Marcel-Paul Schützenberger

ŒUVRES COMPLÈTES

éditées par
Jean Berstel, Alain Lascoux et Dominique Perrin

*

Tome 7 : 1969–1970

Institut Gaspard-Monge, Université Paris-Est
2009

Introduction
Tome VII : 1969–1970

Samuel Eilenberg vient à Paris en 1967. Il fait un cours de théorie des automates, où il défend la vision catégorique. Une longue collaboration s'établit avec Schützenberger. L'un des premiers produits de cette coopération est l'article *Rational sets in commutative monoids* [1969-3] sur les parties rationnelles dans les monoïdes commutatifs. Le résultat principal est que les parties rationnelles des monoïdes commutatifs sont non ambiguës. Le fait que les parties rationnelles des monoïdes commutatifs forment une algèbre de Boole avait été prouvé antérieurement par Ginsburg et Spanier [2]. Cette référence aux « bounded ALGOL-like languages » a dû paraître bien mystérieuse à plus d'un lecteur, puisque l'article ne mentionne pas le lien entre les parties rationnelles des monoïdes commutatifs et les langages algébriques (d'après le théorème de Parikh, l'image commutative d'un langage algébrique est rationnelle).

La coopération entre Schützenberger et Eilenberg se poursuivra pendant toute la durée de préparation du livre *Automata, Languages and Machines*. Dans l'introduction du volume A, paru en 1974 [1], Eilenberg dit : « The reader will find that the name of M.-P. Schützenberger is often mentioned as author (or coauthor with me) of many of the new results or proofs that appear in this volume; most have not been previously published. However, his contribution went much beyond that; virtually every phase of the development presented here was endlessly discussed with him. I am, however, the only one to blame for any shortcomings ». On verra plus loin que les relations deviendront moins bonnes par la suite.

L'article *A combinatorial problem in the theory of free monoids* [1969-2] est écrit avec André Lentin (les taupins se souviennent-ils encore des manuels d'exercices de Lentin et Rivaud ?). Il contient un résultat très important de la combinatoire des mots qui donne une caractérisation des sous-monoïdes engendrés par deux mots qui sont apériodiques. La lecture en est très difficile et se ressent peut-être d'une écriture à deux mains, dont l'une pose des définitions peu utilisées dans la suite et l'autre annonce vers la fin (p. 143) « For $q = 3$, the calculation offers no difficulty in principle, but it is very long. For brievity, we shall not give it here ».

La monographie *Théorie géométrique des polynômes eulériens* [1970-1] est une illustration de la méthode bijective, appelée géométrique, et appliquée de manière systématique dans ce texte. Dominique Foata a édité une nouvelle impression, avec quelques corrections, de ce traité. L'une des notions importantes introduites dans cette monographie si souvent citée est la *transformation fon-*

damentale qui est une généralisation sur les mots de la correspondance entre les représentations par produit de cycles et par table d'une permutation. Elle permet en particulier de mettre en bijection les permutations de n objets ayant k descentes et celles qui ont k excédances. Le nombre de ces permutations est précisément le nombre eulérien $A_{n,k}$. Cette transformation et les résultats qu'elle permet d'obtenir sera par la suite généralisée par divers auteurs et reprise dans de nombreux ouvrages dont le volume 3 du livre de Knuth [3] et dans le chapitre du volume de Lothaire [4] rédigé par Dominique Foata. On trouve également dans cette monographie le premier exposé sur le composé partitionnel, qui sert de cadre ensembliste pour la démonstration de nombreuses identités de la combinatoire ou des fonctions spéciales.

[1] Samuel Eilenberg. *Automata, Languages, and Machines. Vol. A.* Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York, 1974. Pure and Applied Mathematics, Vol. 58.

[2] Seymour Ginsburg and Edwin H. Spanier. Bounded ALGOL-like languages. *Trans. Amer. Math. Soc.*, 113 :333–368, 1964.

[3] Donald E. Knuth. *The Art of Computer Programming. Vol. 3.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973. Sorting and searching, Addison-Wesley Series in Computer Science and Information Processing.

[4] M. Lothaire. *Combinatorics on words.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997. Corrected reprint of the 1983 original.

Année 1969

Bibliographie

[1] Jacques Besson, Pierre Gavaudan, and Marcel-Paul Schützenberger. Sur l'existence d'une certaine corrélation entre le poids moléculaire des acides aminés et le nombre de triplets intervenant dans leurs codages. *C. R. Acad. Sci. Paris*, 268 :1342–1344, 1969.

[2] André Lentin and Marcel-Paul Schützenberger. A combinatorial problem in the theory of free monoids. In *Combinatorial Mathematics and its Applications (Proc. Conf., Univ. North Carolina, Chapel Hill, N.C., 1967)*, pages 128–144. Univ. North Carolina Press, Chapel Hill, N.C., 1969.

[3] Samuel Eilenberg and Marcel-Paul Schützenberger. Rational sets in commutative monoids. *J. Algebra*, 13 :173–191, 1969.

[4] Marcel-Paul Schützenberger. Langages formels et monoïdes finis. In *Séminaire Dubreil-Pisot, année 1969-70*, Exposé No. 3, 3 pages. Inst. H. Poincaré, Paris, 1969.

C. R. Acad. Sc. Paris, t. 268, p. 1342-1344 (3 mars 1969) Série D

BIOCHIMIE. — *Sur l'existence d'une certaine corrélation entre le poids moléculaire des acides aminés et le nombre de triplets intervenant dans leur codage.* Note (*) de MM. **Marcel-Paul Schutzenberger, Pierre Gavaudan** et **Jacques Besson,** présentée par M. Jean Roche.

On observe une relation globale de proportionnalité inverse entre le poids moléculaire des acides aminés et le nombre de triplets les codant ainsi qu'une relation de même sens dans l'étude statistique de la composition des protéines.

Des études théoriques antérieures consacrées à la structure du code génétique [(1), (2)] nous ont conduits à rechercher au sein des classes de codage définies par les nombres de triplets affectés aux acides aminés qu'elles contiennent, des paramètres susceptibles d'individualiser simplement les acides aminés. En effet, quoique de composition connue les classes ne donnent que des informations statistiques faibles puisqu'au sein de chacune d'elles l'identité des acides aminés est ignorée. L'analyse par segments (2) indique dans certaines chaînes des fluctuations sensibles de la classe II, mais là encore l'anonymat des acides aminés fait perdre de précieuses informations. Nous avons pensé qu'il serait peut-être instructif de considérer ce paramètre très simple qu'est le poids moléculaire des acides aminés qui, à notre connaissance, a été laissé de côté jusqu'à ce jour.

Certes l'idée de choisir un paramètre de nature aussi triviale n'est-elle pas très attirante *a priori*, car il est bien certain qu'il fait abstraction de la forme et de la constitution des molécules ainsi que de leurs propriétés physiques et chimiques. Toutefois, représentant une sorte d'indice global de complexité, l'utilisation du poids moléculaire, facile du point de vue statistique, conduit à des résultats appréciables.

Nous disposons tout d'abord les poids moléculaires des acides aminés sur une échelle où ils croissent de gauche à droite ; ils figurent entre parenthèses auprès du numéro de la classe à laquelle appartient l'acide aminé correspondant. GLY (IV, 75) ; ALA (IV, 89) ; SER (VI, 105) ; PRO (IV, 115) ; VAL (IV, 117) ; THR (IV, 119) ; CYS (II, 121) ; LEU (VI, 131) ; ILE (III, 131) ; ASN (II, 132) ; ASP (II, 133) ; LYS (II, 146) ; GLN (II, 146) ; GLU (II, 147) ; MET (I, 149) ; HIS (II, 155) ; PHE (II, 165) ; ARG (VI, 174) ; TYR (II, 181) ; TRP (I, 204).

Sur cette échelle les neuf premiers acides aminés appartiennent aux classes VI, IV et III, respectivement codées par 6, 4 et 3 triplets à l'exception de CYS codée par 2 ; les onze acides aminés suivants appartiennent aux classes II et I à l'exception de ARG codée par 6.

L'attribution à chaque classe d'un poids moléculaire moyen permet de rejoindre et de préciser les précédentes conclusions.

TABLEAU

Classes	I	II	III	IV	VI
P. M. moyen	175	147	131	103	136

(2)

Dans chaque classe, à l'exception de la VI, le poids moléculaire moyen est en gros inversement proportionnel au nombre de triplets. La classe VI s'étend entre 105 (SER) et 174 (ARG) en passant par 131 (LEU) et sa moyenne est fortement alourdie à cause de l'Arginine.

Calculant d'abord la corrélation entre le poids moléculaire et le nombre des codons de chacun des vingt acides aminés on trouve un coefficient de corrélation d'une valeur absolue égale à 0,57 et significative au seuil de 99 %. Considérant ensuite les ordres des fréquences des acides aminés dans un échantillon de plus de 4 000 résidus appartenant à une trentaine de protéines on observe une corrélation encore plus forte égale à 0,77. De plus, à l'intérieur de chaque classe il existe une tendance des acides aminés les plus légers à être les plus fréquents. Enfin, à la nette dérive statistique de la classe IV qui, fortement majorée par rapport à la théorie, tend à être aussi abondante dans les protéines que la classe II, s'associe la différence sensible des poids moléculaires des deux classes. De plus, l'opposition des classes II et IV déjà remarquée [2], pour la redondance, les fréquences des acides aminés et le poids moléculaire, est soulignée par les différences des proportions dominantes des bases des codons (II : A + U et IV : G + C).

Après ces premières remarques que nous développerons ultérieurement on est en droit de se préoccuper de la raison pour laquelle on peut observer une telle corrélation d'ensemble entre le poids moléculaire, le nombre des codons et la fréquence des acides aminés dans les protéines.

A l'évidence le poids moléculaire n'est qu'un paramètre grossier définissant seulement une sorte d'encombrement général, aussi ne voit-on pas comment ce facteur purement pondéral pourrait prescrire directement les nombres de codons affectés aux acides aminés. En dépit de la netteté de la corrélation on n'entrevoit pas la possibilité de l'existence d'une relation immédiate de cause à effet entre les deux séries de valeurs, l'explication devant être probablement de nature beaucoup plus complexe.

Sans ignorer la difficulté que posent les différences que l'on peut supposer devoir exister entre les systèmes biochimiques primitifs et actuels, l'examen des mécanismes modernes de synthèse des acides aminés pourrait fournir un modèle de raisonnement. Les synthèses de la Glycine, de la Sérine (famille du Triose) et de l'Alanine (famille du pyruvate) sont sans doute parmi les plus simples ; celles de la Méthionine, de l'Histidine, de la Phénylalanine, de la Tyrosine et du Tryptophane (à travers l'acide shikimique pour les trois derniers) sont parmi les plus complexes [3]. On pourrait supposer que les premiers acides aminés synthétisés par les chaînes réactionnelles les plus simples auraient exercé une préemption sur un plus grand nombre de codons ; ceux de poids moléculaires plus élevés, et partant plus complexes, venus plus tard à la suite de la mise en place de nouveaux systèmes biochimiques, n'auraient alors disposé que d'un nombre réduit de codons, à la limite un seul pour la Méthionine et le Tryptophane. Notons que Jukes [4] a supposé qu'à l'origine il n'aurait existé qu'une quinzaine d'acides aminés primitifs engendrés par les synthèses abiogéniques. Codés par un code à 16 doublets ils auraient été ultérieurement complétés au cours de l'évolution biochimique par la Méthionine, le Tryptophane et la Tyrosine et les deux amides, avec passage au système de code à triplets. Bien que fondée aussi sur la supposition d'une augmentation évolutive du nombre des acides aminés, notre conception est indépendante de celle de Jukes puisque nous considérons le code à triplets comme primitif ([5]). Nous ne présentons ce modèle que comme une simple hypothèse sans prétendre lui attribuer une valeur explicative d'un caractère exclusif.

Sans doute, en effet, à travers l'indice de complexité de synthèse qu'il symbolise, le poids moléculaire n'est-il qu'un des facteurs ayant contribué à fixer les valeurs de redondance et à organiser les classes de triplets ; d'autres facteurs ont pu interférer avec celui de la complexité des synthèses, masquer son rôle et provoquer d'importants remaniements ; notamment l'organisation du code a pu tendre vers une minimisation des erreurs par groupement de triplets codant des acides aminés de propriétés analogues ([6]). Ces réserves étant faites, notre hypothèse sur la signification de la relation signalée entre poids moléculaire et codage offre l'intérêt de ne pas limiter les possibilités de compréhension de la structure du code à ses seules données internes et de lier son évolution à celle de l'ensemble des systèmes biochimiques.

(*) Séance du 24 février 1969.

([1]) P. Gavaudan et J. Besson, *Comptes rendus*, 264, Série D, 1967, p. 1919-1922.

([2]) P. Gavaudan et J. Besson, *Comptes rendus*, 268, Série D, 1969, p. 173-175.

([3]) P. Bernfeld, *Biogenesis of natural compounds*, Pergamon Press, London, New York, 1963, 929 pages.

([4]) Th. H. Jukes, *Molecules and evolution*, Columbia Univ. Press, New York and London, 1966, 282 pages, *Loc. cit.*, p. 64-69.

([5]) A propos de l'hypothèse du changement de système de codage voir F. H. C. Crick, *Journ. Mol. biol.*, 38, 1968, p. 367-379.

([6]) A. L. Goldberg et R. E. Wittes, *Science*, 153, 1966, p. 420-424.

(*Faculté des Sciences de Poitiers, Station Biologique de Beau-Site,*
25, *Faubourg Saint-Cyprien*, 86-*Poitiers*, *Vienne* ;
Faculté des Sciences de Paris, Département de Mathématiques.)

CHAPTER 8

A Combinatorial Problem in the Theory of Free Monoids[1]

A. LENTIN and M. P. SCHÜTZENBERGER,
Faculté des Sciences, Paris, France

1. INTRODUCTION

The combinatorial properties of free monoids play a role in several lemmas which are used in the theory of free groups (or free Lie algebras), formal languages, automata, variable length codes, and elsewhere.

The purpose of the present article is to determine a property of this type (Theorem 5 below).

2. SUMMARY OF PREVIOUS WORK

2.1. Freedom ; Primitiveness

We take a set $X = \{x, y, \ldots\}$, which generates the *free monoid* X^* whose elements are called *words*. The *length* of a word f is designated by $|f|$, the word of length zero by e. Every subset A of X^* generates a sub-monoid, written A^*, for which it forms a *system of generators*.

[1] This work has been supported by contract AF61(052)945 with the United States Air Force.

Example. Let $X = \{x, y\}$ and let

$$A_1 = \{xyx, x\},$$

$$A_2 = \{xy, yx, x\},$$

$$A_3 = \{x, y, xy\}.$$

It is seen that A_1^*, $A_2^* \subset X^*$ (strict inclusion),

whereas $$A_3^* = X^*.$$

Definition 1. *$B \subset X^*$ is called a base of A (or a code) if there exists a set X' and a surjection φ of X' on B which can be extended to a monomorphism of X'^* in X^*.*

More intuitively, this is equivalent to saying that B is a base iff every word of A^* has a unique factorization in terms of the words of B.

Example. A_1 is a base for A_1^*. A_2 is not a base for A_2^* since xyx is capable of two factorizations, namely, $(xy)x$ and $x(yx)$. A_3 is not a base for A_3^*; but $\{x, y\}$ is.

Definition 2. *The submonoid A^*, generated by the system of generators A, is called free if it has a base.*

Example. A_1^* is free and has the base A_1. A_3^* is free and has the base X. A_2^* is not free: in effect, if it had a base the latter would contain x, but not y, and would therefore necessarily contain xy and yx.

Remark. We have $A_2^* \subset X^*$, and $\mathbf{card}\,(X) < \mathbf{card}\,(A_2)$; note that this can be generalized. In what follows, whenever A^* is a free submonoid, A will always be its base.

Theorem 1. *A necessary and sufficient condition for A^* to be free and of base A is the following (condition L):*

$$(L): \quad \textit{For all} \quad h \in X^* \setminus A^*, \qquad hA^* \cap A^* \cap A^*h = \phi$$

Proof. We will prove the equivalent proposition

$$(A \text{ is not a base}) \Longleftrightarrow (\bar{L}).$$

We have

$$(\bar{L}): \quad \exists\, h \in X^* \setminus A^* \quad \text{such that} \quad hA^* \cap A^* \cap A^*h \neq \phi\,.$$

$$(A \text{ is not a base}) \Longrightarrow (\bar{L})\,.$$

Since A is not a base, A^* contains a set of words which are capable of at least two distinct factorizations, and in that set there is a subset consisting of minimal length words. Let

$$(1) \qquad m = a_{i_1} \ldots a_{i_n} = a_{j_1} \ldots a_{j_p}$$

be such a word. The minimality of m implies that $a_{i_1} \neq a_{j_1}$. We can then take $|a_{i_1}| < |a_{j_1}|$; whence

$$(2) \qquad a_{j_1} = a_{i_1}h\,, \qquad h \in X^*\,, \qquad |h| \neq 0\,.$$

Then

$$(3) \qquad a_{i_2} \ldots a_{i_n} = ha_{j_2} \ldots a_{j_p}\,.$$

But

$$(2) \Longrightarrow A^* \cap A^*h \neq \phi\,,$$

$$(3) \Longrightarrow hA^* \cap A^* \neq \phi\,,$$

whereas

$$\text{minimality and} \quad (3) \Longrightarrow h \in X^* \setminus A^*\,;$$

whence follows $(\bar{L})$.

$$(\bar{L}) \Longrightarrow (A \text{ is not a base})$$

$$(\bar{L}) \Longrightarrow \exists\, h \in X^* \setminus A^* \quad \text{and} \quad g, m, g' \in A^* \quad \text{such that} \quad hg = m = g'h\,.$$

By virtue of its belonging to $X^* \setminus A^*$, h is not empty, and so neither is m. By simplifying on the left in A (if necessary) we can arrange for m and g' not to have the same initial letter in A^*.

The double equality $g'hg = mg = g'm$ (or the one which is left after simplifying) proves that A is not a base.

Definition 3. *If A^* is a free submonoid of base A, we define the set $\pi(A^*)$ of A-primitive elements by the equivalence:*

$$f \in \pi(A^*) \Longleftrightarrow f \in A^* \quad \text{and} \quad f \neq g^p \quad \text{for any}$$
$$g \in A^* \quad \text{and any} \quad p \neq 0, 1 .$$

Instead of *X-primitive* we shall simply say *primitive.*

Remark. It is clear that A-imprimitiveness implies imprimitiveness, but the converse is not true. For instance:

$$A = \{xyx, y\} . \qquad f = xyxy , \qquad f \in \pi(A^*) , \qquad f \notin \pi(X^*) .$$

Every element of a submonoid A^* having the base A can be represented uniquely as the power of an A-primitive element.

Proof. Clearly, there exists such a representation. Suppose then that we have

$$f \in A^* ; \; f = g^p = h^q ; \; p, q \geqslant 1 ; \; g, h \in \pi(A^*) .$$

According to the hypothesis, f, g, and h each have a unique factorization. Proceeding by identification, one shows that $g = h$, whence $p = q$.

The following theorem, as well as its corollaries, are related to the concept of primitiveness (cf. [1]).

Theorem 2. *A necessary and sufficient condition for two words a, $b \in X^*$ to be two powers of the same word (which one can always suppose to be primitive) is that a power a^p of a and a power b^q of b contain a common left (right) factor of length*

$$|a| + |b| - (|a| \cap |b|) ,$$

where $|a| \cap |b|$ stands for the greatest common divisor.

Proof. Set $|a| = \alpha$, $|b| = \beta$. We first treat the case where $\alpha \cap \beta = 1$. Let $a = x_1 \ldots x_\alpha$ and $b = y_1 \ldots y_\beta$; we can take $\beta < \alpha$.

Sufficiency. We have the relations:

$$1 : x_1 = y_1 ,$$

$$2 : x_2 = y_2 ,$$

$$\alpha + \beta - 1 : x_\lambda = y_\mu .$$

Let us scan this set of relations in the following way:

i) If possible, add β to the number of the line one has just read;

ii) otherwise add $-\alpha + \beta$.

This scan is possible, for it is equivalent to uniting the vertices of a polygon in steps of β. Since β and α are relative primes, we exhaust the vertices. We have therefore $a = x_1^\alpha$, $b = x_1^\beta$.

Necessity. We have above a system of $\alpha + \beta - 1$ homogeneous equations in $(\alpha + \beta)$ unknowns. Let us add the relation:

$$\sum_i \lambda_i x_i + \sum_{j'} \mu_{j'} y_j = k$$

where $k \neq 0$ and the coefficients are not all zero. The system is then determined.

It is easily seen that one can form a determined system of the same rank by replacing the last relation with

$$x_\lambda = k_1 , \qquad y_\mu = k_2 .$$

The words a and b can be written with two types of letters, and since $\alpha \cap \beta = 1$, they are not powers of a same third word.

For $\alpha \cap \beta = \delta$, we take sections of length δ and apply the previous result.

Corollary 1. *A necessary and sufficient condition for $a, b \in X^*$ to be powers of the same word is that ab and ba contain a common left factor of length*

$$|a| \pm |b| - (|a| \cap |b|) .$$

Proof. (same notations). The theorem is trivially true for $\alpha = \beta$. Let us suppose $\beta < \alpha$. The hypothesis implies that ab is a left factor a^2, ba a word of the form $b^i b_1$, hence a left factor of b^{i+1}. We apply the theorem.

Corollary 2. *A necessary and sufficient conditon for $a, b \in X^*$ to be powers of the same word is that there exist in $\{a, b\}^*$ two distinct elements having no common factor in $\{a, b\}^*$ and having in X^* a common left factor of degree $|a| + |b| - (|a| \cap |b|)$.*

The proof is by case, in a way analagous to the preceding proof.

Corollary 3. *a, b is a base of $\{a, b\}^*$, and $\{a, b\}^*$ is free iff a and b are not powers of the same word.*

In later applications, we shall frequently use the following corollary.

Corollary 4. *For f, $g \in \pi(X^*)$, $h \in X^*$, $p, q > 1$, the hypothesis $f^p = g^q h$ implies that either $g = f$ and $h = f^{p-q}$, or else $(q - 1) \cdot |g| < |f|$.*

To state the last corollary, we must give finally the definition of a fundamental concept.

Definition. *We shall call sesquipower on X^* a word f of the form :*

$$f = (uv)^k u\ , \quad k > 0\ , \quad uv \in \pi(X^*)\ , \quad v \neq e\ .$$

A sesquipower such that $k \geqslant 2$ will be called a strong sesquipower.

Corollary 5. *For $k \geqslant 2$, a strong sesquipower $(uv)^k u$ has a unique representation as a strong sesquipower.*

Proof. Let $(uv)^k u = (wz)^j w$; then $(uv)^{k+1}$ and $(wz)^{j+1}$ are two powers of primitive words and have a common left factor of length $k|uv| + |u| = j|wz| + |w|$. If we subtract from the length of this common factor the sum of the lengths, we obtain :

$$k|uv| + |u| - |uv| - |wz| = (k-1)|uv| + |u| - \frac{k|uv| + |u|}{j + \theta}\ ,$$
$$0 < \theta < 1\ .$$

This difference has the sign of :

$$[(k-1)(j+\theta) - k]|uv| + (j + \theta - 1)|u|\ .$$

The coefficient of $|uv|$ is :

$$k(j + \theta - 1) - (j + \theta)\ .$$

For $k = 2$ it becomes $j - 2 + \theta$, which is positive. Theorem 2 is now applicable.

2.1. Conjugacy

Definition. *If* A^* *is a free submonoid of base* A, *we define the relation of* A*-conjugacy by the equivalence*

$$f \ A\text{-conj. } g \Longleftrightarrow \exists\, h, h' \in A^* \quad \text{such that}$$
$$f = hh' \quad \text{and} \quad g = h'h .$$

Instead of X*-conjugacy* we shall simply speak of *conjugacy*.

Remark. It is clear that A-conjugacy implies conjugacy, but the converse is not true. For instance:

$A = \{xy, yx\}$; then xy and yx are conjugate, but not A-conjugate. It follows immediately from this definition that

1. f A-conjugate $g \Longrightarrow f, g \in A^*$.
2. i) A-conjugacy is *reflexive* (take $e = h'$).
   ii) A-conjugacy is *symmetric* (evident).
   iii) A-conjugacy is *transitive*.

Take

$$f = hh' , \qquad g = h'h ; \qquad g = kk' , \qquad m = k'k .$$

Then we have in A^*

$$g = h'h = kk' .$$

Utilizing the uniqueness of the factorization in A^* (where A^* is free), we obtain, for example,

$$h' = k_1 , \qquad h = k_2 k' ,$$

where $k = k_1 k_2$: whence

$$f = k_2(k'k_1) , \qquad m = (k'k_1)k_2 .$$

The relation of A*-conjugacy is an equivalence.*

3. A-conjugacy is compatible with the power mapping:

$$f \to f^p$$

In effect,

$$f = hh' , \qquad\qquad g = h'h ,$$

$$f^p = h[(h'h)^{p-1}h'] , \qquad g = [(h'h)^{p-1}h']h .$$

These different results can be synthesized in the following theorem :

Theorem 3. *For* $f,\ g \in AA^*$, *set*

$$C_A(f, g) = \{h \in A^* : fh = hg\} .$$

Then f *A-conjugate* $g \Longleftrightarrow C_A(f, g) \neq \phi$. *Furthermore, for two different A-conjugate words there exists a unique positive integer* p *and a unique ordered pair* $u, v \in A^*$ *such that:*

$$v \neq e ; \qquad uv,\ vu \in \pi(A^*) ; \qquad f = (uv)^p ; \qquad g = (vu)^p .$$

$$C_A(f, g) = u(vu)^* ; \qquad C_A(g, f) = v(uv)^*$$

Proof. *Necessity.*

$$f \text{ A-conjugate } g \Longrightarrow f = hh' , \qquad g = h'h ; \qquad h, h' \in A^*$$

$$fh = hg = hh'h ; \qquad h \in C_A(f, g) ;$$

$$C_A(f, g) \neq \phi .$$

Sufficiency. Let us suppose that $C_A(f, g)$ contains at least one word h ; then we have

$$fh = hg ,$$

$$fhg = hgg ,$$

$$ffh = hgg .$$

More generally, for all $m \geqslant 1$,

$$f^m h = hg^m .$$

However, there exists a unique integer n such that

$$n|f| \leqslant h < (n+1)|f| .$$

We have then :

$$f = f_1 f_2 , \qquad h = f^n f_1 ;$$

$$f^{n+1}f_1 = f^n f_1 g\,;$$

$$f_2 f_1 = g\,.$$

It can be immediately verified that this solution, obtained from necessary conditions, verifies

$$fh = hg$$

Representation. We know that every word of A^* is a power of an A-primitive word, so that we have

$$f = f_1 f_2 = f_0^p\,; \qquad p \geqslant 1\,;$$

$$f_0 = uv\,, \qquad f_1 = (uv)^i u\,, \qquad f_2 = v(uv)^j$$

$$i + j + 1 = p\,,$$

and this uniquely. It follows that

$$f = f_1 f_2 = (uv)^p\,; \qquad uv \in \pi(A^*)\,,$$

$$g = f_2 f_1 = (vu)^p\,; \qquad vu \in \pi(A^*)\,.$$

For $f \neq g$, we have $uv \neq vu$; hence $v \neq e$. The rest of the conclusion is evident.

Corollary 1. *For every $f \in AA^*$, the following properties are equivalent:*

(1) *f is A-primitive;*
(2) *The class of A-conjugates of f containts an A-primitive word;*
(3) *$C_A(f, f) = f^*$ and any relation*

$$f' f^p f'' = f^q \quad \textit{implies that} \quad f', f'' \in f^*\,.$$

(4) *If $f \in A^k$, the class of A-conjugates of A contains exactly k words.*

The proof presents no difficulties.

Finally, Theorem 2 yields the following theorem immediately by a " shift ":

Theorem 4. *A necessary and sufficient condition for the words f and g to be conjugate is that two powers f^p and g^q of these words contain a common factor of length $|f| + |g| - (|f| \cap |g|)$.*

2.3 Relation to other theories

To begin with, it is clear that, for $A = X$, the concepts of primitiveness and conjugacy originate, by restriction to the monoid X^*, in analogous concepts relative to the *free group* generated by X. They can be extended immediately to a base A with the help of the monoid X'^* and the monomorphism which were defined at the beginning. Furthermore, some concepts and results can be extended to other monoids. In order to better visualize these extensions, we give first a " geometrical " interpretation.

To each $f \in X^*$, let us associate the mapping $\hat{f}$ of the segment $[1, \ldots, |f|]$ in X which sends i onto the ith letter of f. Then with the product $h = fg$ (in the monoid) there is associated the mapping $\hat{h} = \widehat{fg} = \hat{f} \cdot \hat{g}$.

$$\hat{h}(i) = \begin{cases} \hat{f}(i)\,, & \text{for} \quad i \in [1, \ldots, |f|]\,; \\ \hat{h}(i - |f|)\,, & \text{for} \quad i \in [|f|, \ldots, |f| + |g|]\,. \end{cases}$$

In this construction, f and g are conjugate iff $\hat{g}$ can be deduced from $\hat{f}$ by a cyclic shift. In other words, there exists a fixed j such that :

$$\hat{f}(i) = \begin{cases} \hat{g}(i + j)\,, & \text{for} \quad i \in [1, \ldots, |g| - j]\,; \\ \hat{g}(i + j - |g|)\,, & \text{for} \quad i \in [|g| - j + 1, \ldots, |g|]\,. \end{cases}$$

In the same way, f is the pth power of g iff

$$|f| = p|g| \text{ and } \hat{f}(i + k|g|) = \hat{g}(i) \text{ for}$$
$$i \in [1, \ldots, |g|] \text{ and } k \in [0, 1, \ldots, p - 1]\,.$$

Thus, the *primitiveness* of a word is equivalent to the *aperiodicity* of the associated mapping onto its interval of definition.

Fine and Wilf have shown that most of these results can be extended to more general monoids consisting of continuous mappings in a topological set X of intervals of the real line, when these mappings are compounded by the product "." This is true in particular of Theorem 4 : its extension shows that two periodic mappings are equal on the necessary and sufficient condition that they coincide on an interval whose length is equal to the sum of the lengths of their respective periods.

3. MAIN RESULTS

3.1. Statement

Theorem 5. *Let $A = \{a, b\}$ be a base such that each word of $a^*b \cap ab^*$ is primitive; then each A-primitive word of A^* is primitive.*

Actually, as we shall see, it suffices that each word of $a^*b \cap ab^*$ of length less than $3|ab|$ be primitive in order to guarantee the conclusion of the property. Also, at most one word of $a^*ab \cap abb^*$ can be imprimitive.

3.2. Terminology

We consider $A = \{a, b\} \subset X^*$, $a \neq b$. In view of the hypotheses, we have that a and b, elements of $a^*b \cup ab^*$ are primitive. For the sake of definiteness, we take $|b| \leqslant |a|$.

We introduce the following terminology:

$$\text{For} \quad d = d_1 d_2 \ldots d_k \in A^k \ (\text{i.e. } d_1, d_2, \ldots, d_k \in A)\,,$$

we call an *A-factor of d* any product $d_i d_{i+1} \ldots d_j$ $(1 \leqslant i \leqslant j \leqslant k)$ occurring in d. Furthermore, we say that $d' = d'_1 d'_2 \ldots d'_{k'} \in A^{k'}$ is a *principal segment of d* iff there exists f, $f' \in X^*$ such that $fd'\,f' = d$ with $|f| < |d_1|$; $|f'| < |d_k|$.

Further we say that c is *disjoint from d* iff

$$f, f' \neq e \quad \text{and for all} \quad j, j', \qquad fd'_1 \ldots d_{j'} \neq d_{1'} \ldots d_j\,.$$

Thus, if d' is a principal disjoint segment of d, any A-factor of d' (or of d, with the exception of $d_1, d_2, \ldots, d_k$, or $d_1 d_2 \ldots d_{k-1}$) is again a principal disjoint segment of a well defined A-factor of d (or of d').

3.3. Preliminary results

(1) *Let c, $d \in A^*$ be conjugate but not A-conjugate. Any A-factor of $c^n (n < 1)$ is a principal disjoint factor of an A-factor of d^n.*

Proof. We have $hc = dh$; hence for all positive integers n, $hc^n = d^n h$ with $h \in X^* \setminus A^*$. The hypothesis that an A-factor of c^n is not disjoint from d would imply that

$$c^n = c_1 c_2\,, \qquad d^n = d_1 d_2\,; \qquad c_1, c_2, d_1, d_2 \in A^*\,, \qquad hc_1 = d_1\,;$$

$$c_2 = d_2 h \,.$$

Thus we would have

$$hc_1c_2 = d_1c_2 = d_1d_2h \quad \text{with} \quad c_1c_2,\, d_1c_2,\, d_1d_2 \in A^* \,;$$

hence $h\, A^* \cap A^*\, h \cap A^* \neq 0$ in contradiction to $h \notin A^*$ and, according to Theorem 1, the hypothesis that A^* is free.

(2) *For $p > 0$ and $q > 2$, $c = a^q$ cannot be a disjoint principal segment of $d = ab^p a$.*

Proof. Let $ab^pa = fa^qf'$ where $|f|, |f'| < |a|$. Since $q > 2$, at least one A-factor a of a^q is a principal disjoint segment of b^p, hence $p \geqslant 2$.

Now either b^p and a^q have a common segment of length $\geqslant |a| + |b|$ or they do not. In the first case, by Theorem 3, a and b are conjugate; then we have

$$a = uv\,, \qquad b = vu\,; \qquad uv(vu)^p uv = f(uv)^q f'\,,$$

where the segments uv of $(vu)^p$ and of $(uv)^q$ must coincide because $p \geqslant 2$ and, by the hypothesis, a, $b \in \pi(X^*)$. Thus

$$uvvu = fuv\,; \qquad (vu)^{p-2} = (vu)^{q-2}\,; \qquad vuuv = uvf'\,.$$

The first (or the third) relation shows that $vu = uv$, i.e., $a = b$ in contradiction to the hypothesis $a \neq b$.

In the second case $|b^p| < |a| + |b|$. Because $|a| \geqslant |b|$ this implies $q = 3$ and we can write

$$a = fg = gh = h'g' = g'f'\,,$$

so that

$$a^q = ghah'g'\,; \qquad b^p = hah' \quad \text{with} \quad |h| + |h'| < b\,.$$

Thus at least one of h or h' (say h) has length $< |b|/2$; hence $1 < |a|/2$. By Theorem 3, the relation $a = fg = gh$ implies

$$f = u'v'\,; \qquad h = v'u'\,, \qquad a = (u'v)^{r'}u'\,,$$

where $r' \geqslant 2$ since $|h| < 1/2|a|$.

Thus a is a strong sesquipower and by Corollary 5, we can write in a unique manner

$$h = (vu)^s\,; \qquad a = (uv)^r u \qquad (s > 0,\ r \geqslant r')$$
$$\text{with} \quad vu \in \pi(X^*)\,.$$

Then

$$b^p = (vu)^s (uv)^r u (uv)^{s'} g'\,; \qquad (uv)^{s'} g' = h'\,; \qquad |g'| < |uv|\,.$$

Again since $vu \in \pi(X^*)$, any segment of b^p equal to vu or to uv is in fact a $\{u, v\}$-factor. The inequality $|h| < |b|/2$ shows that $(vu)^s uv$ is a left factor of b. However, b^p has no other segment $vuuv$ except at its end, where it occurs in $uv(uv)^{s'} g'$. Now this last word is strictly shorter than b and the hypothesis $a \neq b$ implies that $vu \neq uv$. Thus there is a contradiction because $p \geqslant 2$ (as has been shown above).

(3) *For $p > 1$ the word $ab^p a$ cannot be a principal disjoint segment of $b^r a^2 b^s$. For $p = 1$, it is so only if $a^2 b$ is imprimitive.*

Proof. The hypothesis implies

$$b^r = b_1 b_2\,; \qquad b^s = b_3 b_4\,; \qquad ab^p a = b_2 a^2 b_3 \quad \text{with} \quad b_3 b_2 = b^p\,.$$

Thus we have $ab_3 b_2 a = b_2 aab_3$, showing that $ab^p a$, hence $a^2 b^p$, is imprimitive. For $p = 1$, the proposition is proved. For $p > 1$, the proposition will be proved by showing that $a^2 b^p = c^q$, $q > 1$, is incompatible with the hypothesis that a and b are not powers of the same word (Theorem 2).

According to Corollary 4 of Theorem 2, the conclusion is established for

$$|a| \geqslant |c| \quad \text{or} \quad (p-1)|b| \geqslant c\,.$$

Let us suppose that $|a| < |c|$ and $(p-1)|b| < |c|$. Then, in view of the equality

$$2|a| + p|b| = q|c|\,,$$

these inequalities require that

$$2 + \frac{p}{p-1} > q\,;$$

hence, $q = 2$ or 3. Let $q = 2$. For even p, the conclusion follows at once. For odd p, $|b|$ is necessarily even. We have $b = b_1b_2$ with $|b_1| = |b_2|$, which allows us to segment the equation and arrive at the conclusion. For $q = 3$, the calculation offers no difficulties in principle, but it is very long. For brevity we shall not give it here.

(4) *For $p > 0$, ab^pa cannot be a principal disjoint segment of aad', nor of $d'aa$ $(d' \in A^*)$, or of $ab^{p'}a$ or $b^{p'}$.*

Proof. Suppose $fab^paf' = aad'$. The hypothesis of disjointness implies that a is a principal disjoint segment of aa; hence by Corollary 1, $a \notin \pi(X^*)$, which is a contradiction. The same applies to $d'aa$.

In the two other cases, the same argument applies for b and bb.

(5) *Let $c = ab^pa$ $(p > 0)$ be a principal disjoint segment of $d \in A^*$ and suppose that d has no A-factor $ab^{p'}a$ with $p' < p$. Then either d has an A-factor of the form b^rab^s $(r + s = p)$ which is a principal disjoint segment of c or else $p = 1$ and $d = ba^2b$.*

Proof. Assume $d \neq ba^2b$. The case of $d \in b^*a^2ab^*$ is excluded by (2) and (3) above.

For $d = b^{r'}ab^{s'}$ the hypothesis of disjointness implies $r', s' > 1$; we must have

$$|b^{r'-1}ab^{s'-1}| < |ab^pa| < |b^{r'}ab^{s'}|;$$

hence $r' + s' > p + 1$, and $r' - 1 + s' - 1 \geqslant p$: The result is verified.

If $d \notin b^*aa^*b^*$, the case of $d \in b^*$ is excluded by (4) and d must have an A-factor of the form $ab^{p'}a$ where $p' \geqslant p$ by hypothesis. Again by (4), $d \neq ab^{p'}a$ so that either $d = ab^{p'}ad'$ or $d = d'ab^{p'}a$ with $d' \neq e$. The result is verified by taking b^pa or ab^p.

3.4. Conclusion of the proof

We consider $g, g' \in A^*$, conjugate but not A-conjugate, such that $rg^n = g'^nr$, $r \notin A^*$, for all n. Such a situation necessarily obtains when $g \in A^*$ is A-primitive without being primitive; $g = f^m$ $(f \in X^* \setminus A^*,\ m > 1)$ since $f^{m+1} = fg = gf$. According to (1), every A-factor of $g^n(g'^n)$ is a principal disjoint segment of some A-factor of $g'^n(g^n)$. Since a and b are primitive, we cannot have either $g, g' \in a^*$ or $g, g' \in b^*$. If $g \in a^*$ ($g \in b^*$), (5) shows that the only remaining possibility is $g' \in b^*$ ($g' \in a^*$). Thus we can assume now $g, g' \notin a^* \cup b^*$, and suppose that g^2 has an A-factor ab^pa with p positive such that g'^2 has no A-factor $ab^{p'}a$ with $p' < p$.

We shall show that under these conditions at least one word of $a^*ab \cup abb^*$ is imprimitive.

By (5) the principal disjoint segment d of g' that covers ab^pa has an A-segment ab^p or b^pa (unless $p = 1$ and $d = bab$ in which case we know already by (3) that a^2b is imprimitive). Then ab^p (or b^pa) is a proper principal segment of ab^pa; hence of ab^pab^p (b^pab^pa). Thus it is imprimitive.

3.5. Additions

From our proof it now follows that if the set $\{a, b, a^2b\} \cup abb^*$ consists only of primitive words, the only word pairs (if there are any) which are conjugate without being A-conjugates are of the form (a^n, b^n). We can establish the following more accurate result: if a and b are conjugate, $a^*b \cup ab^* \subset \pi(X^*)$; otherwise $a^*b \cup ab^*$ contains at most one imprimitive word.

For the first part of this result, one is led to examine

$$(uv)(vu)^\lambda = c^\mu \,; \qquad \lambda \geqslant 1 \,; \qquad \mu \geqslant 2 \,.$$

The case where $\lambda = 1$, $\mu = 2$ evidently contradicts the hypothesis of primitiveness. For $(\lambda - 1)|uv| > |c|$, the hypothesis of primitiveness is contradicted by Corollary 4 of Theorem 2. There remains the case $(\lambda - 1)|uv| < |c|$. From the equality $(\lambda + 1)|uv| = \mu c$, one obtains $2 > (\mu - 1)(\lambda - 1)$ and the conclusion follows easily.

We give an outline of the proof of the second part of the result.

(i) The following lemma is useful (we have already proven particular cases of it; cf. [2] and [3]):

The condition $a^m b^m = c^q$, $m, n, q \geqslant 2$ implies that a, b and c are imprimitive and powers of the same word.

By Theorem 2, we have only to consider the case where:

$$(m - 1)\alpha < \gamma - (\alpha \cap \gamma)\,, \qquad (n - 1)\beta < \gamma - (\beta \cap \gamma)\,,$$

with $\alpha = |a|$, $\beta = |b|$, $\gamma = |c|$. From the equality

$$m\alpha + n\beta = q\gamma \,,$$

we obtain the condition

$$2 + \frac{1}{m-1} + \frac{1}{n-1} - \frac{m}{m-1}\,\frac{\alpha \cap \gamma}{\gamma} - \frac{n}{n-1}\,\frac{\beta \cap \gamma}{\gamma} > q \,,$$

which characterizes the cases to be studied. We treat them directly.

(ii) We have seen that for $|a| \geqslant |b|$, the only imprimitive word of a^*a^2b is a^2b. By Corollary 4, $a^2b = f^m$ implies that $|a| > |f|$; hence $m = 2$. Solving $a^2b = f^2$ gives $a = (uv)^{k+1}u$, $b = vuuv$, and the technique of (2) above applies.

(iii) The lemma in (i) is applicable to the case $ab^m = f^p$, $ab^{m'} = g^q$ $(p, q \geqslant 2\,;\ m' > m \geqslant 1)$. We have only to consider $m' = m + 1$; then $f^pb = g^q$.

We have the equalities

$$|a| + m|b| = p|f|\,;\ |a| + (m+1)|b| = q|g|\,;$$

and the inequalities

$$(m-1)|b| < |f|\,; \qquad m|b| < |g|\,;$$

and by Corollary 4

$$(p-1)|f| < |g|\,.$$

This system of equalities and inequalities has only the following solutions:

$$m = 1\,, \qquad p = 2\,, \qquad q = 3\,;$$

or

$$m = 2\,, \qquad p = 3\,, \qquad q = 2\,;$$

$$m = 1\,, \qquad q = 2\,, \qquad p \text{ arbitrary}\,;$$

or

$$m \text{ arbitrary}, \qquad p = q = 2\,.$$

The first two solutions contradict the hypothesis $|a| \geqslant |b|$.

There remain the following cases to consider:

$$ab = f^p\,; \qquad ab^2 = g^2 \quad \text{and} \quad ab^m = f^2\,; \qquad ab^{m+1} = g^2\,.$$

The first case can be treated in the same way as the case $a^2b = f^2$ above. In the second case we can assume $m > 1$. Set $f = cb$, $g = db$, giving

$$ab^{m-1} = cbc\,, \qquad ab^m = dbd\,.$$

The relation

$$cbcb = dbd$$

cuts b into two words of equal length which we can show to be equal, and this contradicts the hypothesis of primitiveness.

References

1. Fine, N. J. and Wiff, H. S. "Uniqueness Theorems for Periodic Functions," *Proc. American Math. Soc.*, **16** (1965), 109–114.
2. Lentin, A. "Sur l'Equation $a^M = b^N c^P d^Q$ dans un Monoide Libre," *C. R. Academie Sci.*, **260** (1965), 3242–3244.
3. Lyndon, R. C. and Schützenberger, M. P. "On the Equation $a^M = b^N c^P$ in a Free Group," *Michigan Math. J.*, **9** (**1962**), 289–298.

Reprinted from JOURNAL OF ALGEBRA Vol. 13, No. 2, October 1969

Printed in Belgium

Rational Sets in Commutative Monoids*

SAMUEL EILENBERG

Department of Mathematics, Columbia University, New York, New York 10027

AND

M. P. SCHÜTZENBERGER

Faculté des Sciences de Paris, Paris, France

Communicated by Saunders MacLane

Received August 3, 1969

1. RATIONAL SETS

Let M be a monoid, i.e. a set with an associative multiplication and a two sided unit. The class of *rational* subsets of M is the least class $\mathscr{E}$ of subsets of M satisfying the following conditions:

(1R) The empty set is in $\mathscr{E}$;

(2R) Each single element set is in $\mathscr{E}$;

(3R) If $X, Y \in \mathscr{E}$ then $X \cup Y \in \mathscr{E}$;

(4R) If $X, Y \in \mathscr{E}$ then $XY \in \mathscr{E}$;

(5R) If $X \in \mathscr{E}$ then $X^* \in \mathscr{E}$.

We recall that

$$XY = \{m \mid m = xy, x \in X, y \in Y\},$$
$$X^* = \text{submonoid of } M \text{ generated by } X.$$

Kleene's theorem asserts that if M is free and finitely generated, then the rational sets are precisely the subsets of M recognizable by finite state automata.

Inspired by the notion of unambiguous context-free languages as introduced by Chomsky, we define the smaller class of *unambiguously rational* subsets of M by leaving conditions (1R) and (2R) as they are but replacing conditions (3R)–(5R) by stronger conditions (3UR)–(5UR) as follows:

(3UR) If $X, Y \in \mathscr{E}$ and $X \cap Y = \varnothing$ then $X \cup Y \in \mathscr{E}$;

(4UR) If $X, Y \in \mathscr{E}$ and the product XY is unambiguous (i.e.,

* Work supported in part by contracts NONR 266(57) and AF 61(052)-945.

$x_1y_1 = x_2y_2$ for $x_1, x_2 \in X, y_1, y_2 \in Y$ implies $x_1 = x_2, y_1 = y_2$), then $XY \in \mathscr{E}$;

(5UR) If $X \in \mathscr{E}$ and X is the basis of a free submonoid X^* of M, then $X^* \in \mathscr{E}$.

