\alpha_2}{(\alpha_1 + \alpha_2)[(1 - \gamma)^2 - \gamma_1\gamma_2]} = \frac{(1 - \gamma)(\alpha_1 + \alpha_2) - \alpha_1\alpha_2}{(1 - \gamma)^2 - \gamma_1\gamma_2}$$

On développe alors encore une fois le numérateur en écrivant $(\alpha_1 + \alpha_2)$ sous la forme $2(1 - \gamma) - (\gamma_1 + \gamma_2)$.

On obtient :

$$2(1 - \gamma)^2 - (\gamma_1 + \gamma_2)(1 - \gamma) - (1 - \gamma^2) - \gamma_1\gamma_2 + (\gamma_1 + \gamma_2)(1 - \gamma)$$

qui se simplifie immédiatement et donne le résultat cherché.

CHERCHONS MAINTENANT LES PROPORTIONS LIMITES DES PRÉ-BRANCHES COMPTANT c PRÉ-ARÊTES AVEC $c \geq 1$

Ces proportions se déduisent des proportions de pré-branches sans pré-arêtes et comptant le nombre de n de pré-anti-nœuds par multiplication par le facteur :

$$\frac{(1 - p)^c}{p^c} \binom{n+1}{c}$$

On voit que l'on peut obtenir facilement de cette façon la composition des

différentes sortes de pré-branches du graphe obtenu à partir d'un ensemble stœchiométrique de monomères.

Pour avoir, en effet, par exemple la proportion de pré-branches (1, 1) comportant $2n$ pré-anti-nœuds indépendamment de son nombre de pré-arêtes, il faudra multiplier :

$$\frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma_2 \sum_{i=0}^{n-1} A_{2(n-i)-1}^{2i+2} \gamma^{2(n-i)-1} (\gamma_1 \gamma_2)^i$$

par un facteur qui ne dépend que de p et de $2n$:

$$p^{2n+1} \binom{2n+1}{0} + p^{2n} (1-p) \binom{2n+1}{1} + \dots + p^{2n-j+1} (1-p)^j \binom{2n+1}{j} \dots (1-p)^{2n+1} \binom{2n+1}{2n+1} \\ = ((1-p) + p)^{2n+1} = 1$$

En conséquence, quand on calcule, d'après la somme des résultats partiels, la proportion de pré-branches comptant un nombre quelconque de pré-arêtes, on trouve forcément 1, comme plus haut, puisque le résultat, pour chaque proportion limite de branches comportant n pré-anti-nœuds peut être formellement obtenu en égalant p à l'unité.

ÉTUDE DES PROPORTIONS LIMITES POUR LES DIFFÉRENTES PRÉ-BRANCHES DANS LE CAS D'UN ENSEMBLE DE MONOMÈRES NON STŒCHIOMÉTRIQUE

Rappelons que :

$$\frac{\text{card } E_1}{\text{card } E_2}$$

est choisi conventionnellement supérieur à l'unité. Pour mener à bien le calcul, on utilisera π_3 en ajoutant au système des pré-anti-nœuds fictifs jusqu'à ce qu'il devienne stœchiométrique. Ces pré-anti-nœuds sont composés de deux sites d'espèce E_2 . Cette transformation effectuée, soit donc :

α'_2 le nouveau rapport limite du nombre de sites pré-nodaux d'espèce E_2 au nombre total de sites d'espèce E_2 ;

γ'_2 le nouveau rapport limite du nombre de sites d'espèce E_2 appartenant à des pré-anti-nœuds formé de deux sites d'espèce E_2 au nombre total de sites d'espèce E_2 ;

$\gamma' = \beta_1 = \beta'_1$ où β'_1 désigne le nouveau rapport du nombre de sites d'espèce E_2 situés sur des pré-anti-nœuds formés de sites d'espèces différentes au nombre total de sites d'espèce E_2 .

On a les relations :

$$\gamma' + \gamma'_2 + \alpha'_2 = 1$$

$\frac{\alpha'_2}{\alpha_2}$ = rapport limite de l'ancien nombre de sites d'espèces E_2 au nouveau

$$\beta'_1 < \beta_1 \quad , \quad \gamma'_2 < \gamma_2$$

On commence donc par calculer les proportions limites obtenues dans le nouveau système de monomères suivant les formules établies en posant $p = 1$ et en changeant γ en γ' , γ_2 en γ'_2 , α_2 en α'_2 . Pour obtenir les proportions limites des pré-branches du graphe « non stœchiométrique » :

1° On retirera tout d'abord les pré-anti-nœuds fictifs;

2° On ajoutera la distinction entre arêtes et pré-arêtes.