UNAMBIGUITY THEOREM. *In a free monoid M every rational set is unambiguously rational.*

This theorem is stated here for background only as it will not be used in the sequel.

The conclusion of the theorem is false without the assumption of freeness. Indeed let M be the monoid obtained from the free monoid on three generators by collapsing to a single point the ideal $I = \{uwwv \mid w \neq 1\}$. In M every cyclic submonoid is finite and therefore every unambiguously national set is finite. However $X \backslash I$ is known to be infinite [*1*, p. 30, Satz 18] so that M is infinite. Since M is finitely generated, it is rational without being unambiguously so.

For future use, we tabulate here some elementary properties of rational sets.

(1.1) If X is a rational subset of M, then there exists a finitely generated submonoid M' of M containing X.

(1.2) M is a rational subset of itself if and only if it is finitely generated.

(1.3) If $\varphi : M' \to M$ is a morphism of monoids and X' is a rational subset of M', then $X = \varphi X'$ is a rational subset of M.

(1.4) If $\varphi : M' \to M$ is a surjective morphism of monoids and if X is a rational subset of M, then there is a rational subset X' of M' such that $X = \varphi X'$.

(1.5) If X_1, X_2 are rational subsets of M_1, M_2 respectively, then $X_1 \times X_2$ is a rational subset of $M_1 \times M_2$.

2. COMMUTATIVE MONOIDS

The objective of this paper is to study rational subsets in commutative monoids. We shall use additive notation throughout. In line with this in conditions (4R) and (4UR), XY is to be replaced by $X + Y$.

The main result of this paper is that in all commutative monoids, rational sets are unambiguously rational (Theorem IV below).

The study of rational sets in a commutative monoid M is simplified by the following notions. A subset

$$X = a + B^* \tag{2.1}$$

with $a \in M$, $B \subset M$, B finite, is called *linear*. Here and in the sequel we write $a + B^*$ instead of $\{a\} + B^*$. If further the sum in (2.1) is unambiguous and the elements of B are linearly independent (i.e., B^* is a free commutative monoid with basis B), then X is called *simple*. If $B = \{b_1, ..., b_r\}$ is a set of r elements, then every element $x \in X$ may be written as

$$x = a + n_1 b_1 + \cdots + n_r b_r$$

with $n_i \in N$ (i.e., $n_i \geqslant 0$). If X is simple, then $n_1, ..., n_r$ are unique.

A finite union of linear sets is called *semi-linear*. A finite *disjoint* union of simple sets is called *semi-simple*.

Clearly every semi-linear set is rational and every semi-simple set is unambiguously rational. The converse is also true. To see that, let $\mathscr{E}$ (respectively $\mathscr{E}'$) denote the class of semi-linear (respectively semi-simple) subsets of M. Clearly $\mathscr{E}$ satisfies conditions (1R), (2R) and (3R) while $\mathscr{E}'$ satisfies conditions (1R), (2R), and (3UR). Next assume that

$$X = \cup (a_i + B_i^*), \qquad Y = \cup (c_j + D_j^*) \tag{2.2}$$

the unions being finite as well as the sets B_i, D_j. Then

$$X + Y = \bigcup_{i,j} [a_i + c_j) + (B_i \cup D_j)^*] \tag{2.3}$$

which shows that $X + Y$ is semi-linear. If in (2.2) X and Y are given in semi-simple decompositions and if the sum $X + Y$ is unambiguous, then the union in (2.3) is disjoint and the sets in brackets are simple. Thus $X + Y$ is semi-simple.

Next note that $X^* = E^*$ where $E = \bigcup [\{a_i\} \cup B_i]$. Thus $X^* \in \mathscr{E}$ so that $\mathscr{E}$ satisfies condition (5R).

Suppose now that the decomposition of X given in (2.2) is semi-simple and that X is the basis of a free submonoid X^* of M. Since M is commutative, it follows that X is a single point and X^* is simple. This concludes the argument.

3. The Main Results

We recall that a congruence Q in a monoid M is an equivalence relation in M which when viewed as a subset of $M \times M$ is a submonoid.

Theorem I. *Every congruence Q in a finitely generated commutative monoid M has a rational cross-section*; i.e., *a rational set containing exactly one element from each equivalence class* mod Q.

THEOREM II. *Every congruence Q in a finitely generated commutative monoid M is a rational subset of $M \times M$.*

THEOREM III. *If X and Y are rational subsets of a commutative monoid M, then their intersection $X \cap Y$ and difference $Y \backslash X$ also are rational subsets of M.*

THEOREM IV. *In a commutative monoid M every rational set is unambiguously rational.*

Sections 4 and 5 are devoted to preparations. Theorem I is proved in Section 6. In Section 7 the important notion of a slice is introduced and Theorem II is proved in Section 8. After more preparation in Section 9, Theorems III and IV are proved in Section 10. The proofs of these theorems for finitely generated free monoids are independent of Theorems I and II. Theorems I and II are used to pass to arbitrary commutative monoids. Sections 11–14 are devoted to corollaries, counterexamples, and other applications.

Theorem III, in the case of finitely generated free commutative monoids, was proved by Ginsburg and Spanier [2]. Some of their arguments are reproduced here in order to make this paper entirely self-contained.

4. ORDER PROPERTIES OF N^k

We denote by Z^k the free commutative group on k letters. The elements of Z^k are then n-tuples $x = (x_1, ..., x_k)$ of integers. The conditions $0 \leqslant x_i$, $i = 1, ..., k$, determine the submonoid N^k which is the free commutative monoid on k letters.

In N^k we define the (partial) order $x \leqslant y$ by the condition $x_i \leqslant y_i$ for $i = 1, ..., k$. We shall write $x < y$ if $x \leqslant y$ and $x \neq y$.

Let $X \subset N^k$ and let $y \in N^k$. The sets

$$X^y = \{x \mid x \in X, y \leqslant x\}, \qquad X_y = \{x \mid x \in X, y \textit{ non} \leqslant x\}$$

will be called the *upper* and the *lower part* of X relative to y. Clearly

$$X^y = y + (X - y) \tag{4.1}$$

where $X - y = \{z \mid z \in N^k, z + y \in X\}$. The lower part X_y will be decomposed into disjoint components as follows.

Consider all pairs

$$(i, s), 1 \leqslant i \leqslant k, 0 \leqslant s < y_i \tag{4.2}$$

and define

$$y_{is} = (y_1, \ldots, y_{i-1}, s, 0, \ldots, 0) \in N^k.$$

If $x \in X_y$, then there exists exactly one pair (i, s) such that

$$y_t \leqslant x_t \quad \text{for} \quad t < i, \qquad x_i = s < y_i$$

or equivalently

$$x = y_{is} + x' \qquad \text{with} \qquad x' \in N^k, x_i' = 0.$$

If we denote by N_i^k the submonoid of N^k determined by the condition $x_i = 0$, then we find that X_y is the disjoint union of the sets

$$y_{is} + (X - y_{is}) \cap N_i^k. \tag{4.3}$$

These are the *components* of X_y.

This decomposition of X according to an element $y \in N^k$ will be used systematically as a tool in the proofs. As a first example, we prove (the well known)

PROPOSITION 4.1. *Every set X in N^k of mutually incomparable elements is finite.*

Proof. Let $y \in X$. Since the elements of X are incomparable, we have $X - y = \{0\}$ and thus

$$X^y = y + (X - y) = \{y\}.$$

For any (i, s), the set $X - y_{is}$ is composed of mutually incomparable elements and thus the same holds for $(X - y_{is}) \cap N_i^k$. Thus by recursion, this set is finite. Consequently, all the components of (4.3) are finite and so is X.

PROPOSITION 4.2. *For any subset X of N^k, the set V of minimal elements of X is finite and $X \subset V + N^k$.*

The finiteness of V follows from Proposition 4.1. The inclusion is clear.

5. IDEALS

A subset I of a commutative monoid M is called an *ideal* if $I + M \subset I$.

PROPOSITION 5.1. *In a finitely generated monoid M every ideal I has the form $F + M$ where F is a finite subset of M.*

Proof. Since M is finitely generated, there exists a surjective morphism $\varphi : N^k \to M$. For every ideal I in M the set $\varphi^{-1}I$ is an ideal in N^k. Thus it suffices to consider $M = N^k$.

Let then I be an ideal in N^k and let F be the set of minimal elements in I. Then by Proposition 4.2, F is finite and $I \subset F + N^k$. Since I is an ideal, we have $F + N^k \subset I + N^k \subset I$. Thus $I = F + N^k$.

PROPOSITION 5.2. *In a finitely generated commutative monoid M the ideals satisfy the ascending chain condition.*

Proof. Let $I_1 \subset I_2 \subset \cdots \subset I_n \subset \cdots$ be ideals in M, and let $I = \bigcup I_n$. Then I is an ideal in M. Thus $I = F + M$ where F is a finite subset. Consequently, $F \subset I_n$ for some integer n. Thus $I \subset I_n$ and $I = I_n$.

6. PROOF OF THEOREM I

PROPOSITION 6.1. *Every congruence Q in N^k has a cross-section A such that $N^k \backslash A$ is an ideal.*

Proof. In N^k we consider the lexicographic order $x \prec y$ given by $x = y$ or

$$x_1 = y_1, \ldots x_{i-1} = y_{i-1}, \quad x_i < y_i$$

for some $i = 1, \ldots, k$. We note that this is a well-ordering of N^k satisfying

$$x \prec y \Leftrightarrow x + z \prec y + z$$

for any $x, y, z \in N^k$.

Given a congruence Q in N^k and given $x \in N^k$ let ρx denote the smallest element in the lexicographic order such that $\rho x \sim x \pmod{Q}$. We thus have

$$\rho x \prec x, \; x \sim \rho x, \quad \rho\rho x = \rho x.$$

Let $x, y \in N^k$. Then $\rho(x + y) \sim x + y \sim \rho x + \rho y$ and therefore

$$\rho(x + y) \prec \rho x + \rho y. \tag{6.1}$$

Let

$$A = \{x \mid x \in N^k, x = \rho x\}.$$

Clearly A is a cross-section for Q. Let $x \in N^k \backslash A, y \in N^k$. Then $\rho x \prec x, \rho x \neq x$. Consequently

$$\rho(x + y) \prec \rho x + \rho y \prec x + \rho y \prec x + y.$$

Since $\rho x \neq x$, we have $\rho x + \rho y \neq x + \rho y$. Therefore, $\rho(x + y) \neq x + y$ so that $x + y \in N^k \backslash A$. Thus $N^k \backslash A$ is an ideal.

Proof of Theorem I. Let Q be a congruence in a finitely generated monoid M. Choose a surjective morphism $\varphi : N^k \to M$ and let Q' be the congruence in N^k defined by

$$x \sim y \Leftrightarrow \varphi x \sim \varphi y.$$

Let A be the cross-section for Q' as given by Proposition 6.1. Then clearly φA is a cross-section for Q. To show that φA is rational, it suffices to show that A is a rational subset of N^k. Since $I = N^k \backslash A$ is an ideal, it follows from Proposition 5.1 that I is rational. Since $A = N^k \backslash I$ the rationality of A follows from Theorem III. The reader will have to be careful to note that Theorem I is not used until *after* Theorem III has been proved.

7. Subtractive Submonoids and Slices

Let S be a submonoid of a commutative monoid M. We shall say that S is *subtractive*, if $x, x + y \in S$, $y \in M$, imply $y \in S$. This may equivalently be rephrased as $S - S \subset S$ or $S - S = S$.

Proposition 7.1. *Every subtractive submonoid S of a finitely generated monoid M is itself finitely generated.*

Proof. Let $\varphi : N^k \to M$ be a surjective morphism and let $S' = \varphi^{-1}S$. Then S' is a subtractive submonoid of N^k and if S' is finitely generated, then so is $S = \varphi S'$. Thus we may assume that $M = N^k$.

Let A be the set of all minimal elements in $S \backslash \{0\}$. Then A is finite and $A^* \subset S$. Assume $A^* \neq S$ and let x be a minimal element of the set $S \backslash A^*$. Then $a \leqslant x$ for some element $a \in A$, so that $x = a + y$ with $y \in N^k$. Since S is subtractive, we have $y \in S$. Since $y < x$ we have $y \in A^*$. Thus $x = a + y \in A^*$, a contradiction.

A *slice* in a commutative monoid M is a subset S such that $s, s + x$, $s + y \in S$ imply $s + x + y \in S$. Equivalently, S is a slice if and only if for every $s \in S$ the set $S - s$ is a submonoid of M. An element s of a slice S is called *stable* if the submonoid $S - s$ is subtractive.

Proposition 7.2. *Every slice S in N^k has a stable element.*

Proof. For every $s \in S$, consider the ideal

$$I_s = [(S - s) \backslash \{0\}] + N^k.$$

Let $s \leqslant s'$, $s' \in S$, and let $y \in S - s$. Then $s' = s + x$ and $s + y \in S$. Thus $s' + y = s + x + y \in S$ or equivalently $y \in S - s'$. Thus $S - s \subset S - s'$ so that $I_s \subset I_s'$.

Since the ideals in N^k satisfy the ascending chain condition (Proposition 5.2), there exists $s \in S$ such that $I_s' = I_s$ for every $s' \in S$, $s \leqslant s'$. We shall show that such an s is stable.

Indeed, let $x, x + y \in S - s$. If $y = 0$ then $y \in S - s$ and we are finished. Thus, we may assume $y \neq 0$. We have $s + x \in S$ and $s + x + y \in S$ so that $y \in S - (s + x)$. Since $y \neq 0$ we have $y \in I_{s+x} = I_s$. Consequently, $y = u + w$ with $u \in S - s$, $w \in N^k$, $u \neq 0$. Further, we may choose such a decomposition of y with a shortest possible w. If $w = 0$ then $y \in S - s$ and we are finished. Thus, we may assume $w \neq 0$. Since $y + s \in S$, we have $u + w + s \in S$; i.e., $w \in S - (s + u)$. Since $w \neq 0$ and $s + u \in S$ we have $w \in I_{s+u} = I_s$. Thus $w = u' + w'$ with $u' \in S - s$, $u' \neq 0$. Then

$$y = u + u' + w' \qquad \text{with} \qquad u + u' \in S - s.$$

Thus contradicts the assumption that w was the shortest possible.

Proposition 7.3. *A slice S in a finitely generated monoid M is a rational subset of M.*

Proof. Let $\varphi : N^k \to M$ be a surjective morphism. Then $\varphi^{-1}S$ is a slice in N^k, and it suffices to prove the rationality of $\varphi^{-1}S$. Thus we may assume $M = N^k$. By Proposition 7.2, S contains a stable element y. We decompose S according to the element y. The upper part is

$$y + (S - y). \tag{7.1}$$

Since $S - y$ is a subtractive submonoid of N^k, it is finitely generated by Proposition 7.1. Thus $S - y$ is rational and so is (7.1). The components of the lower part are

$$y_{is} + (S - y_{is}) \cap N_i^k. \tag{7.2}$$

Each of the sets $S - y_{is}$ is a slice. Indeed, if $s, s + x, s + z \in S - y_{is}$, then $s + y_{is}, s + y_{is} + x, s + y_{is} + z \in S$. Then $s + y_{is} + x + z \in S$ so that $s + x + z \in S - y_{is}$. Consequently $(S - y_{is}) \cap N_i^k$ also is a slice. This slice being in N^{k-1} we may assume by induction that it is rational. Thus (7.2) is rational, and therefore S is rational.

8. Proof of Theorem II

The theorem follows directly from Proposition 7.3 in view of

Proposition 8.1. *A congruence Q in a commutative monoid M is a slice in $M \times M$.*

Proof. Let $(x_1, x_2), (x_1 + y_1, x_2 + y_2), (x_1 + z_1, x_2 + z_2) \in Q$. Then

$$x_1 + y_1 + z_1 \sim x_2 + y_2 + z_1 \sim x_1 + y_2 + z_1 \sim x_2 + y_2 + z_2 .$$

Thus $(x_1 + y_1 + z_1, x_2 + y_2 + z_2) \in Q$.

9. Preparation for Theorems III and IV

Proposition 9.1 (Ginsburg–Spanier). *If $X = a + B^*$ is a linear subset of Z^k, then X is the finite union of simple sets $c + D^*$ with $D \subset B$.*

Proof. Without loss of generality, we may assume $a = 0$. Let $B = \{b_1, ..., b_p\}$. If the elements of B are linearly independent, then X is simple and there is nothing to prove. Thus we may assume that

$$t_1 b_1 + \cdots + t_q b_q = t_{q+1} b_{q+1} + \cdots + t_p b_p$$

for some

$$0 < q < p, \quad (t_1, ..., t_p) \in N^p, \quad t_1 > 0.$$

For $j = 1, ..., q$ define

$$A_j = \{s b_j \mid 0 < s < t_j\}, \qquad B_j = B \backslash \{b_j\}$$

$$Y_j = A_j + B_j^*, \qquad Y = \bigcup Y_j$$

Arguing by induction on p, it suffices to show that $X = Y$. Clearly $A_j \subset X$ and $B_j \subset X$. Since X is a submonoid of Z^k, it follows that $Y_j \subset X$ and thus $Y \subset X$.

Let $d \in X$, $d = \sum r_i b_i$, $r_i \geqslant 0$. If $t_j \leqslant r_j$ for all $j = 1, ..., q$ then we may rewrite d as

$$d = \sum_{j \leqslant q} (r_j - t_j) b_j + \sum_{q < l} (r_l + t_l) b_l$$

and thereby diminish the sum $\sum_{j \leqslant q} r_j$. Thus we may assume that $r_j < t_j$ for some $j \leqslant q$. Then $d = r_j b_j + d'$ for some $d' \in B_j^*$. Thus $d \in Y_j$.

Lemma 9.2. *Let M_i, $i = 1, ..., p$, be commutative monoids and let $M = M_1 \times \cdots \times M_p$. Let $X_i \subset M_i$ be semi-simple subsets of M_i such that $Y_i = M_i \backslash X_i$ also are semi-simple. Then $X = X_1 \times \cdots \times X_p$ is a semi-simple subset of M and $M \backslash X$ also is semi-simple.*

Proof. By induction, it suffices to consider the case $p = 2$. We regard M_1 and M_2 as submonoids of M in the obvious way. Then for

$$a_1 + B_1^* \subset M_1, \qquad a_2 + B_2^* \subset M_2$$

we have

$$(a_1 + B_1^*) \times (a_2 + B_2^*) = (a_1 + a_2) + (B_1 \cup B_2)^*.$$

This shows that $X = X_1 \times X_2$ is semi-simple. Further $M\backslash X$ is the disjoint union

$$Y_1 \times Y_2 \cup X_1 \times Y_2 \cup Y_1 \times X_2$$

so that $M\backslash X$ also is semi-simple.

Since $N = 1^*$ and $Z\backslash N = (-1) + (-1)^*$ we see that N and $Z\backslash N$ are simple subsets of Z. Also $N\backslash\{0\} = 1 + N$ is simple. Thus Lemma 9.2 yields the semi-simplicity of the following subsets of Z^k:

$$Z^k, \quad Z^k\backslash N^k, \quad N^p \times N^q\backslash N^p \times 0 \qquad \text{for} \quad p + q = k. \tag{6.1}$$

LEMMA 9.3. *If X is a simple subset of Z^k, then $Z^k\backslash X$ is semi-simple.*

Proof. Let $X = a + B^*$ with B a linearly independent subset of Z^k. Since $Z^k\backslash X = a + (Z^k\backslash B^*)$ it suffices to consider the case $a = 0$, $X = B^*$.

Let $B = \{b_1, ..., b_p\}$. First consider the case $p = k$. Let $B°$ be the subgroup of Z^k. generated by B. Then

$$Z^k\backslash B^* = (Z^k\backslash B°) \cup (B°\backslash B^*).$$

Since the union is disjoint, it suffices to show that each component is semi-simple. Since $b_1, ..., b_k$ are linearly independent, $B°$ is isomorphic with Z^k under an isomorphism mapping B^* onto N^k. Since by (6.1), $Z^k\backslash N^k$ is semi-simple, it follows that $B°\backslash B^*$ is semi-simple.

Next consider $Z^k\backslash B°$. Since $p = k$, the quotient group $Z^k\backslash B°$ is finite. Therefore $Z^k\backslash B°$ is a disjoint finite union of cosets $c + B°$. Thus it suffices to show that $c + B°$ is semi-simple. For this it is enough to show that $B°$ is semi-simple. However, $B° \approx Z^k$, so the conclusion follows from (6.1).

Next assume $k = p + q$, $q > 0$. We can then find elements $b_{p+1}, ..., b_k$ so that the set $C = \{b_1, ..., b_k\}$ is linearly independent. Then

$$Z^k\backslash B^* = (Z^k\backslash C^*) \cup (C^*\backslash B^*).$$

By the above, $Z^k\backslash C^*$ is semi-simple. The monoid C^* is isomorphic with $N^k = N^p \times N^q$ under an isomorphism carrying B^* onto $N^p \times 0$. Thus by (6.1), $C^*\backslash B^*$ is semi-simple.

LEMMA 9.4. *Given a morphism $\varphi : N^k \to Z^m$ and an element $c \in Z^m$ the set*

$$X = \{x \mid x \in N^k, \varphi x = c\}$$

is semi-simple.

Proof. If X is finite, then it clearly is semi-simple. If X is infinite, then by Proposition 4.1, there exists elements $x, x' \in X$ with $x < x'$. Thus $x + y = x'$ for some $y \in N^k$, $y \neq 0$. Thus we have $y \neq 0$, $\varphi y = 0$. For any $x \in X$ we have a unique representation

$$x = ny + z, \quad n \in N, \quad z \in X, \quad y \text{ non} \leqslant z.$$

We thus have the unambiguous sum $X = y^* + X_y$ so it suffices to prove that X_y is semi-simple. The (disjoint) components of the lower part of X are

$$y_{is} + (X - y_{is}) \cap N_i^k.$$

Since

$$(X - y_{is}) \cap N_i^k = \{x \mid x \in N_i^k, \varphi x = c + \varphi y_{is}\}$$

these sets are semi-simple by recursion. Thus X_y is semi-simple as required.

LEMMA 9.5. *If X and Y are semi-simple subsets of Z^k then so is $X \cap Y$.*

Proof. We may assume that X and Y are simple subsets of Z^k. Then

$$X = a + \alpha N^p, \qquad Y = b + \beta N^q$$

where $\alpha : N^p \to Z^k, \beta : N^q \to Z^k$ are injective morphisms. Define the morphisms

$$\varphi : N^p \times N^q \to Z^k \qquad \varphi(x, y) = \alpha x - \beta y$$
$$\tau : N^p \times N^q \to Z^k \qquad \tau(x, y) = \alpha x$$

and let

$$\begin{aligned} W &= \{(x, y) | (x, y) \in N^p \times N^q, \varphi(x, y) = b - a\} \\ &= \{(x, y) | (x, y) \in N^p \times N^q, a + \alpha x = b + \beta y\}. \end{aligned}$$

By Lemma 9.4, W is semi-simple. Further, $a + \tau W = X \cap Y$. Since α and β are injective, it follows that τ is injective on the set W. Therefore $a + \tau W$ is semi-simple.

10. PROOF OF THEOREMS III AND IV

We first consider the case $M = Z^k$. Let then X be a rational subset of Z^k. Then X is semi-linear and by Proposition 9.1, X is a finite union

$X_1 \cup \cdots \cup X_p$ of (not necessarily disjoint) simple sets. By Lemma 9.3 each of the sets $Z^k \backslash X_i$ is semi-simple. Therefore, by Lemma 9.5 the set

$$X' = Z^k \backslash X = \bigcap (Z^k \backslash X_i)$$

is semi-simple. Since X' also is rational, by the above $X = Z^k \backslash X'$ is semi-simple. If Y is another rational subset of Z^k then Y is semi-simple and by Lemma 9.5 the sets

$$Y \cap X, \;\; Y \backslash X = Y \cap X'$$

are semi-simple.

The next case to consider is $M = N^k$. This follows from the case $M = Z^k$ in virtue of the following observation: If $a + B^* \subset N^k$ for $a \in Z^k$, $B \subset Z^k$, then $a \in N^k$ and $B \subset N^k$. Clearly $a \in N^k$. Let $b \in B$. Then $a + nb \in N^k$ for all positive integers n. This implies that all the coordinates of b are non-negative and thus $b \in N^k$.

Next, consider an arbitrary commutative monoid M and let X and Y be rational subsets of M. There exists then a finitely generated submonoid M' of M such that X, $Y \subset M'$. Hence we may assume that M is finitely generated.

Let $\varphi : N^k \to M$ be a surjective morphism and let Q be the congruence in N^k defined by

$$x \sim y \Leftrightarrow \varphi x = \varphi y.$$

Given any rational subset X of M, choose a rational set R in N^k such that $\varphi R = X$. Then note that

$$\varphi^{-1} X = \pi[Q \cap (N^k \times R)]$$

where $\pi : N^k \times N^k \to N^k$ is given by $\pi(x, y) = x$. Since by Theorem II, Q is a rational subset of $N^k \times N^k$, it follows that $\varphi^{-1}X$ is rational. If Y is another rational subset of M, then $\varphi^{-1}Y$ is rational. Since φ is surjective, we have

$$X \cap Y = \varphi[\varphi^{-1}X \cap \varphi^{-1}Y], \qquad Y \backslash X = \varphi[\varphi^{-1}Y \backslash \varphi^{-1}X]$$

so that $X \cap Y$ and $Y \backslash X$ are rational. This concludes the proof of Theorem III in full generality. We note that Theorem I has not been employed.

To complete the proof of Theorem IV we need one more step. Conserving the notation above, we apply Theorem I to obtain a rational cross-section A for Q. Then for any set $X \subset M$ we have $X = \varphi W$ where $W = A \cap \varphi^{-1}X$. If X is rational then so is $\varphi^{-1}X$. Then W is a rational subset of N^k and therefore W is semi-simple. Since $W \subset A$, φ is injective on W and thus φW also is semi-simple.

11. Corollaries of Theorem III

Corollary III.1. *In a finitely generated monoid M the class of rational sets is closed under Boolean operations.*

Indeed, M is then a rational set and therefore $M \backslash X$ is rational for every rational set X.

Corollary III.2. *If $\varphi : M' \to M$ is a morphism of commutative monoids, M' is finitely generated and X is a rational subset of M, then $\varphi^{-1}X$ is a rational subset of M'.*

Indeed, consider the morphisms

$$\mu : M' \to M' \times M, \qquad \mu x = (x, \varphi x),$$

$$\gamma : M' \times M \to M', \qquad \pi(x, y) = x.$$

Then $\varphi^{-1}X = \pi Y$ where $Y = \mu M' \cap (M' \times X)$. Since M' is a rational subset of itself, $\mu M'$ and $M' \times X$ are rational subsets of $M' \times M$. Thus, Y is rational and so is πY.

Corollary III.3. *If M' is a finitely generated submonoid of a commutative monoid M and if $X \subset M'$ is a rational subset of M, then X is also a rational subset of M'.*

In the previous Corollary, choose $\varphi : M' \to M$ to be the inclusion morphism.

Given subset X, Y of a commutative monoid M, we define

$$Y - X = \{m \mid x + m \in Y \text{ for some } x \in X\}.$$

Corollary III.4. *If X and Y are rational subsets of a finitely generated commutative monoid M, then $Y - X$ is rational.*

Indeed, consider the morphisms

$$\varphi : M \times M \to M, \qquad \varphi(x, y) = x + y,$$

$$\pi : M \times M \to M, \qquad \pi(x, y) = y.$$

Then

$$Y - X = \pi[(\varphi^{-1}Y) \cap (X \times M)],$$

which shows that $Y - X$ is rational.

It should be noted that Theorem II is itself a corollary of Corollary III.2. Indeed, if Q is a congruence in M, consider the natural morphism $\varphi : M \to R = M/Q$, and define $\psi : M \times M \to R \times R$ by $\psi(m_1, m_2) = (\varphi m_1, \varphi m_2)$. Then $Q = \psi^{-1}\Delta$ where Δ is the diagonal submonoid of $R \times R$.

Since M is finitely generated, so is R and thus also Δ. Therefore, Δ is rational and by Corollary III.2 so is $\psi^{-1}\Delta = Q$.

12. Ascending Chain Conditions

Proposition 12.1. *A rational slice S in a commutative monoid M is finitely defined*; i.e., *there exists a finite subset F of M such that S is the least slice in M containing F.*

Proof. Since S is rational, it is the finite union of sets

$$a_i + B_i{}^*,\ a_i \in M,\ B_i \subset M,\ B_i \text{ finite.}$$

Let

$$F = \bigcup \{a_i\} \cup (a_i + B_i).$$

Then $F \subset S$. Let S' be any slice in M containing F. For each index i we then have $a_i \in S'$ and $a_i + B_i \subset S'$. Thus $B_i \subset S - a_i$ and since $S - a_i$ is a submonoid, we have $B_i{}^* \subset S' - a_i$; i.e., $a_i + B_i{}^* \subset S'$. Consequently, $S \subset S'$.

This, combined with Proposition 7.3, yields

Corollary 12.2. *In a finitely generated commutative monoid every slice is finitely defined.*

An equivalent statement is

Corollary 12.3. *The slices in a finitely generated commutative monoid satisfy the ascending chain condition.*

In particular, for congruences, we obtain, by Proposition 8.1,

Theorem V. *Every congruence Q in a finitely generated commutative monoid M is finitely defined*; i.e., *there exists a finite subset F of $M \times M$ such that Q is the least congruence containing F.*

Corollary V.1. *The congruences in a finitely generated commutative monoid M satisfy the ascending chain condition.*

Given a congruence Q in a commutative monoid M and given $m \in M$, define the congruence Q_m in M by setting

$$Q_m = \{(x, y) | (x + m, y + m) \in Q\}.$$

Clearly

$$Q \subset Q_m \subset Q_{m+m'} = (Q_m)_{m'}.$$

Q is called *cancellative* if $Q = Q_m$ for all $m \in M$.

COROLLARY V.2. *Let Q be a congruence in a finitely generated commutative monoid M. The class of congruences $\{Q_m\}$, $m \in M$, contains exactly one cancellative congruence Q'. Further, $Q_m \subset Q'$ for every $m \in M$.*

The existence of Q' follows from the ascending chain condition. Since $Q \subset Q'$, we have $Q_m \subset Q_{m}' = Q'$. This implies the uniqueness of Q'.

Theorem V and its corollaries were proved by *L.* Redei [3].

Theorem V may also be deduced from the Hilbert basis theorem as follows. Let R be any commutative ring (with $0 \neq 1$). Writing the monoid M multiplicatively, construct the R-algebra $R[M]$. Given a congruence Q in M, let $I(Q)$ denote the ideal in $R[M]$ generated by elements $x - y$ with $(x, y) \in Q$. We assert that

$$Q = \{(x, y) \mid x, y \in M, x - y \in I(Q)\}. \tag{12.1}$$

Indeed, let Q' be the right-hand side of (12.1). Then Q' is a congruence in M, $Q \subset Q'$ and $I(Q') = I(Q)$. There results a commutative triangle of surjective morphisms

$$\begin{tikzcd} & M \arrow[dl, "\pi"'] \arrow[dr, "\pi'"] & \\ M/Q \arrow[rr, "\varphi"'] & & M/Q' \end{tikzcd}$$

from which we obtain the commutative triangle of surjective R-algebra morphisms

$$\begin{tikzcd} & R[M] \arrow[dl, "R[\pi]"'] \arrow[dr, "R[\pi']"] & \\ R[M/Q] \arrow[rr, "R[\varphi]"'] & & R[M/Q'] \end{tikzcd}$$

Since $I(Q) = I(Q')$ is both the kernel of $R[\pi]$ and $R[\pi']$, it follows that $R[\varphi]$ is an isomorphism. It follows that φ also is an isomorphism and thus $Q = Q'$.

It follows from (12.1) that the set of all congruences in M is mapped by $Q \to IQ$ injectively into the set of ideals in $R[M]$. If M is finitely generated and R is noetherian and $R[M]$ also is noetherian and the ideals in $R[M]$

satisfy the ascending chain condition. Thus the congruences in M also satisfy the ascending chain condition.

The above proof was known to a number of mathematicians including Peter Freyd and Michael O. Rabin.

13. Other Applications

Call a commutative monoid M *cancellative* if $x + y = x + z$ implies $y = z$.

Theorem VI. *The intersection $M_1 \cap M_2$ of two finitely generated submonoids M_1, M_2 of a cancellative commutative monoid M is a finitely generated submonoid of M.*

Proof. Choose morphisms

$$\varphi_i : N^{k_i} \to M \qquad i = 1, 2$$

such that $\varphi_i N^{k_i} = M_i$ for $i = 1, 2$. Consider the product $N^k = N^{k_1} \times N^{k_2}$, $k = k_1 + k_2$ and define

$$\psi_i : N^k \to M \qquad i = 1, 2$$

by

$$\psi_i(x_1, x_2) = \varphi_i x_i .$$

Define

$$S = \{x \mid x \in N^k, \psi_1 x = \psi_2 x\}.$$

Note that $M_1 \cap M_2 = \psi_1 S = \psi_2 S$. Thus it suffices to show that S is finitely generated. By Proposition 7.1 it, therefore, suffices to show that S is a subtractive submonoid of N^k. Let then $x, x + y \in S$. Then

$$\varphi_1 x_1 = \varphi_2 x_2 , \varphi_1 x_1 + \varphi_1 y_1 = \varphi_2 x_2 + \varphi_2 y_2 .$$

Since M is cancellative, it follows that $\varphi_1 y_1 = \varphi_2 y_2$; i.e., $y \in S$.

The conclusion of Theorem VI is false without the assumption that M is cancellative. Indeed, consider the "simplest" example of a noncancellative monoid M given by three generators x, y, z and the single relation $x + y = x + z$.

We may regard M as the quotient monoid of N^3 by the relation Q defined by the single pair $(1, 1, 0), (1, 0, 1)$. The congruence Q may be made explicit as follows:

$$(a, b, c) \sim (a', b', c')$$

if and only if either

$$a = a', \qquad b + c = b' + c'$$

or

$$a = a' = 0, \quad b = b', c = c'.$$

This implies that x, y, z are pairwise linearly independent.

Let $V = \{x, y\}^* \cap \{x, z\}^*$. Viewed as a submonoid of $\{x, y\}^*$, V may be identified with the submonoid of N^2 given as follows:

$$V = \{(a, b) | (a, b, 0) \sim (a', 0, c')\}.$$

Inspecting the congruence we see that we must have $a = a'$, $b = c'$. Thus

$$V = \{(a, b) | (a, b, 0) \sim (a, 0, b)\}.$$

Again, going back to the congruence, we see that

$$V = \{(a, b) |\ a = b = 0 \quad \text{or} \quad a > 0\}.$$

This submonoid of N^2 is not finitely generated since any generating set must contain the sequence $\{(1, n)\}$.

THEOREM VII. *Let M be a finitely generated commutative monoid, X a rational subset of M, and P a set of strictly positive integers. Then the set*

$$P^{-1}X = \{m \in M \mid pm \in X \text{ for some } p \in P\}$$

is rational

Proof. Let $\varphi : N^k \to M$ be a surjective morphism. Then $\varphi^{-1}(P^{-1}X) = P^{-1}(\varphi^{-1}X)$. Thus by Corollary IV.2, it suffices to consider the case $M = N^k$.

Since $P^{-1}(X \cup Y) = P^{-1}X \cup P^{-1}Y$, it suffices to consider the case when $X = a + B^*$ is simple with $B = \{b_1, ..., b_k\}$ linearly independent. Let

$$C = \{y \mid py = a + \sum n_i b_i\,, \quad \text{for some} \quad p \in P \quad \text{and} \quad 0 \leqslant n_i < p\}.$$

For $y \in C$ we have $y \leqslant a + \sum b_i$ and therefore the set C is finite.

Since $C \subset P^{-1}X$ and $pB^* \subset B^*$ for every p, it follows that $C + B^* \subset P^{-1}X$. Conversely, let $y \in P^{-1}X$. Then for some $p \in P$ we have $py = a + \sum q_i b_i$, $0 \leqslant q_i$. Write $q_i = n_i + r_i p$ with $0 \leqslant n_i < p$. Then

$$py = a + \sum n_i b_i + p\,(\sum r_i b_i).$$

Thus setting $b = \sum r_i b_i$, $c = y - b$ we have $c \in C, b \in B^*$, $y = c + b$. Thus $P^{-1}X = C + B^*$ and $P^{-1}X$ is rational.

The conclusion of Theorem VII is false without the assumption of finite generation. Indeed, let M be an infinite set with distinguished elements 0, w, $0 \neq w$. Define $0 + x = x = x + 0$ and $x + y = w$ if $x \neq 0 \neq y$. Then M is a commutative monoid and $X = \{w, 0\}$ is a rational subset of M. Taking $P = \{2\}$ we have $P^{-1}X = M$ which is not rational.

14. Some Counterexamples

We first show that the hypothesis of finite generation is essential in all the theorems and corollaries in which it is made.

In connection with Theorems I and II, consider a monoid M. Then the only cross-section for the congruence $Q = \{(x, x)\}$ is M. Thus Q does not have a rational cross-section. On the other hand, the congruence $Q = M \times M$ is not rational.

If M is a commutative monoid which is not finitely generated, then $\varnothing$ is rational subset of M while $M = M \backslash \varnothing$ is not. This shows that Corollary III.1 fails.

In N^2 consider the submonoid

$$Q = \{(0, 0)\} \cup [(1, 1) + N^2].$$

Clearly Q is a rational subset of N^2. However, Q is not finitely generated, as indeed any generating set for Q must contain all the elements $(n, 1)$ and $(1, n)$ for $n = 1, 2, \ldots$. Therefore Q is not a rational subset of itself. Therefore, Corollaries III.3 and III.2 fail. Incidentally, Q is a congruence in N and is defined by the single pair $(1, 2)$. This shows that for a congruence Q, "finitely generated" is a much stronger notion than "finitely defined."

For Corollary III.4, consider a commutative monoid M which is not finitely generated and which contains an element w such that $M + w = w$. Then taking $X = Y = \{w\}$ we have $Y \backslash X = M$. Thus X and Y are rational while $Y \backslash X$ is not.

For Theorem V and its corollaries, consider the free commutative monoid M generated by the letters x_i, y_i, z_i, $i = 1, 2, \ldots$, and consider the congruence Q defined by the set of pairs $(x_i + z_i, y_i + z_i)$, $i = 1, 2, \ldots$. Then for any $m \in M$ the congruence Q_m is not cancellative. Therefore, Corollary V.2 fails in M, and therefore, also, Corollary V.1 and Theorem V.

To conclude, we give an example of a submonoid M of N^2 which is not rational. Let

$$M = \{(x, y) \mid y \leqslant x^2\}.$$

Consider the "slope" function defined on M by setting

$$\varphi(0, 0) = 0$$

$$\varphi(x, y) = \frac{y}{x} \quad \text{if} \quad x \neq 0.$$

Then φ is unbounded on M. However, on every set $a + B^*$ with $a \in N^2$, $B \subset N^2$, B finite and $a + B^* \subset M$ the function φ is bounded. Thus φ is bounded on any rational subset of N^2 which is in M. Consequently, M is not rational.

In the same way, we can show that for every irrational number $r > 0$ the submonoid

$$M = \{(x, y) \in N^2, y \leqslant rx\}$$

is not rational.

REFERENCES

1. Thue, Axel. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Videnskapsselskapets Skrifter I Mat. Naturv. Klasse* **1912**, 1–67.
2. Ginsburg and Spanier. Bounded ALGOL-like languages. *Trans. Am. Math. Soc.* **113** (1964), 333–368.
3. Redei, L. "The Theory of Finitely Generated Commutative Semi-Groups," Theorems 72, 95 and 60. Oxford–Edinburgh–New York, 1965.

PRINTED IN BRUGES, BELGIUM, BY THE ST. CATHERINE PRESS, LTD.

SÉMINAIRE DUBREIL. ALGÈBRE ET THÉORIE DES NOMBRES

MARCEL P. SCHÜTZENBERGER

Langages formels et monoïdes finis

Séminaire Dubreil. Algèbre et théorie des nombres, tome 23, n° 2 (1969-1970), exp. n° DG 3, p. DG 1-DG 3.

<http://www.numdam.org/item?id=SD_1969-1970__23_2_A3_0>

Séminaire DUBREIL-PISOT 3-01
(Algèbre et Théorie des nombres)
23e année, 1969/70
Demi-groupes, n° 3, 3 p.

LANGAGES FORMELS ET MONOÏDES FINIS

par Marcel P. SCHÜTZENBERGER

L'une des problématiques de ce que l'on appelle la théorie des automates et des langages formels, consiste en l'étude des relations qu'entretiennent certaines parties d'un monoïde libre, X^*, et les monoïdes abstraits définis par le monoïde syntactique de ces dernières. (Rappelons ici que si A est une partie d'un monoïde M, le monoïde syntactique $M//A$ est le quotient de M par la plus grande congruence dont A est une union de classes.)

La légitimité de ce projet repose sur le théorème de Kleene qui caractérise de façon très remarquable les parties dont le monoïde syntactique est fini (les parties "reconnaissables" au sens de S. EILENBERG). Pour l'énoncer, appelons rationnelle la plus petite famille de parties $\mathcal{R} = \mathrm{Rat}(M)$ d'un monoïde M qui satisfasse les trois conditions suivantes :

(1) $\{\emptyset\} \in \mathcal{R}\ ;\quad m \in M \Longrightarrow \{m\} \in \mathcal{R}$.

(2) $A\ ,\ B \in \mathcal{R} \Longrightarrow A \cup B \in \mathcal{R}$ et $AB \in \mathcal{R}$.

(3) $A \in \mathcal{R} \Longrightarrow A^* \in \mathcal{R}$,

où A^* désigne le sous-monoïde engendré par A .

Ceci posé, on a :

THÉORÈME (KLEENE). - *Soit X^* un monoïde libre finiment engendré. $\mathrm{Rat}(X^*)$ est la famille des parties reconnaissables de X^** .

Il est important de souligner le rôle critique joué par l'hypothèse que X^* est un monoïde libre (finiment engendré). Sans celle-ci, le théorème n'est plus valable, et il en résulte que la théorie générale des objets rationnels et reconnaissables, qu'a développée S. EILENBERG, n'a que bien peu à faire de la notion de monoïde syntactique. Pour cette raison, je me limiterai ici au cas du monoïde libre X^* .

Mise à part cette direction de recherches, deux séries de questions se posent selon que l'on part du monoïde fini $X^*//A$, ou au contraire des parties $A \subset X^*$.

La première voie est la plus évidente : elle consiste à étudier comment les propriétés algébriques classiques (idéaux, groupes, etc.) d'un monoïde fini M , se reflètent dans la structure des parties $A \subset X^*$ dont M est le monoïde syntactique. Par exemple, McNAUGHTON et TRACHTENBROT ont considéré celles dont tous les groupes de M sont triviaux. On peut vérifier que cette sous-famille $\mathcal{R}' \subset \underline{Rat}(X^*)$ est définie par les conditions (1) et (2) ci-dessus et la nouvelle condition :

(4) $$A \in \mathcal{R}' \implies X^* \setminus A \in \mathcal{R}' ,$$

remplaçant (3).

De même, l'on peut se demander quelles sont les parties de X^* dont le monoïde syntactique est un groupe, ou un monoïde de Clifford, ou un monoïde inverse, etc. Il n'est pas surprenant que les concepts fondamentaux introduits par P. DUBREIL, dans son mémoire de 1940, jouent un rôle important dans ces recherches. Il est d'ailleurs intéressant de noter que les monoïdes infinis les plus remarquables (groupe libre, monoïde bicyclique) se rencontrent comme monoïdes syntactiques de certaines parties de X^* qui jouent un rôle de base dans la théorie de la famille $\underline{Alg}(X^*)$ des parties dites <u>algébriques</u> de X^*, généralisation naturelle de $\underline{Rat}(X^*)$.

A ces questions se rattache, je crois, la discussion des décompositions de M par produit direct et produit en couronne, qu'a si brillamment effectuée J. RHODES.

Je passe maintenant à la deuxième des voies de recherches, que j'indiquais plus haut, et qui offre peut-être des points de vue plus originaux par rapport à la théorie algébrique classique.

En effet, la structure de monoïde libre de X^* privilégie certaines notions, dont la contre-partie dans les monoïdes abstraits est assez cachée.

Par exemple, il paraît naturel d'étudier les monoïdes syntactiques des sous-monoïdes de X^* qui sont <u>finiment engendrés</u>, et il est curieux que ceux-ci ne soient des groupes que s'ils sont cycliques.

D'autre part, on peut faire intervenir la notion de partie (rationnelle) <u>non ambigüe</u> en restreignant les unions aux parties disjointes, les produits AB au cas où

$$a , a' \in A , \quad bb' \in B , \quad ab = a'b' \implies a = a' , \quad b = b' ,$$

et les sous-monoïdes A^* au cas où A^* est <u>libre</u> et librement engendré par A.

Le résultat fondamental, dû à N. CHOMSKY, est que toute partie rationnelle de X^* peut être présentée comme une partie rationnelle non ambigüe. Le même énoncé vaut pour les monoïdes commutatifs (S. EILENBERG et M. P. SCHÜTZENBERGER).

Combinant maintenant les deux conditions d'engendrement fini et de non-ambigüité, l'on se trouve amené à étudier en profondeur les sous-monoïdes libres finiment engendrés de X^* et leurs monoïdes syntactiques. Des résultats forts surprenants ont été obtenus par J.-F. PERROT, qui a montré que les groupes de Suschkewitsch de ces derniers étaient extrêmement particuliers, et il y a là, je crois, tout un domaine important de recherches.

Dans une direction voisine, G. VIENNOT a considéré les factorisations non ambigües de X^* en produits de sous-monoïdes libres, $X^* = A^* B^*$; ou plus généralement, $X^* = A^* B^* \ldots C^*$.

Cette question apparait de façon naturelle dans bien d'autres structures algébriques. Par exemple, en utilisant une méthode due à M. LAZARD, on peut la rattacher au problème de la recherche des bases monomiales des algèbres de Lie libres (et réciproquement).

Elle joue de façon évidente un rôle important dans la théorie des décompositions de Rhodes, et elle semble sous-jacente à l'étude de certaines relations d'ordre sur les représentations des monoïdes syntactiques, c'est-à-dire, suivant la terminologie d'usage, des automates finis.

Marcel P. SCHÜTZENBERGER
Prof. Fac. Sc. Paris
9 villa Poirier
75 - PARIS 15

(Texte reçu le 3 novembre 1970)

Année 1970

Bibliographie

[1] Dominique Foata and Marcel-Paul Schützenberger. *Théorie géométrique des polynômes eulériens.* Lecture Notes in Mathematics, Vol. 138. Springer-Verlag, Berlin, 1970.

[2] Marcel-Paul Schützenberger. Une application de la théorie de la décomposition des monoïdes. In *Séminaire M. P. Schützenberger, A. Lentin et M. Nivat, 1969/70 : Problèmes Mathématiques de la Théorie des Automates*, Exposé No. 5, 9 décembre 1969, 4 pages. Secrétariat mathématique, Paris, 1970.

[3] Dominique Foata and Marcel-Paul Schützenberger. On the rook polynomials of Ferrers relations. In *Combinatorial Theory and its Applications, II (Proc. Colloq., Balatonfüred, 1969)*, pages 413–436. North-Holland, Amsterdam, 1970.

[4] Marcel-Paul Schützenberger. Sur les contraintes définissant certains modèles formels de langage. In *Cahiers mathématiques*, volume 3, algèbre et combinatoire, pages 41–48. Gauthier-Villars, Paris, 1970. reprise de l'article de 1963.

[5] Marcel-Paul Schützenberger. Health aspects of chemical and biological weapons. Report of a WHO group of consultants, World Health Organisation, 1970.

[6] Marcel-Paul Schützenberger. Santé publique et armes chimiques et biologiques. Rapport d'un groupe de consultants de l'OMS, Organisation Mondiale de la Santé, 1970. Version française.

Lecture Notes in Mathematics

A collection of informal reports and seminars
Edited by A. Dold, Heidelberg and B. Eckmann, Zürich

Series: Institut de Mathématique, Université de Strasbourg
Advisers: P. A. Meyer and M. Karoubi

138

Dominique Foata
Marcel-P. Schützenberger

Théorie Géométrique des Polynômes Eulériens

Springer-Verlag
Berlin · Heidelberg · New York

TABLE DES MATIÈRES.

Work supported in part by contract USAF 61 (052) - 945.

- IV -

*

CHAPITRE 0

INTRODUCTION ET HISTORIQUE DES NOMBRES D'EULER.

1. Bref historique sur les nombres d'Euler.

On sait depuis Euler que la relation

$$\sum_{0\leq n} (u^n/n!)\ A_n(t) = (1-t)\ /\ (-t + \exp(u(t-1))) \qquad (1)$$

définit des polynômes symétriques

$$A_0(t) = 1 \qquad \text{et}$$

$$A_n(t) = t^{n-1}\ A_n(t^{-1}) = \sum_{0\leq k<n} A_{n,k}\ t^k \qquad (n > 0)$$

de degré $n-1$ dont les coefficients $A_{n,k}$ sont des entiers positifs de somme $A_n(1) = n!$.

Worpitzky [31] a donné la formule

$$x^n = \sum_{0\leq k<n} A_{n,k} \begin{bmatrix} x+k \\ n \end{bmatrix} \qquad (2)$$

et Frobenius [13] qui a appelé les $A_n(t)$ polynômes eulériens a indiqué l'identité

$$A_n(t) = \sum_{0\leq k<n} (n-k)!\ (t-1)^k\ S(n,n-k) \qquad (3)$$

dans laquelle $S(n,j)$ désigne les nombres de Stirling de deuxième espèce, c'est-à-dire le nombre de partitions en j classes d'un ensemble de n éléments. Une bibliographie de cette question a été rassemblée par Carlitz [4] .

De par ailleurs, les polynômes eulériens apparaissent dans divers problèmes d'énumération concernant le groupe symétrique $\mathfrak{S}_n$ sur l'ensemble totalement ordonné $[n] = \{1, 2, \ldots, n\}$ ($= \emptyset$ pour $n = 0$) . Ainsi, d'après (1) , la somme alternée $(-1)^{p-1} A_{2p-1}(-1)$ est le coefficient de $u^{2p-1}/(2p-1)!$ dans le développement de $\operatorname{tg} u$ (voir chapitre V § 4 du présent article) et Désiré André [1] , [2] (voir aussi [21] chap. 4) a découvert que ce dernier nombre est celui des permutations $\sigma \in \mathfrak{S}_{2p-1}$ qui sont <u>alternées</u>, c'est-à-dire telles que pour chaque $j \in [p-1]$, on ait à la fois $\sigma(2j) < \sigma(2j-1)$ et $\sigma(2j) < (2j+1)$.

Mac Mahon [19] a étudié le nombre $A'_{n,k}$ des $\sigma \in \mathfrak{S}_n$ telles que $j < \sigma(j)$ pour exactement k éléments $j \in [n]$ et, en application de son "Master Theorem", il a montré que $A'_{n,k}$ est aussi le nombre des $\sigma \in \mathfrak{S}_n$ qui possèdent k "<u>montées</u>", c'est-à-dire qui satisfont à $\sigma(j) < \sigma(j+1)$ pour exactement k valeurs $j \in [n-1]$. Carlitz et Riordan [7] ont reconnu que les $A'_{n,k}$ sont précisément les coefficients des polynômes eulériens et Riordan a généralisé ceux-ci au moyen de sa théorie des "rook polynomials". Appelons r-<u>excédance</u> de $\sigma \in \mathfrak{S}_n$ tout $j \in [n]$ tel que $j + r \leq \sigma(j)$ $(0 \leq r)$ et soit ${}^rA_{n,k}$ le nombre des $\sigma \in \mathfrak{S}_n$ ayant k r-excédances $({}^1A_{n,k} = A_{n,k})$; J. Riordan, dans le dernier chapitre de son livre [24] , considère avec des notations un peu différentes les polynômes

$$ {}^{r}A_{n}(t) = \sum_{0 \leq k} {}^{r}A_{n,k}\, t^{k} \qquad (4) $$

et à leur sujet établit des extensions des formules (1) et (3) . D'autres généralisations sont dues à Carlitz et son école ([5] , [8] , [10]) .

Soit $|\Delta' M\sigma|$ le nombre des <u>montées</u> de $\sigma \in \mathfrak{S}_n$. Roselle [25] a calculé le polynôme $B_n(t)$ défini comme la somme de $t^{|\Delta' M\sigma|}$ étendue au sous-ensemble $\mathfrak{G}_n$ des $\sigma \in \mathfrak{S}_n$ tels que $1 \neq \sigma(1)$ et $1 + \sigma(j) \neq \sigma(j+1)$ pour chaque $j \in [n-1]$. Il note que $B_n(1)$ est le nombre de permutations sans points fixes de $\mathfrak{S}_n$ et il constate que la somme alternée $(-1)^p\, B_{2p}(-1)$ est égale au coefficient de $u^{2p}/(2p)!$ dans le développement de $1/\cos u$, c'est-à-dire au 2p-ème <u>nombre d'Euler</u>. Or, on sait depuis longtemps que ce coefficient est le nombre des $\sigma \in \mathfrak{S}_{2p}$ tels que $\sigma(2p-1) < (2p)$ et $\sigma(2j-1) < \sigma(2j) > \sigma(2j+1)$ $(j \in [p-1])$, ([1] , [14]) , ceci étant d'ailleurs un cas particulier d'une formule plus générale due à Entringer [11] et dans une direction assez différente de la théorie des "runs up and down" développée par David et Barton à des fins statistiques [3] .

Enfin dans la théorie dite du "problème de Newcomb", on considère au lieu de l'ensemble <u>ordonné</u> $[n]$ un ensemble <u>préordonné</u> quelconque X . Les énoncés y dépendent donc de façon cruciale de la structure de X , ce qui conduit à une problématique sensiblement différente, sauf si l'on réintroduit des hypothèses particulières sur X comme par exemple dans le cas des polynômes de Shanks [27] , des polynômes de Poussin [22] ou dans celui de la "spécification $(1^r(n-r))$" qui fait apparaître directement les entiers ${}^rA_{n,k}$. Hormis ce dernier cas, nous avons entièrement laissé de côté le problème de Newcomb qui nous eût entrainé fort loin des polynômes eulériens. Au demeurant, les mêmes techniques de base ont été employées récemment par l'un de nous ([9]) pour traiter le cas général et certaines de ses applications.

2. <u>Résumé du mémoire</u>.

Les théorèmes qui viennent d'être rappelés ont été, en règle générale, établis en utilisant conjointement quelques propriétés géométriques ("combinatoires") des permutations et les méthodes plus expéditives du calcul différentiel et intégral. En particulier, aucune connexion sauf la coïncidence de l'aboutissement de deux séries de calculs ne semble avoir été vue entre les sommes alternées $A_n(-1)$ et $B_n(-1)$ et la signification des polynômes $A_n(t)$ et $B_n(t)$ en termes d'excédances ou de montées.

Le but du présent mémoire est au contraire de développer la théorie géométrique sous-jacente et c'est de façon subsidiaire que nous en déduisons des identités entre séries ou polynômes. Nous nous sommes cependant attachés à toujours retrouver les résultats classiques. Dans de nombreux cas, cette approche évacue pratiquement tous les calculs : c'est ce qui se produit par exemple en ce qui concerne les formules reliant les polynômes eulériens et les nombres de Stirling. Dans d'autres cas, nous obtenons des séries d'identités nouvelles (par exemple les "formules sommatoires" généralisant celle de Worpitzky données dans la section 6 du chapitre II ou le théorème 5.6). Les méthodes du chapitre III contiennent implicitement l'énumération du nombre des excédances pour les permutations dont les longueurs des cycles satisfont à des conditions de divisibilité données.

Plus important nous semble la démonstration du fait que toutes les identités classiques concernant les polynômes eulériens sont seulement la traduction de propriétés très simples des morphismes d'ensembles totalement ordonnés finis. Pour l'essentiel, elles dérivent soit de méthodes élémentaires courantes comme l'inversion de Möbius ou la formule exponentielle, soit d'une opération unique nouvelle $\sigma \to \hat{\sigma}$ appelée ici <u>transformation fondamentale</u>, déjà introduite par l'un de nous ([12]) dans le cadre général du problème de Newcomb. Simultanément, les énoncés que nous proposons, expriment, en règle, des bijections entre ensembles. Ils sont donc plus riches que les identités énumératives classiques auxquelles ils se réduisent quand, en fin de calcul, on substitue à ces ensembles le nombre de leurs éléments.

Le chapitre I est consacré à l'étude détaillée des propriétés de la transformation fondamentale. Celle-ci est une bijection de $\mathfrak{S}_n$ sur lui-même ayant la propriété que l'ensemble des "excédances" de σ est envoyé, de façon biunivoque, sur celui des "descentes" de $\hat{\sigma}$. Elle permet ici d'établir que la distribution du nombre des excédances sur les permutations de $\mathfrak{S}_n$ est la même que sur le sous-ensemble $\mathfrak{C}_{n+1}$ des <u>permutations circulaires</u> de $\mathfrak{S}_{n+1}$. Ce résultat est étendu dans le paragraphe 4 , où nous employons les <u>bi-excédances</u> (c'est-à-dire les $j \in [n]$ tels que j soit <u>à la fois</u> strictement plus petit que $\sigma(j)$ et que $\sigma^{-1}(j)$) et les <u>creux</u> (c'est-à-dire les $j \in [n]$ tels que j soit <u>à la fois</u> strictement plus petit que $\sigma(j-1)$ et $\sigma(j+1)$) pour étudier les sommes alternées. Cette dualité pourrait être étendue à des constructions plus complexes sur lesquelles nous reviendrons peut-être dans un autre travail.

Dans le chapitre II, nous retrouvons et généralisons diverses formules de récurrence de Riordan et l'identité de Worpitzky, en application des résultats précédents et de la considération des <u>morphismes</u> $[n] \to [m]$ c'est-à-dire, puisque $[n]$ et $[m]$ sont des ensembles totalement ordonnés, des applications $\varphi : [n] \to [m]$ telles que $i \leq j$ implique $\varphi(i) \leq \varphi(j)$.

Dans le chapitre III, nous croyons utile de donner d'abord une théorie systématique de la formule exponentielle classique de Cauchy exprimant le groupe symétrique en fonction des permutations circulaires. Cette formule est un cas particulier d'une construction très générale permettant de ramener divers problèmes d'énumération à un problème analogue sur une sous-famille "génératrice" constituée par des objets "connexes". Afin de clarifier ces notions, nous donnons quelques énoncés sous une forme qui permettrait de traiter les énumérations d'arborescences. Retournant aux permutations, une application de cette formule et des résultats du chapitre I nous permet d'obtenir dans le chapitre IV la fonction génératrice exponentielle du nombre des r-excédances pour les permutations ayant une composition en cycles donnée.

Dans le chapitre V , nous établissons un théorème sur la distribution du nombre des montées pour les permutations ayant un nombre de creux fixé. De façon plus explicite, soit $\mathfrak{S}'_{n,k}$ l'ensemble des permutations $\sigma \in \mathfrak{S}_n$ ayant k creux et telles que $\sigma(1) = n$ $(0 < 2k \leq n)$; alors le nombre de $\sigma \in \mathfrak{S}'_{n,k}$ ayant j montées est donné par le coefficient binomial $\begin{bmatrix} n-2k \\ j \end{bmatrix}$ $(0 \leq j \leq n-2k)$. Nous déduisons de ce résultat les expressions des sommes alternées $A_n(-1)$ et $B_n(-1)$ en fonction du nombre des permutations alternées. En particulier, les développements de $\operatorname{tg} u$ et $1/\cos u$ sont obtenus sans calcul à partir de l'expression des fonctions génératrices exponentielles des polynômes $A_n(t)$ et $B_n(t)$ données dans le chapitre IV .

Dans tout ce travail, étant donnés deux ensembles A et B finis et une application $\varphi : A \to B$, nous appellerons ensemble pondéré l'application $\varphi^{\#} : B \to \mathbb{N}$ définie pour chaque $b \in B$ par $\varphi^{\#}(b) = \operatorname{Card} \varphi^{-1}(b)$. Par abus de notation, on désignera par φA l'ensemble pondéré $\varphi^{\#}$ et il sera commode d'identifier φA et $\varphi^{\#}$ à l'élément $\Sigma\{\varphi^{\#}(b) \cdot b : b \in B\}$ du $\mathbb{Q}$-module libre de base B .

Nous sommes reconnaissants au Professeur J. Riordan de nous avoir fait bénéficier de ses conseils et de son érudition. La dactylographie de ce mémoire est due à Mademoiselle Cler, du Département de Mathématique de Strasbourg, que nous tenons à remercier.

CHAPITRE I

PROPRIÉTÉS GÉNÉRALES DES SYSTÈMES D'EXCÉDANCES ET DE MONTÉES.

1. Excédances.

Dans tout ce chapitre, nous utilisons la notation x_+ pour désigner la partie positive $x_+ = \text{Max}\,\{0,x\}$ de tout $x \in \mathbb{Z}$ et pour $\sigma \in \mathfrak{S}_n$, nous désignons pas σw le n-uple (ou vecteur) $(\sigma(1), \sigma(2), \ldots, \sigma(n)) \in \mathbb{N}^n$. Par conséquent $\sigma w = \emptyset$ pour $n = 0$.

DÉFINITION 1.1.

Pour $\sigma \in \mathfrak{S}_n$, le système des 0-excédances de σ est le n-uple $E\sigma = (E\sigma(1), E\sigma(2), \ldots, E\sigma(n)) \in \mathbb{N}^n$, où pour chaque $k \in [n]$, on pose

$$E\sigma(k) = (\sigma(k) - (k-1))_+$$

Par exemple, avec $n = 6$ et $\sigma w = (6,4,1,2,5,3)$, on a
$E\sigma = ((6-0)_+, (4-1)_+, (1-2)_+, (2-3)_+, (5-4)_+, (3-5)_+) = (6,3,0,0,1,0) \in \mathbb{N}^6$

DÉFINITION 1.2.

Quelque soit l'entier $p > 0$, on note Δ, Δ' et Δ'' les applications de $\mathbb{N}^p$ dans $\mathbb{N}^{p-1}$ envoyant respectivement chaque vecteur $x = (x_1, x_2, \ldots, x_p) \in \mathbb{N}^p$ sur

$$\Delta x = (\ (x_1 - 1)_+ \ ,\ (x_2 - 1)_+ \ ,\ \ldots,\ (x_{p-1} - 1)_+ \ )$$

$$\Delta' x = (x_2,\ x_3,\ \ldots,\ x_p) \qquad \text{et}$$

$$\Delta'' x = (x_1,\ x_2,\ \ldots,\ x_{p-1}) \qquad .$$

Il est immédiat que les trois opérateurs Δ , Δ' , Δ'' ainsi définis commutent deux à deux.

Prenant le même exemple que ci-dessus, on obtient

$E\sigma = (6,3,0,0,1,0)$ $(= \Delta^{0}E\sigma = \Delta'^{0}E\sigma = \Delta''^{0}E\sigma)$;

$\Delta E\sigma = (5,2,0,0,0)$; $\Delta' E\sigma = (3,0,0,1,0)$; $\Delta'' E\sigma = (6,3,0,0,1)$;

$\Delta^2 E\sigma = (4,1,0,0)$; $\Delta\Delta' E\sigma = \Delta'\Delta E\sigma = (2,0,0,0)$; $\Delta'^2 E = (0,0,1,0)$,

etc... .

On notera que $\Delta E\sigma$ est simplement la suite $(\ (\sigma(1) - 1)_+ \ ,\ (\sigma(2) - 2)_+ \ ,\ \ldots,\ (\sigma(n-1) - (n-1))_+ \ )$ et que plus généralement le vecteur $\Delta^r E$ décrit les <u>r-excédances</u> de σ mentionnées dans l'introduction. L'une des raisons motivant l'introduction de Δ' est contenue dans le lemme 1.4 ci-dessous. L'opérateur Δ'' permettra dans le deuxième chapitre de formuler une intéressante propriété de symétrie des polynômes eulériens (propriété 2.3) .

<u>Remarque</u> 1.3.

Pour $\sigma \in \mathfrak{S}_n$ et $k \in [n]$, on a $E\sigma(k) = 1$ si et seulement si $(\sigma(k) - k + 1)_+ = 1$, c'est-à-dire si $k = \sigma(k)$ est un <u>point fixe</u> de σ .

- 10 -

LEMME 1.4.

Soit $\zeta \in \mathfrak{S}_n$ *la permutation circulaire envoyant* n *sur* 1 *et chaque* $k < n$ *sur* $k+1$ *ou encore la permutation définie par* $\zeta w = (2, 3, \ldots, n, 1)$. *Pour chaque* $r \geq 0$ *et chaque* $\sigma \in \mathfrak{S}_n$ *on a*

$$\Delta'^r E\sigma = \Delta^r E\sigma\, \zeta^r \quad .$$

PREUVE.

Posons $\sigma' = \sigma\zeta^r$. Pour chaque $k \in [n-r]$, on a $\sigma'(k) = \sigma(r+k)$. Donc on obtient $\Delta'^r E\sigma(k) = E\sigma(r+k) = (\sigma(r+k)+1-r-k)_+ = (\sigma'(k)+1-k-r)_+ =$ $= ((\sigma'(k)+1-k)_+ - r)_+ = \Delta^r E\sigma'(k)$.

Q.E.D.

Ce simple résultat a la conséquence immédiate suivante qui nous servira fréquemment par la suite.

THÉORÈME 1.5.

Quelque soit le monôme Γ *de degré* $r \geq 0$ *en les applications* Δ *et* Δ', *les ensembles pondérés* $\Delta^r E\,\mathfrak{S}_n$, $\Delta'^r E\,\mathfrak{S}_n$ *et* $\Gamma E\,\mathfrak{S}_n$ *sont égaux.*

PREUVE.

Puisque $\sigma \to \sigma\zeta^r$ est une bijection de $\mathfrak{S}_n$ sur lui-même, l'égalité $\Delta^r E\,\mathfrak{S}_n = \Delta'^r E\,\mathfrak{S}_n$ résulte immédiatement du lemme 1.4 . Comme Δ et Δ' commutent, on peut écrire $\Gamma = \Delta^s \Delta'^{r-s}$, d'où $\Gamma E\,\mathfrak{S}_n = \Delta^r E\,\mathfrak{S}_n$.

Q.E.D.

2. Descentes et montées.

En parallèle avec les excédances, nous introduisons le système des 0-descentes, $D\sigma$, et des 0-montées, $M\sigma$, de $\sigma \in \mathfrak{S}_n$ par la définition suivante, dans laquelle on convient que $\sigma(0) = \sigma^{-1}(0) = \sigma(n+1) = 0$.

DÉFINITION 1.6.

Pour $\sigma \in \mathfrak{S}_n$, on pose

$$D\sigma = (D\sigma(1) , D\sigma(2) , \ldots, D\sigma(n)) \in \mathbb{N}^n$$

$$M\sigma = (M\sigma(1) , M\sigma(2) , \ldots, M\sigma(n)) \in \mathbb{N}^n$$

où pour chaque $k \in [n]$

$$D\sigma(k) = (\sigma(-1 + \sigma^{-1}(k)) - (k-1))_+$$

$$M\sigma(k) = (\sigma(1 + \sigma^{-1}(k-1)) - (k-1))_+ \quad .$$

Par exemple, prenant encore $\sigma w = (6,4,1,2,5,3)$, on obtient

$$D\sigma = ((4-0)_+ , (1-1)_+ , (5-2)_+ , (6-3)_+ , (2-4)_+ , (0-5)_+) = (4,0,3,3,0,0)$$

et

$$M\sigma = ((6-0)_+ , (2-1)_+ , (5-2)_+ , (0-3)_+ , (1-4)_+ , (3-5)_+) = (6,1,3,0,0,0) .$$

Par construction, tous les termes de $D\sigma$ sont nuls ou supérieurs ou égaux à 2 . D'autre part, $D\sigma(n)$ est toujours nul. Par conséquent, $D\sigma$ et $\Delta D\sigma$ ont le même nombre de termes (strictement) positifs. Il est clair que $\Delta^r D\sigma$ et $\Delta^{r-1}M\sigma$ décrivent les différences supérieures ou égales à r $(r > 0)$ entre termes consécutifs de σw , la connexion entre D et M étant explicitée dans le lemme 1.7 ci-dessous. Il est encore utile de noter que les termes positifs de $\Delta'^r D\sigma$ (ou de $\Delta'^r \Delta D\sigma$) correspondent aux paires $(j-1\ ,\ j)$ $(0 \leq j-1)$ telles que $\sigma(j-1) > \sigma(j) > r$.

LEMME 1.7.

<u>Soit</u> $\sigma \to \tilde{\sigma}$ <u>la bijection de</u> $\mathfrak{S}_n$ <u>sur lui-même définie par l'identité</u> $\tilde{\sigma}(k) = \sigma(n+1-k)$ $(k \in [n])$. <u>On a</u> $M\tilde{\sigma}(1) = \sigma(n)$ <u>et</u> $M\tilde{\sigma}(k+1) = \Delta D\sigma(k)$ <u>pour chaque</u> $k \in [n-1]$.

PREUVE.

Par définition $M\tilde{\sigma}(1) = (\tilde{\sigma}(1) - 0)_+ = \tilde{\sigma}(1)$ et $\tilde{\sigma}(1) = \sigma(n+1-1) = \sigma(n)$ Soit $k = \tilde{\sigma}(j)$ avec $k \in [n-1]$; on a alors $\tilde{\sigma}^{-1}(k) = j$ et $\sigma(n+1-j) = k$. D'où il vient

$$\begin{aligned} M\tilde{\sigma}(k+1) &= (\tilde{\sigma}(1+j)-k)_+ = (\sigma(n-j)-k)_+ = (\ \sigma(-1+(n+1-j))-k\ )_+ \\ &= (\ \sigma(-1+\sigma^{-1}(k)) - (k-1)-1\ )_+ = \Delta D\sigma(k) \quad . \end{aligned}$$

Q.E.D.

Prenant σ comme dans l'exemple ci-dessus, on trouve $\Delta D\sigma = (3,0,2,2,0)$, $\tilde{\sigma}w = (3,5,2,1,4,6)$ et $M\tilde{\sigma} = (3,3,0,2,2,0)$.

La construction d'une bijection reliant E et M est l'objet des sections suivantes.

3. La transformation fondamentale.

Etant donnée une permutation $\sigma \in \mathfrak{S}_n$, l'ensemble $\sigma^*(k) = \{\sigma^p(k) : p \in \mathbb{N}\}$ est l'orbite contenant k $(k \in [n])$. On note $z(\sigma)$ le nombre des orbites de σ ou, de façon équivalente, le nombre des cycles de σ . A chaque $k \in [n]$, nous attachons la paire $\Pi_\sigma(k) = (\bar{k}, q_k)$ où $\bar{k}$ est l'élément maximum de l'orbite $\sigma^*(k)$ et où $q_k = \min\{p \in \mathbb{N} : \sigma^p(k) = \bar{k}\}$

DÉFINITION 1.8.

Pour $\sigma \in \mathfrak{S}_n$, $\hat{\sigma}$ est la permutation telle que pour chaque k , $\Pi_\sigma(\hat{\sigma}(k))$ est le k-ième terme de la suite $(\Pi_\sigma(j))_{(j \in [n])}$ ordonnée par ordre lexicographique.

Par exemple, en prenant encore $\sigma w = (6,4,1,2,5,3)$, on a $z(\sigma) = 3$, les trois orbites étant $\{4,2\}$, le point fixe $\{5\}$ et $\{6,3,1\}$. Comme $4 = \sigma^0(4) = \sigma^1(2)$, puis $5 = \sigma^0(5)$ et enfin $6 = \sigma^0(6) = \sigma^1(1) = \sigma^2(3)$ la suite des $\Pi_\sigma(j)$, ordonnée suivant l'ordre lexicographique, est $((4,0) , (4,1) , (5,0) , (6,0) , (6,1) , (6,2))$, d'où $\hat{\sigma}w = (4,2,5,6,1,3)$.

Rappelons qu'un élément x_k d'une suite $(x_1, x_2, \ldots, x_n) \in \mathbb{N}^n$ est dit saillant si et seulement s'il n'existe aucun élément d'indice inférieur qui soit supérieur ou égal à lui, c'est-à-dire si l'on a $x_{k'} < x_k$ pour tout $k' < k$. Donc par définition x_1 est toujours saillant.

LEMME 1.9.

L'élément k de $[n]$ est maximum dans son orbite $\sigma^*(k)$ (c'est-à-dire $\Pi_\sigma(k) = (k,0)$) si et seulement si k est saillant dans $\hat{\sigma}w$.

De plus, k est un point fixe de σ si et seulement si $k = \hat{\sigma}(j)$ avec soit $j = n$, soit $j < n$ et $\sigma(j+1)$ un autre élément saillant de $\hat{\sigma w}$.

PREUVE.

Soit $\Pi_\sigma(k) = (\bar{k},q)$. Si et seulement si q est différent de 0 l'entier k n'est pas l'élément maximum $\bar{k}$ de son orbite et k n'est pas saillant dans $\hat{\sigma w}$ puisque $\Pi_\sigma(\bar{k}) = (\bar{k},0)$ précède $\Pi_\sigma(k)$ dans l'ordre lexicographique, donc $\bar{k}$ précède k dans la suite $\hat{\sigma w}$.

Réciproquement, soient $q = 0$ et $\hat{\sigma}(j) = k$. Ou bien on a $j = 1$ auquel cas k est saillant, ou bien $j \geq 2$, auquel cas pour tout $j' < j$, l'élément $\Pi_\sigma(\hat{\sigma}(j'))$ est avant l'élément $\Pi_\sigma(\hat{\sigma}(j))$ $(=(k,0))$ pour l'ordre lexicographique ; ce qui implique, en posant $\hat{\sigma}(j') = k'$, que l'on a $\hat{\sigma}(j') = k' < \bar{k}' < k = \hat{\sigma}(j)$ et prouve l'équivalence entre les éléments maximaux des orbites de σ et les éléments saillants de $\hat{\sigma w}$.

Supposons maintenant $k = \bar{k} = \hat{\sigma}(j)$. Si et seulement si k n'est pas un point fixe de σ , il est immédiatement suivi dans $\hat{\sigma w}$ d'un élément k' tel que $\Pi_\sigma(k') = (k,1)$. Par conséquent, on a $j < n$ et $\hat{\sigma}(j+1) = k'$ n'est pas saillant.

Q.E.D.

Dans ce qui suit $\mathfrak{S}'_n$ désigne l'ensemble des $\sigma \in \mathfrak{S}_n$ telles que $\sigma(1) = n$ et $\mathfrak{C}_n$ est le sous-ensemble des permutations circulaires.

PROPRIÉTÉ 1.10.

<u>L'application</u> $\sigma \to \hat{\sigma}$ <u>est bijective et satisfait à</u> $\sigma(n) = \hat{\sigma}(n)$. <u>Sa restriction</u> à $\mathfrak{S}_n$ <u>est une bijection sur</u> $\mathfrak{S}'_n$.

PREUVE.

Etant donné $\tau \in \mathfrak{S}_n$, il existe un et un seul $\sigma \in \mathfrak{S}_n$ tel que $\tau = \hat{\sigma}$, car d'une part, les éléments saillants de τw livrent les éléments maximaux des orbites de σ d'après le précédent lemme, d'autre part, les restrictions de σ à chacune de ses orbites sont déterminées par la succession même des éléments de τw . Ceci établit le caractère bijectif de $\sigma \to \hat{\sigma}$. Comme $\tau(1)$ et n sont toujours des éléments saillants de τw , il est clair que $z(\sigma) = 1$ si et seulement si $\tau(1) = n$. Enfin, si l'on a $k = \sigma(n)$, c'est-à-dire $n = \sigma^{-1}(k)$, l'entier n est l'élément maximum de $\sigma^*(k)$ et par conséquent, $\Pi_\sigma(k) = (n, q_k)$ est le dernier terme de la suite $(\Pi_\sigma(j))_{(j \in [n])}$ ordonnée par ordre lexicographique, c'est-à-dire $k = \hat{\sigma}(n)$.

Q.E.D.

4. <u>Relations entre les excédances et les descentes</u>.

Nous allons donner une première application de la transformation fondamentale $\sigma \to \hat{\sigma}$. Aux 1-excédances de σ correspondent les 1-descentes de $\hat{\sigma}$, mais les 0-descentes de $\hat{\sigma}$ ne correspondent pas aux 0-excédances de σ . Ceci amène à introduire un vecteur $D'\hat{\sigma}$ tel que $D\hat{\sigma} + D'\hat{\sigma}$ ($=(D + D')\hat{\sigma}$) rende compte à la fois des descentes et des éléments saillants dans la suite $\hat{\sigma}w$ et qu'en outre l'identité $E\sigma = (D + D')\hat{\sigma}$ soit vérifiée.

- 16 -

DÉFINITION 1.11.

Pour $\tau \in \mathfrak{S}_n$, le vecteur $D'\tau$ est le n-uple $(D'\tau(1), \ldots, D'\tau(n)) \in \mathbb{N}^n$, où pour $j \in [n]$, on pose

$D'\tau(j) = +1$ si d'une part, j est saillant et d'autre part, soit $j = n$, soit $j < n$ et $\tau(1 + \tau^{-1}(j))$ est aussi saillant ;

$D'\tau(j) = 0$ dans tous les autres cas.

THÉORÈME 1.12.

On a identiquement

$$E\sigma = (D + D')\hat{\sigma} \qquad \underline{\text{et}} \qquad \Delta E\sigma = \Delta D\hat{\sigma} \quad .$$

PREUVE.

Quelque soit $\tau \in \mathfrak{S}_n$, remarquons d'abord que $j < n$ est saillant dans τw seulement si $\tau^{-1}(j) < n$.

Nous vérifions le résultat pour chaque $k = \hat{\sigma}(j) \in [n]$ en posant $\Pi_\sigma(k) = (\bar{k}, q)$ et en distinguant les cas suivants :

i) $q > 0$; on a alors $\hat{\sigma}(j-1) = k'$ où $\Pi_\sigma(k') = (\bar{k}, q-1)$, c'est-à-dire $k' = \sigma^{-(q-1)}(\bar{k}) = \sigma^{-(q-1)}(\sigma^q(k)) = \sigma(k)$. Donc par définition, on a $D\hat{\sigma}(k) = (k' - (k-1))_+ = (\sigma(k) - (k-1))_+ = E\sigma(k)$ avec $D\hat{\sigma}(k) = (D + D')\hat{\sigma}(k$ puisque k n'est pas saillant.

ii) $q = 0$, c'est-à-dire $k = \bar{k}$. On a toujours $D\hat{\sigma}(k) = 0$ car soit $j-1 = 0$, soit $j-1 > 0$ avec $k' = \hat{\sigma}(j-1)$ appartenant à une orbite dont l'élément maximum est strictement plus petit que $k = \bar{k}$. D'autre part,

$E\sigma(k) = 0$ ou 1 selon que $\sigma(k) \neq k$ ou non car $E\sigma(k) = (\sigma(k) - (k-1))_+$ où $\sigma(k) \leq k$ puisque $k = \bar{k}$. D'après la deuxième partie du lemme 1.9 et la définition de $D + D'$, on a donc $E\sigma(k) = 1$ si et seulement si $1 = (D + D')\hat{\sigma}(k) \neq D\hat{\sigma}(k) = 0$.

Q.E.D.

On remarquera que l'on a $(D + D')\tau(j) \neq D\tau(j)$ si et seulement si $D\tau(j) = 0$ et $(D + D')\tau(j) = 1$.

Exemple.

On a vu dans les exemples précédents que si $\sigma w = (6,4,1,2,5,3)$, on avait $E\sigma = (6,3,0,0,1,0)$ et $\hat{\sigma}w = (4,2,5,6,1,3)$. Comme les éléments successifs $\hat{\sigma}(3) = 5$ et $\hat{\sigma}(4) = 6$ sont saillants dans $\hat{\sigma}w$, on a $(D + D')\hat{\sigma}(5) = 1$. Il vient donc $(D + D')\hat{\sigma} = (6,3,0,0,1,0) = E\sigma$ et aussi $\Delta D\hat{\sigma} = (5,2,0,0,0) = \Delta E\sigma$.

5. Applications aux permutations alternées.

Une permutation $\sigma \in \mathfrak{S}_n$ est alternée si l'on a $\sigma(2j) < \sigma(2j-1)$, $\sigma(2j+1)$ pour tout entier j tel que $0 < 2j < n$ et si l'on a encore $\sigma(n) < \sigma(n-1)$ lorsque n est pair. D'autre part, elle est biexcédée si pour tout $j \in [n]$, on a $j < \sigma(j)$, $\sigma^{-1}(j)$ ou $j > \sigma(j)$, $\sigma^{-1}(j)$.
Les ensembles des permutations alternées et biexcédées sont notés respectivement $\mathfrak{T}_n$ et $\mathfrak{B}_n$. Nous allons, dans cette section, utiliser la transformation fondamentale pour construire entre $\mathfrak{T}_n$ et $\mathfrak{B}_n$ une bijection qui servira au chapitre V.

LEMME 1.13.

Soit $k = \hat{\sigma}(j)$ $(j \in [n])$; on a $k < \sigma(k)$, $\sigma^{-1}(k)$ si et seulement si $j \neq 1$ et soit $j < n$ et $\hat{\sigma}(j) < \hat{\sigma}(j-1)$, $\hat{\sigma}(j+1)$, soit $j = n$ et $\hat{\sigma}(j) < \hat{\sigma}(j-1)$.

PREUVE.

L'hypothèse $k < \sigma(k)$ entraîne $k \neq \bar{k}$ où $\bar{k}$ est le maximum de l'orbite $\sigma^*(k)$. Donc $j \neq 1$ et $\hat{\sigma}(j-1) = \sigma(k)$ entraînant l'équivalence de $k < \sigma(k)$ et de $\hat{\sigma}(j) < \hat{\sigma}(j-1)$, puisque l'on a $\hat{\sigma}(j) > \hat{\sigma}(j-1)$ quand $k = \bar{k}$. Distinguons deux cas selon que $j = n$ ou non.

Si $j = n$, on a $\sigma^{-1}(k) = \bar{k} = n$, donc $k < \sigma^{-1}(k)$ et le résultat est prouvé.

Si $j \neq n$, ou bien $\sigma^{-1}(k) = \bar{k}$, auquel cas l'hypothèse $k < \sigma^{-1}(k)$ équivaut à $k = \hat{\sigma}(j) < \hat{\sigma}(j+1)$, puisque $\hat{\sigma}(j+1)$ est le maximum d'une autre orbite ; ou bien $\sigma^{-1}(k) \neq \bar{k}$, auquel cas $\hat{\sigma}(j+1) = \sigma^{-1}(k)$ et $k < \sigma^{-1}(k)$ équivaut à $\hat{\sigma}(j) < \hat{\sigma}(j+1)$.

Q.E.D.

PROPOSITION 1.14.

Toute permutation $\sigma \in \mathfrak{B}_n$ a tous ses cycles de longueur paire ; donc $\mathfrak{B}_n = \emptyset$ si n est impair. La transformation fondamentale $\sigma \to \hat{\sigma}$ établit, lorsque n est pair, une bijection de $\mathfrak{B}_n$ sur $\mathfrak{T}_n$.

PREUVE.

Dire que σ est biexcédée équivaut à dire que dans toute orbite de σ , d'élément maximum $\bar{k}$, on a identiquement $\sigma^{2p+1}(\bar{k}) < \sigma^{2p}(\bar{k})$,

$\sigma^{2p+2}(\bar{k})$ $(0 \le p)$. Ces inégalités impliquent d'abord que l'orbite n'est pas réduite à un seul élément ; supposons maintenant qu'elle ait un nombre impair d'éléments, disons $2q+1$ (avec $q \ge 1$) . Ces mêmes inégalités appliquées à $p = q$ entraînent que l'on a $\bar{k} = \sigma^{2q+1}(\bar{k}) < \sigma^{2q}(\bar{k})$, ce qui est impossible puisque $\bar{k}$ est l'élément maximum dans son orbite. La permutation σ n'a donc que des cycles de longueur paire et par conséquent $\mathcal{B}_n = \emptyset$ si n est impair.

La dernière partie de la proposition résulte du lemme 1.13 et des deux équivalences suivantes, qui découlent immédiatement de ce qui précède :

i) σ est dans $\mathcal{B}_{2p}$ si et seulement si les inégalités $k < \sigma(k)$, $\sigma^{-1}(k)$ sont satisfaites pour exactement p indices k ;

ii) τ est dans $\mathfrak{T}_{2p}$ si et seulement si, en posant $\tau(2p+1) = 2p+1$, les inégalités $\tau(j) < \tau(j-1)$, $\tau(j+1)$ sont vraies pour exactement p indices $j > 1$.

Q.E.D.

<u>Exemple.</u>

Nous donnons ci-dessous le tableau des cinq permutations biexcédées de $\mathfrak{S}_4$ et en face de chacune d'elles, la permutation alternée qui lui correspond par l'application fondamentale.

$\sigma(i)$ \ i	1 2 3 4	1 2 3 4	i / $\hat{\sigma}(i)$
	2 1 4 3	2 1 4 3	
	3 4 1 2	3 1 4 2	
$\mathcal{B}_4$	4 3 2 1	3 2 4 1	$\mathfrak{T}_4$
	4 3 1 2	4 1 3 2	
	3 4 2 1	4 2 3 1	

6. Relations entre les excédances et les montées.

Toujours au moyen de la transformation fondamentale, nous construisons une bijection $\sigma \to \bar{\sigma}$ telle que $E\sigma = M\bar{\sigma}$.

Soit en effet $\sigma \in \mathfrak{S}_n$ et notons ζ la permutation circulaire définie dans le lemme 1.4 $(\zeta w = (2, 3, \ldots, n, 1))$. On pose successivement

$$\sigma_1 = \sigma\zeta$$

puis $\sigma_2 = \hat{\sigma}_1$ (transformation fondamentale) ,

et enfin $\bar{\sigma} = \tilde{\sigma}_2$ (où $\sim$ est défini dans le lemme 1.7) .

On a alors :

THÉORÈME 1.15.

L'application $\sigma \to \bar{\sigma}$ est une bijection de $\mathfrak{S}_n$ sur lui-même satisfaisant à

$$E\sigma = M\bar{\sigma} \quad .$$

PREUVE.

Le caractère bijectif est évident d'après ce qui précède. Ensuite, on a $E\sigma(1) = \sigma(1) = \sigma_1(n)$ et on vérifie que $E\sigma(j+1) = \Delta E\sigma_1(j)$ pour chaque $j \in [n-1]$. D'après la propriété 1.10 , on a $\sigma_2(n) = \sigma_1(n)$ et d'après le théorème 1.12 , il vient $\Delta E\sigma_1 = \Delta D\sigma_2$. Par conséquent, en utilisant le lemme 1.7 , on a bien $E\sigma = M\bar{\sigma}$.

Q.E.D.

Par exemple, partant de $\sigma w = (6,4,1,2,5,3)$, on a $E\sigma = (6,3,0,0,1,0)$, puis $\sigma_1 w = (4,1,2,5,3,6)$ et $\sigma_2 w = (5,4,1,2,3,6)$; enfin $\bar{\sigma} w = (6,3,2,1,4,5)$. Comme on a $\Delta E\sigma_1 = \Delta D\sigma_2 = (3,0,0,1,0)$, on vérifie bien que $M\bar{\sigma} = E\sigma$.

On a noté que $E\sigma(k) = 1$ si et seulement si $\sigma(k) = k$. Par conséquent, la restriction de $\sigma \to \bar{\sigma}$ au sous-ensemble $\mathcal{D}_n$ des permutations sans points fixes est une bijection sur le sous-ensemble des $\bar{\sigma} \in \mathfrak{S}_n$ telles que $M\bar{\sigma}(j) \neq 1$, c'est-à-dire des $\bar{\sigma}$ telles que $\bar{\sigma}(1) \neq 1$ et $1 + \bar{\sigma}(j) \neq \bar{\sigma}(j+1)$ pour chaque $j \in [n-1]$. L'ensemble de ces permutations $\bar{\sigma}$ n'est autre que la classe, que nous noterons $\mathcal{G}_n$ des permutations sans successions. On a donc

COROLLAIRE 1.16.

La restriction de $\sigma \to \bar{\sigma}$ à l'ensemble $\mathcal{D}_n$ des permutations sans points fixes est une bijection sur l'ensemble $\mathcal{G}_n$ des permutations sans successions telles que $E\sigma = M\bar{\sigma}$.

7. Relations avec les permutations circulaires.

Pour terminer ce chapitre, il nous reste à étudier la distribution des vecteurs-excédances sur l'ensemble $\mathfrak{C}_n$. Combinant d'abord les résultats de la propriété 1.10 et du théorème 1.12 , nous avons déjà la

PROPRIÉTÉ 1.17.

La restriction de la transformation fondamentale $\sigma \to \hat{\sigma}$ à l'ensemble $\mathfrak{C}_n$ est une bijection sur $\mathfrak{S}'_n$ telle que $\Delta E\sigma = \Delta D\hat{\sigma}$.

Nous construisons ensuite une bijection $\sigma \to \sigma'$ de $\mathfrak{S}''_n = \{\sigma \in \mathfrak{S}_n : \sigma(n) = 1\}$ sur $\mathfrak{S}'_n$ telle que $\Delta E\sigma = \Delta D\sigma'$. Si σ est dans $\mathfrak{S}''_n$, on pose $i = \hat{\sigma}^{-1}(n)$ et σ' est défini comme l'unique permutation telle que

$$\sigma' w = (\ \hat{\sigma}(i), \hat{\sigma}(i+1), \ldots, \hat{\sigma}(n), \hat{\sigma}(1), \ldots, \hat{\sigma}(i-1)\ ) \quad .$$

LEMME 1.18.

<u>L'application</u> $\sigma \to \sigma'$ <u>est une bijection de</u> $\mathfrak{S}''_n$ <u>sur</u> $\mathfrak{S}'_n$ <u>satisfaisant à</u> $\Delta E\sigma = \Delta D\sigma'$.

PREUVE.

D'après la propriété 1.10 , on a $\sigma(n) = \hat{\sigma}(n)$ et par conséquent $\hat{\sigma}$ est dans $\mathfrak{S}''_n$. Il est clair que $\sigma \to \sigma'$ est bijectif. En outre, $\Delta E\sigma = \Delta D\hat{\sigma}$ d'après le théorème 1.12 . Il suffit donc de vérifier $\Delta D\sigma'(k) = \Delta D\hat{\sigma}(k)$ pour chaque $k \in [n]$.

Distinguons d'abord deux cas particuliers :

i) $k = \hat{\sigma}(1)$. On a $k = \sigma'(n-i+2)$ avec $\sigma'(n-i+1) = \hat{\sigma}(n) = 1$. Donc $\Delta D\hat{\sigma}(k) = (0-k)_+$ et $\Delta D\sigma'(k) = (1-k)_+$ sont nuls.

ii) $k = \hat{\sigma}(i) = n$. On a encore $\Delta D\hat{\sigma}(k) = (\hat{\sigma}(i-1)-n)_+ = 0$ et aussi, puisque $n = \sigma'(1)$, $\Delta D\sigma'(k) = (0-n)_+ = 0$.

Dans le cas général où $k = \hat{\sigma}(j+1) \neq n$ $(j \in [n-1])$, on a $k = \sigma'(j' + 1)$ avec $j' = j-i+1$ ou $n+j-i+1$ selon que $j \geq 1$ ou non.

Dans ces deux cas, on a $\hat{\sigma}(j) = \sigma'(j')$ et par conséquent, $(\hat{\sigma}(j) - k)_+$ est la valeur commune de $\Delta D\hat{\sigma}(k)$ et $\Delta D\sigma'(k)$.

Q.E.D.

Ce lemme permet la construction suivante :

Soit $\sigma \in \mathfrak{S}_{n-1}$; on définit $\sigma_1 \in \mathfrak{S}''_n$ en posant

$\sigma_1(n) = 1$ et $\sigma_1(k) = 1 + \sigma(k)$ pour chaque

$k \in [n-1]$; puis l'on pose

$\sigma_2 = \sigma'_1$ où $\tau \to \tau'$ est la bijection définie dans le lemme 1.18 ;

enfin,

σ'' est la permutation définie par $\hat{\sigma}'' = \sigma_2$.

THÉORÈME 1.19.

<u>L'application</u> $\sigma \to \sigma''$ <u>est une bijection de</u> $\mathfrak{S}_{n-1}$ <u>sur</u> $\mathfrak{C}_n$ <u>telle que</u> $E\sigma = \Delta E\sigma''$.

PREUVE.

Tout d'abord la bijectivité de $\sigma \to \sigma''$ est évidente. Si σ est dans $\mathfrak{S}_{n-1}$, on a ensuite $\sigma_1 \in \mathfrak{S}''_n$ et $E\sigma(k) = (\sigma(k) - (k-1))_+ =$ $= (\sigma_1(k)-k)_+ = \Delta E\sigma_1(k)$ pour chaque $k \in [n-1]$ d'où il résulte $E\sigma = \Delta E\sigma_1$. D'autre part, d'après le lemme 1.18 , on a $\sigma_2 \in \mathfrak{S}'_n$ et $E\sigma = \Delta E\sigma_1 = \Delta D\sigma_2$. Enfin, en vertu de $\sigma_2 \in \mathfrak{S}'_n$, la propriété 1.17 montre que l'on a $\sigma'' \in \mathfrak{C}_n$ et $\Delta E\sigma'' = \Delta D\sigma_2$.