Pré-branches (1, 1), sans pré-anti-nœud, avec une non-arête.

$$\mu \frac{\alpha_1^2}{\alpha_1 + \alpha_2} \quad \text{en posant} \quad \mu = \frac{\alpha_2 \gamma'_2 - \alpha'_2 \gamma_2}{\alpha_2 \gamma'_2}$$

Pré-branches (1, 1), comptant $2n$ pré-anti-nœuds, et une non-arête.

$$\frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma'_2 \sum_{i=0}^{n-1} A_{2(n-i)-1}^{2i+2} \gamma'^{[2(n-i)-1]} (\gamma_1 \gamma'_2)^i \mu (1 - \mu)^i (i + 1) \quad (1)$$

Pré-branches (1, 1), comptant $2n+1$ pré-anti-nœuds et une non-arête.

$$\frac{\alpha_1^2}{\alpha_1 + \alpha_2} \gamma'_2 \sum_{i=0}^n A_{2(n-i)}^{2(i+1)} \gamma'^{2(n-i)} (\gamma_1 \gamma'_2)^i \mu (1 - \mu)^i (i + 1) \quad (2)$$

Dans de telles pré-branches, on voit que, respectivement, $2n - 1$ et $2n$ arêtes sont réalisables. En conséquence, pour avoir la proportion limite des pré-branches (1, 1) comptant $2n$ pré-anti-nœuds, m pré-arêtes et une non-arête, il suffit de multiplier l'expression (1) par le facteur

$$\binom{2n+1-2}{m} p^{2n+1-2-m} (1-p)^m = \binom{2n-1}{m} p^{2n-1-m} (1-p)^m$$

On retrouve évidemment ici les termes du binôme

$$(p + (1-p))^{2n-1}$$

— pour avoir la proportion limite des pré-branches (1, 1) comptant

$2n + 1$ pré-anti-nœuds, m pré-arêtes et une non-arête, il suffit de multiplier l'expression (2) par le facteur :

$$\binom{2n+2+1-2}{m} p^{2n+2+1-m} (1-p)^m = \binom{2n+1}{m} p^{2n+1-m} (1-p)^m,$$

terme de rang m du binôme

$$(p + (1 - p))^{2n+1}$$

Pour obtenir d'une façon générale, les proportions des pré-branches (1,1) comptant k pré-anti-nœuds, m pré-arêtes et l non-arêtes, on pourra utiliser de la même façon les formules qui se compliquent « simplement ». En appelant la proportion limite ainsi définie

$$P_{k,l,m}^{(1,1)}$$

on aura par exemple les formules :

$$P_{2n,l,m}^{(1,1)} = \frac{\alpha_1^2}{\alpha_1 + \alpha_2} p^{(2n-m+1)} (1-p)^m \binom{2n}{m} \gamma_2' \sum_{i=-1}^{n-1} A_{2(n-i)-1}^{2(i+1)} \gamma'^{2(n-i)-1} (\gamma_1 \gamma_2')^i \mu^l (1-\mu)^{i-l} \binom{i+1}{l}$$

et

$$P_{2n+1,l,m}^{(1,1)} = \frac{\alpha_1^2}{\alpha_1 + \alpha_2} p^{(2n-m+2)} (1-p)^m \binom{2n}{m} \gamma_2' \sum_{i=l-1}^n A_{2(n-i)}^{2(i+1)} \gamma'^{2(n-i)} (\gamma_1 \gamma_2')^i \mu^l (1-\mu)^{i-l} \binom{i+1}{l}$$

La même méthode conduit à des expressions similaires pour les

$$P_{k,l,m}^{(2,2)} \quad \text{et les} \quad P_{k,l,m}^{(1,2)}$$

CONCLUSION

L'utilité de ces calculs pour les chimistes vient de ce que l'on peut, dans certains « processus de polycondensation » assimiler une partie des molécules aux pré-nœuds, et l'autre partie aux pré-anti-nœuds d'un des graphes aléatoires que nous venons d'étudier.