Q.E.D.

8. Tableau des bijections utilisées.

Il paraît intéressant de rappeler les propriétés des bijections construites dans le premier chapitre et d'indiquer leur référence.

La bijection	envoie	sur	propriété	référence
$\sigma \to \sigma\zeta^r$	$\mathfrak{S}_n$	$\mathfrak{S}_n$	$\Delta'^r E\sigma = \Delta^r E\sigma\, \zeta^r$ $(r \geq 0)$	Lemme 1.4.
$\sigma \to \tilde{\sigma}$	$\mathfrak{S}_n$	$\mathfrak{S}_n$	$M\sigma(1) = \sigma(n)$ et $M\sigma(k+1) = \Delta D\sigma(k)$ $(k \in [n-1])$	Lemme 1.7.
$\sigma \to \hat{\sigma}$	$\mathfrak{S}_n$	$\mathfrak{S}_n$		Définition 1.8.
	$\mathfrak{C}_n$	$\mathfrak{S}'_n$		Proposition 1.10.
	$\mathfrak{S}_n$	$\mathfrak{S}_n$	$E\sigma = (D+D')\hat{\sigma}$ $\Delta E\sigma = \Delta D\hat{\sigma}$	Théorème 1.12.
	$\mathfrak{B}_n$ (n pair)	$\mathfrak{T}_n$ (n pair)		Proposition 1.14.
$\sigma \to \bar{\sigma}$	$\mathfrak{S}_n$	$\mathfrak{S}_n$	$E\sigma = M\bar{\sigma}$	Théorème 1.15.
	$\mathfrak{D}_n$	$\mathfrak{G}_n$		Corollaire 1.16.
$\sigma \to \sigma''$	$\mathfrak{S}''_n$	$\mathfrak{S}'_n$	$\Delta E\sigma = \Delta D\sigma'$	Lemme 1.18.
$\sigma \to \sigma''$	$\mathfrak{S}_{n-1}$	$\mathfrak{C}_n$	$E\sigma = \Delta E\sigma''$	Théorème 1.19.

9. Notations générales.

Nous réunissons dans cette section toutes les notations utilisées pour les sur- et sous-ensembles de $\mathfrak{S}_n$. Pour $n > 0$ et $0 < k,r \leq n$, on considère les sous-ensembles suivants de $\mathfrak{S}_n$:

- $\mathfrak{C}_n$ l'ensemble des permutations circulaires ; $(\mathfrak{C}_0 = \phi)$.
- $\mathfrak{G}_n$ l'ensemble des permutations sans successions, c'est-à-dire des $\sigma \in \mathfrak{C}_n$ telles que $1 \neq \sigma(1)$ et $1 + \sigma(j) \neq \sigma(j+1)$ pour $j \in [n-1]$;
- $\mathfrak{D}_n$ l'ensemble des permutations sans points fixes ;
- $\mathfrak{I}_n$ l'ensemble des permutations alternées, c'est-à-dire des $\sigma \in \mathfrak{S}_n$ telles que $\sigma(2j) < \sigma(2j-1)$, $\sigma(2j+1)$ pour tout entier j tel que $0 < 2j < n$ et en plus si n est pair, telles que $\sigma(n) < \sigma(n-1)$;
- $\mathfrak{B}_n$ l'ensemble des permutations biexcédées, c'est-à-dire des $\sigma \in \mathfrak{S}_n$ telles que pour chaque $j \in [n]$, on ait $j < \sigma(j)$, $\sigma^{-1}(j)$ ou $j > \sigma(j)$, $\sigma^{-1}(j)$;
- $\mathfrak{S}_{n,k}$ l'ensemble des permutations $\sigma \in \mathfrak{S}_n$ telles que $\sigma(1) = k$; $\mathfrak{S}'_n = \mathfrak{S}_{n,n}$;
- $\mathfrak{S}''_n$ l'ensemble des permutations $\sigma \in \mathfrak{S}_n$ telles que $\sigma(n) = 1$.
- ${}_r\mathfrak{S}_n$ l'ensemble des permutations $\sigma \in \mathfrak{S}_n$ telles que $\sigma^{-1}(n-r+1) < \sigma^{-1}(n-r+2) < \ldots < \sigma^{-1}(n)$.

On posera également $\mathfrak{S} = \bigcup_{0 \leq n} \mathfrak{S}_n$; puis $\mathfrak{C} = \bigcup \mathfrak{C}_n$, $\mathfrak{G} = \bigcup \mathfrak{G}_n$, $\mathfrak{I} = \bigcup \mathfrak{I}_n$, $\mathfrak{B} = \bigcup \mathfrak{B}_n$ et $\mathfrak{D} = \bigcup \mathfrak{D}_n$ où la réunion est étendue à l'ensemble des entiers $n > 0$.

Enfin, on utilise les notations courantes suivantes :

$\mathbb{N}$ l'ensemble des entiers naturels 0, 1, 2, ...

$\mathbb{Z}$ l'ensemble des entiers

$\mathbb{Q}$ l'ensemble des nombres rationnels.

CHAPITRE II

LES POLYNÔMES EULÉRIENS

1. Interprétation des polynômes eulériens.

Les théorèmes 1.12 , 1.15 , 1.19 et la propriété 1.17 permettent d'établir immédiatement l'égalité des cinq ensembles pondérés :

$$E\,\mathfrak{S}_n \ , \quad (D + D')\mathfrak{S}_n \ , \quad M\,\mathfrak{S}_n \ , \quad \Delta E\,\mathfrak{C}_{n+1} \ , \quad \Delta D\,\mathfrak{S}'_{n+1} \ ;$$

et le théorème 1.12 donne encore :

$$\Delta E\,\mathfrak{S}_n = \Delta D\,\mathfrak{S}_n \quad .$$

La propriété 2.1 suivante résulte alors du théorème 1.5 :

PROPRIÉTÉ 2.1.

Soit Γ un monôme en Δ et Δ' de degré $r \geq 0$. On a

$$\Gamma E\,\mathfrak{S}_n = \Gamma(D + D')\mathfrak{S}_n = \Gamma M\,\mathfrak{S}_n = \Gamma\Delta E\,\mathfrak{C}_{n+1} = \Gamma\Delta D\,\mathfrak{S}'_{n+1}$$

où en outre

$$\Gamma E\,\mathfrak{S}_n = \Gamma D\,\mathfrak{S}_n$$

si et seulement si Γ a au moins un Δ comme facteur.

Nous notons $|x|$ le nombre de termes positifs de tout $x \in \underset{\approx}{N}^P$ et, introduisant une indéterminée t, nous posons $\theta x = t^{|x|}$. Si K est une application dans $\underset{\approx}{N}^P$ d'une partie $\mathcal{P}$ de $\mathfrak{S}_n$, l'ensemble pondéré $\theta K \mathcal{P} = \Sigma \{\theta K\sigma : \sigma \in \mathcal{P}\} = \underset{0 \leq k}{\Sigma} t^k$. Card $\{\sigma \in \mathcal{P} : |K\sigma| = k\}$ sera appelé, par abus de langage, <u>fonction génératrice</u> de K. Pour $\mathcal{P}$ fini, $\theta K \mathcal{P}$ sera donc un polynôme en t à coefficients dans $\underset{\approx}{N}$ et nous dirons que $(\mathcal{P}, K)$ est une <u>interprétation.</u>

Les polynômes eulériens $A_n(t)$ et leurs généralisations ${}^rA_n(t)$ selon Riordan sont définis par :

$$ {}^rA_n(t) = \theta \Delta^r E\, \mathfrak{S}_n \qquad \text{pour} \quad 0 \leq r < n \quad . $$

On conviendra que ${}^rA_n(t) = n!$ pour $r \geq n$. On posera $A_0(t) = 1$ et $A_n(t) = {}^1A_n(t)$ pour $n > 0$. Par construction, les ${}^rA_n(t)$ sont des polynômes de degré au plus égal à $n-r$. L'énoncé suivant en donne plusieurs interprétations par simple application de la propriété précédente.

PROPRIÉTÉ 2.2.

<u>Soit</u> Γ <u>un monôme en</u> Δ <u>et</u> Δ' <u>de degré</u> r. <u>On a</u>

$$ {}^rA_n(t) = \theta \Gamma E\, \mathfrak{S}_n = \theta \Gamma (D+D')\mathfrak{S}_n = \theta \Gamma M\, \mathfrak{S}_n = \theta \Gamma \Delta E\, \mathfrak{S}_{n+1} = \theta \Gamma \Delta D\, \mathfrak{S}'_{n+1} $$

<u>où en outre,</u> ${}^rA_n(t) = \theta \Gamma D\, \mathfrak{S}_n$ <u>si et seulement si</u> Γ <u>a au moins un</u> Δ <u>comme facteur</u>.

De même, d'après le corollaire 1.16 , on a l'égalité entre les ensembles pondérés $E\,\mathcal{D}_n$ et $M\,\mathcal{C}_n$, d'où encore

$$\theta E\,\mathcal{D}_n = \theta M\,\mathcal{C}_n \quad \text{pour} \quad n > 0 \qquad .$$

La valeur commune de ces deux derniers polynômes sera désignée par $B_n(t)$. L'interprétation $(\mathcal{C}_n , M)$ de ces polynômes est due à Roselle [25] . L'interprétation $(\mathcal{D}_n , E)$ servira à établir au chapitre IV l'expression de la fonction génératrice exponentielle des $B_n(t)$.

2. Propriétés de symétrie.

Les identités (1) et (4) ci-dessous sont bien connues (Cf. Riordan [24]). Nous en donnons ici des démonstrations élémentaires. Les polynômes ${}^rA_n(t)$ pour $r > 1$ n'ont pas de propriété de symétrie évidente. En revanche, si l'on forme les polynômes réciproques $t^{n-r}\,{}^rA_n(t^{-1})$, on obtient plusieurs interprétations (Cf. les relations (2) et (3) ci-dessous) qui nous serviront effectivement dans les sections 2.4 et 2.6 , pour établir des connexions avec le problème de Simon Newcomb et pour démontrer de nouvelles identités sur les polynômes eulériens.

PROPRIÉTÉ 2.3.

Pour $n > 0$, *on a* :

$$ {}^0A_n(t) = t\,A_n(t) \qquad . \tag{1}$$

De plus, si Γ *est un monôme de la forme* $\Delta''^{r-1}\Delta$ *ou* $\Delta''^{r-1}\Delta'$ $(r > 0)$, on a

$$t^{n-r}\ {}^{r}A_n(t^{-1}) = \theta\Gamma E\,\mathfrak{S}_n = \theta\Gamma(D+D')\mathfrak{S}_n = \theta\Gamma M\,\mathfrak{S}_n \quad . \tag{2}$$

On a encore

$$t^{n-r}\ {}^{r}A_n(t^{-1}) = \theta\Delta''^{r-1}\Delta D\,\mathfrak{S}_n \quad , \tag{3}$$

d'où en particulier, pour $r = 1$

$$t^{n-1}\ A_n(t^{-1}) = A_n(t) \quad . \tag{4}$$

PREUVE.

Soit $\sigma \in \mathfrak{S}_n$; définissant $\check{\sigma}$ par l'identité $\check{\sigma}(k) = n+1 - \sigma(n+1-k)$ l'on a $E\check{\sigma}(k) = 1$ si et seulement si $E\sigma(n+1-k) = 1$ et $E\check{\sigma}(k) > 1$ si et seulement si $E\sigma(n+1-k) = 0$. Dans ces conditions, il vient $|E\check{\sigma}| + |\Delta E\sigma| = n$, ce qui établit

$${}^{0}A_n(t) = t^n\ A_n(t^{-1}) \quad . \tag{5}$$

D'autre part, la relation entre les vecteurs $E\check{\sigma}$ et $E\sigma$ peut encore s'exprimer par la condition

$$\Delta E\check{\sigma}(k) > 0 \text{ si et seulement si } \Delta' E\sigma(n-k) = 0 \text{ pour } 1 \le k < n. \tag{6}$$

Mais la condition (6) est encore équivalente à

$$|\Delta'^{r-1}\Delta E\check{\sigma}| + |\Delta''^{r-1}\Delta' E\sigma| = n-r$$

ou encore à

$$|\Delta''^{r-1}\Delta E\check{\sigma}| + |\Delta'^{r}E\sigma| = n-r \quad .$$

On en déduit :

$$\theta\Delta''^{r-1}\Delta E\,\mathfrak{S}_n = t^{n-r} \sum_{0\le k} t^{-k}\,\mathrm{Card}\,\{\sigma \in \mathfrak{S}_n \,:\, |\Delta'^{r}E\sigma| = k\}$$

$$= t^{n-r}\,{}^{r}A_n(t^{-1})$$

d'après la propriété 2.2 .

Les relations (2) et (3) résultent alors de la propriété 2.1 et faisant $r = 1$ dans (3), on obtient $t^{n-1}\,A_n(t^{-1}) = \theta\Delta D\,\mathfrak{S}_n = A_n(t)$, c'est-à-dire la relation (4). Enfin, l'identité (2) résulte à la fois de (4) et de (5) .

Q.E.D.

Remarque 2.4.

D'après la définition 1.6 , pour $1 \le r \le n$ et $\sigma \in \mathfrak{S}_n$, il y a exactement $|\Delta'\Delta''^{r-1}M\sigma|$ indices i tels que $1 \le i < n$, $\sigma(i) < \sigma(i+1)$ et $\sigma(i) < n-r$. Comme on a posé d'autre part

$${}^{r}A_n(t) = \sum_{0\le k\le n-r} {}^{r}A_{n,k}\; t^k$$

et que d'après la propriété 2.3 , on a

$$\theta\Delta'\Delta''^{r-1}M\,\mathfrak{S}_n = t^{n-r}\;{}^{r}A_n(t^{-1}) = \sum_{0 \le s \le n-r} {}^{r}A_{n,n-r-s}\; t^s \qquad ,$$

il vient :

$${}^{r}A_{n,n-r-s} = \mathrm{Card}\,\{\sigma \in \mathfrak{S}_n : |\Delta'\Delta''^{r-1}M\sigma| = s\} \quad \text{pour} \quad 0 \le s \le n-r \quad .$$

Cette remarque nous servira au paragraphe 6 du présent chapitre.

3. Relations de récurrence.

Dans cette section, nous établissons une relation de récurrence sur les polynômes eulériens qui généralise l'identité (1) ci-dessus et redémontrons la relation de récurrence trouvée par Riordan [24] .

PROPRIÉTÉ 2.5.

Pour $0 \le r \le n$, *on a l'identité*

$$t\;.\;{}^{(r+1)}A_n(t) = {}^{r}A_n(t) + r(t-1)\;.\;{}^{r}A_{n-1}(t) \qquad .$$

PREUVE.

Pour $r = 0$, l'identité se réduit à $t\,A_n(t) = {}^{0}A_n(t)$. Pour $r = n$, elle est encore vraie avec la convention que nous avons faite que ${}^{r}A_n(t) = n!$ quand $r \ge n$. Nous supposons donc $0 < r < n$. Pour chaque $\sigma \in \mathfrak{S}_n$, on a

$$\Delta^r E\sigma = (\;(\sigma(1)-r)_+\;,\;(\sigma(2)-r-1)_+\;,\;\ldots,\;(\sigma(n-r)-(n-1))_+\;)$$

et $\Delta'\Delta^r E\sigma$ est formé des $n-r-1$ derniers termes de $\Delta^r E\sigma$. Donc $|\Delta^r E\sigma| - |\Delta'\Delta^r E\sigma| = 0$ ou 1 selon que $(\sigma(1)-r)_+ = 0$ ou non, c'est-à-dire selon que $\sigma(1) \leq r$ ou non. Posant

$$\mathfrak{S}_{n,s} = \{\sigma \in \mathfrak{S}_n : \sigma(1) = s\} \quad ,$$

il en résulte que

$$\begin{aligned} \theta\Delta^r E\,\mathfrak{S}_{n,s} &= \theta\Delta'\Delta^r E\,\mathfrak{S}_{n,s} && \text{si} \quad s \leq r \\ &= t\;\theta\Delta'\Delta^r E\,\mathfrak{S}_{n,s} && \text{si} \quad s > r \quad . \end{aligned}$$

Utilisant ces deux relations ainsi que les égalités

$${}^rA_n(t) = \theta\Delta^r E\,\mathfrak{S}_n = \sum_s \theta\Delta^r E\,\mathfrak{S}_{n,s}$$

et

$${}^{(r+1)}A_n(t) = \theta\Delta'\Delta^r E\,\mathfrak{S}_n = \sum_s \theta\Delta'\Delta^r E\,\mathfrak{S}_{n,s} \quad ,$$

on obtient

$${}^rA_n(t) - P = t({}^{(r+1)}A_n(t) - P)\,,$$ où P désigne la valeur commune des sommes sur $s \leq r$ de $\theta\Delta^r E\,\mathfrak{S}_{n,s}$ et $\theta\Delta'\Delta^r E\,\mathfrak{S}_{n,s}$.

Attachons maintenant à chaque $\sigma \in \mathfrak{S}_{n,s}$, la permutation $\sigma' \in \mathfrak{S}_{n,1}$ telle que

$$\sigma'(1) = 1 \quad ; \quad \sigma'(\sigma^{-1}(1)) = s \quad ; \quad \sigma'(k) = \sigma(k) \quad \text{autrement.}$$

Si et seulement si $s \leq r$, on a $\Delta^r E\sigma' = \Delta^r E\sigma$. Par conséquent, on a $P = r\,\theta\Delta^r E\,\mathfrak{S}_{n,1} = r\,\theta\Delta^r\Delta' E\,\mathfrak{S}_{n,1} = r\,{}^rA_{n-1}(t)$ d'après la propriété 2.2 et l'on obtient enfin l'identité cherchée sous la forme équivalente

$${}^rA_n(t) - r\,{}^rA_{n-1}(t) = t(\,{}^{(r+1)}A_n(t) - r\,{}^rA_{n-1}(t)\,) \quad .$$

Q.E.D.

Remarque 2.6.

Riordan ([24] p. 214) a trouvé une autre relation de récurrence, à savoir

$${}^rA_n(t) = [r + (n-r)t] \;.\; {}^rA_{n-1}(t) + t(1-t) \;.\; {}^rA'_{n-1}(t) \tag{7}$$

$(0 \leq r \leq n \;;\; 1 \leq n)$

où ${}^rA'_{n-1}(t)$ désigne la dérivée du polynôme ${}^rA_{n-1}(t)$. Comme on a posé

$${}^rA_n(t) = \sum_{0\leq k\leq n-r} {}^rA_{n,k}\, t^k \quad ,$$

cette relation de récurrence est encore équivalente aux $(n-r+1)$ relations suivantes (Cf. [24] p. 215)

$${}^rA_{n,k} = (k+r) \;.\; {}^rA_{n-1,k} + (n+1-k-r) \;.\; {}^rA_{n-1,k-1} \tag{8}$$

$(0 \leq k \leq n-r)$

où l'on a posé ${}^rA_{n,k} = 0$ si $k < 0$ ou $k > n-r$.

Comme l'a noté Welschinger [30] , on peut rédémontrer facilement (8) et par suite (7) en prenant les polynômes eulériens dans l'interprétation

$$ {}^{r}A_n(t) = \theta\Delta'^{\,r-1}\Delta D\,\mathfrak{S}_n \quad . $$

En effet, on vérifie tout d'abord que les relations (8) sont vraies pour $n = r$, en notant que ${}^{n}A_{n,0} = n!$ et ${}^{n}A_{n,k} = 0$ pour $k \neq 0$. On suppose ensuite $0 \leq r < n$ et l'on pose pour $i \in [n]$ et $\sigma \in \mathfrak{S}_{n-1}$

$$ \eta_i(\sigma) = (\ \sigma(1),\ \ldots,\ \sigma(i-1),\ n\ ,\ \sigma(i),\ \ldots,\ \sigma(n-1)\ ) \quad . $$

Il est clair que l'on a

$$ \mathfrak{S}_n = \{\eta_i(\sigma) : i \in [n]\ ,\ \sigma \in \mathfrak{S}_{n-1}\} \quad . $$

Soit $\sigma \in \mathfrak{S}_{n-1}$; on a $|\Delta'^{\,r-1}\Delta D\sigma| = k$ [en abrégé : $\sigma \in {}^{r}\mathfrak{a}_{n-1,k}$] si et seulement s'il existe k couples $(j-1,j)$ tels que $1 \leq j-1$ et $\sigma(j-1) > \sigma(j) \geq r$.

Prenons σ dans ${}^{r}\mathfrak{a}_{n-1,k}$; on observe alors que $\eta_i(\sigma)$ appartient à ${}^{r}\mathfrak{a}_{n,k}$ pour les seuls indices i suivants

(i) $1 \leq i-1$ et $\sigma(i-1) > \sigma(i) \geq r$;

(ii) $i \in [n-1]$ et $\sigma(i) < r$;

(iii) $i = n$,

c'est-à-dire pour exactement $k + (r-1) + 1 = k + r$ indices $i \in [n]$. Pour les autres $n - (k+r)$ indices i ne satisfaisant à aucune des conditions (i) , (ii) , (iii) , on a

$$\eta_i(\sigma) \in {}^r\mathfrak{a}_{n,k+1} \quad .$$

On constate donc que l'ensemble ${}^r\mathfrak{a}_{n,k}$ est contenu dans la réunion $\cup \{\eta_i({}^r\mathfrak{a}_{n-1,k} \cup {}^r\mathfrak{a}_{n-1,k-1}) : i \in [n]\}$. On voit ensuite que la relation

$$\eta_i(\sigma) \in {}^r\mathfrak{a}_{n,k}$$

est vérifie pour exactement $(k+r)$ indices i si σ est dans ${}^r\mathfrak{a}_{n-1,k}$ et pour exactement $n - (k-1+r) = n+1-k-r$ indices i si σ est dans ${}^r\mathfrak{a}_{n-1,k-1}$. Les relations (8) sont ainsi démontrées.

4. Relations avec le "problème de Simon Newcomb".

Nous allons exploiter maintenant le lien entre les polynômes ${}^rA_n(t)$ et les polynômes générateurs que l'on définit pour le "problème de Simon Newcomb avec une spécification $(1^r(n-r))$" (voir Mac Mahon [20] , vol. 1, chap. 4 et 5) . Dans la preuve de la propriété qui suit, nous considérons σw comme le mot $\sigma(1)\ \sigma(2) \ldots \sigma(n)$ dont les lettres sont des éléments de $[n]$.

PROPRIÉTÉ 2.7.

Soit $r \geq 2$; on a

$$t^{n-r}\ {}^rA_n(t^{-1}) = r!\ \theta\Delta D\ {}_r\mathfrak{S}_n$$

où

$$ {}_r\mathfrak{S}_n = \{\sigma \in \mathfrak{S}_n : \sigma^{-1}(n-r+1) < \sigma^{-1}(n-r+2) < \dots < \sigma^{-1}(n)\} \quad . $$

PREUVE.

D'après la propriété 2.3 , il nous suffit d'établir

$$ \theta\Delta''^{r-1}\Delta D\,\mathfrak{S}_n = r!\ \theta\Delta D\ {}_r\mathfrak{S}_n $$

ou encore de trouver une surjection $\sigma \to \sigma'$ de $\mathfrak{S}_n$ sur ${}_r\mathfrak{S}_n$ telle que

i) $|\Delta''^{r-1}\Delta D\sigma| = |\Delta D\sigma'|$

ii) l'application $\sigma \to \sigma'$ soit homogène de degré $r!$, c'est-à-dire que l'image inverse de chaque $\sigma' \in {}_r\mathfrak{S}_n$ ait $r!$ éléments.

Prenons $\sigma \in \mathfrak{S}_n$ et soit $\sigma w = g_1\ i_{n-r+1}\ g_2 \dots g_r\ i_n\ g_{r+1}$ où $\{i_{n-r+1}, \dots, i_n\} = \{n-r+1, \dots, n\}$. On pose $\sigma' w = g_1(n-r+1)g_2 \dots g_r\ n\ g_{r+1}$. Il est clair que σ' est dans ${}_r\mathfrak{S}_n$ et que $\sigma \to \sigma'$ est homogène de degré $r!$.

D'autre part, puisque les éléments $n-r+1$, $n-r+2, \dots, n$ se présentent dans cet ordre dans le mot $\sigma' w$, on a $\Delta D\sigma'(j) = 0$ pour $j = n-r+1$, $n-r+2$, $\dots$, $n-1$, ou encore $|\Delta D\sigma'| = |\Delta''^{r-1}\Delta D\sigma'|$. Enfin, l'identité $|\Delta''^{r-1}\Delta D\sigma| = |\Delta''^{r-1}\Delta D\sigma'|$ résulte du fait que l'on a $\sigma'(j) = \sigma(j)$ si $\sigma(j) \le n-r$ et que $\sigma'(j) > n-r$ si et seulement si $\sigma(j) > n-r$.

Q.E.D.

Si l'on envoie tout $\sigma' w = g_1(n-r+1)\, g_2(n-r+2) \dots g_r\, n\, g_{r+1}$ de ${}_r\mathfrak{S}_n$ sur le mot $f = g_1(n-r+1)\, g_2(n-r+1) \dots g_r(n-r+1)\, g_{r+1}$, on définit une bijection de ${}_r\mathfrak{S}_n$ sur une classe de mots de spécification $(1^r(n-r))$, c'est-à-dire des mots de longueur n , qui contiennent $n-r+1$ lettres distinctes dont l'une d'entre elles est répétée r fois. Si l'on définit maintenant $|\Delta D f|$ comme le <u>nombre de descentes</u>, c'est-à-dire le nombre de couples de lettres successives dans f qui vont en décroissant, on voit que l'on a $|\Delta D f| = |\Delta D \sigma'|$. Ainsi $t^{n-r}\, {}^rA_n(t^{-1})$ est le <u>polynôme générateur du nombre des descentes pour un ensemble de spécification</u> $(1^r(n-r))$.

5. <u>Relations avec les nombres de Stirling</u>.

Rappelons que pour $0 < q \leq p$, le <u>nombre de Stirling de deuxième espèce</u> $S(p,q)$ est le nombre de partitions de $[p]$ en q parties non vides. Le résultat suivant est obtenu par Riordan ([24] p. 213) au moyen de calculs assez complexes :

PROPRIÉTÉ 2.8.

<u>L'entier</u> $S(p,q)$ <u>est le nombre de parties</u> $W \subset [p] \times [p]$ <u>de</u> $p-q$ <u>éléments qui satisfont aux conditions suivantes</u> :

i) W <u>est une quasi-permutation, c'est-à-dire qu'il existe au moins un</u> $\sigma \in \mathfrak{S}_p$ <u>tel que</u> $W \subset \{(k,\sigma(k)) \subset [p] \times [p] : k \in [p]\}$,

ii) W <u>est supra-diagonale, c'est-à-dire que</u> $(k,k') \in W$ <u>implique</u> $k < k'$.

PREUVE.

Soit $\{E_1, E_2, \ldots, E_q\}$ une partition de $[p]$ que nous pouvons considérer comme formée des classes d'une relation d'équivalence $E \subset [p] \times [p]$. A chaque $E_j = \{i_1 < i_2 < \ldots < i_{n_j}\}$ comprenant $n_j > 0$ éléments, nous associons la quasi-permutation supra-diagonale $E'_j = \{(i_1, i_2), (i_2, i_3), \ldots, (i_{n_j-1}, i_{n_j})\}$ contenant n_j-1 (éventuellement zéro) éléments de $[p] \times [p]$. Posant $E' = \bigcup_{1\leq j\leq q} E'_j$, on voit que E' est une quasi-permutation supra-diagonale ayant $\Sigma(n_j-1) = p-q$ éléments et que E est la plus fine des équivalences sur $[p]$ qui contienne E'.

Réciproquement, étant donnée une quasi-application supra-diagonale W ayant $p-q$ éléments, soit E la plus fine de toutes les équivalences sur $[p]$ qui contienne W. On a $W = E'$ et la bijection désirée est établie.

Q.E.D.

La formule de Riordan ([24] p. 214)

$$ {}^rA_n(s+1) = \sum_{0\leq k\leq n-r} s^k(n-k)!\ S(n+1-r, n+1-r-k) $$

étant obtenue par celui-ci au moyen de la méthode algébrique classique d'inversion de Möbius, nous pensons pouvoir nous dispenser d'en reproduire ici la preuve.

6. Les identités de Worpitzky.

Soient m,n,s trois entiers tels que $0 \leq s < n \leq m+s$ et $(i_1, i_2, \ldots, i_s)$ une suite strictement croissante d'entiers compris entre 1 et $n-1$ que nous nommerons indices distingués. Nous allons d'abord dénombrer l'ensemble $\Phi_{m,n,s}$ de tous les morphismes $\varphi : [n] \to [m]$ tels que $\varphi(i) \neq \varphi(i+1)$ si i n'est pas un indice distingué. Le dénombrement de $\Phi_{m,n,s}$ permet non seulement d'obtenir l'identité (9) ci-dessous due à Worpitzky, mais une généralisation de celle-ci au cas des polynômes ${}^rA_n(t)$ $(r > 1)$. La proposition 2.9 ci-dessous est bien connue.

PROPOSITION 2.9.

On a :

$$\operatorname{Card} \Phi_{m,n,s} = \begin{bmatrix} m+s \\ n \end{bmatrix} .$$

PREUVE.

Il suffit de faire correspondre, de façon bijective, à tout $\varphi \in \Phi_{m,n,s}$ un morphisme injectif $\psi : [n] \to [m+s]$ (c'est-à-dire une application strictement croissante). Dans ce but, désignons pour tout entier $k \in [n]$, par $\theta(k)$ le nombre d'indices distingués avant k, à savoir le nombre d'indices j tels que $i_j < k$. On a $\theta(1) = 0$ et $\theta(n) = s$. La bijection $\varphi \to \psi$ est alors définie de la façon suivante. Pour tout $k \in [n]$, on pose $\psi(k) = \varphi(k) + \theta(k)$. On a $1 \leq \psi(n) \leq m+s$ et ψ est strictement croissante, car si $k-1$ est distingué, on a $\theta(k-1) < \theta(k)$ et si $k-1$ ne l'est pas on a $\varphi(k-1) < \varphi(k)$. Dans les deux cas, il vient

$\psi(k-1) < \psi(k)$. Enfin, l'application $\varphi \to \psi$ est trivialement injective. Elle est aussi surjective, puisque φ est uniquement déterminé par les relations $\varphi(k) = \psi(k) - \theta(k)$ $(1 \le k \le n)$.

Q.E.D.

L'identité de Worpitzky sur les nombres d'Euler s'obtient par simple application de cette proposition. L'ensemble de toutes les applications de $[n]$ dans $[m]$ étant noté $H_{m,n}$, soit $\varphi \in H_{n,m}$; on définit $\delta\varphi$ comme l'unique $\sigma \in \mathfrak{S}_n$ telle que la suite des paires $(\varphi\sigma(1)\ ,\ \sigma(1))$, $(\varphi\sigma(2)\ ,\ \sigma(2))$, ..., $(\varphi\sigma(n)\ ,\ \sigma(n))$ soit croissante pour l'ordre lexicographique. Par conséquent, $\delta^{-1}\sigma$ est l'ensemble des applications $\varphi : [n] \to [m]$ telles que $\varphi\sigma(i) \le \varphi\sigma(i+1)$ pour $i \in [n-1]$ et telles que l'égalité $\varphi\sigma(i) = \varphi\sigma(i+1)$ ne soit possible que si $\sigma(i) < \sigma(i+1)$. D'après la remarque 2.4 , il y a exactement $s = |\Delta' M\sigma|$ indices i vérifiant une telle inégalité ; donc, d'après la précédente proposition, Card $\delta^{-1}\sigma = \begin{bmatrix} m+s \\ n \end{bmatrix}$. Comme le nombre de permutations $\sigma \in \mathfrak{S}_n$ satisfaisant à $|\Delta M\sigma| = s$ est donné par le nombre d'Euler $A_{n,s}$, il vient enfin

$$m^n = \sum_{0 \le s \le n-1} \begin{bmatrix} m+s \\ n \end{bmatrix} A_{n,s} \quad . \tag{9}$$

Cette identité est un cas particulier de l'identité (10) ci-dessous.

PROPRIÉTÉ 2.10.

<u>Pour</u> $r \in [n]$, <u>on a</u>

$$\sum_{0 \le s \le n-r} {}^{r}A_{n,n-r-s} \begin{bmatrix} m+s \\ n \end{bmatrix} = m^{n-r}\ m! \ / \ (m-r)! \quad . \tag{10}$$

PREUVE.

Notons d'abord que pour $r = 1$, on retrouve bien l'identité (9) , puisque l'on a $A_{n,n-1-s} = A_{n,s}$ d'après la propriété 2.3 . D'autre part, l'identité est vraie pour $r = n$ avec les conventions que nous avons adoptées. On prendra donc $r \in [n-1]$. Soit $H_{m,n,r}$ l'ensemble des applications $\varphi : [n] \to [m]$ dont la restriction à $\{n-r+1, n-r+2, \dots, n\}$ est <u>injective</u>. Il est immédiat que l'on a $\mathrm{Card}\, H_{m,n,r} = m^{n-r}\, m! \,/\, (m-r)!$. Pour $\varphi \in H_{m,n,r}$, on définit $\delta\varphi$ comme étant l'unique $\sigma \in \mathfrak{S}_n$ telle que la suite $(\varphi\sigma(1)\ ,\ \sigma(1))\ ,\ (\varphi\sigma(2)\ ,\ \sigma(2))\ ,\ \dots,\ (\varphi\sigma(n)\ ,\ \sigma(n))$ soit croissante pour l'ordre lexicographique. Comme précédemment, on a $\varphi\sigma(i) \le \varphi\sigma(i+1)$ pour $i \in [n-1]$, mais l'égalité n'est possible que si l'on a $\sigma(i) < \sigma(i+1)$ et $\sigma(i) \le n-r$ puisque la restriction de φ à l'ensemble $\{n-r+1, n-r+2, \dots, n\}$ est injective. Or d'après la remarque 2.4 , il y a exactement $|\Delta'\Delta''^{r-1}M\sigma|$ indices i tels que $1 \le i < n$, $\sigma(i) < \sigma(i+1)$ et $\sigma(i) \le n-r$. D'après la précédente proposition, on a $\mathrm{Card}\, \delta^{-1}\sigma = \begin{bmatrix} m+s \\ n \end{bmatrix}$ avec $|\Delta'\Delta''^{r-1}M\sigma| = s$ et comme le nombre de $\sigma \in \mathfrak{S}_n$ satisfaisant à $|\Delta'\Delta''^{r-1}M\sigma| = s$ est égal à ${}^{r}A_{n,n-r-s}$, on obtient l'identité désirée.

Q.E.D.

Pour terminer cette section, nous donnons l'espression explicite des coefficients ${}^{r}A_{n,k}$ obtenue par un simple calcul traduisant de nouveau l'inversion de Möbius (Cf. par exemple [26]) à partir des identités (9) et (10) .

En effet, pour $0 < n+r$, on a

$$1/(1-t)^{n+r} = \sum_{0\le k} t^k \begin{bmatrix} n+r-1+k \\ n+r-1 \end{bmatrix} .$$

Donc :

$$\begin{aligned} {}^rA_{n-1+r}(t) \,/\, (1-t)^{n+r} &= \sum_{0\le k} t^k \begin{bmatrix} n+r-1+k \\ n+r-1 \end{bmatrix} \sum_{0\le s\le n-1} {}^rA_{n+r-1,s}\, t^s \\ &= \sum_{0\le j} t^j \sum_{0\le s\le \min(j,n-1)} {}^rA_{n+r-1,s} \begin{bmatrix} n+r-1+j-s \\ n+r-1 \end{bmatrix} \\ &= \sum_{0\le j} t^j \sum_{0\le s\le n-1} {}^rA_{n+r-1,n-1-s} \begin{bmatrix} j+r+s \\ n+r-1 \end{bmatrix} \\ &= \sum_{0\le j} t^j\, (j+r)^{n-1}\, (j+r)!\,/\,j! \end{aligned}$$

en utilisant le fait que $\begin{bmatrix} n+r-1+j-s \\ n+r-1 \end{bmatrix} = 0$ si $j < n-1$ et $s = j+1, \ldots, n-1$ et, pour la dernière étape, en se servant de l'identité (10) . Il en résulte

$${}^rA_{n-1+r}(t) \,/\, (r!(1-t)^{n+r}) = \sum_{0\le j} t^j\, (j+r)^{n-1} \begin{bmatrix} j+r \\ r \end{bmatrix} .$$

On a donc pour $0 \le k \le n-1$, l'expression explicite :

$${}^rA_{n-1+r,k} = r! \sum_{0\le i\le k} (-1)^i\, (k-i+r)^{n-1} \begin{bmatrix} n+r \\ i \end{bmatrix} \begin{bmatrix} k-i+r \\ r \end{bmatrix} . \qquad (11)$$

7. Table des polynômes eulériens.

Le paragraphe 4 du présent chapitre ou l'identité (11) ci-dessus montrent que tous les coefficients des polynômes ${}^rA_n(t)$ sont divisibles par $r!$ (on verra une autre démonstration de ce résultat à la fin du chapitre IV) . Comme on a posé

$${}^rA_n(t) = \sum_{0 \le k \le n-r} {}^rA_{n,k}\, t^k \qquad (0 \le r \le n) \quad ,$$

on peut écrire

$${}^rA_{n,k} = r!\ {}^ra_{n,k}$$

où ${}^ra_{n,k}$ est un entier $(0 \le k \le n-r)$.

Le présent tableau donne les premières valeurs des coefficients ${}^ra_{n,k}$ pour $r = 1 , 2 , 3 , 4 , 5 ,\ r \le n \le 8$ et $0 \le k \le n-r$.

$\boxed{r = 1}$

n \ k	0	1	2	3	4	5	6	7
1	1							
2	1	1						
3	1	4	1					
4	1	11	11	1				
5	1	26	66	26	1			
6	1	57	302	302	57	1		
7	1	120	1191	2416	1191	120	1	
8	1	247	4293	15619	15619	4293	247	1

$\boxed{r = 2}$

n \ k	0	1	2	3	4	5	6
2	1						
3	2	1					
4	4	7	1				
5	8	33	18	1			
6	16	131	171	41	1		
7	32	473	1208	718	88	1	
8	64	1611	7197	8422	2682	183	1

$\boxed{r = 3}$

n \ k	0	1	2	3	4	5
3	1					
4	3	1				
5	9	10	1			
6	27	67	25	1		
7	81	376	326	56	1	
8	243	1909	3134	1314	119	1

$\boxed{r = 4}$

n \ k	0	1	2	3	4
4	1				
5	4	1			
6	16	13	1		
7	64	113	32	1	
8	256	821	531	71	1

- 46 -

$\boxed{r = 5}$

n \ k	0	1	2	3
5	1			
6	5	1		
7	25	16	1	
8	125	171	39	1

*

CHAPITRE III

LA FORMULE EXPONENTIELLE.

Les trois premières sections de ce chapitre contiennent la définition et quelques propriétés d'une construction très générale que nous appelons "composé partitionnel". La motivation de cette notion apparaît dans les sections suivantes, ainsi qu'au chapitre IV , où nous appliquons toutes ces techniques aux polynômes eulériens. Comme nous l'avons déjà mentionné, ces résultats ont été fréquemment étudiés en liaison avec divers problèmes d'énumération et tout particulièrement dans [29] , [14] et [15] .

Dans ce chapitre, si Z est un ensemble non vide, on note Z^* et Z^+ les monoïdes libre et abélien libre engendrés par Z . On appellera mots les éléments de Z^* , qu'on présentera comme des suites $g = z_1 z_2 \dots z_r$ où $z_1, z_2, \dots, z_r$ appartiennent à Z ; l'entier r est la longueur du mot g . On appellera monômes les éléments de Z^+ . Si α est le morphisme canonique de Z^* sur Z^+ , on prendra dans chaque classe $\alpha^{-1}(f)$ où $f \in Z^+$ un mot canonique qu'on identifiera à f . Le monôme f sera dit de degré r si le mot f est de longueur r .

1. La formule de Hurwitz.

Dans cette section, Y est un ensemble non vide, muni d'une application $\lambda : Y \to \mathbb{N}$. Le même symbole désignera les morphismes dans $\mathbb{N}$ étendant cette application aux monoïdes Y^* et Y^+ . Le morphisme canonique de Y^* sur Y^+ sera noté α .

Soit $\mathcal{P}$ l'ensemble des parties *finies* (y compris la partie vide) de $\mathbb{N}$; on considère le sous-ensemble Y_λ du produit cartésien $Y \times \mathcal{P}$ composé de tous les couples (y, I) satisfaisant à la condition

$$\text{Card } I = \lambda y \quad .$$

On forme ensuite le monoïde libre Y_λ^* engendré par Y_λ .

Soit

$$h = (y_1, I_1)\ (y_2, I_2)\ \dots\ (y_r, I_r)$$

un mot de Y_λ^* de longueur $r > 0$. On pose

$$\beta h = g = y_1\ y_2\ \dots\ y_r \in Y^* \quad \text{et} \quad \lambda h = \lambda g \quad . \tag{1}$$

DÉFINITION 3.1.

Pour chaque entier $r > 0$, le *composé partitionnel marqué* de Y *de degré* r est le sous-ensemble $Y^{((r))}$ de Y_λ^* formé de tous les mots $h = (y_1, I_1)\ (y_2, I_2)\ \dots\ (y_r, I_r)$ de longueur r satisfaisant aux conditions

i) $I_j \cap I_{j'} = \emptyset$ si $j \neq j'$;

ii) $\cup \{I_j : j \in [r]\} = [\lambda h]$.

On notera que si λy_j est strictement positif pour tout $j \in [r]$, la *famille* $\{I_1, I_2, \dots, I_r\}$ est une *partition* de l'ensemble $[\lambda h]$, aucun de ces sous-ensembles n'étant vide.

Par convention, on supposera l'existence d'un ensemble $Y^{((r))}$ pour $r = 0$ contenant un seul élément, à savoir l'élément neutre de Y . Ce dernier est envoyé par β sur l'élément neutre de Y^* . On identifiera, d'autre part, le composé partitionnel marqué $Y^{((1))}$ de degré 1 avec Y .

THÉORÈME 3.2. (Formule de Hurwitz).

<u>Soit</u> $E(Y) = \Sigma\ \{y/\lambda y! : y \in Y\}$ <u>la fonction génératrice exponentielle de</u> Y <u>(par rapport à</u> λ) . <u>Pour tout</u> $r \geq 0$, <u>on a dans la</u> $\underset{\approx}{Q}$-<u>algèbre large de</u> Y^* <u>l'identité</u>

$$(E(Y))^r = \Sigma\ \{\beta h/\lambda h! : h \in Y^{((r))}\} \quad . \tag{2}$$

PREUVE.

Avec nos conventions sur β , il n'y a rien à prouver pour $r = 0$. Pour chaque mot $g = y_1\ y_2 \ldots y_r$ de longueur r $(r > 0)$ de Y^* , le nombre de mots $h \in Y^{((r))}$ tels que $\beta h = g$ est égal au nombre de suites $(I_1, I_2, \ldots, I_r)$ de parties de $\underset{\approx}{N}$ satisfaisant aux conditions i) et ii) de la définition 3.1 ainsi qu'à la condition $\text{Card } I_j = \lambda y_j$ pour chaque $j \in [r]$. Or le nombre de telles suites est évidemment donné par le coefficient multinomial

$$\underset{\approx}{\text{Mult}}\ (\lambda, g) = \lambda g!/(\lambda y_1!\ \lambda y_2! \ldots \lambda y_r!) \quad .$$

Donc

$$\Sigma\ \{\beta h/\lambda h! : h \in Y^{((r))} , \beta h = g\}$$

est égal au produit $(1/(\lambda y_1!\ \lambda y_2!\ \dots\ \lambda y_r!))g$ puisque $\lambda h = \lambda g$ pour $\beta h = g$. Or le facteur $(1/(\lambda y_1!\ \lambda y_2!\ \dots\ \lambda y_r!))$ est simplement le coefficient de g dans le développement de $(\ E(Y)\ )^r$ et le résultat s'en déduit par sommation sur tous les mots de longueur r de Y^* .

Q.E.D.

2. Le composé partitionnel.

Nous conservons les mêmes notations que dans la section 1 , mais nous supposons cette fois que $\lambda y > 0$ pour tout $y \in Y$. Soit $h = (y_1,I_1)\ (y_2,I_2)\ \dots\ (y_r,I_r)$ un mot de $Y^{((r))}$. L'hypothèse ci-dessus entraîne, puisque l'on a $\mathrm{Card}\ I_j = \lambda y_j$ pour chaque $j \in [r]$, que les sous-ensembles $I_1, I_2, \dots, I_r$ sont non vides, donc tous distincts. Il en résulte que h est multilinéaire, c'est-à-dire a toutes ses lettres distinctes. En notant δ le morphisme canonique de Y_λ^* sur Y_λ^+ , on voit donc que la classe abélienne $\delta^{-1}\delta h$ contient exactement $r!$ mots. De plus, les conditions i) et ii) de la définition 3.1 ne faisant pas intervenir l'ordre des lettres de h , il s'en suit que $Y^{((r))}$ contient toute une classe abélienne $\delta^{-1}\delta h$ dès qu'il contient h .

DÉFINITION 3.3.

On appelle composé partitionnel de Y de degré r $(r \geq 0)$, l'ensemble

$$Y^{(r)} = \delta Y^{((r))}$$

et l'union

$$Y^{(+)} = \bigcup_{0\leq r} Y^{(r)}$$

est le composé partitionnel de Y .

Les éléments de $Y^{(r)}$ sont donc des monômes $f = (y_1,I_1)\,(y_2,I_2)\,\ldots\,(y_r,I_r)$ de Y^+_λ . On désigne par γf le monôme $m = y_1\,y_2\,\ldots\,y_s$ de Y^+ et l'on pose encore $\lambda f = \lambda m$. On a donc l'identité

$$\alpha\beta = \gamma\delta$$

où α est le morphisme $\alpha : Y^* \to Y^+$ et où β a été défini en (1) . On peut rassembler les remarques ainsi faites dans un lemme.

LEMME 3.4.

Pour tout $f \in Y^{(r)}$, on a

$$\mathrm{Card}\,\{h \in Y^{((r))} : \delta h = f\} = r!$$

et si $\delta h = f$, on a

$$\gamma f = \alpha\beta h \quad .$$

Venons-en à la formule fondamentale de ce chapitre.

THÉORÈME 3.5. (Formule exponentielle).

Dans la $\mathbb{Q}$-algèbre large de Y^+ , on a l'identité

$$\Sigma\;\{\gamma f/\lambda f! \;:\; f \in Y^{(+)}\} = \exp\,E(Y) \quad .$$

PREUVE.

On a $\Sigma \{\gamma f/\lambda f! : f \in Y^{(0)}\} = 1$. D'autre part, pour $f \in Y^{(r)}$ $(r > 0)$, on a

$$\Sigma \{\alpha\beta h/\lambda h! : \delta h = f\} = r! \quad (\gamma f/\lambda f!)$$

d'après le lemme précédent. D'où, il résulte

$$\Sigma \{\gamma f/\lambda f! : f \in Y^{(r)}\} = (1/r!)\, \Sigma \{\alpha\beta h/\lambda h! : h \in Y^{((r))}\}$$

pour chaque $r > 0$. Utilisant le théorème 3.2 , on obtient donc par sommation sur tous les $r \in \mathbb{N}$

$$\Sigma \{\gamma f/\lambda f! : f \in Y^{(+)}\} = \sum_{0 \leq r} (1/r!)\,(\, E(Y)\,)^r$$

$$= \exp E(Y) \quad .$$

Q.E.D.

<u>Remarque</u> 3.6.

Dans la $\mathbb{Q}$-algèbre large de Y^+ , on a pris la topologie des séries formelles induite par l'ordre o suivant : si $\underset{\approx}{a} = \Sigma \{m\, a_m : m \in Y^+\}$ est une série formelle, son ordre $o(\underset{\approx}{a})$ est défini par

$$o(\underset{\approx}{a}) = \inf \{n > 0 : \lambda m = n \,,\, a_m \neq 0\} \quad .$$

Utilisons les notations abrégées

$$\gamma\{Y^{(+)} \cap \lambda^{-1}n\} = \Sigma\ \{\gamma f : f \in Y^{(+)} , \lambda f = n\} \quad \text{pour} \quad n \geq 0$$

et $$\{Y_n\} = \Sigma\ \{y : y \in Y , \lambda y = n\} \quad \text{pour} \quad n > 0 .$$

Les séries formelles $\gamma\{Y^{(+)} \cap \lambda^{-1}n\}$ et $\{Y_n\}$ sont d'ordre égal à n , ce qui permet d'écrire la formule exponentielle sous la forme

$$\sum_{0\leq n} (1/n!)\ \gamma\{Y^{(+)} \cap \lambda^{-1}n\} = \exp\Big[\sum_{0\leq n} (1/n!)\ \{Y_n\} \Big] \quad . \tag{3}$$

3. Une formule d'inversion pour les séries exponentielles.

Désignons par $z(f)$ pour $f \in Y^{(+)}$ l'unique $r \in \mathbb{N}$ tel que $f \in Y^{(r)}$; l'entier $z(f)$ n'est autre que le degré de f . Posons

$$\bar{\gamma}\{Y^{(+)} \cap \lambda^{-1}n\} = \Sigma\ \{\gamma f \,.\, (-1)^{z(f)+n} : f \in Y^{(+)} , \lambda f = n\}$$

pour $n \geq 0$.

PROPRIÉTÉ 3.7.

Dans la $\mathbb{Q}$-algèbre large de Y^{+} , on a l'identité

$$\Big(\sum_{0\leq n} (1/n!)\ \gamma\{Y^{(+)} \cap \lambda^{-1}n\} \Big)^{-1} = \sum_{0\leq n} ((-1)^{n}/n!)\ \bar{\gamma}\{Y^{(+)} \cap \lambda^{-1}n\} \quad . \tag{4}$$

PREUVE.

D'après le théorème 3.5 , le membre de gauche de l'identité à établir, soit U , est égal à $(\exp E(Y))^{-1}$, c'est-à-dire à $\exp(-E(Y))$. Notons φ le morphisme envoyant sur $-y$ chaque $y \in Y$, ceci équivaut à $U = \varphi \exp E(Y)$, donc de nouveau d'après le théorème 3.5 , à

$$U = \varphi \sum_{0 \le n} (1/n!)\ \gamma\{Y^{(+)} \cap \lambda^{-1}n\}$$

$$= \sum_{0\le n} (1/n!) \sum_{0 \le r} (-1)^r\ \gamma\{Y^{(+)} \cap \lambda^{-1}n \cap z^{-1}r\}$$

$$= \sum_{0 \le n} ((-1)^n/n!)\ \bar{\gamma}\{Y^{(+)} \cap \lambda^{-1}n\} \quad .$$

Q.E.D.

Ceci termine l'établissement des formules que nous utiliserons par la suite. Le théorème 3.2 avec une interprétation adéquate des objets en cause exprime que la <u>transformation de Borel</u>

$$\Sigma\ \{y : y \in Y\} \to \Sigma\ \{y/\lambda y! : y \in Y\}$$

est un morphisme dans l'algèbre large de Y^+ de l'algèbre large de base Y par rapport au "<u>produit d'intercalement</u>" ("<u>shuffle</u>" de Chen, Fox et Lyndon). Le théorème 3.5 est appelé "formule de Cauchy" dans les problèmes concernant le groupe symétrique. Sous une forme ou sous une autre, elle a été retrouvée et utilisée souvent dans diverses questions d'énumération. Nous l'appellerons simplement <u>formule exponentielle</u>. En prenant le logarithme des deux membres on obtiendrait évidemment la fonction génératrice exponentielle $E(Y)$ de Y en fonction de la fonction génératrice exponentielle du composé partitionnel $Y^{(+)}$ de Y.

Nous utiliserons par la suite la

DÉFINITION 3.8.

Soit A un monoïde abélien ; une application $\mu : Y^{(+)} \to A$ sera dite <u>multiplicative</u> ssi il existe un morphisme $\mu' : Y^{+} \to A$ tel que le diagramme

$$\begin{tikzcd} Y^{(+)} \arrow[r, "\gamma"] \arrow[d, "\mu"'] & Y^{+} \arrow[dl, "\mu'"] \\ A & \end{tikzcd}$$

soit commutatif.

4. <u>Le composé partitionnel des applications</u>.

Il existe de nombreuses familles de structures qui peuvent être considérées comme le composé partitionnel d'une de leurs sous-familles. Nous examinerons ici, à titre d'exemple, la famille des applications avec le but d'introduire les notions nécessaires pour traiter le cas particulier des permutations.

DÉFINITION 3.9.

Soit $f : I \to I$ une application d'un ensemble fini I dans lui-même. L'<u>équivalence</u> f^* de f est la relation d'équivalence dans $I \times I$ telle que $i , i' \in I$ appartiennent à la même classe ssi il existe des itérées f^p et $f^{p'}$ de f satisfaisant à $f^p(i) = f^{p'}(i')$.

Nous appellerons <u>sous-domaines</u> de f les classes de cette équivalence et leur nombre sera désigné par $z(f)$. L'application f sera <u>connexe</u> si $z(f) = 1$. Ainsi les sous-domaines d'une permutation sont les orbites de celle-ci ; les permutations circulaires sont les permutations connexes.

Dans la suite, on notera F_n l'ensemble des applications de $[n]$ dans lui-même $(n \geq 0)$ et l'on posera

$$F = \bigcup_{0 \leq n} F_n \; .$$

Désignons par $I_1, I_2, \ldots, I_r$ $(r = z(f))$ les sous-domaines d'une application $f \in F_n$ $(n > 0)$. Pour tout $j \in [r]$, on note ω_j l'unique morphisme (d'ensembles ordonnés) $\omega_j : [\mathrm{Card}\ I_j] \to [n]$ qui a pour image I_j et f_j' la restriction de f à I_j . Par définition de l'équivalence f^* on voit que $f_j'(I_j) \subset I_j$ et il est licite de poser $f_j = \omega_j^{-1} f_j' \omega_j$ $(j \in [r])$. Les applications f_j envoient $[\mathrm{Card}\ I_j]$ dans lui-même et sont toutes connexes $(j \in [r])$. Enfin, il est clair que toute application $f \in F_n$ détermine, de façon biunivoque le <u>monôme</u> (appartenant au monoïde

$(F \times \mathcal{P})^+$, où $\mathcal{P}$ désigne toujours l'ensemble des parties finies de $\mathbb{N}$)

$$(f_1,I_1)\,(f_2,I_2)\,\ldots\,(f_r,I_r)$$

que l'on appellera sa <u>factorisation canonique</u>, les f_j eux-mêmes étant les <u>facteurs</u> de f . Par commodité, on identifiera tout $f \in F$ avec sa factorisation canonique et $f \in F_0$ avec le monôme unité.

Le raccordement avec les trois premières sections se fait de la façon suivante. Soit donnée une famille $\mathfrak{F}$ d'applications connexes dont les domaines sont des ensembles de la forme $[n]$ $(n \in \mathbb{N})$. Posons $Y = \mathfrak{F}$ et prenons pour λ l'application qui envoie sur n chaque $f \in \mathfrak{F}$ de domaine $[n]$ $(n \in \mathbb{N})$. Formons ensuite le composé partitionnel $\mathfrak{F}^{(+)}$. On constate alors que la factorisation canonique d'une application $f \in F$ appartient au composé partitionnel $\mathfrak{F}^{(+)}$ si et seulement si les facteurs de f appartiennent à $\mathfrak{F}$. Avec l'identification faite ci-dessus, on a ainsi la proposition suivante :

PROPOSITION 3.10.

<u>Soit</u> $\mathfrak{F} \subset F$ <u>une famille d'applications connexes</u>. <u>Le composé partitionnel</u> $\mathfrak{F}^{(+)}$ <u>est l'ensemble des applications</u> $f \in F$ <u>dont les facteurs appartiennent à</u> $\mathfrak{F}$.

La propriété suivante découle immédiatement de la définition du composé partitionnel d'un ensemble d'applications. Elle exprime le fait que la factorisation canonique d'une application f conserve les excédances et les points fixes de f . De façon précise, et en conservant les notations ci-dessus, on a

PROPRIÉTÉ 3.11.

Soit $(f_1,I_1)\ (f_2,I_2)\ \dots\ (f_r,I_r)\quad (r > 0)$ la factorisation canonique d'une application f . Pour tout $j \in [r]$, le morphisme $\tau_j = \omega_j^{-1}$ est une bijection de I_j sur $[\mathrm{Card}\ I_j]$ telle que pour tout $i \in I_j$ on ait les équivalences

$$i < f(i) \Leftrightarrow \tau_j(i) < f_j\,\tau_j(i)$$

$$i = f(i) \Leftrightarrow \tau_j(i) = f_j\,\tau_j(i)$$

$$i > f(i) \Leftrightarrow \tau_j(i) > f_j\,\tau_j(i) \quad .$$

PREUVE.

En effet, si l'entier i est dans le sous-domaine I_j , on a $f_j(i) = \omega_j^{-1}\ f_j{}'\,\omega_j(i) = \tau_j\ f_j{}'\,\tau_j^{-1}(i)$ où $f_j{}'$ est la restriction de f à I_j . Les équivalences ci-dessus résultent alors du fait que $\tau_j : I_j \to [\mathrm{Card}\ I_j]$ est un morphisme strictement croissant.

Q.E.D.

Réécrivons la formule exponentielle (3) et la formule d'inversion (4) dans ce cas particulier du composé partitionnel des applications. On a d'abord

$$\mathfrak{F}^{(+)} \cap \lambda^{-1}n = F_n \cap \mathfrak{F}^{(+)} \qquad \text{pour} \quad n \geq 0$$

et

$$\mathfrak{F} \cap \lambda^{-1}n = F_n \cap \mathfrak{F} \qquad \text{pour} \quad n > 0$$

et les deux identités (3) et (4) se présentent ainsi

$$\sum_{0\le n} (1/n!)\ \gamma\{F_n \cap \mathfrak{F}^{(+)}\} = \exp\Big[\sum_{0<n} (1/n!)\ \{F_n \cap \mathfrak{F}\}\Big] \tag{5}$$

$$\Big(\sum_{0\le n} (1/n!)\ \gamma\{F_n \cap \mathfrak{F}^{(+)}\}\Big)^{-1} = \sum_{0\le n} ((-1)^n/n!)\ \bar{\gamma}\{F_n \cap \mathfrak{F}^{(+)}\} \ . \tag{6}$$

En fait, ces deux identités seront appliquées ci-après sous la forme suivante. On suppose donnée une application multiplicative $\mu : \mathfrak{F}^{(+)} \to \Omega$ (Cf. définition 3.8) . On forme ensuite l'algèbre sur $\mathbb{Q}$ du monoïde Ω , soit $\bar{\Omega}$ et l'on considère l'algèbre $\bar{\Omega}[[u]]$ des séries formelles à coefficients dans $\bar{\Omega}$ et à une indéterminée u . On a ainsi la

PROPOSITION 3.12.

<u>Soit</u> $\mu : \mathfrak{F}^{(+)} \to \Omega$ <u>une application multiplicative</u>. <u>Dans l'algèbre des séries formelles</u> $\bar{\Omega}[[u]]$, <u>on a les identités</u>

$$\sum_{0\le n} (u^n/n!)\ \mu\{F_n \cap \mathfrak{F}^{(+)}\} = \exp\Big[\sum_{0<n} (u^n/n!)\ \mu\{F_n \cap \mathfrak{F}\}\Big] \tag{7}$$

$$\Big(\sum_{0\le n} (u^n/n!)\ \mu\{F_n \cap \mathfrak{F}^{(+)}\}\Big)^{-1} = \sum_{0\le n} ((-u)^n/n!)\ \bar{\mu}\{F_n \cap \mathfrak{F}^{(+)}\} \tag{8}$$

<u>où</u> $\bar{\mu}f = (-1)^{z(f)+n}\,\mu f$ <u>pour tout</u> $f \in F_n \cap \mathfrak{F}^{(+)}$ $(n \ge 0)$.

PREUVE.

Soit μ' le morphisme de $\mathfrak{F}^{+}$ dans Ω tel que $\mu = \mu'\gamma$. Désignons par φ l'application envoyant tout $f \in F_n \cap \mathfrak{F}$ sur le monôme

$u^n \mu' f$ $(n > 0)$. Comme μ est multiplicative, on a $\varphi\gamma f = u^n \mu f$ pour tout $f \in F_n \cap \mathfrak{F}^{(+)}$ $(n \geq 0)$. Tous les monômes $\varphi\gamma f$ où $f \in F_n \cap \mathfrak{F}^{(+)}$ sont donc de degré n (en u) . On peut donc prolonger φ en un morphisme continu de la $\underset{\approx}{Q}$-algèbre large de $\mathfrak{F}^+$ dans $\overline{\Omega}[[u]]$. Appliquant ainsi φ aux deux membres des deux identités (5) et (6) , on obtient les identités (7) et (8) .

Q.E.D.

5. Applications.

Il est évident que si $\mathfrak{F}$ est l'ensemble $\mathfrak{C}$ des permutations circulaires, le composé partitionnel $\mathfrak{C}^{(+)}$ est exactement l'ensemble $\mathfrak{S} = \bigcup_{0 \leq n} \mathfrak{S}_n$. On a de plus

$$F_n \cap \mathfrak{F}^{(+)} = \mathfrak{S}_n \quad \text{pour} \quad n \geq 0$$

et

$$F_n \cap \mathfrak{F} = \mathfrak{C}_n \quad \text{pour} \quad n > 0 \quad .$$

Enfin, si σ est dans $\mathfrak{S}_n$, se rappelant que $z(\sigma)$ est le nombre des orbites de σ , on voit que le coefficient $(-1)^{z(\sigma)+n}$ est la signature $\epsilon(\sigma)$ de σ . Dans ces conditions les deux identités (7) et (8) s'écrivent :

$$\sum_{0 \leq n} (u^n/n!)\, \mu\{\mathfrak{S}_n\} = \exp\left[\sum_{0<n} (u^n/n!)\, \mu\{\mathfrak{C}_n\}\right] \tag{9}$$

$$\left(\sum_{0 \leq n} (u^n/n!)\, \mu\{\mathfrak{S}_n\}\right)^{-1} = \sum_{0 \leq n} ((-u)^n/n!)\, \overline{\mu}\{\mathfrak{S}_n\} \tag{10}$$

où $\overline{\mu}\sigma = \epsilon(\sigma)\mu\sigma$ pour tout $\sigma \in \mathfrak{S}$.

Donnons quelques exemples d'application des formules (9) et (10) .

Soit $(x_n)_{n\geq 1}$ une suite d'indéterminées commutatives. Si σ est dans $\mathfrak{S}_n$, posons $\mu\sigma = x_1^{m_1} x_2^{m_2} \dots x_n^{m_n}$ où pour tout $k \in [n]$, l'entier m_k est le nombre de cycles de longueur k dans la permutation σ . La fonction μ ainsi définie est multiplicative. Dans ce cas, $\mu\{\mathfrak{S}_n\}$ est le <u>polynôme indicateur de cycles</u> de $\mathfrak{S}_n$ ou polynôme de Bell (Cf. [24] p. 68) et $(u^n/n!)\ \mu\{\mathfrak{C}_n\}$ se réduit à $u^n x_n/n$ puisque Card $\mathfrak{C}_n = (n-1)!$. La formule exponentielle permet donc de retrouver l'expression explicite de la fonction génératrice exponentielle de ces polynômes, à savoir

$$\sum_{0\leq n} (u^n/n!)\ \mu\{\mathfrak{S}_n\} = \exp\left[\sum_{0<n} u^n x_n/n\right] \quad .$$

Un autre exemple de composé partitionnel est donné par l'ensemble U des applications <u>ultimement idempotentes</u>, c'est-à-dire, pour tout $n \geq 0$, des applications $f \in F_n$ telles que $f^n = f^{n-1}$. Considérons, en effet, pour tout $n > 0$, l'ensemble V_n des applications $f \in F_n$, dont l'image de la $(n-1)$-ième itérée f^{n-1} soit réduite à un seul point. Les éléments de V_n sont encore appelés "<u>arborescences</u>". Posons $V = \bigcup_{0<n} V_n$; il est alors clair que le composé partitionnel de V est l'ensemble U . Posant $U_n = F_n \cap U$ pour $n \geq 0$, on obtient donc pour toute application multiplicative $\mu : U \to \Omega$, deux identités analogues à (9) et (10) en substituant U_n à $\mathfrak{S}_n$, V_n à $\mathfrak{C}_n$ et en posant $\bar{\mu}f = (-1)^{z(f)+n}$ pour tout $f \in U_n$ $(n > 0)$. On pose naturellement $\mu U_0 = \bar{\mu} U_0 = 1$.

En particulier, prenons pour μ l'application qui envoie sur 1 tout $f \in U$. Comme on a, de façon évidente,

$$\text{Card } V_n = n \text{ Card } U_{n-1} \quad \text{pour} \quad n > 0 \qquad ,$$

il vient

$$\sum_{0 \leq n} (u^n/n!)\ \mu\{V_n\} = u \sum_{0 \leq n} (u^n/n!) \text{ Card } U_n$$

et on retrouve, en appliquant la formule (7) , ce résultat bien connu : que la série formelle $w = \sum_{0 \leq n} (u^n/n!) \text{ Card } U_n$ est solution dans $\underset{\approx}{Q}[[u]]$ de l'équation

$$w = \exp(uw) \qquad .$$

Enfin, prenons pour $\mathfrak{F}$ l'ensemble G de toutes les applications connexes de F . Dans ce cas, le composé partitionnel de G est F tout entier et l'on obtient encore, pour toute application multiplicative $\mu : F \to \Omega$ donnée, deux identités analogues à (9) et à (10) .

6. <u>Une identité entre déterminants et permanents</u>.

Dans l'énoncé qui suit, Ξ est une matrice infinie $\Xi = (\xi_{i,j})\ (i,j=1,2,\ldots)$ à coefficients dans un anneau commutatif $\bar{\Omega}$; on désigne pour tout $n > 0$, par Ξ_n la matrice $(\xi_{i,j})\ (1 \leq i,j \leq n)$, par $\underset{\approx}{\text{Det}}\ \Xi_n$ son déterminant et par $\underset{\approx}{\text{Per}}\ \Xi_n$ son permanent.

THÉORÈME 3.13.

Soient a , b *et* c *trois éléments de* $\bar{\Omega}$ *et* Ξ *une matrice infinie ayant ses coefficients supradiagonaux* (*resp. diagonaux, infradiagonaux*) *égaux à* a (*resp.* b , c) . *On a l'identité*

$$(1 + \sum_{0<n} (u^n/n!) \operatorname{Per} \Xi_n)^{-1} = 1 + \sum_{0<n} ((-u)^n/n!) \operatorname{Det} \Xi_n \quad . \tag{11}$$

PREUVE.

Désignons par $\xi_{i,j}$ les éléments de la matrice Ξ , puis posons $\mu\sigma = \xi_{1,\sigma(1)}\, \xi_{2,\sigma(2)} \cdots \xi_{n,\sigma(n)}$ pour tout $\sigma \in \mathfrak{S}_n$.

Avec les notations de la propriété 3.11 (en prenant $f = \sigma$), on voit que pour un entier i appartenant au sous-domaine (i.e. à l'orbite) I_j , on a

$$\xi_{i,\sigma(i)} = \xi_{\tau_j(i), f_j \tau_j(i)} \quad .$$

On a ainsi $\mu\sigma = \prod_j \prod_i \xi_{i,\sigma(i)}$ où j varie dans $[r]$, et où i , pour j fixé, varie dans I_j . Si i est dans I_j , $i' = \tau_j(i)$ est dans $[\operatorname{Card} I_j]$ et l'on a, d'après ce qui précède

$$\mu\sigma = \prod_j \prod_{i'} \xi_{i', f_j(i')} = \prod_j \mu f_j \quad .$$

L'application μ est donc multiplicative et l'on peut écrire

$$\mu\{\mathfrak{S}_n\} = \sum_\sigma \xi_{1,\sigma(1)} \cdots \xi_{n,\sigma(n)} = \operatorname{Per} \Xi_n$$

$$\bar{\mu}\{\mathfrak{S}_n\} = \sum_{\sigma} \epsilon(\sigma)\, \xi_{1,\sigma(1)} \cdots \xi_{n,\sigma(n)} = \mathrm{Det}\, \Xi_n \quad .$$

Le théorème 3.13 résulte alors de l'identité (10) .

Q.E.D.

Remarque 3.14.

Notons que l'identité (11) n'est pas vraie pour toutes les matrices, mais comme l'a remarqué Kittel [18] , on peut construire d'autres matrices infinies Ξ que celles considérées dans l'énoncé du théorème 3.13 pour lesquelles l'identité (11) est vérifiée. Par exemple, si la première colonne de la matrice Ξ n'a que des zéros, on a $\mathrm{Per}\, \Xi_n = \mathrm{Det}\, \Xi_n = 0$ pour tout $n > 0$ et l'identité (11) est trivialement vérifiée.

De même considérons la matrice Ξ définie par

$$\begin{aligned} \xi_{i,j} &= 1 && \text{si} \quad 1 \le i \le j \\ &= -(i-1) && \text{si} \quad 1 \le i-1 = j \\ &= 0 && \text{si} \quad 1 \le j < i-1 \quad ; \end{aligned}$$

on obtient facilement $\mathrm{Per}\, \Xi_1 = 1$ et $\mathrm{Per}\, \Xi_n = 0$ pour tout $n > 1$, ainsi que $\mathrm{Det}\, \Xi_n = n!$ pour tout $n \ge 1$. L'identité (11) est encore vérifiée ; on retrouve en fait l'identité

$$(1 + u)^{-1} = \sum_{0 \le n} (-u)^n \quad .$$

Notons enfin le résultat élémentaire $(n > 0)$

$$\underset{\sim\sim\sim}{\text{Det}}\ \Xi_n = (\ c(b-a)^n - a(b-c)^n\ )\ /\ (c-a) \qquad \text{si} \quad c \neq a$$

$$= (b-a)^{n-1}\ (b + (n-1)a) \qquad \text{si} \quad c = a \quad .$$

Portant ces valeurs dans la formule (11) , on est conduit à l'identité :

$$(1 + \sum_{0<n} (u^n/n!)\ \underset{\sim\sim\sim}{\text{Per}}\ \Xi_n)^{-1} = c\ \exp(\ (a-b)u\ ) - a\ \exp(\ (c-b)u\ )\ )\ )/(c-a)\ (\text{si} \quad c \neq a\ )$$

$$= (1 - au)\ \exp(\ (a-b)u\ ) \qquad (\text{si} \quad c = a)\ . \qquad (12)$$

*

CHAPITRE IV

FONCTIONS GÉNÉRATRICES DES POLYNÔMES EULÉRIENS.

1. Fonction génératrice exponentielle de ${}^0A_n(t)$, $A_n(t)$, $B_n(t)$.

Pour $\sigma \in \mathfrak{S}_n$ $(n > 0)$, nous définissons $E'\sigma \in \{0,1\}^n$ par la condition que $E'\sigma(k) = 1$ ou 0 selon que k est ou non un point fixe de σ . De par la définition de E et ΔE on a donc immédiatement :

$$|E\sigma| = |E'\sigma| + |\Delta E\sigma| \quad .$$

Introduisant une nouvelle indéterminée t' , nous posons $\theta'\sigma = t'^{|E'\sigma|}\, t^{|\Delta E\sigma|}$ et $\bar{A}_n(t,t') = \theta'\mathfrak{S}_n$ (=1 pour $n=0$) . Par conséquent on a

$$ {}^0A_n(t) = \bar{A}_n(t,t) \qquad ; $$

$$ A_n(t) = \bar{A}_n(t,1) \qquad \text{et} $$

$$ \bar{A}_n(t,0) = \theta\Delta E\{\sigma \in \mathfrak{S}_n : |E'\sigma| = 0\} = \theta\Delta E\,\mathcal{D}_n = \theta E\,\mathcal{D}_n = B_n(t) $$

(voir la fin du paragraphe 1 du chapitre II).

THÉORÈME 4.1.

<u>On a</u>

$$ \bar{A}(t,t',u) = \sum_{0\le n} (u^n/n!)\, \bar{A}_n(t,t') = \exp(ut' + C(t,u)) \quad (1) $$

où

$$C(t,u) = \sum_{2\leq n} (u^n/n!)\ t\ A_{n-1}(t) \quad . \tag{2}$$

PREUVE.

Que θ' soit multiplicative découle de la propriété 3.11 et des définitions des vecteurs $E'\sigma$ et $\Delta E\sigma$ $(\sigma \in \mathfrak{S}_n)$. Par conséquent, le membre de droite de l'identité (9) du chapitre III devient $\exp(\sum_{0<n} (u^n/n!)\ \theta'\{\mathfrak{C}_n\})$. Pour $n = 1$, on a $\theta'\{\mathfrak{C}_n\} = t'$ et pour $n > 1$, on a d'après les propriétés 2.2 et 2.3, $\theta'\{\mathfrak{C}_n\} = \theta\Delta E\{\mathfrak{C}_n\} = t\ A_{n-1}(t)$.

Q.E.D.

Le théorème 4.1 nous a donné une identité sur les polynômes $\bar{A}_n(t,t')$. Nous allons maintenant trouver une formule explicite pour la fonction génératrice $\bar{A}(t,t',u)$ en utilisant les résultats de la section 6 du chapitre III.

THÉORÈME 4.2.

On a

$$\bar{A}(t,t',u) = \sum_{0\leq n} (u^n/n!)\ \bar{A}_n(t,t') = (1-t)/(\exp((t-t')u) - t\ \exp((1-t')u)) \ . \tag{3}$$

En particulier :

$$\bar{A}(t,t,u) = \sum_{0\le n} (u^n/n!)\ {}^0A_n(t) = (1-t)\ /\ (1 - t\ \exp((1-t)u)) \qquad (4)$$

$$\bar{A}(t,1,u) = \sum_{0\le n} (u^n/n!)\ A_n(t) = (1-t)\ /\ (-t + \exp((t-1)u)) \qquad (5)$$

$$\bar{A}(t,0,u) = \sum_{0\le n} (u^n/n!)\ B_n(t) = (1-t)\ /\ (\exp(ut) - t\ \exp(u)) \quad . \qquad (6)$$

PREUVE.

Avec les notations du théorème 3.13 , si l'on pose $a=t$, $b=t'$ et $c=1$, on a pour $\sigma \in \mathfrak{S}_n$ $(n > 0)$ l'égalité $\xi_{1,\sigma(1)} \cdots \xi_{n,\sigma(n)} = \theta'\sigma$; soit $\bar{A}_n(t,t') = \underset{\approx}{\mathrm{Per}}\ \Xi_n$. La première identité résulte donc de la formule (11) du chapitre III . En posant successivement $t'=t$, puis $t'=1$, enfin $t'=0$, on obtient les trois suivantes.

Q.E.D.

Remarque 4.3.

Ces formules peuvent aussi s'obtenir par le procédé suivant. D'après la propriété 2.2 , on a ${}^0A_n(t) = t\ A_n(t)$ pour tout $n > 0$; on en tire

$$\bar{A}(t,t,u) = \sum_{0\le n} (u^n/n!)\ {}^0A_n(t) = 1 + t \sum_{0<n} (u^n/n!)\ A_n(t) \quad , \quad \text{soit}$$

$$\bar{A}(t,t,u) = 1 + t(\bar{A}(t,1,u) - 1) \quad . \qquad (7)$$

D'autre part, d'après le théorème 3.9 on a

$$\exp\,[C(t,u)] = \bar{A}(t,1,u)\ \exp(-u) \tag{8}$$

ou encore

$$\bar{A}(t,t,u) = \exp(ut - u)\ \bar{A}(t,1,u) \quad . \tag{9}$$

Du système formé par les deux équations (7) et (9) , on déduit immédiatement les identités (4) et (5) . On calcule ensuite $C(t,u)$ en utilisant la formule (8) et l'on en tire l'identité (3) en se servant du théorème 4.1 .

<u>Remarque</u> 4.4.

Les formules (4) et (5) sont connues (cf. Riordan [24] p. 215 & 39). La formule (6) a été obtenue par Roselle [25] , par les méthodes traditionnelles du calcul différentiel et intégral, dans le cas particulier où $B_n(t) = \theta M\mathfrak{G}_n$ $(n > 0)$.

Notons encore que du théorème 4.1 résulte immédiatement, par simple dérivation, que la fonction génératrice $\underset{\approx}{A} = \bar{A}(t,1,u)$ est solution de l'équation différentielle de Bernoulli :

$$\frac{\partial}{\partial u}\underset{\approx}{A} = \underset{\approx}{A}(1 + t(\underset{\approx}{A} - 1)\,) \quad .$$

On peut aussi prouver ce résultat directement et pour ce faire, nous ferons la convention suivante que nous utiliserons encore dans la section 2 : si σ est dans $\mathfrak{S}_n$ $(n > 0)$, on considère σw comme le mot $\sigma(1)\ \sigma(2)\ \ldots\ \sigma(n)$ dont les lettres sont les éléments de $[n]$; lorsque σ est l'élément unique $\sigma_0 \in \mathfrak{S}_0$, σw est le <u>mot vide</u> $\sigma_0 w$. Si $f = y_1\ y_2\ \ldots\ y_m$ est un mot dont les lettres $y_1, y_2, \ldots, y_m$ sont des entiers tous distincts, on désigne par ω l'unique morphisme surjectif $\omega : \{y_1, y_2, \ldots, y_m\} \to [m]$ et l'on note ωf le mot $\omega y_1\ \omega y_2\ \ldots\ \omega y_m$. On pose encore $\omega\sigma_0 w = \sigma_0 w$.

Prenons alors un mot $\sigma w \in \mathfrak{S}_{n+1}$ $(n \geq 0)$; il s'écrit univoquement $\sigma w = f(n+1)f'$. Considérons l'application $\sigma w \to (\omega f\ ,\ \omega f')$; on a $\omega f \in \mathfrak{S}_m$ et $\omega f' \in \mathfrak{S}_{n-m}$ pour un certain m tel que $0 \leq m \leq n$ et puisque ω est un morphisme strictement croissant, on a encore :

$$|\Delta D\sigma| + \delta_{n,m} = |\Delta D\omega\, f| + |\Delta D\omega\, f'| + 1$$

où, comme d'usage, $\delta_{n,m} = 1$ ou 0 selon que $m=n$ ou $m \neq n$. D'autre part, l'image réciproque par l'application ci-dessus du couple $(\tau w\ ,\ \tau' w)$ où $\tau \in \mathfrak{S}_m$ et $\tau' \in \mathfrak{S}_{n-m}$, contient exactement $\begin{bmatrix} n \\ m \end{bmatrix}$ éléments. On en déduit

$$A_{n+1}(t) = A_n(t) + t \sum_{0 \leq m < n} \begin{bmatrix} n \\ m \end{bmatrix} A_m(t)\ A_{n-m}(t) \qquad (n \geq 0) \quad .$$

Il en résulte que la fonction génératrice $\underset{\approx}{A} = \bar{A}(t,1,u)$ est bien solution de l'équation différentielle précédente. Ce résultat a été établi pour la première fois par Riordan [23] .

2. Fonction génératrice exponentielle des polynômes ${}^{r}A_n(t)$.

Posons pour $r > 0$,

$${}^{r}A(t,u) = \sum_{r-1\leq n} (u^{n-r+1} / (n-r+1)!)\ {}^{r}A_n(t) \quad .$$

On a en particulier ${}^{1}A(t,u) = \bar{A}(t,1,u)$, dont on connaît déjà la formule explicite (cf. (5)) . Le but de la présente section est d'établir l'identité remarquable suivante, due à Riordan ([24] p. 235)

$${}^{r}A(t,u) = (r-1)!\ ({}^{1}A(t,u))^{r} \quad .$$

Construction d'une bijection de $\bigcup_{0\leq n} \mathfrak{S}_{n+r-1}$ sur le produit cartésien $\mathfrak{S}_{r-1} \times \mathfrak{S}^{((r))}$ de $\mathfrak{S}_{r-1}$ par le composé partitionnel marqué $\mathfrak{S}^{((r))}$ $(r \geq 1)$.

Tout d'abord, pour définir le composé partitionnel marqué $\mathfrak{S}^{((r))}$, il faut munir l'ensemble $\mathfrak{S}$ d'une application λ et nous prenons naturellement pour λ l'application définie par

$$\lambda\sigma = n \Leftrightarrow \sigma \in \mathfrak{S}_n \qquad (n \in \mathbb{N}) \quad .$$

Notons que l'élément unique $\sigma_0 \in \mathfrak{S}_0$ appartient à $\mathfrak{S}$, satisfait à $\lambda\sigma_0 = 0$ et est distinct de l'élément neutre e du monoïde $\mathfrak{S}^*$ pour lequel on a aussi $\lambda e = 0$.

Soit maintenant $\sigma \in \mathfrak{S}_{n+r-1}$ $(0 \leq n)$; le mot σw s'écrit univoquement $\sigma w = g_1\ i_1\ g_2\ i_2 \cdots g_{r-1}\ i_{r-1}\ g_r$ où $\{i_1, i_2, \ldots, i_{r-1}\} = [r-1]$ et où $g_1, g_2, \ldots, g_r$ sont des mots (éventuellement vides) dont les lettres sont des entiers. Soient $I_1, I_2, \ldots, I_r$ les sous-ensembles de $\underset{\approx}{N}$ dont les éléments sont respectivement les lettres des mots $g_1, g_2, \ldots, g_r$. Posant $\sigma_j w = \omega g_j$ pour chaque $j \in [r]$, (où ω est le morphisme défini à la fin de la section précédente et où l'on a $\sigma_j = \sigma_0$ si le mot g_j est vide), on voit immédiatement que le mot

$$h = (\sigma_1, I_1)\ (\sigma_2, I_2)\ \ldots\ (\sigma_r, I_r)$$

est un élément du composé partitionnel marqué $\mathfrak{S}^{((r))}$.

On note $\beta'\sigma$ ($=\beta h$ dans les notations du chapitre III) le mot $\sigma_1\ \sigma_2 \ldots \sigma_r \in \mathfrak{S}^*$ de longueur r . Soit ensuite $\bar{\sigma}$ la permutation définie par

$$\bar{\sigma} w = i_1\ i_2 \ldots i_{r-1} \qquad .$$

On a $\bar{\sigma} \in \mathfrak{S}_{r-1}$ et comme

$$\lambda h = \lambda\beta'\sigma = \lambda\sigma_1 + \ldots + \lambda\sigma_r = \lambda\sigma - (r-1) = n \qquad ,$$

on voit que l'application $\sigma \to (\bar{\sigma}, h)$ envoie $\mathfrak{S}_{n+r-1}$ dans $\mathfrak{S}_{r-1} \times \mathfrak{S}^{((r))} \cap \lambda^{-1} n$. Il est d'autre part immédiat de vérifier que cette application est bijective. Ceci achève la construction de la bijection cherchée.

Maintenant, puisque Card $\mathfrak{S}_{r-1}$ est égal à $(r-1)!$, on peut écrire

$$\Sigma\,\{\beta'\sigma/(\lambda\sigma-r+1)! : \sigma\in\mathfrak{S}_{n+r-1}\} =$$

$$= (r-1)!\ \Sigma\,\{\beta h/\lambda h! : h\in\mathfrak{S}^{((r))}\}$$

$$= (r-1)!\ (\,\Sigma\,\{\sigma/\lambda\sigma! : \sigma\in\mathfrak{S}\}\,)^r \quad ,$$

d'après le théorème 3.2 . Notant $\varpi_r\sigma$ l'image abélienne de $\beta'\sigma$, on obtient l'identité suivante valable dans l'algèbre large sur $\underset{\approx}{Q}$ du monoïde abélien $\mathfrak{S}^+$

$$\Sigma\,\{\varpi_r(\sigma)/(\lambda\sigma-r+1)! : \sigma\in\bigcup_{0\le n}\mathfrak{S}_{n+r-1}\} = (r-1)!\ (\,\Sigma\,\{\sigma/\lambda\sigma! : \sigma\in\mathfrak{S}\})^r \quad . \tag{10}$$

Soient enfin $\mu : \mathfrak{S}^+ \to \Omega$ un morphisme dans un monoïde abélien Ω et u une indéterminée. Comme déjà vu au chapitre III, on vérifie que l'application $\sigma \to u^{\lambda\sigma}\,\mu\sigma$ $(\sigma\in\mathfrak{S})$ peut être prolongée en un morphisme continu φ de la $\underset{\approx}{Q}$ - algèbre large de $\mathfrak{S}^+$ dans $\bar{\Omega}\,[[u]]$. Appliquant φ aux deux membres de l'identité (10), on trouve

$$\sum_{0\le n} (u^n/n!)\mu\varpi_r\,\{\mathfrak{S}_{n+r-1}\} = (r-1)!\ (\,\sum_{0\le n}(u^n/n!)\mu\{\mathfrak{S}_n\}\,)^r \quad . \tag{11}$$

THÉORÈME 4.5.

<u>On a pour</u> $r > 0$

$${}^{r}A(t,u) = (r-1)!\ ({}^{1}A(t,u))^r \quad .$$

PREUVE.

Pour $r = 1$, il n'y a rien à démontrer. Supposons $r > 1$ et prenons pour μ le morphisme prolongeant l'application $\sigma \to \theta\Delta D\sigma = t^{|\Delta D\sigma|}$ au monoïde $\mathfrak{S}^+$. Si l'on a $\sigma \in \mathfrak{S}_{n+r-1}$ et $\sigma w = g_1 \, i_1 \, g_2 \, i_2 \, \ldots \, g_{r-1} \, i_{r-1} \, g_r$ avec $\{i_1, i_2, \ldots, i_{r-1}\} = [r-1]$, il vient $|\Delta D g_j| = |\Delta D \omega g_j|$ $(j \in [r])$, puisque ω est un morphisme injectif. D'autre part, puisque le mot $g_1 \, g_2 \, \ldots \, g_r$ contient toutes les lettres du mot σw supérieures ou égales à r, on a $|\Delta'^{r-1}\Delta D\sigma| = \sum_j |\Delta D\omega \, g_j|$, d'où $\mu\varpi_r \, \sigma = \theta\Delta'^{r-1}\Delta D\sigma$. On obtient alors d'après la propriété 2.2 $\mu\varpi_r \{\mathfrak{S}_{n+r-1}\} = {}^rA_{n+r-1}(t)$. Comme on a d'autre part $\mu\{\mathfrak{S}_n\} = A_n(t)$, le théorème 4.5 résulte de l'identité (11).

Q.E.D.

3. <u>Autres interprétations des polynômes eulériens</u>.

Les techniques du chapitre précédent pourraient être appliquées à d'autres problèmes d'énumération. Par exemple, au lieu d'introduire la fonction multiplicative θ' du paragraphe 1, on peut poser pour chaque $\sigma \in \mathfrak{S}_n$ $(n > 0)$

$$\mu\sigma = \theta'\sigma \, . \, r^{z(\sigma)}$$

où r est une indéterminée et où $z(\sigma)$ est le nombre de cycles de σ.

La fonction μ est évidemment multiplicative et l'on peut appliquer la proposition 3.12. De plus, on a, comme dans la preuve du théorème 4.1 ,

$$\mu\{\mathfrak{C}_n\} = t'\, r \qquad \text{pour} \quad n = 1$$

$$= r\ \theta\Delta E\ \mathfrak{C}_n = r\ t\ A_{n-1}(t) \qquad \text{pour} \quad n > 1 \quad .$$

La proposition 3.12 conduit donc à l'identité, dans laquelle $\mu\{\mathfrak{S}_0\} = 1$,

$$\sum_{0\leq n} (u^n/n!)\ \mu\{\mathfrak{S}_n\} = \exp\left[r(ut' + \sum_{2\leq n} (u^n/n!)\ t\ A_{n-1}(t)\ )\right] \quad . \qquad (12)$$

Le membre de gauche de cette dernière identité est la fonction génératrice exponentielle des permutations classées <u>à la fois par nombre de cycles et par nombre d'excédances</u>. Maintenant les identités (3) et (12), ainsi que le théorème 4.1 permettent d'écrire, lorsque r est un entier positif

$$\sum_{0\leq n} (u^n/n!)\ \mu\{\mathfrak{S}_n\} = (\bar{A}(t,t',u))^r \quad . \qquad (13)$$

On obtient donc d'après (3) , la formule explicite de cette fonction génératrice exponentielle.

Si nous posons identiquement $t' = 1$, nous obtenons $\mu\{\mathfrak{S}_n\} = \Sigma\ \{t^{|\Delta E\sigma|}\ r^{z(\sigma)} : \sigma \in \mathfrak{S}_n\}$ que nous allons désigner par $Q_n(t,r)$ $(n > 0)$. On posera également $Q_0(t,r) = 1$. Il résulte de (13) que l'on a, lorsque r est un entier positif,

$$(r-1)!\ \sum_{0\leq n} (u^n/n!)\ Q_n(t,r) = (r-1)!\ ({}^1A(t,u))^r$$

$$= {}^rA(t,u) \qquad \text{d'après le théorème 4.5}$$

$$= \sum_{0\leq n} (u^n/n!)\ {}^rA_{n+r-1}(t) \quad .$$

On en déduit une nouvelle interprétation des polynômes ${}^rA_n(t)$, à savoir

$$ {}^rA_{n+r-1}(t) = (r-1)!\ Q_n(t,r) \qquad (r>0) \quad . \tag{14}$$

Enfin, désignons par $s(\sigma)$ le nombre des éléments saillants de la suite σw où $\sigma \in \mathfrak{S}_n$ $(n>0)$. Comme l'on a $|M\hat{\sigma}| + |\Delta E\sigma| = n$ et $s(\hat{\sigma}) = z(\sigma)$, on obtient

$$\sum \{t^{|M\sigma|}\ r^{s(\sigma)} : \sigma \in \mathfrak{S}_n\} = t^n \sum \{t^{-|\Delta E\sigma|}\ r^{z(\sigma)} : \sigma \in \mathfrak{S}_n\}$$

$$= t^n\ Q_n(t^{-1},r) \quad . \tag{15}$$

Le premier membre de l'identité (15) est le polynôme générateur des permutations $\sigma \in \mathfrak{S}_n$ classées à la fois suivant <u>leur nombre d'éléments saillants et leur nombre de montées</u>. Ce polynôme générateur a été considéré pour la première fois par Dillon et Roselle [10] qui ont à son sujet prouvé un certain nombre d'identités, qu'on pourrait retrouver à partir des formules (14) et (15) et des résultats de ce chapitre.

Enfin, notons que la propriété 2.6 fait apparaître que les coefficients des polynômes ${}^{r}A_{n}(t)$ sont tous divisibles par $r!$ (ce que ne fait pas apparaître le théorème 4.5). Si donc on pose

$$ {}^{r}A_{n}(t) = r!\ {}^{r}P_{n}(t) \quad , $$

il semble intéressant d'obtenir une interprétation pour les polynômes ${}^{r}P_{n}(t)$ $(0 \leq r \leq n)$.

D'abord, si $r = n$, on a ${}^{r}P_{n}(t) = 1$ et pour $r = 0$ et 1 , on a ${}^{r}P_{n}(t) = {}^{r}A_{n}(t)$. On fait donc l'hypothèse $1 < r < n$.

La restriction de tout $\sigma \in \mathfrak{S}_{n}$ à $[n-r]$ est une injection de $[n-r]$ dans $[n]$ que nous noterons $\varphi\sigma$. L'application φ est évidemment une surjection de $\mathfrak{S}_{n}$ sur l'ensemble $\mathcal{J}_{n-r,n}$ des injections de $[n-r]$ dans $[n]$ telle que l'image inverse de tout $\tau \in \mathcal{J}_{n-r,n}$ a $r!$ éléments. Introduisons l'application Λ de $\mathbb{N}^{p}$ $(p > 0)$ dans lui-même envoyant chaque vecteur $(x_1, x_2, \ldots, x_p)$ sur $((x_1-1)_+, (x_2-1)_+, \ldots, (x_p-1)_+)$. On a ainsi $\Delta = \Delta''\Lambda = \Lambda\Delta''$ et $\Delta^{r} = \Delta''^{r}\Lambda^{r}$ pour $r > 0$. Définissant le vecteur-excédance d'une injection de façon évidente, on a ainsi

$$ \begin{aligned} \Delta^{r}E\sigma &= (\ (\sigma(1)-r)_+, \ \ldots, \ (\sigma(n-r)-r)_+\ ) \\ &= \Lambda^{r}E\varphi\sigma \quad . \end{aligned} $$

D'où l'on déduit $\Delta^r E\,\mathfrak{S}_n = r!\ \Lambda^r E\,\mathcal{J}_{n-r,n}$ et par suite

$$ {}^rA_n(t)/r! = {}^rP_n(t) = \theta\Lambda^r E\,\mathcal{J}_{n-r,n} \quad . $$

Cette dernière interprétation des polynômes ${}^rA_n(t)/r!$ est due à Strosser [28] .