Voici un exemple. Quand, à partir d'un mélange liquide de molécules

n'ayant pas réagi, on favorise les réactions chimiques de façon que le maximum de liaisons chimiques soit réalisé, le corps peut prendre alors un aspect solide. Il est gélifié. Mais quand les sites réactifs portés par les molécules de départ ne forment pas un ensemble stœchiométrique, il reste toujours dans le « gel » une certaine proportion de molécules solubles en chaînes.

C. M. Bruneau émet l'hypothèse suivante : dans les graphes que nous venons de décrire, il suppose que la proportion de molécules solubles est formée des suites de pré-anti-nœuds reliés par des arêtes mais situées entre deux non-arêtes. Cette hypothèse permet de calculer la distribution des différents types de chaînes de molécules de la « fraction soluble », la prévision de leur masse moléculaire moyenne, de leurs longueurs, etc. Ajoutons que C. M. Bruneau a obtenu un très bon accord entre théorie et expérimentation.

Du point de vue mathématique, nous pensons que de nombreuses questions, au sujet de ces graphes, restent à résoudre. *A propos des vidanges et des procédures*, citons par exemple :

— Comment caractériser l'issue de tirages exhaustifs avec une contrainte donnée par un graphe R ?

— Si π_1 , π_2 et π_3 ne sont plus équivalentes, peuvent-elles néanmoins conduire à des graphes équivalents en pré-branches ? A quelles conditions ?

— Quelles sont les conditions nécessaires et suffisantes pour qu'une vidange soit une vidange forte ?

— Que dire si les couples à partir desquels on effectue la procédure π_1 ne sont plus équiprobables ? π_3 reste-t-elle équivalente ?

— Que dire en général des vidanges non stationnaires ?

Et à propos des graphes proprement dits ?

— Que peut-on dire de leur nombre cyclomatique ?

— Que peut-on dire de leur connexité et de son évolution au fur et à mesure que l'on ajoute des arêtes ?

— Peut-on calculer sur ces graphes des probabilités exactes de percolation ?

BIBLIOGRAPHIE

[1] C. BERGE, *Théorie des graphes et ses applications*. Dunod.

[2] C. M. BRUNEAU, Description statistique de graphes aléatoires pouvant servir de modèle à des processus « polyagrégation ». *Revue Française du Traitement de l'Information. Chiffres*, t. 9, n° 3, 1966, p. 201.

- [3] C. M. BRUNEAU, Théorie des graphes stochastiques appliquée à la synthèse et à la dégradation aléatoire des composés moléculaires multifonctionnels. Thèse de Doctorat ès Sciences Physiques, Paris, 1965. *Annales de Chimie*, n° 7, 1966.
- [4] FELLER, *An introduction to probability theory and its applications*. Second edition J. Wiley and Sons.
- [5] E. N. GILBERT, Random Graphs. *Ann. Math. Statist.*, t. 30, 1959, p. 1141-1144.
- [6] E. N. GILBERT, Random minimal trees. *J. Soc. Indust. Appl. Math.*, vol. 13, n° 2, June 1965.
- [7] J. M. HAMMERSLEY et H. L. FRISCH, Percolation processes and related topics. *J. Soc. Indust. Appl. Math.*, vol. 11, n° 4, décembre 1963.
- [8] M. LOEVE, *Probability Theory* (3^e édition). N. Y., 1963. Van Nostrand.
- [9] A. M. MOOD, The distribution theory of runs. *Ann. of Math. Stat.* Wilks, Baltimore, vol. XI, p. 367-392.
- [10] H. RAYNAUD, Sur le comportement asymptotique de certains graphes aléatoires. *C. R. Acad. Sc. Paris*, t. 262, 2 mai 1966, p. 1008-1011, série A.
- [11] H. RAYNAUD, *Sur une famille de graphes aléatoires*. Actes des Journées d'études sur la Théorie des Graphes. I. C. C., Rome, 1966. Dunod, p. 339-348.
- [12] A. RENYI et P. ERDÖS, *On the evolution of random graphs*. A. Magyar Tud. Akadémia Matematikai Kutató Intézetének Közleményei 5, 1960, A. 17-61.

Manuscrit reçu le 23 avril 1968.