*

CHAPITRE V

LES SOMMES ALTERNÉES $A_n(-1)$ ET $B_n(-1)$.

1. Distribution du nombre des descentes sur $\mathfrak{S}'_n$.

Nous attachons à chaque $\sigma \in \mathfrak{S}'_n$ $(= \{\sigma \in \mathfrak{S}_n : \sigma(1) = n\})$ un mot $V(\sigma) = v_1 v_2 \ldots v_n$ dans les lettres de l'alphabet $X = \{m, \bar{m}, d, \bar{d}\}$ par les règles suivantes, où, par définition, $\sigma(n+1) = \sigma(1)$ $(=n)$.

(1) Pour chaque $j \in [n]$, on a $v_j \in \{d,\bar{d}\}$ ou $v_j \in \{m,\bar{m}\}$ selon que $\sigma(j) > \sigma(j+1)$ ou $\sigma(j) < \sigma(j+1)$;

(2) si $v_j \in \{d,\bar{d}\}$, $j \in [n-1]$, on a $v_j = d$ ou $\bar{d}$ selon que v_{j+1} est dans $\{d,\bar{d}\}$ ou dans $\{m,\bar{m}\}$;

(3) si $v_j \in \{m,\bar{m}\}$ $(1 < j \leq n)$, on a $v_j = m$ ou $\bar{m}$ selon que $v_{j-1} \in \{m,\bar{m}\}$ ou $v_{j-1} \in \{d,\bar{d}\}$.

En raison de $\sigma(1) = \sigma(n+1) = n$, on a toujours $v_1 \in \{d,\bar{d}\}$, $v_n \in \{m,\bar{m}\}$ et les seules occurences des lettres $\bar{d}$ et $\bar{m}$ se rencontrent dans les facteurs $v_j v_{j+1} = \bar{d}\bar{m}$ correspondant aux indices $j \in [n-1]$ tels que $\sigma(j) > \sigma(j+1) < \sigma(j+2)$. Par exemple, pour $\sigma(w) = (7,1,4,6,3,2,5)$, on aurait $V(\sigma) = \bar{d}\ \bar{m}\ m\ d\ \bar{d}\ \bar{m}\ m$.

Introduisons maintenant pour toute lettre x et tout mot g la dérivation $(g \frac{\partial}{\partial x})$ envoyant chaque mot $f = x_1 x_2 \ldots x_p$ sur l'ensemble pondéré formé de tous les mots obtenus en remplaçant dans f chaque occurrence de la lettre x par le mot g. Formellement $g\frac{\partial}{\partial x}$ est l'opérateur linéaire défini par sa restriction à X, à savoir $(g\frac{\partial}{\partial x})x' = g$ ou x' selon que $x' = x$ ou $x' \in X \setminus \{x\}$ et par l'identité $(g\frac{\partial}{\partial x})\,ff' = (g\frac{\partial}{\partial x})f \,.\, f' + f \,.\, (g\frac{\partial}{\partial x})f'$. Donc si $f = f_1 x\, f_2 x \ldots f_{r-1} x\, f_r$, où les f_i ne contiennent pas la lettre x, l'on aura :

$$(g\frac{\partial}{\partial x})\,f = f_1 g f_2 x \ldots f_{r-1} x f_r + f_1 x f_2 g \ldots f_{r-1} x f_r + \ldots + f_1 x f_2 x \ldots f_{r-1} g f_r \quad .$$

Par exemple, on a :

$$(\bar{d}\,\bar{m}\frac{\partial}{\partial m})\,(\bar{d}\,\bar{m}\,m\,d\,\bar{d}\,\bar{m}\,m) = \bar{d}\,\bar{m}\,\bar{d}\,\bar{m}\,m\,d\,\bar{d}\,\bar{m}\,m + \bar{d}\,\bar{m}\,m\,d\,\bar{d}\,\bar{m}\,m\,\bar{d}\,\bar{m}$$

LEMME 5.1.

<u>Soit</u>

$$\nabla = (d\,\bar{d}\frac{\partial}{\partial \bar{d}}) + (\bar{m}\,m\frac{\partial}{\partial \bar{m}}) + (\bar{d}\,\bar{m}\frac{\partial}{\partial d}) + (\bar{d}\,\bar{m}\frac{\partial}{\partial m}) \quad .$$

<u>On a identiquement</u> $V\mathfrak{S}'_n = \nabla\, V\mathfrak{S}'_{n-1} \qquad (n \geq 2)$.

PREUVE.

Il existe une bijection de $\mathfrak{S}'_{n-1} \times [n]$ sur $\mathfrak{S}'_n$ envoyant chaque $(\sigma', k) \in \mathfrak{S}'_{n-1} \times [n]$ sur la permutation $\sigma \in \mathfrak{S}'_n$ telle que σw soit obtenue en ajoutant 1 à tous les chiffres de $\sigma' w$ et en insérant 1 entre le k-ième et le $(k+1)$-ième terme de $\sigma' w$. Soit $V(\sigma') = v'_1\, v'_2 \ldots v'_{n-1}$

et supposons $v'_k \in \{d,\bar{d}\}$ c'est-à-dire $1 \leq k < n-1$ et $\sigma'(k) > \sigma'(k+1)$. On a $\sigma(k) = 1 + \sigma'(k) > \sigma(k+1) = 1 < \sigma(k+2) = 1 + \sigma'(k+1)$, donnant dans $V(\sigma)$ le facteur $v_k\, v_{k+1} = \bar{d}\,\bar{m}$. Maintenant :

i) si $v'_k = \bar{d}$, c'est-à-dire si $v'_{k+1} = \bar{m}$ et $\sigma'(k+1) < \sigma'(k+2)$, on a $v_{k+2} = m$ puisque $\sigma(k+2) = 1 + \sigma'(k+1) < \sigma(k+3) = 1 + \sigma'(k+2)$ et toute l'opération équivaut au remplacement de $v'_{k+1} = \bar{m}$ par $v_{k+1}\, v_{k+2} = \bar{m}\, m$. Remarquons qu'avec nos conventions, si $k = n-2$, on a $\sigma'(k+2) = \sigma'(n) = \sigma'(1) = n-1$;

ii) si $v'_k = d$, c'est-à-dire si $\sigma'(k+1) < \sigma'(k+2)$, on a encore $v_{k+2} \in \{d,\bar{d}\}$ et $V(\sigma)$ est déduit de $V(\sigma')$ en remplaçant $v'_k = d$ par $v_k\, v_{k+1} = \bar{d}\,\bar{m}$.

Un raisonnement analogue s'applique si $v'_k = m$ ou $\bar{m}$.

Q.E.D.

Notons maintenant α le morphisme canonique envoyant le monoïde libre engendré par $\{m,\bar{m},d,\bar{d}\}$ sur le monoïde commutatif libre de même base.

THÉORÈME 5.2.

<u>Il existe des entiers positifs</u> $c_{n,k}$ <u>tels que</u>

$$\alpha V\, \mathfrak{S}'_n = \sum_{0<2k\leq n} c_{n,k}\, (\bar{d}\,\bar{m})^k\, (d + m)^{n-2k} \qquad (n \geq 2) .$$

PREUVE.

Pour $n=2$, on a $V\mathfrak{S}'_n = \bar{d}\,\bar{m}$ et le résultat s'en déduit par induction sur n puisque ∇ commute avec α .

Q.E.D.

Remarque 5.3.

Les coefficcients $c_{n,k}$ des polynômes $\alpha V\mathfrak{S}'_n$ obéissent à des relations de récurrence qu'il est facile d'établir. Posons, par convention, $c_{n,k} = 0$ si $k < 1$ ou si $2k > n$; on a alors les relations :

$$c_{2,1} = 1 \quad \text{et, pour} \quad n \geq 3 \ , \ k \geq 1$$

$$c_{n,k} = k\, c_{n-1,k} + 2(n+1-2k)\, c_{n-1,k-1} \quad .$$

Remarque 5.4.

La fonction génératrice des nombres $c_{n,k}$ est donnée par Barton & David ([3] p. 180, voir aussi [17]). La théorie de ces auteurs se rattache aux considérations présentes en utilisant l'observation suivante dont la démonstration est laissée au lecteur.

Pour $\sigma \in \mathfrak{S}_{n-1}$, soit $\sigma' \in \mathfrak{S}'_n$ définie par $\sigma'(1) = n$, $\sigma'(1+j) = n+1-\sigma(n-j)$. Le nombre des facteurs d ou $\bar{d}$ de $V(\sigma')$ surpasse de 1 le nombre des $j \in [n-1]$ tels que $\sigma(j) > \sigma(j-1)$ et le nombre des facteurs $\bar{d}\,\bar{m}$ de $V(\sigma')$ est égal au nombre des $j \in [n-2]$ tels que $\sigma(j) > \sigma(j+1) < \sigma(j+2)$, augmenté d'une unité.

2. Applications aux polynômes eulériens.

Le théorème 5.2 va nous permettre de donner une interprétation combinatoire aux nombres $A_n(-1)$. Il est commode, tout d'abord, de noter la relation suivante sur les cardinaux des ensembles $\mathfrak{T}_n$ des permutations alternées (Cf. chap. I, § 9) .

PROPRIÉTÉ 5.5.

Pour $p > 0$, on a

$$\text{Card } \mathfrak{T}_{2p-1} = \text{Card } \mathfrak{T}_{2p} \cap \mathfrak{S}'_{2p} \quad .$$

PREUVE.

En effet, l'application qui envoie chaque $\sigma' \in \mathfrak{S}'_n$ telle que $V(\sigma') = (\bar{d}\ \bar{m})^p$ $(n = 2p > 1)$ sur l'élément $\sigma \in \mathfrak{S}_{n-1}$ défini par $\sigma(j) = \sigma'(j+1)$ $(j \in [n-1])$, est une bijection sur $\mathfrak{T}_{n-1}$. D'autre part, il est clair que l'on a :

$$\mathfrak{T}_{2p} \cap \mathfrak{S}'_{2p} = \{\sigma' \in \mathfrak{S}'_{2p} : V(\sigma') = (\bar{d}\ \bar{m})^p\} \quad .$$

Q.E.D.

THÉORÈME 5.6.

On a pour $n \geq 2$ l'identité

$$t\ A_{n-1} = \sum_{0<2k\leq n} c_{n,k}\ t^k (1+t)^{n-2k} \quad . \tag{1}$$

On a, de plus, pour $p \geq 1$

$$A_{2p}(-1) = 0 \qquad \text{et}$$

$$(-1)^{p-1} A_{2p-1}(-1) = \text{Card } \mathfrak{X}_{2p-1} \quad .$$

PREUVE.

Pour $\sigma' \in \mathfrak{S}'_n$, le nombre des occurences des lettres d ou $\bar{d}$ dans $V(\sigma')$ est égal à $|\Delta D\sigma'|$, puisque l'on a $v_n \in \{m,\bar{m}\}$. Or d'après les propriétés 2.2 et 2.3 , on a $\theta\Delta D\, \mathfrak{S}'_n = {}^0A_{n-1}(t) = t\, A_{n-1}(t)$. On en déduit que $t\, A_{n-1}(t)$ est obtenu en faisant $d = \bar{d} = t$ et $m = \bar{m} = 1$ dans $\alpha V\, \mathfrak{S}'_n$. La formule (1) résulte alors du théorème 5.2 .

D'autre part, le second membre de la formule (1) admet le facteur $(1 + t)$ si n est impair. On en conclut que $A_{2p}(-1)$ est nul pour $p \geq 1$. Au contraire, pour $n = 2p \geq 2$, on voit que $c_{2p,p} = (-1)^{p-1} A_{2p-1}(-1)$. Or

$$c_{2p,p} = \text{Card } \{\sigma' \in \mathfrak{S}'_{2p} : V(\sigma') = (\bar{m}\ \bar{d})^p\} = \text{Card } \mathfrak{X}_{2p} \cap \mathfrak{S}'_{2p} = \text{Card } \mathfrak{X}_{2p-1}$$

d'après la propriété 5.5 . Le théorème 5.6 en résulte.

Q.E.D.

3. Applications aux polynômes $B_n(t)$.

Nous donnons enfin des identités analogues à celles du théorème 4.6 , concernant les polynômes $B_n(t) = \theta E\, \mathcal{D}_n = \theta M\, \mathcal{G}_n$, où comme précédemment

$$\mathcal{D}_n = \{\sigma \in \mathfrak{S}_n : \sigma(j) \neq j\} \quad ,$$

$$\mathcal{G}_n = \{\sigma \in \mathfrak{S}_n : 1 \neq \sigma(1) \quad ; \quad 1 + \sigma(j) \neq \sigma(j+1)\} \quad .$$

Pour démontrer le théorème 5.9 ci-dessous, nous allons de nouveau appliquer la formule exponentielle et utiliser les propriétés élémentaires des permutations et de la transformation fondamentale du chapitre I.

Pour $n = 2p > 0$ et $\sigma \in \mathfrak{S}_n$, nous posons $\mu\sigma = 1$ si et seulement si σ est biexcédée, c'est-à-dire $\sigma \in \mathcal{B}$ (Cf. chap. I, § 9) et $\mu\sigma = 0$ dans les autres cas.

LEMME 5.7.

L'application μ est multiplicative. En d'autres termes, σ est une permutation biexcédée, si tous les termes de sa factorisation canonique sont aussi des permutations biexcédées.

PREUVE.

Ce lemme résulte encore de la propriété 3.11 . Soit $\sigma_1\, \sigma_2 \dots \sigma_r$ la décomposition en produit de cycles disjoints d'une permutation $\sigma \in \mathcal{B}$ et

$(f_1,I_1)\ (f_2,I_2)\ \ldots\ (f_r,I_r)$ sa factorisation canonique. Avec les mêmes notations que dans la propriété 3.11 , on a $f_j = \tau_j\, \sigma_j\, \tau_j^{-1}$ $(j \in [r])$.

Par suite

$$i < \sigma(i) \Leftrightarrow \tau_j(i) < f_j\, \tau_j(i) \quad \text{et}$$

$$i < \sigma^{-1}(i) \Leftrightarrow \sigma\sigma^{-1}(i) < \sigma^{-1}(i) \Leftrightarrow f_j\, \tau_j\, \sigma^{-1}(i) < \tau_j\, \sigma^{-1}(i)$$

$$\Leftrightarrow \tau_j(i) < f_j^{-1}\, \tau_j(i)$$

puisque les entiers i et $\sigma^{-1}(i)$ appartiennent à la même orbite. On a les mêmes équivalences en remplaçant le symbole $<$ par $>$.

Q.E.D.

LEMME 5.8.

On a pour $p > 0$

$$\mu\{\mathfrak{S}_{2p}\} = \text{Card}\ \mathfrak{B}_{2p} \qquad \text{et} \qquad \mu\{\mathfrak{S}_{2p-1}\} = \mu\{\mathfrak{C}_{2p-1}\} = 0 \tag{2}$$

et

$$\mu\{\mathfrak{C}_{2p}\} = \text{Card}\ \mathfrak{B}_{2p} \cap \mathfrak{C}_{2p} = \text{Card}\ \mathfrak{T}_{2p} \cap \mathfrak{S}'_{2p} \quad . \tag{3}$$

PREUVE.

Les relations (2) résultent de la définition de μ et de la proposition 1.14 .

D'après la propriété 1.10 et la proposition 1.14 , la transformation fondamentale $\sigma \to \hat{\sigma}$ est une bijection de $\mathfrak{C}_{2p} \cap \mathfrak{B}_{2p}$ sur $\mathfrak{S}'_{2p} \cap \mathfrak{T}_{2p}$.

La relation (3) est ainsi vérifiée.

Q.E.D.

Pour la démonstration du théorème ci-dessous, l'utilisation des nombres complexes est une simple commodité d'écriture évitant de recourir au produit de Hadamard.

THÉORÈME 5.9.

<u>Pour</u> $p > 0$, <u>on a</u>

$$B_{2p-1}(-1) = 0$$

$$(-1)^p B_{2p}(-1) = \text{Card}\, \mathfrak{T}_{2p} \quad .$$

PREUVE.

On applique la formule (9) du chapitre III avec l'application multiplicative μ du lemme 5.7 . Le premier membre de cette formule s'écrit d'après (2)

$$1 + \sum_{0<n} (u^n/n!) \text{ Card } \mathfrak{B}_n \quad .$$

- 88 -

Maintenant, on a pour $p > 0$

$$\begin{aligned}
\mu\{\mathfrak{C}_{2p}\} &= \mathrm{Card}\ \mathfrak{T}_{2p} \cap \mathfrak{S}'_{2p} && \text{(d'après (3))} \\
&= \mathrm{Card}\ \mathfrak{T}_{2p-1} && \text{(d'après la propriété 5.5)} \\
&= (-1)^{p-1} A_{2p-1}(-1) && \text{(d'après le théorème 5.6)} \quad .
\end{aligned}$$

Compte-tenu de la relation (2) le second membre de la formule (9) du chapitre III s'écrit donc

$$\exp \Big[\sum_{0<p} (u^{2p} / (2p)!)\ (-1)^{p-1} A_{2p-1}(-1) \Big] \quad .$$

En utilisant le fait que $A_n(-1) = 0$ si n est pair, on en déduit :

$$\sum_{0\leq n} (u^n/n!)\ \mathrm{Card}\ \mathfrak{B}_n = \exp \Big[\sum_{2\leq n} ((iu)^n/n!)\ (-1)\ A_{n-1}(-1) \Big]$$

où i est le nombre complexe de module 1 et d'argument $\pi/2$. Mais le nombre de droite de cette dernière équation est la valeur pour $t = -1$ de l'expression $\exp \big[\sum_{2\leq n} ((iu^n/n!)\ t\ A_{n-1}(t) \big]$, qui d'après le théorème 4.1 est égale à $\sum_{0\leq n} ((iu)^n/n!)\ B_n(t)$. En identifiant terme à terme, il en résulte que l'on a $B_n(-1) = 0$ si n est impair et que pour $n = 2p$, on a $(-1)^p B_{2p}(-1) = \mathrm{Card}\ \mathfrak{B}_{2p} = \mathrm{Card}\ \mathfrak{T}_{2p}$ d'après la proposition 1.14 .

Q.E.D.

4. Les développements de $tg\ u$ et de $1/\cos u$.

De l'identité (5) du théorème 4.2 , on tire :

$$\sum_{0\leq n} ((iu)^n/n!)\ A_n(-1) = 2 / (1 + e^{-2iu}) \quad ,$$

qu'on peut réécrire :

$$\sum_{0<n} (u(iu)^{n-1}/n!)\ A_n(-1) = (1 - e^{-2iu}) / (i(1 + e^{-2iu})) = tg\ u$$

soit, en utilisant le théorème 5.6 ,

$$tg\ u = \sum_{0<p} (u^{2p-1} / (2p-1)!)\ \mathrm{Card}\ \mathfrak{T}_{2p-1} \quad . \tag{4}$$

De même, d'après l'identité (6) du théorème 4.2 , on a

$$\sum_{0\leq n} ((iu)^n/n!)\ B_n(-1) = 2 / (e^{-iu} + e^{iu}) = 1 / \cos u \quad .$$

D'après le théorème 5.9 , on déduit donc :

$$1 / \cos u = 1 + \sum_{0<p} (u^{2p} / (2p)!)\ \mathrm{Card}\ \mathfrak{T}_{2p} \quad . \tag{5}$$

Les identités (4) et (5) sont dues à Désiré André [1] . Nous avons pu les établir ici sans recourir aux méthodes traditionnelles du calcul différentiel et intégral, en n'utilisant que l'identité de Cauchy et des constructions sur la catégorie des ensembles totalement ordonnés finis.

Nous laissons au lecteur l'amusement de vérifier par les mêmes techniques la formule élémentaire

$$1 / \cos u = \exp [\int \operatorname{tg} u \, du]$$

en utilisant une définition appropriée de l'intégrale.

5. Table des nombres d'Euler.

On a souvent appelé nombres d'Euler les coefficients du développement de $\operatorname{tg} u$ et de $1 / \cos u$. Les valeurs numériques de ces premiers coefficients ont déjà été obtenues par Euler lui-même (voir [16] pp. 299-301). Nous reproduisons ci-dessous ces premières valeurs. Rappelons que pour $n > 0$, $n > 0$, on note $\mathfrak{T}_n$ le sous-ensemble de $\mathfrak{S}_n$ formé par les permutations alternées, qu'on a ensuite les identités

$$(-1)^{p-1} A_{2p-1}(-1) = \operatorname{Card} \mathfrak{T}_{2p-1}$$

$$(-1)^p B_{2p}(-1) = \operatorname{Card} \mathfrak{T}_{2p} \qquad (p > 0) \quad .$$

Le tableau des quantités

$$t_n = \operatorname{Card} \mathfrak{T}_n \quad \text{pour} \quad n = 1, 2, \ldots, 14$$

est alors le suivant :

n	t_n
1	1
2	1
3	2
4	5
5	16
6	61
7	272
8	1385
9	7936
10	50521
11	353792
12	2702765
13	22368256
14	199360981

*

RÉFÉRENCES.

1. D. ANDRÉ Développements de sec x et de tang x , C.R. Acad. Sc. Paris, **88** (1879), p. 965-967.
2. D. ANDRÉ Mémoire sur le nombre des permutations alternées, Journ. de Math. **7** (1881), p. 167.
3. D.E. BARTON & F.N. DAVID - Combinatoriel Chance, Griffin, London, (1962).
4. L. CARLITZ Eulerian numbers and polynomials, Math. Magazine **32** (1959), p. 247-260.
5. L. CARLITZ Eulerian numbers and polynomials of higher order, Duke Math. J. **27** (1960), p. 401-423.
6. L. CARLITZ A note on Eulerian numbers, Arch. Math. **14** (1963), p. 383-390.
7. L. CARLITZ & J. RIORDAN - Congruences for Eulerian Numbers, Duke Math. J. **20** (1953), p. 339-343.
8. L. CARLITZ, D.P. ROSELLE & R.A. SCOVILLE - Permutations and sequences with repetitions by number of increases, J. of Combinatorial Theory **1** (1966), p. 350-374.
9. P. CARTIER & D. FOATA - Problèmes combinatoires de commutation et réarrangements. Lecture Notes in Math., n° 85, Springer-Verlag, Berlin (1969).
10. J.F. DILLON & D.P. ROSELLE - Eulerian numbers of higher order, Duke Math. J. **35** (1968), p. 247-256.

11. R.C. ENTRINGER - A combinatorial interpretation of the Euler and Bernoulli numbers, Nieuw Arch. V. Wiskunde, **14** (1966), p. 241-246.

12. D. FOATA Etude algébrique de certains problèmes d'analyse combinatoire et du calcul des probabilités. Publ. Inst. Statist. Univ. Paris, **14** (1965), p. 81-241.

13. G. FROBENIUS Über die Bernoullischen Zahlen und die Eulerschen Polynome, Sitz. Berichte Preuss. Akad. Wiss., (1910), p. 808-847.

14. R. FRUCHT A combinatorial approach to the Bell polynomials and their generalisations. Recent Progress in Combinatorics (W.T. Tutte, Ed.), Academic Press, London & New York (1964), p. 69-74.

15. R. FRUCHT y G.-C. ROTA - Polinomios de Bell y partitiones de conjuntos finitos, Scientia, **126** (1965), p. 5-10.

16. Ch. JORDAN Calculus of finite differences, Röttig & Romwalter, Budapest (1939).

17. W.O. KERMACK & A.G. Mc KENDRICK - Some distributions associated with a randomly arranged set of numbers, Proc. Roy. Soc. Edinburgh Sect. A., **57** (1937), p. 332-376.

18. B. KITTEL Communication privée.

19. P.A. MAC MAHON - Second memoir on the composition of numbers, Phil. Trans. Royal Soc. London, A **207** (1908), p. 65-134.

20. P.A. MAC MAHON - Combinatory Analysis, Cambridge Univ. Press (1915-1916).

21. E. NETTO Lehrbuch der Combinatorik, B.G. Teubner, Leipzig (1900).

22. F. POUSSIN Sur une propriété arithmétique de certains polynômes associés aux nombres d'Euler, C.R. Acad. Sc. Paris, **266** (1968), p. 392-393.

23. J. RIORDAN Triangular permutations numbers, Proc. Amer. Math. Soc. **2** (1951), p. 404-407.

24. J. RIORDAN An Introduction to Combinatorial Analysis, Wiley, New York, (1959).

25. D.P. ROSELLE Permutations by number of rises and successions, Proc. Amer. Math. Soc. **19** (1968), p. 8-16.

26. G.-C. ROTA On the foundations of Combinatorial Theory, J. Wahrscheinlichkeitstheorie, **2** (1966), p. 340-368.

27. E.B. SHANKS Iterated sums of powers of binomial coefficients, Amer. Math. Monthly, **58** (1951), p. 404-407.

28. R. STROSSER Séminaire de Théorie Combinatoire, I.R.M.A., Université de Strasbourg, 1969-70.

29. G.E. UHLENBECK & G.W. FORD - The theory of graphs with applications to the virial development of the properties of gases. Studies in Statistical Mechanics, vol. I (J. de Boer & G.E. Uhlenbeck, Ed.), North-Holland, Amsterdam (1962), p. 119-211.

30. Ph. WELSCHINGER - Séminaire de Théorie Combinatoire, I.R.M.A., Université de Strasbourg, 1969-70.

31. J. WORPITZKY Studien über die Bernoullischen und Eulerschen Zahlen, J. für die reine und angewandte Math. **94** (1883), p. 203-232.

MARCEL P. SCHÜTZENBERGER

Une application de la théorie de la décomposition des monoïdes

Séminaire Schützenberger, tome 1 (1969-1970), exp. n° 5, p. 1-4.
<http://www.numdam.org/item?id=SMS_1969-1970__1__A5_0>

Séminaire SCHÜTZENBERGER-LENTIN-NIVAT 5-01
(Problèmes mathématiques
de la Théorie des automates) 9 décembre 1969
Année 1969/70, n° 5, 4 p.

UNE APPLICATION DE LA THÉORIE DE LA DÉCOMPOSITION DES MONOÏDES

par Marcel P. SCHÜTZENBERGER

1. Introduction.

Etant donnés un alphabet (fini ou non) X, et une famille $\mathfrak{M}$ de monoïdes, désignons par $\mathfrak{M}(X)$ la famille de toutes les parties de X^* de la forme $Y^* \cap M' \mu^{-1}$, où Y est une partie finie de X, μ un morphisme de X^* dans un monoïde $M \in \mathfrak{M}$, et M' une partie de M.

On se propose de caractériser $\mathfrak{M}(X)$ par une traduction à peu près directe de la théorie de la décomposition des monoïdes, dans le cas où $\mathfrak{M}$ est la famille de tous les monoïdes finis tels que leurs groupes appartiennent à une famille donnée $\mathcal{G} \neq \emptyset$ de groupes finis, fermée par produit en couronne, et contenant les diviseurs de ses membres.

Pour ce faire, on supposera X infini, et, étant donnés une famille $\mathfrak{A}$ de parties de X^*, et deux parties finies Y, $Z \subset X$, on désignera par $\Lambda(\mathfrak{A})$ la famille des substitutions $\lambda : Z^* \to Y^*$ satisfaisant les deux conditions suivantes :

- $Z\lambda$ est non ambigu ;
- Il existe une partition $Y = Y_1 \cup Y_2$ telle que, pour chaque $z \in Z$, $z\lambda$ soit une union de parties de la forme $A_y\, y$, où $y \in Y_1$, $A_y \subset Y_2^*$, $A_y \in \mathfrak{A}$.

Ceci posé, on vérifiera la propriété :

Propriété. - $\mathfrak{M}(X)$ est la plus petite famille $\mathfrak{A}$ de parties de X^* telle que :

(i) Toute partie finie non ambigue de X^* appartient à $\mathfrak{A}$;

(ii) $\mathcal{G}(X) \subset \mathfrak{A}$;

(iii) $\mathfrak{A}$ est fermée par les opérations polynomiales non ambigues (i. e., si A, $B \in \mathfrak{A}$,

$$A \cup B \text{ non ambigu} \implies A \cup B \in \mathfrak{A} ,$$

$$AB \text{ non ambigu} \implies AB \in \mathfrak{A}) ;$$

(iv) Si Y et Z sont deux parties finies de X, $\lambda \in \Lambda(\mathfrak{A})$, $B \in \mathfrak{A}$, $B \subset Z^*$, alors $B\lambda \in \mathfrak{A}$.

2. Vérification de $\mathfrak{A} \subset \mathfrak{M}(X)$.

Soit $A \in \mathfrak{A}$. Il résulte immédiatement de la définition de $\mathcal{S}(X)$ et de $\mathfrak{A}$ qu'il existe une partie finie $V \subset X$ telle que $A \subset V^*$. De plus, A est non ambigu par définition, si A est fini, s'il appartient à $\mathcal{S}(X)$, ou s'il est union ou produit de deux autres parties de $\mathfrak{A}$. Si A est obtenu par substitution $A = B\lambda$ ($\lambda : Z^* \to Y^*$, $\lambda \in \Lambda(\mathfrak{A})$, $B \subset Z^*$, non ambigu), A est encore non ambigu, puisque $Z\lambda \subset Y_2^* Y_1$, $Y_1 \cap Y_2 = \emptyset$, implique que tout mot de $Z^* \lambda$ ait au plus un facteur gauche dans $Z\lambda$.

Considérant le monoïde syntactique M de A dans V^* , il suffit donc de vérifier que tout groupe dans M appartient à $\mathcal{S}$. Ceci est clair, si A est fini ou appartient à $\mathcal{S}(X)$. Le même résultat découle d'un énoncé connu, si $A = B \cup C$ ou $= BC$, où B , $C \in \mathfrak{M}(X)$. Il suffit donc de considérer le cas où $A = B\lambda$, avec $\lambda : Z^* \to Y^*$, $\lambda \in \Lambda(\mathfrak{M}(Y))$, $B \in \mathfrak{M}(Z)$.

Adjoignant au besoin une lettre supplémentaire à Z , il est loisible de supposer que $Z\lambda = Y_2^* Y_1$.

Considérons les parties $A_y \subset Y_2^*$ $(y \in Y_1)$ intervenant dans la définition de λ . Chacune d'elles a son monoïde syntactique dans $\mathfrak{M}$. Comme Y_1 est fini, il existe donc un monoïde $P \in \mathfrak{M}$, et un morphisme $\pi : Y^* \to P$ tel que chaque A_y soit de la forme $P_y \pi^{-1}$ $(P_y \subset P)$.

Soit, d'autre part, $\chi : Z^* \to Q$ le monoïde syntactique de $B \subset Z^*$.

Nous définissons un morphisme ρ de Y^* dans le monoïde des applications $Q \times P \to Q \times P$ en posant, pour chaque $(q , p) \in Q \times P$, $y \in Y$,

$$\begin{aligned} (q , p)y &= (q , p{\cdot}y\pi) , \quad \text{si} \quad y \in Y_2 , \\ &= (q{\cdot}z\chi , 1) , \quad \text{si} \quad y \in Y_1 , \end{aligned}$$

$$p \in P_y , \qquad A_y\, y \subset z\lambda .$$

Par construction, $B\lambda$ est image inverse d'une partie du monoïde $R = Y^* \rho$, et R est un sous-monoïde du produit en couronne $Q \circ \bar{P}$, où $\bar{P}$ est obtenu en "ajoutant les constantes au t. m. (P , P) " (c'est-à-dire, où $\bar{P}$ est isomorphe au quotient du produit libre $P * u = P'$ par les relations $p'u \equiv u$ $(p' \in P')$). Il est connu que Q , $P \in \mathfrak{M}$ implique $\bar{P} \in \mathfrak{M}$ et $Q \circ P \in \mathfrak{M}$, et le résultat est donc établi.

3. Vérification de $\mathfrak{M}(X) \subset \mathfrak{A}$.

Considérons $A \in \mathfrak{M}(X)$. Par hypothèse, il existe une partie finie $Y \subset X$, un monoïde $M \in \mathfrak{M}$, et un morphisme $\mu : Y^* \to M$, tels que A soit une union de parties de la forme $m\mu^{-1}$ $(m \in M)$. Sans perte de généralité, on supposera désormais que $A = m\mu^{-1}$, et on prouvera $A \in \mathfrak{A}$ par induction sur Card M , en éliminant d'abord trois cas particuliers.

(1) M est un groupe. – La conclusion $A \in \mathfrak{A}$ résulte immédiatement de $\mathcal{G}(X) \subset \mathfrak{A}$.

Comme $\{1\}$ est un groupe dans $\mathcal{G}$, ceci couvre le premier cas de l'induction. Il en résulte de plus que $Z^* \in \mathfrak{A}$, pour toute partie finie Z de X , puisque $Z^* = 1\nu^{-1}$, où ν est le morphisme de Z^* sur $\{1\}$.

(2) M est cyclique. – On a $\mu = \psi\mu'$, où $\psi : Y^* \to \underset{\sim}{N}$ et $\mu' : \underset{\sim}{N} \to M$ sont deux morphismes. Donc, il existe $p , q \in \underset{\sim}{N}$ tels que

$$A = (p + q\underset{\sim}{N})\psi^{-1} \quad (= A(p , q)) \quad ,$$

où, en outre, $\underset{\sim}{Z}_q \in \mathcal{G}$, en convenant que $\underset{\sim}{Z}_0 = \{0\}$.

Si $0 = p = q$, ou si $q > 0$, $0 \leqslant p \leqslant q - 1$, on a

$$A(p , q) = p(\psi\pi)^{-1} \quad ,$$

où π est le morphisme naturel de $\underset{\sim}{N}$ sur $\underset{\sim}{Z}_q$, et le résultat découle du (1) ci-dessus.

Dans les autres cas, on a la formule

$$A(p , q) = \sum_{1 \leqslant j} (Y_n \circ \psi^{-1})^* (Y_n\, j\psi^{-1})\, A(p'_j , q) \quad ,$$

où p'_j est le plus petit élément de $\underset{\sim}{N} \cap (p - j + q\underset{\sim}{N})$. Comme le membre de droite est manifestement non ambigu, le résultat découle de l'hypothèse d'induction du cas traîté ci-dessus, puisque $p'_j \geqslant q \implies p'_j < p$.

(3) $M = 1 \cup L$, où L est une $\mathcal{L}$-classe. – On sait qu'il existe un groupe G , et un morphisme $\gamma : M \to G$ tel que, pour chaque $h \in L$, la restriction à hM de γ soit un isomorphisme.

Ceci implique

$$h\mu^{-1} = (Y_n\, 1\mu^{-1})^* \sum\{(Y_n\, h'\mu^{-1}).g(\mu\gamma)^{-1} : h' \in hM , \quad g = (h'\gamma)^{-1}\, (h\gamma) \in G\} \quad ,$$

établissant $h\mu^{-1} \in \mathfrak{A}$ pour $h \in L$. Comme $L^2 = L$, on a $1\mu^{-1} = (Y_n\, 1\mu^{-1})^*$, un cas déjà traîté.

5-04

Nous considérons maintenant le cas où M ne rentre dans aucune des catégories envisagées ci-dessus. D'après le lemme de Krohn et Rhodes, il existe :

- Un idéal à gauche non vide $L \not\subseteq M \setminus 1$,
- Un sous-monoïde $T \neq 1$, M ,

tels que $M \subset L \cup T$.

Posons

$$Y_1 = Y \cap L\mu^{-1} \ ; \qquad Y_2 = (Y \setminus Y_1) \cap T\mu^{-1} \ .$$

L'hypothèse d'induction permet de supposer que Y_1 et Y_2 sont non vides, et comme M est fini, nous pouvons prendre un alphabet fini Z et une bijection $\nu : Z \to L' = L \cap (Y_2^* Y_1)\mu$. Définissons une substitution $\lambda : Z^* \to Y^*$, en posant, pour chaque $z \in Z$,

$$z\lambda = \{g \in Y_2^* Y_1 : g\mu = z\nu\} \ .$$

Par construction, $z\lambda$ est union disjointe de parties de la forme $A_y\, y$, où $y \in Y_1$, $A_y \in Y_2^*$, $Y_2 \cap Y_1 = \emptyset$, et où $A_y \in \mathfrak{A}$, puisque A_y est image inverse, donc partie du sous-monoïde $T \not\subseteq M$. Donc $\lambda \in \Lambda(\mathfrak{A})$.

Maintenant, l'identité $Y^* = (Y_2^* Y_1)^* Y_2^*$ montre que chaque $a \in A$ a exactement une factorisation $a = a_1 a_2$, où $a_1 \in (Y_2^* Y_1)^*$, $a_2 \in Y_2^*$. Par conséquent, A est une union finie disjointe de produits non ambigus de la forme $A_1 A_2$, où

$$A_1 = (Y_2^* Y_1)^* \cap m\mu^{-1} \qquad (m \in 1 \cup L) \ ,$$

$$A_2 = Y_2^* \cap m\mu^{-1} \qquad (m \in T) \ .$$

La deuxième relation implique $A_2 \in \mathfrak{A}$, d'après $T \not\subseteq M$ et l'hypothèse d'induction, et, puisque $\mathfrak{A}$ est fermé, par les opérations polynomiales non ambigues, il suffit de montrer $A_1 \in \mathfrak{A}$. Or la première relation équivaut à $A_1 = B\lambda$, où $B = m\nu^{-1}$, et où $\nu : Z^* \to 1 \cup L$ est le morphisme étendant la bijection $\nu : Z \to L'$. Comme $L \not\subseteq M \setminus 1$, on a $\mathrm{Card}(1 \cup L) < \mathrm{Card}(M)$, d'où $B \in \mathfrak{A}$, par induction, et le résultat est établi.

(Texte reçu le 19 janvier 1971)

Marcel P. SCHÜTZENBERGER
9 villa Poirier
75 - PARIS 15

ON THE ROOK POLYNOMIALS OF FERRERS RELATIONS

by

D. Foata & M.P. Schützenberger

Reprinted from

Colloquia Mathematica Societatis János Bolyai, 4
Combinatorial Theory and its Applications, Vol. 2
P. Erdös, A. Rényi & V. T. Sós, Editors, North-Holland Publishing Co., Amsterdam, London, 1970.

On the rook polynomials of Ferrers relations

by

D. Foata and M. P. Schützenberger
Strasbourg, France Paris, France

1. INTRODUCTION

A quasi-permutation on a set I is a relation on I that is contained in a bijective relation. Formally $Q \subset I \times I$ is a quasi-permutation if and only if

$$(i,j),\ (i',j') \in Q \implies (i,j) = (i',j') \quad \text{or} \quad i \neq i' \text{ and } j \neq j'.$$

For Q finite the weight $\lambda(Q)$ of Q is the number of elements $(i,j) \in Q$ and for any relation $R \subset I \times I$ we shall denote $Q_k(R)$ the set of the quasi-permutations of weight k contained in R. Then assuming R finite, the rook polynomial $\rho(R)$ of R is the generating function

$$1 + \sum_{0<k} t^k \operatorname{card} Q_k(R).$$

We refer the reader to the last two chapters of Riordan's book on combinatorial analysis [4] for the general theory of rook polynomials and their applications. Riordan gives several theorems stating conditions for two relations to have the same rook polynomial or, as we shall say, to be rook-equivalent. Our theorem 3 belongs to this type. It generalizes the obvious fact that any relation is rook-equivalent with its transpose.

The case when R is a Ferrers relation plays a role in the applications since it relates to the Laguerre polynomials, the Eulerian polynomials, the Shanks polynomials, the Poussin polynomials and more generally to the so-called "Newcomb's problem for arbitrary specification". Our main result (theorem 11) states that a cross-section ("minimal set of representatives") of the Ferrers relations with respect to the rook-equivalence is provided by those relations which are strictly decreasing, i.e. which correspond to partitions into unequal parts.

In the last section we effectively compute the rook-equivalent decreasing Ferrers relations for those relations which are total preorders deprived for an arbitrary subset of its equivalence classes (theorem 19). Since this family is closed under complementation, one might apply Riordan's theory of "complementary boards" to deduce non-trivial identities by using the rather explicit expressions for the rook polynomial that are given in our property 5.

2. A GENERAL PROPERTY OF ROOK EQUIVALENCE

We use the standard notation $[n]$ for the ordered set $\{1,2,\dots,n\}$ $([0] = \phi)$. Thus we can say in short that α is a (m,n)-injection if and only if it is a map sending each pair $(i,j) \in [m]\times[m]$ onto the pair $(\alpha_1(i), \alpha_2(j)) \in [n] \times [n]$ where α_1 and α_2 are two injections of $[m]$ into $[n]$. We write then $\alpha = \alpha_1 \times \alpha_2$, $A_i = \alpha_i([m])$ $(i = 1,2)$.

Definition 1.

The (m,n)-injection $\alpha : [m]\times[m] \to [n]\times[n]$ is compatible with the relation $R \subset [n] \times [n]$ if and only if there exist subsets $\bar{A}_i \subset [n]\setminus A_i$ $(i=1,2)$ such that $R \cap (A_1 \times [n]) = R \cap (A_1 \times A_2) \cup (A_1 \times \bar{A}_2)$ and symmetrically $R \cap ([n] \times A_2) = R \cap (A_1 \times A_2) \cup (\bar{A}_1 \times A_2)$.

- 414 -

Definition 2.

With the same notations, the α-transpose of R is the relation

$$R' = R \setminus (A_1 \times A_2) \cup \alpha(\tilde{S}) \subset [n] \times [n]$$

where $\tilde{S}$ is the ordinary transpose
$\tilde{S} = \{(i,j) \in [m] \times [m] : (j,i) \in S\} \subset [m] \times [m]$ of the inverse image

$$S = \alpha^{-1}(R \cap (A_1 \times A_2)) \subset [m] \times [m] .$$

THEOREM 3.

If the injection α is compatible with the relation R, then R and its α-transpose R' are rook-equivalent.

Proof.

Note that reciprocally R is the α-transpose of R'. Accordingly, if suffices to construct an injective map of $\underset{\sim}{Q}(R)$ into $\underset{\sim}{Q}(R')$.

Consider any given quasi-permutation Q contained in R. The inverse image $P = \alpha^{-1}(Q \cap (A_1 \times A_2))$ is a quasi-permutation contained in $S = \alpha^{-1}(R \cap (A_1 \times A_2)) \subset [m] \times [m]$.

Let B_1 and B_2 be the least subsets of $[m]$ that satisfy $P \subset B_1 \times B_2$. Since P is a quasi-permutation, we have $\lambda(P) = \underset{\sim}{\mathrm{Card}}\, B_1 = \underset{\sim}{\mathrm{Card}}\, B_2$.

We define a (m,m)-injection $\sigma = \sigma_1 \times \sigma_2$ by the following two conditions where $i, i' = 1,2$ and $i' \neq i$.

(1) The restriction of σ_i to $[m] \setminus B_i$ is the unique order-preserving bijection of this set onto $[m] \setminus B_{i'}$;

(2) For each $(k,k') \in P$ we set $\sigma_1(k) = k'$ and $\sigma_2(k') = k$.

Clearly $\sigma(P)$ is a quasi-permutation contained in the transpose $\tilde{S}$ of S .

We now extend σ to a (n,n)-injection $\tau = \tau_1 \times \tau_2$ by letting

$$\tau_i(j) = j \quad \text{for any} \quad j \in [n] \setminus A_i \,;$$

$$\tau_i(j) = \alpha_i(\sigma_i(\alpha_i^{-1}(j))) \quad \text{for any} \quad j \in A_i \,.$$

Because of the compatibility condition $R \setminus (A_1 \times A_2)$ is invariant under τ. Thus $Q' = \tau(Q)$ is a quasi-permutation contained in R'. Finally $Q \longrightarrow Q'$ is injective because in view of our canonical choice of σ, this injection, hence Q itself, are determined by Q' without ambiguity.

3. APPLICATION TO FERRERS RELATIONS

Let $\mathbb{P} = \mathbb{N} \setminus \{0\}$ denote the set of all positive integers. A Ferrers relation is a relation $F(\varphi) \subset \mathbb{P} \times \mathbb{P}$ that is defined by a non-increasing map $\varphi : \mathbb{P} \longrightarrow \mathbb{N}$ such that the sum $\sum_i \varphi(i)$ is finite and the condition

$$F(\varphi) = \{(i,j) \in \mathbb{P} \times \mathbb{P} : \; j \leq \varphi(i)\} \,.$$

Its transpose $\tilde{F}(\varphi)$ is another Ferrers relation $F(\tilde{\varphi})$ where $\tilde{\varphi} : \mathbb{P} \longrightarrow \mathbb{N}$ is defined by

$$\tilde{\varphi}(i) = \mathrm{Max}\{ j \in \mathbb{P} : \; \varphi(j) \geq i \} \qquad \text{if} \quad i \leq \varphi(i)$$

$$= \quad 0 \qquad \text{otherwise}$$

We shall say shortly that φ is <u>decreasing</u> if and only if it is so on the non-trivial part of its domain, i.e. if and only if $\varphi(i) > 0$ implies $\varphi(i) > \varphi(i+1)$. Then $F(\varphi)$ will be called a <u>decreasing</u> Ferrers relation.

In the next two lemmas we consider a fixed decreasing map φ and for convenience we set

$$p = \tilde{\varphi}(1) \quad (= \mathrm{Max}\{ i \in \mathbb{P} : \; \varphi(i) > 0 \})$$

$$x_i = \varphi(i) \qquad (i \in [p]) ;$$

$$r_0 = 1 ;$$

$$r_k = \sum_{(k)} \prod_{1 \leq j \leq k} (x_{i_j} - k + j) \qquad (k \in [p])$$

$$r_k = 0 \quad \text{for} \quad p > 0$$

- 416 -

where $\sum_{(k)}$ indicates a summation over the set $[p]^{(k)}$ of the $\binom{p}{k}$ strictly increasing sequences $(i_1 < i_2 < \cdots < i_k)$ of length k with elements in $[p]$.

Lemma 4. The rook polynomial of $F(\varphi)$ is

$$\rho(F(\varphi)) = \sum_{0 \le k \le p} t^k r_k .$$

Proof. For any relation $R \subset [p] \times \mathbb{P}$ the set $Q_k(R)$ of the quasi-permutations of weight k contained in R is empty for $k > p$; and for $k \le p$ it is the disjoint union over all sequences (of length k) $I \subset [p]^{(k)}$ of the sets $Q_k((R \cap I) \times \mathbb{P})$.

Further, when $R = F(\varphi)$, each restriction $F(\varphi) \cap (I \times \mathbb{P})$ is rook-equivalent with a decreasing Ferrers relation $F(\varphi')$ where $\varphi'(j) = \varphi(i_j)$ $(I = (i_1, i_2, \ldots, i_k))$ and $\tilde{\varphi}'(1) = k$.

Thus it suffices to consider the special case of $k = p$. Then $r_p = (x_1 - p+1)(x_2 - p+2) \cdots (x_{p-1} - 1) x_p$ and the equality of this quantity with $\operatorname{Card} Q_p(F(\varphi))$ is trivial since this last number is the number of injections $\eta : [p] \longrightarrow \mathbb{P}$ such that $\eta(i) \le \varphi(i)$ identically.

Q.E.D.

PROPERTY 5. For each $k \in [p]$ one has the identity

$$r_k = \sum_{0 \le j \le k} a_j \, S(p-j, p-k) \qquad (0 \le k \le p)$$

where the $S(i,j)$'s are Stirling numbers of the second kind* and where a_j $(0 \le j \le p)$ are the symmetric functions $a_0 = 1$ and $a_j = \sum_{(j)} y_{i_1} y_{i_2} \cdots y_{i_j}$ $(j \in [p])$ in the variables $y_i = x_i - p + i$ $(i \in [p])$.

Proof. Let $p' = p-1$; $x'_i = x+1$ $(i \in [p'])$ and denote by

*We make the usual convention that $S(0,0) = 1$ and $S(i,j) = 0$ if exactly one of i and j is zero.

primed letters the quantities defined with respect to p' and the x'_i's in the same manner as the corresponding quantities $(r, y$ or $a)$ were defined with respect to p and the x_i's.

Thus, by definition

$$r_k = (x_1 - k+1)\, r'_{k-1} + r'_k = (y_1 + p - k)\, r'_{k-1} + r'_k$$

and

$$y_{i+1} = x_{i+1} - p + i + 1 = x'_i - p' + i = y'_i \qquad (i \in [p']).$$

Using induction on p, the first relation gives

$$r_k = (y_1 + p-k) \sum_j a'_j S'(p'-j,\, p'-k+1) + \sum_j a'_j S'(p'-j,\, p'-k) =$$

$$= \sum_j y_1 a'_j S(p'-j, p'-k+1) + \sum_j a'_j ((p-k)S(p'-j, p'-k+1) + S(p'-j,\, p'-k))$$

that is

$$r_k = \sum_j (y_1 a'_{j-1} + a'_j)\, S(p-j, p-k)$$

because of the classical identity

$$S(p-j, p-k) = (p-k)\,S(p-j-1, p-k) + S(p-j-1, p-k-1)$$

and the result follows since the second relation implies the identity $a_j = y_1 a'_{j-1} + a'_j$.

Q.E.D.

Corollary 6. The Ferrers relations $F(\varphi)$ and $F(\varphi')$ defined by two decreasing maps φ and φ' are rook-equivalent only if $\varphi = \varphi'$.

Proof. Let $\tilde{\varphi}(1) = p \geq \tilde{\varphi}'(1) = p'$. If $p \neq p'$, t^p has a positive coefficient in $\rho(F(\varphi))$ and a zero coefficient in $\rho(F(\varphi'))$. Thus we can assume $p = p'$.

Because of the strictly decreasing character of the sequence

- 418 -

$\underset{\sim}{x} = (x_1, x_2, \dots, x_p)$, the sequence $\underset{\sim}{y} = (y_1, y_2, \dots, y_p)$ is non-increasing. Its members are positive because $y_p = x_p > 0$.

Since the correspondence $\underset{\sim}{x} \longrightarrow \underset{\sim}{y}$ is injective for each given p, it follows that $\underset{\sim}{x} \longrightarrow \{y_1, y_2, \dots, y_p\}$ is also injective. Now the set $\{y_1, y_2, \dots, y_p\}$ is unequivocally determined by the symmetric functions a_k and the result follows from lemma 5.

Q.E.D.

Observation 7. Because of the well-known orthogonality relations between the Stirling numbers of the first kind $s(i,j)$ and those of the second kind $S(i,j)$, the formula of lemma 5 is equivalent with

$$a_k = \sum_j r_j \, s(p-j, p-k).$$

As a side remark we may note that the formula of Property 6 does not depend upon the fact that the x_i's form a decreasing sequence of positive integers. By taking all the y_i's equal to 0 (resp. to 1) and by using a straightforward computation this formula gives directly the known identities

$$S(p, p-k) = \sum_{(k)} i_1 (i_2 - 1) \dots (i_k - k + 1)$$

$$S(p+1, p+1-k) = \sum_{0 \le j \le k} \binom{p}{j} S(p-j, p-k).$$

We now return to our main argument. To simplify notations, each Ferrers relation $F(\varphi)$ is considered as a relation in $[n] \times [n]$ where $n = \sum_i \varphi(i)$.

Definition 8.

Let $\varphi : \mathbb{P} \longrightarrow \mathbb{N}$ be non-increasing. The element $(k,k') \in \mathbb{P} \times \mathbb{P}$ is <u>admissible</u> if and only if the following inequalities hold

(1) $$0 \le \varphi(k) - k' \le \tilde{\varphi}(k') - k \le \varphi(k-1) - k'$$

where by convention $\varphi(k-1) - k' = +\infty$ for $k = 1$.

Definition 8 bis.

If (k,k') is admissible, the (k,k')-transform of φ is the map $\varphi' = \Theta(k,k';\varphi)$ defined by $\varphi'(i) = k'-k+\tilde{\varphi}(k'-k+i)$ if $k \le i \le \tilde{\varphi}(k')$ and $\varphi'(i) = \varphi(i)$ otherwise.

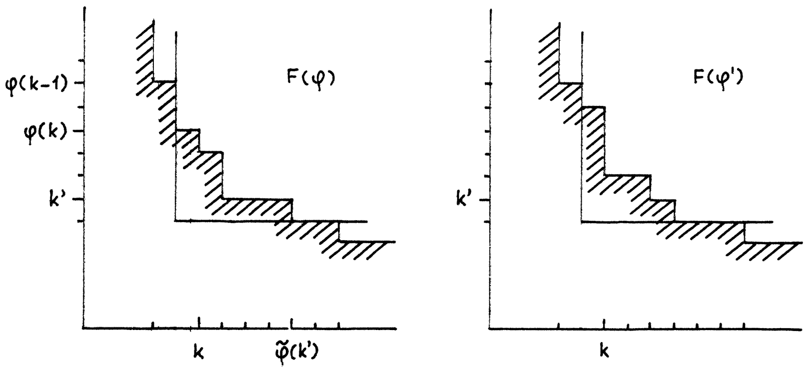


Fig. 1.

As can be seen from figure 1, where (k,k') is admissible, the relation $F(\varphi')$ is obtained from $F(\varphi)$ by transposing the set $F(\varphi) \cap B$ where $B = \{(i,j) \in \mathbb{P} \times \mathbb{P} : i \ge k,\ j \ge k'\}$ and leave $F(\varphi) \setminus B$ invariant. As shown in lemma 5 below, the admissibility conditions will insure that the relation $F(\varphi')$ obtained from $F(\varphi)$ is still a Ferrers relation having the same rook polynomial as $F(\varphi)$.

Lemma 9.

Let φ, (k,k') and $\varphi' = \Theta(k,k';\varphi)$ as above and assume $\sum \varphi(i) \le n$. Then φ' is a non-increasing map such that the Ferrers relations $F(\varphi)$ and $F(\varphi')$ are rook-equivalent.

Proof.

Using inequalities (1) and the fact that φ is non-increasing, we have $k' \le \varphi(i) \le \varphi(k)$ for each i such that $k \le i \le \tilde{\varphi}(k')$. Accordingly the set $F(\varphi) \cap B = \{(i,j) \in F(\varphi) : i \ge k,\ j \ge k'\}$ is entirely contained in $[k, \tilde{\varphi}(k')] \times [k', \varphi(k)]$.

- 420 -

Let $m = \tilde{\varphi}(k') - k + 1$ and consider the (m,n)-injection α such that $\alpha(i,j) = (k-1+i, k'-1+j)$ identically.

Keeping the same notations as in section 2 we have

$$A_1 = \alpha_1([m]) = [k, \tilde{\varphi}(k')] \quad \text{and} \quad A_2 = \alpha_2([m]) = [k', k'-1+m] .$$

Again using inequalities (1) we get $k'-1+m = \tilde{\varphi}(k') + k' - k \geq \varphi(k) \geq k'$. Hence $F(\varphi) \cap B$ is a subset of $A_1 \times A_2$. Therefore $\alpha^{-1}(F(\varphi) \cap (A_1 \times A_2))$ is the Ferrers relation $F(\psi) \subset [m] \times [m]$ where $\psi(i) = \varphi(k-1+i) - k'+1$ for each $i \in [m]$.

For the same reason $F(\varphi) \setminus (A_1 \times A_2)$ is the Ferrers relation $F(\bar{\varphi})$ where $\bar{\varphi}(i) = k'-1$ or $= \varphi(i)$ depending upon $k \leq i \leq k-1+m = \tilde{\varphi}(k')$ or not. It follows that $F(\bar{\varphi}) \cap (A_1 \times [n]) = A_1 \times [k'-1]$.

Because of inequalities (1) we have

$F(\bar{\varphi}) \cap ([n] \times A_2) = [k-1] \times A_2$ since either $[k-1] = \phi$ or $\varphi(k-1) \geq \tilde{\varphi}(k') + k' - k = k'-1+m$.

Thus the compatibility condition is satisfied and $F(\varphi)$ is rook-equivalent with its α-transpose $F' = F(\bar{\varphi}) \cup \alpha(F(\tilde{\psi}))$.

Now since $\bar{\varphi}(i) = k'-1$ for $k \leq i \leq k-1+m$ we have

$$F' = \{(i,j) : j \leq \bar{\varphi}(i) + \psi'(i)\} \quad \text{where} \quad \psi'(i) = \tilde{\psi}(i-k+1) \text{ if } k \leq i \leq k-1+m ;$$
$$= 0 \text{ otherwise.}$$

Direct computation shows that $\bar{\varphi} + \psi'$ is in fact the (k,k')-transform φ' of φ.

Further, inequalities (1) imply $\varphi'(k-1) \geq \varphi'(k)$ if $k > 1$. Since both $\bar{\varphi}$ and ψ' (for $i \geq k$) are non-increasing, this establishes that φ' is also non-increasing, and that, accordingly, $F' = F(\varphi')$ is a Ferrers relation.

Q.E.D.

Remark 10.

Consider any non-increasing map φ. It follows from definition 4 that the pair $(1, \varphi(1))$ is always admissible (with respect to φ) since inequalities (1) reduce to

$$0 \leq \varphi(1) - \varphi(1) \leq \tilde{\varphi}(\varphi(1)) - 1 .$$

Let $m = \tilde{\varphi}(\varphi(1))$. The $(1,\varphi(1))$ -transform φ' of φ satisfies

$$\begin{aligned} \varphi'(1) &= \varphi(1) + m - 1 & \\ \varphi'(i) &= \varphi(i) - 1 & \text{for } 1 \leq i \leq m ; \\ \varphi'(i) &= \varphi(i) & \text{otherwise} . \end{aligned}$$

In particular $\varphi = \varphi'$ if and only if $m = 1$, that is, if and only if $\varphi(1) > \varphi(2)$ since $m = \tilde{\varphi}(\varphi(1))$ is characterized by $\varphi(1) = \varphi(2) = \dots$ $\dots = \varphi(m) > \varphi(m+1)$.

We shall say that $\bar{\varphi}$ is a transform of the non-increasing map φ if and only if $\bar{\varphi}$ can be obtained by a succession of (k,k') -transformations (in which, of course, the admissibility conditions are satisfied).

We now come to our main theorem.

THEOREM 11.

Each Ferrers relation is rook-equivalent with exactly one decreasing Ferrers relation.

Proof. In view of Corollary 6, it suffices to show that any Ferrers relation has a transform which is decreasing.

Consider a Ferrers relation $F(\varphi)$ which is not decreasing and let j be the least value such that $\varphi(j) = \varphi(j+1)$. Setting $k' = \varphi(j)$ and $j' = \tilde{\varphi}(k') = \mathrm{Max}\{ i \in [n] : \varphi(i) = \varphi(j)\}$ we have $j < j'$, hence

$$0 = \varphi(j) - k' < j' - j = \tilde{\varphi}(k') - j .$$

proving that $k = \mathrm{Min}\{ i \in [n] : \varphi(i) - k' < \tilde{\varphi}(k') - i \}$ is positive.

- 422 -

Because of the minimal character of k, the pair (k,k') is admissible. Thus φ admits a (k,k')-transform $\varphi' = \Theta(k,k';\varphi)$. Further φ' precedes φ in the sense that the sequence $(\varphi'(1),\varphi'(2),\dots,\varphi'(n))$ precedes the sequence $(\varphi(1),\varphi(2),\dots,\varphi(n))$ in lexicographic order because by construction

$$\varphi'(i) = \varphi(i) \quad \text{for} \quad i<k \quad \text{and}$$

$$\varphi'(i) = \tilde{\varphi}(k') + k' - k > \varphi(i)\,.$$

Using induction on this order, it concludes the proof.

Q.E.D.

The subsequent results require some further definitions.

For $a>0$ and $b\geq 0$ we shall denote $\eta_{a,b}$ the non-increasing map such that $\eta_{a,b}(i) = b$ or 0 depending upon $i\leq a$ or not. Let φ be a non-increasing map. The transpose of $\tilde{\varphi} + \eta_{a,b}$ will be called the (a,b)-<u>translate</u> of φ and denoted $\Delta(a,b;\varphi)$.

From the geometrical point of view the graph of $\Delta(a,b;\varphi)$ is obtained from $F(\varphi)$ by first considering $F(\varphi)$ as a subset of $\mathbb{P}\times[a]$ and then making a b-length translation of $F(\varphi)$ to the right.

In particular if $a\geq\varphi(1)$, the $(a-b)$-translate of φ is the non-increasing map φ' such that

$$\varphi'(i) = a \qquad \text{for} \quad i\leq b \quad \text{and}$$

$$\varphi'(i) = \varphi(i-b) \quad \text{for} \quad i>b\,.$$

From the proof of Theorem 11 it follows that the unique decreasing map $\hat{\varphi}$ in which the Ferrers relations $F(\varphi)$ and $F(\hat{\varphi})$ are rook-equivalent is a transform of φ; we shall then say that $\hat{\varphi}$ is the <u>decreasing transform</u> of φ.

The following lemma is a special case of a theorem in Riordan ([4] p. 181, theorem 3).

- 423 -

Lemma 12. Let $\bar{\varphi}$ (resp. $\bar{\varphi}'$) be a transform of φ (resp. φ') Then

(i) $\bar{\varphi}+\eta_{a,b}$ is a transform of $\varphi+\eta_{a,b}$ if we have $a \geq \tilde{\varphi}(1)$;

(ii) $\Delta(a,b;\bar{\varphi})$ is a transform of $\Delta(a,b;\varphi)$ if the inequality $a \geq \varphi(1)$ holds;

(iii) $\Delta(a,b;\bar{\varphi})+\bar{\varphi}'$ is a transform of $\Delta(a,b;\varphi)+\varphi'$ if the inequalities $a \geq \varphi(1)$ and $b \geq \tilde{\varphi}'(1)$ both hold.

Proof.

For proving (i) and (ii), is suffices to consider the case when $\bar{\varphi} = \Theta(k,k';\varphi)$. Direct verification shows that

(i) if $a \geq \tilde{\varphi}(1)$, the pair $(k,k'+b)$ is admissible for $\varphi_1 = \varphi+\eta_{a,b}$ and then we have $\bar{\varphi}_1 = \Theta(k,k'+b;\varphi_1)$;

(ii) if $a \geq \varphi(1)$, the pair $(k+b,k')$ is admissible for $\varphi_2 = \Delta(a,b;\varphi)$ and we get $\bar{\varphi}_2 = \Theta(k+b,k';\varphi_2)$.

Now let $\psi = \Delta(a,b;\varphi)+\varphi'$ and $\bar{\psi} = \Delta(a,b;\bar{\varphi})+\bar{\varphi}'$. We already know from part (ii) that $\bar{\varphi}_1 = \Delta(a,b;\bar{\varphi})$ is a transform of $\varphi_1 = \Delta(a,b;\varphi)$. Note also that $\varphi_1(i) = a$ for $i \leq b$ and $\tilde{\varphi}_1(i) = 0$ for $i > a$; this implies

$$\tilde{\psi}(i) = \tilde{\varphi}'(i-a) \quad \text{for} \quad i > a.$$

To prove (iii) it suffices to consider the case when $\bar{\varphi}' = \Theta(k,k';\varphi')$. Since $b \geq \tilde{\varphi}'(1)$, we necessarily have $k \leq b$ and then $\varphi_1(k) = a$. It then follows that

$$\begin{aligned} \psi(k)-(k'+a) &= \varphi_1(k)+\varphi'(k)-(k'+a) \\ &= \varphi'(k)-k'. \end{aligned}$$

We also have

- 424 -

$$\tilde{\psi}(k'+a) - k = \tilde{\varphi}'(k') - k \qquad \text{and}$$

$$\psi(k-1) - (k'+a) = \varphi'(k-1) - k'.$$

These last three equations show that the pair $(k, k'+a)$ is admissible for ψ and then we have $\bar{\psi} = \Theta(k, k'+a; \psi)$.

Q.E.D.

We consider now a very special case needed in the study of Newcomb's problem.

Definition 13.

Let $\underset{\sim}{d} = (d_1, d_2, \ldots, d_r)$ be a sequence of $r > 0$ positive integers and set $\bar{d}_s = d_1 + d_2 + \cdots + d_s$ $(1 \le s \le r)$.

The special map of type $\underset{\sim}{d}$, denoted by $\underset{\sim}{\mathrm{Spec}}\ \underset{\sim}{d}$, is the mapping φ from $\mathbb{P}$ into $\mathbb{N}$ that is defined by the following conditions:

$$\varphi(i) = 0 \qquad \text{for} \qquad i > \bar{d}_r$$

$$\varphi(\bar{d}_r) = 1 \qquad \text{and}$$

$$\varphi(i) = \varphi(i+1) + 2 \quad \text{or} \quad = \varphi(i+1) + 1 \qquad \text{depending upon}$$

$$i \in \{\bar{d}_1, \bar{d}_2, \ldots, \bar{d}_{r-1}\} \qquad \text{or not for} \qquad 1 \le i < \bar{d}_r .$$

The definition implies that any special map is decreasing. One computes easily that the value at 1 of $\underset{\sim}{\mathrm{Spec}}\ \underset{\sim}{d} = \varphi$ is $\bar{d}_r + r - 1$ and that $\tilde{\varphi}(1) = d_r$.

For $\underset{\sim}{a} = (a_1)$ the Ferrers relation $F(\underset{\sim}{\mathrm{Spec}}\,(d))$ is a "triangular board" in Riordan's terminology (see [4] p. 213). Then its rook polynomial is given by Riordan's formula ([4] p. 214) which involves the Stirling numbers $S(i,j)$ as it could be shown directly from our property 5, since it corresponds to the case when $y_i = 1$ identically.

We now give two results which show that "translation" preserves the "special" character of a Ferrers relation.

Lemma 14. Let $\psi = \underset{\sim}{\text{Spec}}(d_1, d_2, \dots, d_r)$ and $q \le r$. The special map $\bar\psi = \underset{\sim}{\text{Spec}}(d_1+1, d_2+1, \dots, d_q+1, d_{q+1}, \dots, d_r)$ is the decreasing transform of $\Delta(d, q; \psi)$ where $d = 1 + \psi(1)$.

Proof.

Let $\varphi = \Delta(d,q;\psi)$; we have $\varphi(i) = d$ for $1 \le i \le q$ and $\varphi(i) = \psi(i-q)$ for $q < i \le \bar d_r + q$. On the other hand $\bar\psi(i) = \psi(i-q)$ for $\bar d_q + q \le i \le \bar d_r + q$.

Consequently, if $q = 1$, it is readily verified that $\Delta(d, 1; \psi) = \bar\psi$ and the lemma is proved in this case.

If $q > 1$, we construct the $(1, \varphi(1))$-transform φ' of φ. We have $\varphi'(1) = \varphi(1) + q - 1 \ (= d')$, $\varphi'(i) = \varphi(1) - 1 = \psi(1)$ for $1 < i \le q$ and $\varphi'(i) = \varphi(i)$ for $i > q$.

We now distinguish two cases:

(i) suppose $d_1 > 1$, and let $\psi' = \underset{\sim}{\text{Spec}}(d_1 - 1, d_2, \dots, d_r)$; one can verify that φ' is obtained by the following two transformations. We first let $\psi'' = \Delta(\psi(1), q; \psi')$; then we have $\varphi' = \Delta(d', 1; \psi'')$. On the other hand let $\bar\psi' = \underset{\sim}{\text{Spec}}(d_1 - 1 + 1, d_2 + 1, \dots, d_q + 1, d_{q+1}, \dots, d_r)$; as we have $d' = 1 + \psi'(1) + q = 1 + \bar\psi'(1)$, we get $\bar\psi = \Delta(d', 1; \bar\psi')$.

By induction on $d_1 + d_2 + \dots + d_r$ the special map $\bar\psi'$ is a transform of ψ'' and by lemma 12 $\bar\psi = \Delta(d', 1; \bar\psi')$ is a transform of $\varphi' = \Delta(d', 1; \psi'')$. This concludes the proof in this case.

(ii) suppose $d_1 = 1$ and construct the $(2, \varphi'(2))$-transform φ'' of φ'. Using the same device as above, let $\psi' = \underset{\sim}{\text{Spec}}(d_2, d_3, \dots, d_r)$; we successively form $\psi'' = \Delta(1 + \psi'(1), q-1; \psi')$ and $\psi''' = \Delta(d'-1, 1; \psi'')$ where $d' = \psi(1) + q$. Then we obtain $\varphi'' = \Delta(d', 1; \psi''')$. On the other hand, let $\bar{\psi}' = \underset{\sim}{\text{Spec}}(d_2+1, d_3+1, \dots, d_q+1, d_{q+1}, \dots, d_r)$; we obtain $\bar{\psi} = \Delta(d', 1; \chi)$ where $\chi = \Delta(d'-1, 1; \bar{\psi}')$.

Again by induction and using lemma 12 as above, we verify that $\bar{\psi}$ is a transform of φ'', hence of φ.

Q.E.D.

Lemma 15.

Let be ψ as above, $q \geq r \geq 0$, $d = \psi(1) + q - r + 1$ and $\varphi = \Delta(d, q; \psi)$. Then the decreasing transform of φ is $\bar{\psi} = \underset{\sim}{\text{Spec}}(d'_1, d'_2, \dots, d'_q)$ where $d'_i = 1$ for $i \leq q-r$ and $d'_i = 1 + d_{i-q+r}$ for $i > q-r$.

Proof.

Lemmas 14 and 15 coincide for $q = r$. Suppose $q > r$, and take the $(1, \varphi(1))$-transform φ' of φ. As in the proof of lemma 14, we verify that φ' is obtained from ψ by the two successive transformations. First let $\psi'' = \Delta(\psi(1)+q-r, q-1; \psi)$; then $\varphi' = \Delta(\bar{\psi}(1), 1; \psi'')$. By induction on q a transform of $\psi'' = \Delta(\psi(1) + (q-1) - r+1, q-1; \psi)$ is given by $\bar{\psi}'' = \underset{\sim}{\text{Spec}}(d''_1, d''_2, \dots, d''_{q-1})$ where $d''_i = 1$ for $i \leq q-1-r$ and $d''_i = 1 + d_{i-q+1+r}$ for $i > q-1-r$. Therefore it follows from lemma 12 that $\bar{\psi} = \Delta(\bar{\psi}(1), 1; \bar{\psi}'')$ is a transform of $\varphi' = \Delta(\bar{\psi}(1), 1; \psi'')$ (hence of φ).

Q.E.D.

It is to be noted that for $r = 0$, i.e. for $\varphi = \eta_{q,q}$, one has simply $\bar{\psi} = \sum(1, 1, \dots, 1)$. This is a special case of Riordan's formula (34) ([4] p. 211).

4. APPLICATION TO NEWCOMB'S PROBLEM

Let $\gamma : [n] \longrightarrow [p]$ be an order-preserving surjection and J a subset of $[p]$. Then the relation $\bar{R}(\gamma, J) \subset [n] \times [n]$ is defined by the conditions $(i,j) \in \bar{R}(\gamma, J)$ if and only if $\gamma(i) < \gamma(j)$ or $\gamma(i) = \gamma(j) \in J$. Thus for $J = [p]$ the relation $\bar{R}(\gamma, J)$ is simply the total <u>preorder</u>

induced by γ, and $\bar{R}(\gamma, J)$ is the complement of a total preorder when $J = \emptyset$. Both extreme cases occur naturally in Newcomb's problem that involves the determination of the rook polynomial of $\bar{R}(\gamma, [p])$ or $\bar{R}(\gamma, \emptyset)$. In the generalization of this problem studied by one of us ([1] & [2]) it appears just as natural to consider the rook polynomial of $\bar{R}(\gamma, J)$ for any subset J of $[p]$. This is the purpose of this last section.

Keeping the same notations we let $\pi : [p] \longrightarrow \mathbb{P}$ be the map defined by

$$\pi(j) = \underset{\sim}{\text{card}}\, \gamma^{-1}(j) \qquad (j \in [p])$$

and prove the following property (Cf. Riordan [4], Ex. 4, p. 185).

Property 16. For any permutation σ of $[p]$, the relations $\bar{R}(\gamma, J)$ and $\bar{R}(\sigma\gamma, \sigma J)$ are rook-equivalent.

Proof.

It suffices to verify the property in the special case where σ is the transposition exchanging two consecutive values q and $q+1$ of $[p]$.

Set
$$m = \pi(q) + \pi(q+1);$$
$$a' = \operatorname{Min}\{i \in [n] : \gamma(i) = q\};$$
$$a = \operatorname{Max}\{i \in [n] : \gamma(i) = q+1\};$$

and define the (m,n)-injection α by $\alpha(i,j) = (a+1-i,\, a'+j-1)$.

The verification of the compatibility condition is trivial and one sees that $\bar{R}(\sigma\gamma, \sigma J)$ is the α-transpose of $\bar{R}(\gamma, J)$. Q.E.D.

Observe now that the relation

$$R(\gamma, J) = \{(i,j) \in [n] \times [n] : (i, n+1-j) \in \bar{R}(\gamma, J)\}$$

is obviously a Ferrers relation which is rook-equivalent to $\bar{R}(\gamma, J)$. We let $\varphi = \varphi(\gamma, J)$ denote the non-increasing map such that $R(\gamma, J)$ is the Ferrers relation $F(\varphi)$. According to property 16 we can assume that J is the interval $[p']$ with $0 \le p' \le p$ and that the two restrictions of π respectively to $[p']$ and $[p] \setminus [p']$ are non-increasing.

Throughout this section φ will designate the non-increasing map $\varphi = \phi(\gamma, J)$ with $\gamma : [n] \longrightarrow [p]$ and $J = [p']$ $(0 \le p' \le p)$.

We first consider the case when $J = [p]$, that is when $\bar{R}(\gamma, [p])$ is the total preorder induced by γ. With our conventions the map $\pi : [p] \longrightarrow \mathbb{P}$ is then non-increasing and we can define the non-increasing map $\bar{\pi}$ its transpose, as defined in the beginning of section 3.

Lemma 17.

The decreasing transform of the defining map $\varphi = \phi(\gamma, [p])$ of the Ferrers relation $R(\gamma, [p])$ is the special map

$$\psi = \underset{\sim}{\mathrm{Spec}}(\tilde{\pi}(q), \tilde{\pi}(q-1), \dots, \tilde{\pi}(1))$$

where $q = \pi(1)$.

Proof.

For $p = 1$, we have on the one hand

$\pi(1) = n = q$, $\tilde{\pi}(1) = \tilde{\pi}(2) = \dots = \tilde{\pi}(n) = 1$, $\tilde{\pi}(n+1) = 0$.

On the other hand, $\varphi = \eta_{n,n}$ since $R(\gamma, [p]) = [n] \times [n]$. Thus the result is covered by Lemma 15 and we can use induction on $p \ge 2$.

Define $\pi' : [p-1] \longrightarrow \mathbb{P}$ by letting $\pi'(i) = \pi(i+1)$ for each $i \in [p-1]$. Thus let $r = \pi'(1)$ $(= \pi(2))$; we have

(1) $$r \le q.$$

Moreover, let

$$d'_1 = \tilde{\pi}'(r),\ d'_2 = \tilde{\pi}'(r-1), \dots, d'_r = \tilde{\pi}'(1)$$

and $\psi' = \underset{\sim}{\mathrm{Spec}}(\underset{\sim}{d'})$; then, by the induction hypothesis, ψ' is the decreasing transform of $\varphi' = \phi(\gamma', [p-1])$ where γ' is the unique order-preserving surjection of $[n']$ $(n' = n - \pi(1))$ onto $[p-1]$ that satisfies

$\pi'(j) = \underset{\sim\sim}{\text{card}}\ \gamma'^{-1}(j)$ $(j \in [p-1])$ identically.

As we know, the value of ψ' at 1 is

$$\psi'(1) = d'_1 + d'_2 + \dots + d'_r + r - 1 = \sum \tilde{\pi}'(i) + r - 1 .$$

On the other hand

$$\sum \tilde{\pi}'(i) = \sum \pi'(i) = \sum \pi(i+1) = n - \pi(1) = n - q \ ;$$

hence

(2) $$n = \psi'(1) + q - r + 1 .$$

We also note that the sequence $(\tilde{\pi}(q), \tilde{\pi}(q-1), \dots, \tilde{\pi}(1))$ is the sequence obtained when putting $q-r$ elements equal to 1 in front of $\underset{\sim}{d}'$ and increasing by 1 each term of $\underset{\sim}{d}'$. In other words

$$\psi = \underset{\sim\sim}{\text{Spec}}\,(1, \dots, 1, 1+d'_1, \dots, 1+d'_r)$$

where the 1 's are repeated $(q-r)$ times in the sequence.

Furthermore φ is seen to be the (n,q)-translate of φ', i.e.

(3) $$\varphi = \Delta(n, q ; \varphi') .$$

As $n \geq \varphi'(1)$, the conditions of lemma 12 are fulfilled and we conclude that φ is a transform of $\bar{\varphi} = \Delta(n, q ; \psi')$.

In view of (1), (2) and (3) we can apply lemma 15 by taking ψ' instead of ψ, $d = n$ and $\bar{\varphi}$ in place of φ and we deduce that ψ is a transform of $\bar{\varphi}$, hence of φ.

Q.E.D.

We now study the case of complements of total preorders. With the

- 430 -

same notations as above, the set J is assumed to be empty. In the following lemma we characterize the decreasing Ferrers relation that is rook-equivalent to $R(\gamma,\phi)$. We can assume $p>1$, since otherwise $R(\gamma,\phi) = \bar{R}(\gamma,\phi) = \phi$. We set $b = \pi(1) - \pi(2)$ and $r = \pi(2)$. Since π is non-increasing, we can consider the transpose $\tilde{\pi}$ of π. We then have $r = \mathrm{Max}\{i : \tilde{\pi}(i) > 1\}$.

Lemma 18.

The decreasing transform of the defining map $\varphi = \phi(\gamma,\phi)$ of $R(\gamma,\phi)$ is $\psi + \eta_{m,b}$ where $m = n - \pi(1)$ and where ψ is the special map $\underset{\sim}{\mathrm{Spec}}\,(\tilde{\pi}(1)-1, \tilde{\pi}(2)-1, \ldots, \tilde{\pi}(r)-1)$.

Proof.

Instead of $\phi(\gamma,\phi)$ and $R(\gamma,\phi)$ we will also write $\phi(\pi)$ and $R(\pi)$. Since the lemma is trivial for $n \leq 2$, we can use induction on $n \geq 3$.

We distinguish two cases.

Case 1: $b = \pi(1) - \pi(2) > 0$.

Let $n' = n-b$ and define $\pi' : [p] \longrightarrow \mathbb{P}$ by letting $\pi'(1) = \pi(1) - b = \pi(2)$; $\pi'(i) = \pi(i)$ otherwise. By construction we have $b' = \pi'(1) - \pi'(2) = 0$ and then

$$\begin{aligned} \tilde{\pi}'(i) &= \tilde{\pi}(i) && \text{for} \quad 1 \leq i \leq r; \\ &= 0 && \text{for} \quad i > r. \end{aligned}$$

Thus since $b' = 0$, the induction hypothesis implies that the special map ψ defined in the theorem is a transform of the defining map $\varphi' = \phi(\pi')$ of the Ferrers relation $R(\pi')$. Now from the definition of $R(\pi)$ and $R(\pi')$ we have that

$$\varphi = \varphi' + \eta_{m,b}$$

where $m = n - \pi(1)$. As $m = n - \pi(1) = n-b-r = \sum\{\tilde{\pi}(i)-1 : 1 \leq i \leq r\} = \tilde{\psi}(1)$, the conditions of lemma 12 are satisfied and $\psi + \eta_{m,b}$ is indeed a transform of φ.

- 431 -

Case 2: $b = \pi(1) - \pi(2) = 0$.

By hypothesis we have $\pi(1) = \pi(2) = \dots = \pi(s) = r$ where $s \geq 2$ and either $s = p$ or $s < p$ and $\pi(s) > \pi(s+1)$. Define maps π' and π'' of $[p]$ into $\mathbb{P}$ by letting $\pi'(s) = \pi(s) - 1$; $\pi'(i) = \pi(i)$ for $i \neq s$;

$\pi''(1) = \pi(1) - 1$; $\pi''(i) = \pi(i)$ for $i \neq 1$.

Thus $\sum \pi'(i) = \sum \pi''(i) = n-1$. The map π' is non-increasing and $\tilde{\pi}'(r) = r-1$, $\tilde{\pi}'(i) = \tilde{\pi}(i)$ for $i < r$ $(= \pi(1) = \pi'(1))$. Hence by the induction hypothesis $\varphi' = \phi(\pi')$ has for transform

$\psi' = \underset{\sim}{\text{Spec}}\,(\tilde{\pi}(1) - 1, \tilde{\pi}(2) - 1, \dots, \tilde{\pi}(r-1) - 1, \tilde{\pi}(r) - 2)$

where eventually $\tilde{\pi}(r) - 2 = 0$.

Now by property 16, $R(\pi')$ and $R(\pi'')$ that is $F(\varphi')$ and $F(\varphi'')$ (where $\varphi'' = \phi(\pi'')$) are rook-equivalent. Further $\varphi = \phi(\pi)$ is equal to $\varphi'' + \eta_{m,1}$. Thus observing that $\tilde{\psi}'(1) = m-1$, we can again conclude from lemma 12 that the special map ψ defined in the lemma is a transform of φ since it is equal to $\psi' + \eta_{m,1}$.

Q.E.D.

We now come to the general case when J is not necessarily equal to $[p]$ or empty.

Theorem 19.

Let $\varphi = \phi(\gamma, J)$ be the defining map of the Ferrers relation $R(\gamma, J)$ where $\gamma: [n] \longrightarrow [p]$ is an order-preserving surjection and $J = [p']$ with $0 \leq p' \leq p$.

Moreover let

$$\pi'(i) = \underset{\sim}{\text{card}}\ \gamma^{-1}(i) \quad \text{for} \quad i \in [p']$$
$$= 0 \quad \text{for} \quad i > p',$$
$$\pi''(i) = \underset{\sim}{\text{card}}\ \gamma^{-1}(p'+i) \quad \text{for} \quad i \in [p-p']$$
$$= 0 \quad \text{for} \quad i > p-p'.$$

- 432 -

Then the decreasing transform of φ is

$$\psi + \eta_{m,b}$$

where $m = n - \pi''(1)$, $b = \pi''(1) - \pi''(2)$ and

$$\psi = \underset{\sim}{\mathrm{Spec}}\,(\tilde{\pi}'(q), \ldots, \tilde{\pi}'(1), \tilde{\pi}''(1)-1, \ldots, \tilde{\pi}''(r)-1)$$

with $q = \pi'(1)$ and $r = \pi''(2)$.

Proof.

The case $J = [p]$ or $J = \emptyset$ has been considered in lemmas 17 and 18. We shall then assume $1 \le p' < p$, and will also use the following notations:

$$n' = \pi'(1) + \cdots + \pi'(p'), \quad p'' = p - p', \quad n'' = n - n',$$

$$\gamma'(i) = \gamma(i) \quad \text{for} \quad i \in [n'] \quad \text{and} \quad \gamma''(i) = \gamma(n'+i) \quad \text{for} \quad i \in [n''].$$

Let us define the two maps φ' and φ'' as follows:

$$\begin{aligned} \varphi'(i) &= \varphi(i) - n'' \quad \text{for} \quad i \in [n'] \\ &= 0 \quad \text{for} \quad i > n' \\ \varphi''(i) &= \varphi(i+n') \quad \text{for} \quad i \in \mathbb{P}. \end{aligned}$$

By construction we have

$$\varphi = \Delta(n'', n'; \varphi'') + \varphi'.$$

Moreover we clearly have $\varphi' = \phi(\gamma', [p'])$ and $\varphi'' = \phi(\gamma'', \emptyset)$ if $p'' > 1$ and $\varphi'' = 0$ if $p'' = 1$. According to lemma 17 the decreasing transform of φ' is then the special map

$$\bar{\varphi}' = \underset{\sim}{\mathrm{Spec}}\,(\tilde{\pi}'(q), \tilde{\pi}'(q-1), \ldots, \tilde{\pi}'(1))$$

where $q = \pi'(1)$. In the same manner lemma 18 asserts that if $p'' > 1$, the decreasing transform of φ'' is $\bar{\varphi}'' = \psi'' + \eta_{m'',b}$ where $m'' = n'' - \pi''(1)$ and $\psi'' = \underset{\sim}{\mathrm{Spec}}\,(\tilde{\pi}''(1)-1, \tilde{\pi}''(2)-1, \ldots, \tilde{\pi}''(r)-1)$ with $r = \pi''(2)$. If $p'' = 1$, we let $\bar{\varphi}'' = \varphi'' = 0$.

- 433 -

Now since $n'' > \varphi''(1)$ and $n' = \tilde{\varphi}'(1)$, the conditions of lemma 12 are fulfilled and we can conclude that

$$\bar{\varphi} = \Delta(n'', n'; \bar{\varphi}'') + \bar{\varphi}'$$

is a transform of $\varphi = \Delta(n'', n'; \varphi'') + \varphi'$. As one has $n' = \pi(1) + \dots + \pi(p') = \tilde{\pi}'(q) + \dots + \tilde{\pi}'(1)$, the value of $\bar{\varphi}'$, at n' is equal to 1. On the other hand the value of $\bar{\varphi}''$ at 1 is equal to 0 if $p'' = 1$ and if $p'' > 1$, we have

$$\begin{aligned}\bar{\varphi}''(1) &= b + \sum_{1 \le i \le r} (\tilde{\pi}''(i) - 1) + r - 1 \\ &= n'' - 1 .\end{aligned}$$

Accordingly we have $\bar{\varphi}(n') = n'' + 1$ and $\bar{\varphi}(n'+1) = n'' - 1$ if $p'' > 1$ and $\bar{\varphi}(n'+1) = 0$ if $p'' = 1$. As $\bar{\varphi}'$ and ψ'' are special, this shows, in the case when $p'' > 1$, that $\bar{\varphi}$ is equal to $\eta_{m,b} + \psi$ where $m = n' + m'' = n - \pi''(1)$, $b = \pi''(1) - \pi''(2)$ and $\psi = \underset{\sim\sim}{\mathrm{Spec}}(\tilde{\pi}'(q), \dots, \tilde{\pi}'(1), \tilde{\pi}''(1) - 1, \dots, \tilde{\pi}''(r) - 1)$ with $r = \pi''(2)$. If $p'' = 1$, one has $\bar{\varphi} = \eta_{n', n''} + \psi'$ where $\psi' = \underset{\sim\sim}{\mathrm{Spec}}(\tilde{\pi}'(q), \dots, \tilde{\pi}'(1))$. In fact, in this last case we can also write $\bar{\varphi} = \eta_{m,b} + \psi$ since then $r = \pi''(2) = 0$ and ψ is reduced to $\psi = \underset{\sim\sim}{\mathrm{Spec}}(\tilde{\pi}'(q), \dots, \tilde{\pi}'(1))$.

Q.E.D.

R e m a r k 20.

The sequence $(\tilde{\pi}'(q), \dots, \tilde{\pi}'(1), \tilde{\pi}''(1) - 1, \dots, \tilde{\pi}''(r) - 1)$ just defined is unimodal, namely it is the juxtaposition of a non-decreasing sequence (of length q) and a non-increasing sequence (of length r). Thus the decreasing transform of a map $\varphi = \phi(\gamma, J)$ is apart from the function $\eta_{m,b}$ a special map $\psi = \underset{\sim\sim}{\mathrm{Spec}}(d_1, \dots, d_s)$ whose defining sequence $(d_1, \dots, d_s)$ is unimodal. Conversely, let $\underset{\sim}{d} = (d_1, \dots, d_s)$ be an unimodal sequence of positive integers and b be a non-negative integer. Then form the decreasing map $\psi = \eta_{m,b} + \underset{\sim\sim}{\mathrm{Spec}}(\underset{\sim}{d})$. It is readily verified that ψ is the decreasing transform of (at least) one non-increasing map φ of the form $\phi(\gamma, [p'])$.

- 434 -

A Ferrers relation $R(\gamma, [p'])$ in fact depends on four parameters n, p, γ and p'. Let us write $R(n,p,\gamma,p')$ instead of $R(\gamma,p')$. Thus it is easily proved that $F(\psi)$ is rook-equivalent to exactly one Ferrers relation $R(n,p,\gamma,p')$ such that $p \geq 2p'$.

Example 21.

Let $n = 14$, $p = 6$, $p' = 3$; $\pi(1) = 4$, $\pi(2) = 3$, $\pi(3) = 2$, $\pi(4) = 2$, $\pi(5) = 2$, $\pi(6) = 1$. The sequence of positive values of $\varphi = \phi(\gamma,[3])$ where γ is determined by the map π just defined, is $(14,14,14,14,10,10,10,7,7,3,3,1,1)$. The decreasing transform of φ is $\bar{\varphi} = \underset{\sim}{\mathrm{Spec}}(1,2,3,3,2,1)+\eta_{12,0}$ and the sequence of positive values of $\bar{\varphi}$ is then $(17,15,14,12,11,10,8,7,6,4,3,1)$.

REFERENCES

[1] P.CARTIER & D.FOATA: Problèmes combinatoires de commutation et réarrangements, Lecture Notes in Math. n°85, Springer-Verlag, Berlin (1969).

[2] D.FOATA: Etude algèbrique de certains problèmes d'analyse combinatoire et du calcul des probabilités, Publ. Inst. Statist. Univ. Paris 14 (1965), 81-241.

[3] F.POUSSIN: Sur une propriété arithmétique de certains polynômes associés aux nombres d'Euler, C.R.Acad. Sc. Paris 266 (1968), 392-393.

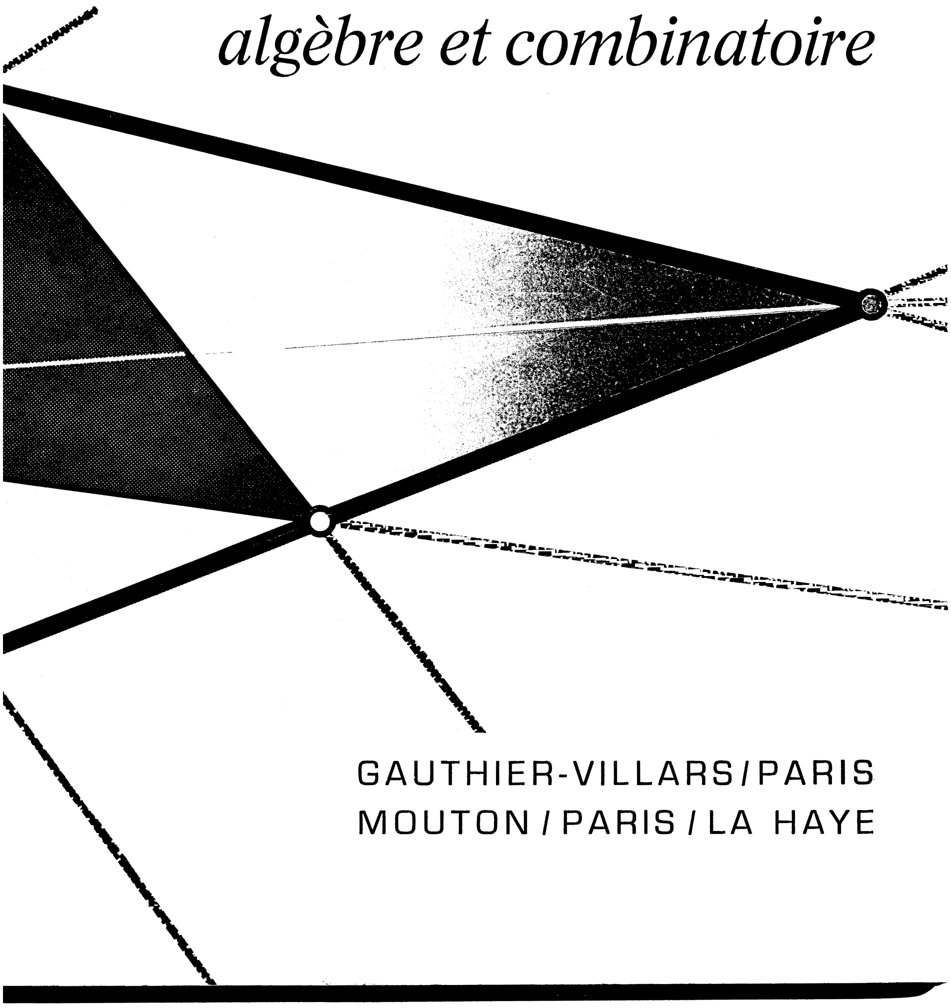
[4] J.RIORDAN: An Introduction to Combinatorial Analysis, Wiley, New York (1958).

MATHEMATIQUES ET SCIENCES DE L'HOMME

CAHIERS MATHEMATIQUES

3

algèbre et combinatoire



GAUTHIER-VILLARS / PARIS
MOUTON / PARIS / LA HAYE

M. P. SCHUTZENBERGER

Sur les contraintes définissant certains modèles formels de langage

L'un des objectifs partiels de la linguistique est l'établissement de grammaires, c'est-à-dire, pour une langue naturelle donnée, l'établissement d'un système explicite de règles qui permettent, en théorie du moins, de décider pour toute suite arbitraire de phonèmes si cette suite appartient ou non à l'ensemble des phrases grammaticalement correctes de la langue considérée. L'expérience semble indiquer qu'à un certain niveau une partie de ces règles est d'une nature assez simple pour pouvoir être discutée par des procédés purement formels. C'est là le cas des règles d'emploi des parenthèses ou, si l'on préfère, le formalisme de l'emboîtement des syntagmes et le but de cet exposé est de présenter quelques résultats obtenus dans cette direction par N. Chomsky, ses amis et leurs élèves; je renvoie aux travaux de cet auteur et G. A. Miller pour une bibliographie complète de la question. Comme les problèmes proprement linguistiques ne seront pas abordés ici, il sera commode d'utiliser les notations suivantes qui ne risquent pas d'entraîner de confusion bien que les termes de « lettres » et de « mots » y soient utilisés d'une façon qui apparaîtra sans doute incongrue aux linguistes.

Soit X un ensemble *fini* dont les éléments sont appelés conventionnellement des *lettres*, F l'ensemble de toutes les séquences finies que l'on peut former avec ces lettres, ces séquences étant elles-mêmes appelées des *mots*. Techniquement, F est le *monoïde libre* engendré par X. Ceci veut dire que le mot vide (noté e) appartient à F et qu'à toute paire de mots f, f' de F est associé leur *produit* ff', c'est-à-dire par définition, le mot formé de f suivi de f'. Nous supposerons désormais que les lettres de X sont indexées par les nombres $\pm\, i$ $(1 \leqq i \leqq n)$ c'est-à-dire que X est l'ensemble $\{x_1, x_{-1}, x_2, x_{-2}, \ldots, x_n, x_{-n}\}$. Un mot f sera dit *réduit* s'il ne contient aucune paire de lettres consécutives ayant des indices opposés, c'est-à-dire pour $f = x_{i_1}x_{i_2} \ldots x_{i_m}$ si $i_1 \neq -i_2$, $i_2 \neq -i_3, \ldots, i_{m-1} \neq -i_m$. Il est clair qu'à chaque mot f on peut associer un mot réduit, disons αf, obtenu en effaçant dans f toutes les paires de lettres consécutives dont les indices sont opposés et en répétant cette opération aussi longtemps qu'elle est possible. Par exemple

$$\alpha(x_1x_{-2}x_3x_{-3}x_2x_2x_{-1}x_1)$$

est égal à x_1x_2 à la suite des opérations suivantes: effacement de x_3x_{-3} et de $x_{-1}x_1$ ce qui donne $x_1x_{-2}x_2x_2$; effacement de $x_{-2}x_2$ ce qui donne finalement x_1x_2. On peut vérifier que l'ordre dans lequel sont effectuées les opérations d'effacement est sans effet sur le résultat final αf qui est donc, pour chaque f, un mot bien défini. Plus généralement, f et f' étant deux mots quelconques, on a l'identité: $\alpha(ff') = \alpha(\alpha f \alpha f')$. Par exemple, pour $f = x_1x_{-2}x_3x_{-3}$, $f' = x_2x_2x_{-1}x_1$, on a:

$$\alpha f = x_1x_{-2}, \qquad \alpha f' = x_2x_2$$

et comme plus haut $\alpha(x_1x_{-2}x_2x_2) = x_1x_2$. Cette construction définit un ensemble D de mots privilégiés: ceux dont la forme réduite est le mot vide e ou, si l'on veut, ceux qui peuvent être réduits à rien par les opérations successives d'effacement. Par exemple le mot

$$f = x_1x_{-2}x_3x_{-3}x_2x_2x_{-1}x_1x_{-2}x_{-1}$$

appartient à D puisqu'avec des notations évidentes l'on a:

$$f = x_1(x_{-2}(x_3x_{-3})x_2)(x_2(x_{-1}x_1)x_{-2})x_{-1}$$

par contre le mot

$$f' = x_1x_{-2}x_3x_{-3}x_2x_2x_{-1}x_1x_{-1}x_{-2}$$

obtenu en permutant les deux dernières lettres de f n'appartient pas à D parce que $\alpha f'$ est le mot non vide $x_1x_2x_{-1}x_{-2}$. D est défini algébriquement comme le noyau de l'homomorphisme β de F sur le groupe libre (engendré par l'ensemble X^+ des $x \in X$ d'indices positifs) qui satisfait pour chaque lettre x_i la condition que βx_{-i} soit l'inverse de βx_i (dans le groupe évidemment).

Soit d'autre part X_1 un sous-ensemble de X, V un ensemble de mots de deux lettres et $R(X_1, V)$ l'ensemble de tous les mots de F qui commencent par une lettre de X_1 et dont aucune paire de lettres consécutives ne forme un mot de V. Par exemple pour $X_1 = X^+$ comme plus haut et

$$V = \{x_ix_j : i < 0 < j\},$$

$R(X_1, V)$ est l'ensemble des mots de la forme ff' où toutes les lettres de f ont des indices positifs (ce que l'on notera $f \in F_+$), et où toutes les lettres de f' ont des indices négatifs. Nous dirons que l'ensemble $L = D \cap R(X_1, V)$ des mots appartenant à la fois à D et à $R(X_1, V)$ *est* un *langage CF* (Context Free) *standard*. Pour l'exemple qui vient d'être donné, L est formé de tous les mots de la forme $f\bar{f}$ où f appartient à F_+ et où $\bar{f}$ est obtenu en « retournant » le mot f, ce qui donne $\tilde{f}$, et en remplaçant dans $\tilde{f}$ chaque lettre x_i par la lettre d'indice opposé x_{-i} (Par exemple pour $f = x_1x_1x_3x_2$, on pose

$$\tilde{f} = x_2x_3x_1x_1 \qquad \text{et} \qquad \bar{f} = x_{-2}x_{-3}x_{-1}x_{-1}).$$

Plus généralement, soit donné un homomorphisme φ de F dans lui-même, c'est-à-dire pour chaque lettre x_i de X un certain mot φx_i (éventuellement vide). Considérant un langage CF standard L, on forme l'ensemble de tous les mots φf où f est dans L et ceci constitue ce que nous appellerons un langage CF *général*. Je rappelle que si

$$f = x_{i_1}x_{i_2}x_{i_3} \ldots x_{i_n}$$

le mot φf est défini comme le produit

$$\varphi x_{i_1}\varphi x_{i_2}\varphi x_{i_3} \ldots \varphi x_{i_n}$$

et φ n'est donc pas autre chose qu'une règle systématique de réécriture. Reprenant l'exemple donné plus haut de $L = \{ f\bar{f} : f \in F_+ \}$ et posant

$$\varphi x_i = \varphi x_{-i} = x_i$$

pour chaque x_i de X^+, on obtient le langage $M = \{ f\tilde{f} : f \in F_+ \}$ formé des *mots en miroir* (ou « palindromes ») $f\tilde{f}$ dont la seconde moitié $\tilde{f}$ est précisément la première moitié « retournée ».

La définition qui vient d'être donnée d'un langage CF général n'est pas la définition originale de Chomsky, mais il est facile de voir que ces deux définitions sont à très peu près équivalentes. Cette définition et les raisons pour lesquelles ces langages sont dits « context free » sont données dans l'ouvrage de Chomsky et Miller indiqué en références.

Ceci dit nous pouvons examiner les contraintes qu'imposent aux lettres d'un mot l'hypothèse que celui-ci appartient à un langage CF donné L et, afin de simplifier, nous supposerons que L est standard. Ces contraintes sont d'une double nature : les premières résultent de l'appartenance à $R(X_1, V)$ et peuvent être considérées comme purement *locales* en ce sens qu'un être dont la mémoire se limiterait à la dernière lettre lue pourrait vérifier pour une séquence arbitraire de lettres si elles sont ou non satisfaites. Il n'est pas utile de rappeler les illustrations linguistiques ou pseudo-linguistiques bien connues de ce type extrêmement banal de contraintes (cf. plus bas).

Par contre, les contraintes imposées par l'appartenance à D sont un peu moins simples : elles peuvent s'étendre aussi loin que l'on veut et elles nécessitent en général une mémoire non bornée. En effet, d'après l'identité

$$\alpha(\alpha f \alpha f') = \alpha f f',$$

deux mots f_1 et f_2 sont tels qu'il existe un mot f'' pour lequel $f_1 f''$ et $f_2 f''$ appartiennent simultanément à D, si et seulement si on a $\alpha f_1 = \alpha f_2$. Réciproquement, quand cette relation est vérifiée pour tout f'' tel que $f_1 f''$ soit dans D on a aussi $f_2 f''$ dans D. Par conséquent, ayant déjà lu un segment initial f_1 d'un mot il est nécessaire de conserver trace en mémoire au moins de sa forme réduite αf_1 afin de pouvoir déterminer si le mot complet $f_1 f''$ appartient ou non à D. Notons qu'il existe $((2n-1)^k - 1)n(n-1)^{-1}$ mots réduits distincts de longueur au plus k ; donc pour vérifier par lecture séquentielle de gauche à droite l'appartenance à D d'un mot quelconque de longueur $\leqq 2k$, il faut disposer d'une mémoire pouvant accumuler approximativement $k(1 + \operatorname{Log}_2 n)$ bits d'informations. En outre, l'exemple donné plus haut de

$$L = x_{i_1} x_{i_2} \ldots x_{i_m} x_{-i_m} \ldots x_{-i_2} x_{-i_1}$$

montre qu'une fraction arbitrairement grande de ces informations doit être gardée en mémoire un temps arbitrairement long, à savoir depuis le début jusqu'à la fin du mot. Par contre, et c'est ce qui fait peut-être l'intérêt des langages CF, les contraintes existant entre les lettres d'un mot de D n'interviennent chacune essentiellement qu'une fois et ne se croisent pas les unes les autres. Soit par exemple un mot $f \in D$ de la forme $f_1 x_i f_2$ où f_1 ne contient ni la lettre x_i ni la lettre x_{-i}. Le mot f_2 est soumis à la condition d'avoir la forme $f_3 x_{-i} f_4$ où f_3 ne contient ni x_i ni x_{-i} et où en outre $\alpha f_3 = e$: c'est en ce sens que nous dirons que la contrainte imposée par la présence de x_i ne joue essentiellement qu'une fois, bien qu'évidemment il puisse y avoir un nombre arbitrairement grand de lettres x_i dont chacune impose l'existence d'une lettre x_{-i} qui en permet l'effacement ultérieur. De plus, dans ce même exemple, si

$$f_3 = f_5 x_j f_6$$

où f_5 ne contient aucune des lettres $x_{\pm j}$, le fait que l'on doive avoir $\alpha f_3 = e$ entraîne que f_6 soit égal à $f_7 x_{-j} f_8$ où cette fois $\alpha f_7 = e$. On a donc nécessairement :

$$f = f_1 x_i f_5 x_j f_7 x_{-j} f_8 x_{-i} f_4$$

avec

$$\alpha f_7 = \alpha f_5 f_8 = \alpha f_1 f_4 = e$$

et, dans un sens assez évident, la contrainte associée à la lettre x_i « passe par dessus » la contrainte associée à la lettre x_j sans se croiser avec elle.

Par contre, l'ensemble des mots de la forme ff, avec f un mot quelconque de F^+, ne peut pas être un langage CF (standard ou non). En effet ici les contraintes consistant en ce que chaque lettre de la seconde moitié du mot correspond à une lettre de la première moitié s'entrecroisent et il peut y avoir un nombre arbitrairement grand de telles intersections (ceci n'est évidemment pas une preuve formelle que $\{ ff : f \in F^+ \}$ n'est pas CF). De même on montrerait que les ensembles de mots :

$$L_1 = \{ x_1^n x_2^n x_3^{n'} : n, n' > 0 \}$$

et

$$L_2 = \{ x_1^n x_2^{n'} x_3^{n'} : n, n' > 0 \}$$

sont tous deux des langages CF mais qu'il n'en est pas de même de leur intersection

$$L_1 \cap L_2 = \{ x_1^n x_2^n x_3^n : n > 0 \}.$$

On pourrait illustrer, à l'aide de règles tirées de la grammaire de langues naturelles, ces deux contre-exemples et ceci montrerait, s'il le fallait, que les contraintes de type CF ne concernent qu'un horizon limité de la réalité linguistique. Une fois encore, je me bornerai à renvoyer aux travaux de Chomsky pour une discussion approfondie de ces problèmes, de la valeur explicative en linguistique des règles CF et enfin de leur rôle dans la construction de grammaires moins formelles. Cependant il convient de rappeler que dès le niveau des langages CF, les questions les plus évidentes sont, en général, *indécidables* au sens technique du terme. Ainsi, Bar-Hillel, Perles et Shamir ont montré à l'aide du contre-exemple classique de Post qu'il n'existe aucun algorithme permettant de décider pour deux homomorphismes φ_1 et φ_2 quelconques si les langages CF

$$L_1 = \{ \varphi_1 f\bar{f} : f \in F_+ \}$$

et

$$L_2 = \{ \varphi_2 f\bar{f} : f \in F_+ \}$$

ont ou non au moins un mot en commun.

D'autre part, il résulte des définitions que la mémoire d'un dispositif destiné à tester séquentiellement l'appartenance des mots à un langage CF standard donné n'est utilisée que d'une manière extrêmement restreinte qui rappelle certains aspects élémentaires de la technique de programmation connue sous le nom de « *push down storage* ». De façon schématique, (en laissant de côté la mémoire bornée qui vérifie l'appartenance à $R(X_1, V)$), si αf est le mot enregistré en mémoire après la lecture du segment initial f, et si la lettre suivante est x_i, il suffit d'ajouter celle-ci à la fin de αf quand αf se termine par x_j où $j \neq -i$ et d'effacer x_j dans le cas contraire où $j = -i$. Cette procédure ne nécessite donc en fait que la considération de l'extrémité finale du mot mis en mémoire, d'où le terme « push down ».

Il existe bien entendu d'autres types de contraintes définissant des familles de langage et les plus étudiées d'entre elles constituent ce que l'on appelle après Kleene les *événements réguliers*, qui contiennent en particulier *tous* les langages n'ayant qu'un nombre fini de mots et les langages de la forme $R(X_1, V)$ utilisés plus haut. Avec nos notations, tout événement régulier P est obtenu en prenant un sous-ensemble X_1 de X^+, un ensemble V contenant tous les mots de deux lettres $x_i x_j$ avec i négatif et j positif, ce qui donne un certain langage CF standard dont, finalement, P sera l'image par un homomorphisme φ tel que φx_i soit le mot vide pour chacun des i négatifs. Plus simplement, on peut définir P comme l'image homomorphique de $R(X'_1, V', X'_2)$ où l'homomorphisme φ n'est soumis à aucune restriction et où l'événement régulier $R(X'_1, V', X'_2)$ est l'ensemble des mots commençant par une lettre du sous-ensemble X'_1 de X, n'ayant aucun facteur de longueur deux dans V′ et se terminant par une lettre du sous-ensemble X'_2 de X. Donc, pour les événements réguliers, il est possible de remplacer toutes les contraintes résultant de l'appartenance à D par la seule contrainte supplémentaire que le mot se termine par une lettre appartenant à un ensemble distingué. De façon plus directe, Kleene avait défini un événement régulier comme un ensemble P de mots tel qu'il soit possible de reconnaître si un mot arbitraire appartient à P par lecture séquentielle de ce mot, en ne gardant en mémoire à chaque lettre qu'une quantité d'information bornée par une valeur finie (ne dépendant que de P).

Par exemple, l'ensemble de tous les mots de longueur paire qui ne contiennent pas plus de trois fois la lettre x_2 forme un événement régulier puisque l'appartenance à P peut être reconnue en ne gardant en mémoire pour chaque segment initial que la parité de la longueur de ce segment (soit 1 bit d'information) et le nombre (au plus égal à 3) de fois où x_2 est déjà apparu (soit 2 bits d'information).

Ceci équivaut à dire que P est un événement régulier si et seulement si il est l'image inverse d'un sous-ensemble de l'image de F par un homomorphisme de ce monoïde dans un monoïde *fini*. Donc, ces questions qui étaient en règle générale indécidables pour les langages CF généraux deviennent ici susceptibles d'une solution algorithmique par exhaustion de tous les cas possibles dans un certain ensemble fini. Cette propriété essentielle qui semble avoir fasciné les esprits épris de finitude, permet de rattacher très simplement aux événements réguliers les « grammaires probabilistes » n'utilisant qu'un nombre fini d'états (ce sont les « *finite state sources* » de Shannon) c'est-à-dire — en gros — les schémas markoviens finis et propose le problème de discuter la façon dont une telle grammaire peut approcher un langage CF propre.

Intuitivement, il semble bien évident qu'un langage tel que l'ensemble $M = \{ f\tilde{f} : f \in F_+)$ des *mots en miroir* qui joue un rôle central dans l'étude des langages CF ne puisse être approché que de façon tout à fait triviale par un événement régulier R. En effet, en dénotant par m_k, r_k et q_k le nombre des mots de longueur k de M, R et $Q = R \cap M$, on voit facilement que :

1) soit Q ne constitue qu'une fraction asymptotiquement nulle de M en ce sens que $\lim_{k \to \infty} q_k m_k^{-1} = 0$ (par exemple, si R est contenu dans M) ;

2) soit, au contraire, Q est une fraction asymptotiquement nulle de R, c'est-à-dire $\lim_{k \to \infty} q_k r_k^{-1} = 0$ (par exemple, si M est contenu dans R).

Plus généralement soit $A = \varphi L$ un langage CF (non nécessairement standard) satisfaisant la condition restrictive que chacun de ses mots soit l'image par φ d'un seul mot du langage standard L et cherchons à l'approximer par une séquence infinie strictement croissante d'événements réguliers $\{ R_i \}$ ($i = 1, 2, \ldots$) tous contenus dans A. Pour chaque i la valeur de cette approximation est fournie par la suite des différences $a_k - r_{i,k}$ (où a_k, $r_{i,k}$ sont, comme plus haut, le nombre des mots de longueur k de A et de R_i), soit encore, de façon condensée par le nombre $\bar{a} - \bar{r}_i$ où

$$\bar{a} = \sum_{k > 0} a_k x^k, \qquad \bar{r}_i = \sum_{k > 0} r_{i,k} x^k$$

avec

$$x = (2n + 1)^{-1}$$

ce qui revient à faire une moyenne (pondérée par les x^k) de toutes ces différences. Or, pour chaque i, $\bar{r}_i$ est une fraction rationnelle dont le dénominateur $\bar{p}_i$ peut être considéré comme une évaluation grossière de la complexité de R_i c'est-à-dire de la quantité d'information qu'il faut pouvoir garder en mémoire afin de tester l'appartenance d'un mot à R_i. Moyennant ces interprétations et le fait aisément vérifié que $\bar{a}$ est un nombre algébrique, des théorèmes classiques d'arithmétique montrent que pour chaque valeur ε de $\bar{a} - \bar{r}_i$, il n'existe qu'un nombre fini de R_i dont la « complexité » $\bar{p}_i$ soit inférieure à une certaine fonction (donnée par le théorème) de ε. C'est là le résultat cherché qui suggère comme une sorte de complémentarité entre les contraintes CF et les contraintes correspondant aux mémoires bornées.

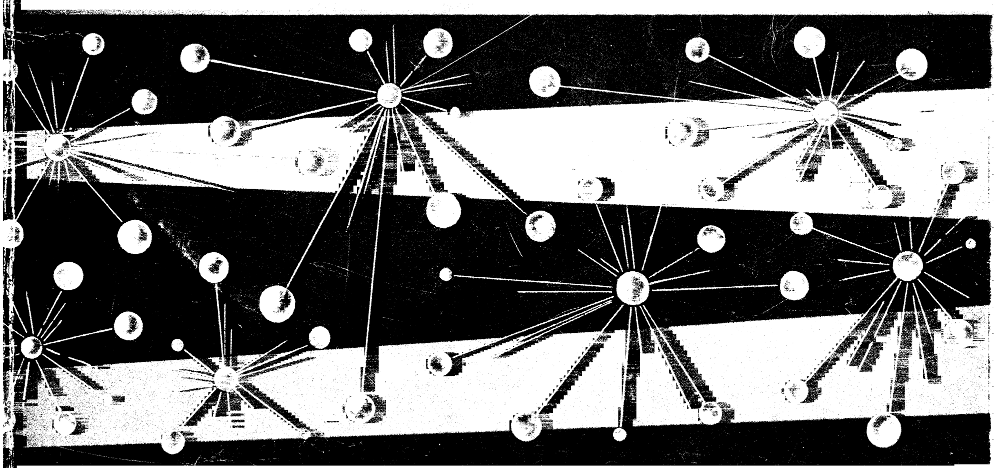
Enfin, sur le plan des faits linguistiques observables, la nécessité d'une approximation des langues naturelles plutôt par des langues CF que par ces modèles à mémoire finie que sont les événements réguliers semble avoir été admise implicitement par les techniciens de la traduction automatique eux-mêmes, puisque tous les programmes présentés jusqu'ici sont basés sur des méthodes qui sont essentiellement du type « push down » évoqué plus haut. Il me semble y avoir là une certaine ironie des choses sur laquelle je concluerai ces remarques.

Références

N. Chomsky et G. A. Miller. *L'Analyse formelle des langues naturelles*. (Collection « Mathématiques et Sciences de l'Homme »), Paris, Gauthier-Villars, La Haye, Mouton, 1968.

A. Lentin et M. Gross. *Notions sur les grammaires formelles*. Paris, Gauthier-Villars, 1967.

health aspects of chemical and biological weapons



WORLD HEALTH ORGANIZATION

1. INTRODUCTION

On 29 January 1969, the Secretary-General of the United Nations requested the Director-General of the World Health Organization to co-operate with the United Nations Group of Consultant Experts on Chemical and Bacteriological (Biological) Weapons in the preparation of a report on this subject. WHO was asked to provide such information as the Organization considered useful for the United Nations report, which was to be transmitted to the Eighteen-Nation Committee on Disarmament,[1] the Security Council and the General Assembly, if possible by 1 July 1969, as requested in Resolution 2454 A (XXIII) adopted by the General Assembly on 20 December 1968 (see Annex 7).

In order to help WHO in this task, the Director-General appointed a number of consultants. In addition, liaison was maintained with the Disarmament Affairs Division of the United Nations (which serviced the Group of Consultant Experts appointed by the Secretary-General), the Food and Agricultural Organization of the United Nations (FAO), the Stockholm International Peace Research Institute (SIPRI), and the Pugwash Organization, in order to avoid unnecessary overlap in their respective contributions.

The possible development and use of chemical and bacteriological weapons and their destructive potentialities have been matters of concern to WHO for several years. In 1967, the Twentieth World Health Assembly, on a recommendation of the WHO Executive Board, adopted a resolution (see Annex 8) welcoming Resolution 2162 (XXI) of the United Nations General Assembly and calling upon all Member States of WHO to exert every effort to implement it. The Director-General was therefore glad to meet the request to assist the United Nations in this matter, and in late May 1969 an interim report was completed and forwarded to the Secretary-General. Some of the information contained in the WHO submission was incorporated into the final report of the United Nations Group of Consultant Experts on Chemical and Bacteriological (Biological) Weapons (hereafter referred to as " the United Nations report "),[2] which was released to the public on 2 July 1969 and transmitted to the Eighteen-Nation Committee on Disarmament for discussion during the summer of 1969 before being

[1] Renamed on 26 August 1969 the " Conference of the Committee on Disarmament ".

[2] United Nations (1969) *Chemical and bacteriological (biological) weapons and the effects of their possible use. Report of the Secretary-General*, New York (United Nations Publication, Sales No : E.69.I.24).

considered at the Twenty-fourth session of the United Nations General Assembly later in the year.

The relatively short period of time available for the preparation of the WHO submission to the United Nations did not permit the health and related scientific aspects of chemical and biological warfare to be covered to the extent and in the depth merited by the importance of the subject. For this reason and in pursuance of resolution WHA22.58 (Annex 9) adopted by the Twenty-second World Health Assembly in July, 1969, a further study of the problem was undertaken with a view to expanding and revising certain sections of the interim report.

2. COMPARISON OF THE WHO AND UNITED NATIONS REPORTS AND THEIR CONCLUSIONS

The United Nations report presents a comprehensive review of the problem, and includes consideration of military aspects, plant and animal diseases, ecology, and economic and security aspects, along with implications to human health. The report was intentionally written in a style that would be easily understood by governments and by the lay non-specialist reader, and it does not attempt to present highly technical information or to provide a detailed analysis of public health considerations and medical effects.

The present WHO report, on the other hand, attempts to deal with the subject of chemical and biological warfare on a more technical level and to make quantitative estimates ; it is addressed particularly to public health and medical authorities. Thus the WHO report and the United Nations report are complementary. Both arrive at essentially the same technical conclusions, although inevitably there are some differences with respect to the choice of emphasis and the assessment of possible effects on public health, which reflect the differing approaches and technical orientations of the groups that prepared the WHO and United Nations reports. It is hoped, therefore, that the present document will provide the Member States of WHO with the technical information that will enable them to appreciate more fully the public health implications of the possible use of chemical and biological weapons.

The following main conclusions emerge from the WHO analysis :

1. *Chemical and biological weapons pose a special threat to civilians. This is because of the often indiscriminate nature of such weapons, and because the high concentrations in which they would be used in military operations could lead to significant unintended involvement of the civilian population within the target area and for considerable distances downwind.*

2. *The large-scale or, with some agents, even limited use of chemical and biological weapons could cause illness to a degree that would overwhelm existing health resources and facilities.*

3. *Large-scale use of chemical and biological weapons could also cause lasting changes of an unpredictable nature in man's environment.*

4. *The possible effects of chemical and biological weapons are subject to a high degree of uncertainty and unpredictability, owing to the involvement of complex and extremely variable meteorological, physiological, epidemiological, ecological, and other factors.*

5. *Although advanced weapons systems would be required for the employment of chemical and biological agents on a militarily significant scale against large civilian targets, isolated and sabotage attacks not requiring highly sophisticated weapons systems could be effective against such targets in certain circumstances with some of these agents.*

These conclusions are in harmony with the conclusions of the United Nations Group of Consultant Experts on Chemical and Bacteriological (Biological) Weapons and with the hope for further action to deal with the threat posed by the existence of these weapons, as expressed by the Secretary-General, U Thant, in the foreword to the United Nations' report.

3. AIM AND SCOPE OF THE WHO REPORT

The present report attempts to analyse the health effects of the possible use of chemical and biological weapons on civilian population groups at different levels of social and economic development, and the resulting implications for WHO and its Member States. The assessment is confined to civilian populations, and no attempt is made to consider the purely military aspects of the problem, except insofar as they may relate to civilian populations as possible targets for attack. The military aspects of chemical and biological warfare are considered in the United Nations report and in a report being prepared by SIPRI. The report also makes qualitative and quantitative estimates of the health effects of selected chemical and biological agents employed under specified hypothetical conditions.

4. WORKING DEFINITIONS OF CHEMICAL AND BIOLOGICAL WEAPONS FOR THE PURPOSES OF THIS REPORT

Chemical agents of warfare include all substances employed for their toxic effects on man, animals, or plants.[1]

Biological agents include those that depend for their effects on multiplication within the target organism, and are intended for use in war to cause disease or death in man, animals or plants.[2]

A lethal agent is one intended to cause death when man is exposed to concentrations well within the capability of delivery for military purposes.[3]

An incapacitating agent is one intended to cause temporary disease or to induce temporary mental or physical disability, the duration of which greatly exceeds the period of exposure.[4]

A harassing agent (or short term incapacitant) is one capable of causing a rapid disablement that lasts for little longer than the period of exposure.[4]

Casualties are deaths or disabilities.

5. SELECTION OF CHEMICAL AND BIOLOGICAL AGENTS AS MODELS FOR QUALITATIVE AND QUANTITATIVE ASSESSMENTS

There are many chemical and biological agents that are potentially suitable for use in war. A selection of some of the more likely candidates for possible use as lethal, incapacitating and harassing agents has been made

---

[1] This definition is intended to exclude chemicals now employed in warfare such as high explosives, smoke, and incendiary substances (e.g., napalm, magnesium, and white phosphorus) that exert their primary effects through physical force, fire, air-deprivation or reduced visibility.

[2] This definition therefore excludes toxins elaborated by some microbes (e.g., botulinal toxin and staphylococcal enterotoxin) when they are preformed outside the target organism. In some discussions of chemical and biological weapons, such toxins are classified as biological agents because the technology of their production resembles that of biological agents rather than that of chemical agents.

[3] In lower doses, such agents can cause severe and sustained disability and certain of them may act predominantly in this way when employed in combat.

[4] No sharp line of demarcation can be drawn between lethal and incapacitating agents used in chemical and biological warfare, because incapacitating agents can be lethal or permanently disabling under certain circumstances (e.g., in the presence of malnutrition or pre-existing disease ; in infants or the aged ; or when there is exposure to unusually high doses, as in enclosed spaces or in close proximity to functioning chemical or biological weapons). For similar reasons, no sharp demarcation line can be drawn between harassing agents and other anti-personnel chemical agents ; furthermore, harassing agents may be used in war in conjunction with high-explosive, fragmentation or other weapons to increase the lethal effectiveness of the latter—as distinct from their employment in riot control in order to reduce injuries and to save lives.

for the purposes of this report, based on published information as well as on theoretical considerations.

The chemical and biological agents described in Annexes 1 and 2 have been selected to illustrate possible varieties of usage and effects : aerosol exposure ; sabotage of communal water supply ; quick or delayed action ; essentially non-spreading and spreading types of infective disease ; vector introduction. Qualitative descriptions are given for all of them, and a number of representative types have been selected for quantitative assessments of the effects of their possible use.

6. BASES OF THE ESTIMATES OF CASUALTIES

The actions and toxicities of the chemical warfare agents considered are reasonably well known. For many, there are generally agreed estimates of lethal doses, although it should be kept in mind that often these are based on assumptions concerning the relative susceptibility of different animal species, and that susceptibility varies between individuals. The clinical symptoms, the prognosis, and the general methods of treatment and prevention, if any, can be assessed from commonly available information.

The agents chosen for discussion as possible biological weapons are those whose clinical effects are well known ; for some of them data on approximate infective doses by inhalation or ingestion are also available, either from published information about laboratory accidents or from studies in human volunteers. However, little is known about the susceptibility of man to artificial aerosols of infective agents and about the consequences of exposure to very high doses. This report is confined to a discussion of known agents with a wide range of infectivity and lethality. Other agents could conceivably be developed. In spite of these uncertainties, it has been considered useful to attempt to predict the range of immediate effects of the agents dealt with in this report.

It is considered unlikely that the general conclusions reached in this study would have been greatly modified if it had been possible to consult classified information concerning chemical and biological agents, although the assessments concerning the feasibility of their use in particular circumstances might have been made more realistic.

The estimates of casualties recorded in Tables 8 and 10, Annex 3, are based on the number of individuals within a segment of a particular population group exposed to a given chemical or biological agent. The assumptions made and the variables taken into account are stated in Annex 3. These assumptions deal with delivery, dissemination, and persistence of the agents ; meteorological conditions ; effective concentrations and doses ; human infective doses, attack and case fatality rates ; and chemotherapy.

In the detailed analysis of hypothetical attack with tularaemia, pneumonic plague, and VX, the effects of the availability and use of health facilities on the outcome are considered (Annex 4). Annex 5 describes the possible effects of sabotage of a communal water supply with botulinal toxin, the typhoid bacillus and LSD.

The estimates of casualties that are given in Annex 3 are applicable to a surprise attack on an urban area that possesses no specific protection against attack with chemical and biological agents, and in which the buildings are freely ventilated. It will be seen that even where the target area is not very densely populated the effects of attack by a single aircraft could create health problems of unprecedented magnitude.

It is certain that the scale and effectiveness of an attack (and also the cost and feasibility) would depend upon military technology. To discuss this aspect in any depth is outside the terms of reference of this report. A limited attack on a civilian population using a pattern favouring the effectiveness of the weapons employed has therefore been chosen for consideration. In this way, it is hoped to illustrate the scale of the danger that could arise from the use of such weapons. Similar results could be produced by attacks on a larger scale if the circumstances were unfavourable for the pattern chosen.

7. LONG-TERM EFFECTS

A few general considerations regarding the possible long-term effects of chemical and biological agents should be noted. First of all, insufficient knowledge is available to allow reliable predictions to be made. In many cases, not much more can be done than to outline the various possibilities needing further study. Beyond that, there are problems of evaluation that, while still having a considerable technical component, also involve value judgements that are clearly beyond the scope of this report. For example, in comparison with the direct effects of a lethal chemical or biological attack, a limited risk of long-term harmful effects to health may seem relatively unimportant. On the other hand, the long-term effects of the military use of agents that are not directly lethal may be considered more important than their immediate effects. In the latter connexion it should be kept in mind that non-military experience with disease-causing organisms and chemicals present in the environment may not be a good guide to the effects of those same agents under the quite different conditions and in the generally higher doses involved in their military employment.

Possible long-term health effects of chemical and biological warfare include (1) chronic illness caused by exposure to chemical and biological agents (see specific descriptions in Annexes 1 and 2); (2) delayed effects in persons directly exposed to chemical and biological agents; (3) creation

14

of new foci of infective disease ; and (4) effects mediated by ecological changes.

Delayed effects of direct human exposure

There is wide concern at present regarding the possibility that exposure to both infective and chemical agents already present in the human environment may cause harmful effects of a delayed nature. The effects of greatest concern are :

(*a*) *Carcinogenesis*. Both viral and chemical agents have been strongly implicated in the causation of cancer in man. Whether infection by any of the viruses contemplated for possible military employment can be carcinogenic in man is not at present known. A limited amount of information is available on the ability of certain classes of chemicals to induce cancer, mainly in experimental animals. For example, many alkylating agents have been found to be carcinogenic. Some compounds of military interest, such as mustard, CS, and others, are alkylating agents. As discussed in Annex 1, there is evidence suggesting a significant increase in respiratory tract cancer among veterans exposed to mustard gas in the First World War, and a large increase in such cancers has been reported among workers engaged in the manufacture of mustard gas in the Second World War.

(*b*) *Teratogenesis*. Certain chemicals and infective agents can cause severe damage to the developing human foetus. Thalidomide and the rubella virus are particularly well known teratogens. It is not known whether any agents likely to be used in chemical or biological warfare would have teratogenic effects at the doses likely to be received by pregnant women in civilian populations under direct attack with or unintentionally exposed to such agents during wartime. In this regard, it may be noted that the use of 2,4,5-trichlorophenoxyacetic acid, an anti-plant chemical that has been extensively used for both military and non-military purposes, has recently been restricted by the United States Government because experiments have shown that relatively high oral doses of this compound are teratogenic in mice and rats.

(*c*) *Mutagenesis*. Until recent years little attention has been given to the possibility that infectious diseases or chemicals in the environment might cause detrimental alterations in the human genome. Several chemicals are known to induce such changes in experimental organisms and in cultured human cells. Infection with certain viruses causes extensive chromosome breakage in man, but it is not known whether any heritable effect results. At least in the case of rubella it can be said that genetic damage is not massive, although the induction of a lower but nevertheless significant frequency of mutations cannot be excluded.

New foci of infective disease

As discussed in Annex 2, biological warfare would entail a risk that new foci of infective disease might be established, either in human populations or in lower animals, including vector arthropods. This possibility has been discussed in the United Nations report:

" A bacteriological (biological) attack might lead to the creation of multiple and densely distributed foci of infection from which, if ecological conditions were favourable, natural foci might develop in regions where they had previously never existed or in areas from which they had been eliminated by effective public health measures."

Effects mediated by ecological change

The possibility of the direct establishment of new foci of disease has been referred to immediately above. New foci might also be established as the result of ecological changes following the use of biological agents infective for man and animals. This possibility has also been discussed in the United Nations report:

" ... the large scale use of bacteriological (biological) weapons might reduce populations of susceptible wild species below the level at which they could continue to exist. The elimination of a species or group of species from an area would create in the ecological community an empty niche which might seriously disturb its equilibrium or which might be filled by another species more dangerous to man because it carried a zoonosis infection acquired either naturally or as a result of the attack. This would result in the establishment of a new natural focus of disease."

As for anti-personnel chemical warfare, it can at least be said that the massive dissemination of several chemical agents during the First World War has not apparently caused any major long-term ecological damage in Europe.

However, new foci of human disease may also be produced as a result of the use of anti-plant agents. Extensive damage to the flora over large areas may create conditions favouring the establishment of new vectors or reservoirs of disease infective to man. One example of the way in which damage to plant life can create new health hazards is cited in the United Nations report:

" When a forest in a state of ecological equilibrium is destroyed by cutting, a secondary forest regenerates, which contains fewer species of plants and animals than were there originally, but larger numbers of those species which survive. If secondary forest is replaced by grassland, these changes are even more marked. If one or more of the animal species which increases in numbers is the host of an infection dangerous to man (a zoonosis) then the risk of human infection is greatly increased. This is exemplified by the history of scrub typhus in south-east Asia, where the species of rat which maintains the infection and the vector mite are much more numerous in secondary forest, and even more so in grassland, so increasing the risk of the disease being transmitted to people as forest is cleared."

Finally a profound long-term adverse effect on human health could result from any major reduction in the quality or quantity of the food supply. This could occur directly from the use of anti-crop agents or indirectly through ecological changes that might result from chemical or biological warfare.

8. SUMMARY

A. *Qualitative considerations*

Chemical and biological agents that might be considered for possible use in warfare are described in Annexes 1 and 2, and pertinent assumptions and other background factors are considered earlier in this text and in Annexes 3, 4, and 5. Also of importance, although more difficult to assess, are the possible long-term effects referred to in section 7, and the psychosocial consequences associated with the problem of chemical and biological weapons (see Annex 6).

The rapid action of the lethal *chemical agents* (see Annex 1) would preclude any large reduction of mortality by specific treatment. Possible protection by gas masks or shelters requires a highly disciplined and prepared population, a condition that is not fulfilled in most countries today, and it would pose serious economic and psychosocial problems if such a defence programme were to be implemented.

The outstanding characteristics of *biological weapons* (see Annex 2) for potential use in warfare are the following:

(*a*) The large variety of biological agents and the possible combinations available for such purposes.

(*b*) The possibilities for manipulating currently circulating strains of micro-organisms for warfare purposes, by producing antigenically modified or antibiotic-resistant types (tularaemia, plague, anthrax, influenza) that would by-pass available prophylactic or therapeutic procedures.[1]

(*c*) The unpredictability of the direct effects. A biological attack intended to be highly lethal might prove relatively ineffective, whereas an attack intended to be merely incapacitating might kill an unexpectedly

[1] Mass immunizations would be of doubtful protective value because of the multiplicity of agents and strains that might be employed, quite apart from the adverse immunological side-reactions to be expected.

large proportion of the target population. Also, certain agents (anthrax, coccidioidomycosis) could persist for long periods in a resistant spore form, which could be spread over very large distances by wind carriage in the course of time.

(*d*) The unpredictability of secondary effects such as the likelihood of contagion and the danger that epidemics might be initiated. There is the additional danger that epidemics might occur unintentionally through escape of virulent strains being purposely sought in laboratories.

(*e*) Although biological agents themselves are easy to produce, complex production and delivery systems are needed if even minimal reliance is to be placed on the outcome of an attack, except perhaps where the intention is simply to produce social disruption by a limited sabotage effort (e.g., the introduction of smallpox).

Of the above characteristics, (*a*) and (*b*) would favour the attacker, whereas (*c*) and (*d*) would reduce the value of biological weapons from a military point of view.

B. *Quantitative estimates* (Tables 8, 9 and 10, Annex 3)

1. Assessments have been made of the primary effects of possible small-scale airborne attacks on cities of 0.5-5 million population in industrially developed and developing countries. The postulated mode of attack consisted of one or a few bombers dispersing specific chemical or biological agents along a 2-km line perpendicular to the direction of the wind. On the basis of the particular assumptions employed, the following conclusions have been reached :

(*a*) Of the known chemical warfare agents, only the nerve gases, and possibly botulinal toxin, have a casualty-producing potential comparable to that of biological agents.

(*b*) Under atmospheric conditions favourable to the attacker, an efficiently executed attack on a city with 4 tons of sarin (requiring some 15-20 tons of weapons) could cause tens of thousands of deaths in an area of about 2 km². Even in unfavourable conditions there could be thousands of deaths. If 4 tons of VX were used in such an attack, the casualties would not be appreciably greater in unfavourable meteorological conditions, but in favourable conditions this small attack would affect an area of about 6 km² and could cause anywhere between 50 000 and 180 000 deaths.

(*c*) If a suitably stabilized botulinal toxin or a fine aerosol of VX (particles of 5μ diameter) were developed and 4 tons were employed, several hundreds of thousands of deaths could result because of the greater coverage possible with such agents—12 km² for botulinal toxin and 40 km² for monodispersed VX aerosol. A larger total weight of weapons, perhaps

2-3 times that needed for the agents in (*b*) above, would have to be used to deliver these forms of botulinal toxin and VX.

(*d*) If a biological agent such as anthrax were used, an attack on a city by even a single bomber disseminating 50 kg of the dried agent in a suitable aerosol form would affect an area far in excess of 20 km², with tens to hundreds of thousands of deaths. A similar attack with any one of a number of other more labile biological agents could affect from 1 km² to more than 20 km², depending upon agent used, with tens to hundreds of thousands of casualties and many thousands of deaths.

2. Limited sabotage of a communal water supply with the typhoid fever bacillus, LSD, or a stable botulinal toxin, could cause considerable disruption and deaths in a large city (see Annex 5), affecting tens of thousands of people.

3. Sabotage-induced or open attacks, causing the secondary spread of epidemics of yellow fever, pneumonic plague, smallpox or influenza, might under certain conditions ultimately result in many millions of illnesses and deaths (see Annex 2).

4. The numbers of potential casualties and deaths recorded in this report represent the possibilities arising out of a very small and limited attack already well within the capabilities of a number of nations, with the possibility that an ever-increasing number of countries will acquire similar capabilities. With technologically advanced weapons and a larger scale of attack, achievable without too much difficulty by militarily advanced powers, the magnitude of destructiveness attendant upon the use of chemical and biological weapons would be considerably increased.

9. IMPLICATIONS FOR THE WORLD HEALTH ORGANIZATION AND ITS MEMBER STATES

According to Art. 2 (*c*) of WHO's Constitution, WHO shall "... furnish appropriate technical assistance and, in emergencies, necessary aid upon the request or acceptance of Governments". The use of chemical and biological weapons would unquestionably result in extensive health and medical emergencies, including mass illnesses, deaths and epidemics, that WHO might be called upon to help overcome. An attempt to assess the magnitude of public health problems with respect to a minimal attack with selected examples of agents (tularaemia, plague, VX) is contained in Annex 4. This limited assessment, supported by analyses made in other parts of this report and in the United Nations report, reveals the very large and essentially wasteful effort that would be involved in undertaking elaborate measures for defence against specific agents. Also, as pointed out in Annex 6, such measures could well add credibility to projected fears of annihilation in other countries. The resultant reciprocal fears between nations might contribute in turn to a proliferation of chemical and biological weapons and an accelerated arms race, resulting in vastly increased danger of accidental or deliberate release of chemical and biological agents.

Certain measures could, however, be taken within the framework of existing needs and resources that would redound to the benefit of health and preventive medical activities currently underway, and would not give rise to fears of this kind. These measures include the improvement of rapid detection and diagnostic facilities for air pollution and for communicable diseases, which would obviously be of value for health and laboratory services in general; improved medical management for natural disasters, including decontamination procedures; and the wider use of safety features in buildings (ventilation filters) and for communal water supply systems (see Annex 5). While such measures might act as a partial deterrent to irresponsible groups and might significantly reduce casualties from a very small attack or from the spreading effects of an attack on a neighbouring country, they cannot be relied upon to afford major protection to a country subjected to a determined attack.

As long as chemical and biological research directed specifically to military use is continued, it will be considered necessary by some countries to continue research towards detection of and protection against such agents. This research could in itself point to agents more destructive than those now existing. In view of the power of existing agents in conditions favourable to their use and the possibility of developing new and even more dangerous weapons, it is imperative to find ways of abolishing any presumed need for this militarily orientated research as soon as possible.

It is therefore clear that in the last analysis the best interests of all Member States and mankind in general will be served by the rapid implementation of the resolutions on chemical and biological warfare adopted by the United Nations General Assembly and the World Health Assembly (Annexes 7 and 8), and by any additional steps that would help ensure outlawing the development and use in all circumstances of chemical and biological agents as weapons of war.

Finally, there is the possibility that WHO might be called upon by the United Nations to help deal with allegations of use of chemical and biological weapons between nations and to assist in the limitation of chemical and biological weapons, and disarmament. The technical resources of WHO [1]

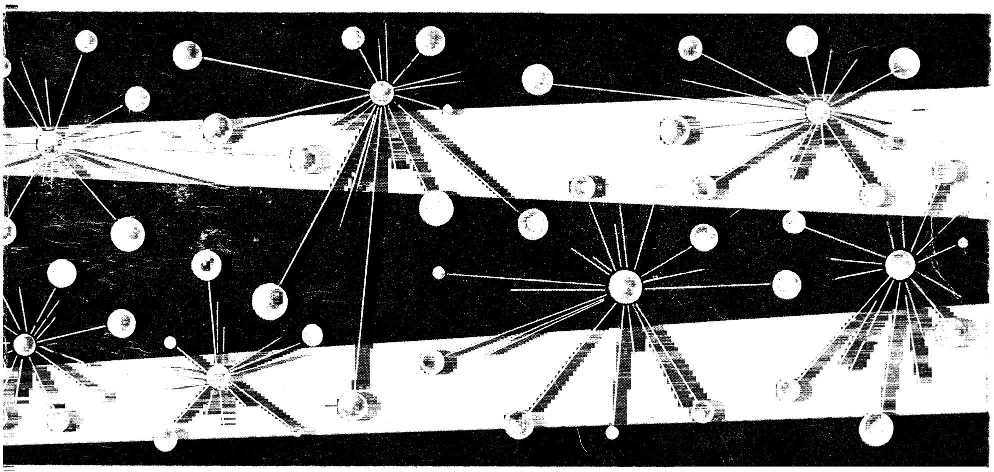
[1] An example of such technical resources is the collection of epidemiological information on communicable diseases that has been made by WHO for many years, through its serum banks and its surveillance programmes involving specific diseases. This information provides an invaluable background and potential for determining changes in communicable disease patterns, as well as for obtaining knowledge of diseases already existing in a community. Apart from its general epidemiological value, expansion in the accumulation of such data could be very useful for investigating any possible future allegations of use of biological weapons.

could contribute greatly to the resolution of many of the difficulties that are associated with these problems and are now being discussed within the framework of the the United Nations.

ACKNOWLEDGEMENTS

The Group expresses its thanks and appreciation to the many staff members of WHO who gave valuable help during the preparation of this report. Acknowledgement is also made to the Stockholm International Peace Research Institute, the Pugwash Organization, and the World Meteorological Organization for their collaboration and for supplying much useful information.

santé publique et armes chimiques et biologiques



ORGANISATION MONDIALE DE LA SANTÉ

CONSULTANTS

Professeur O. V. Baroyan, Directeur de l'Institut Gamaleja d'Epidémiologie et de Microbiologie, Académie des Sciences médicales de l'URSS, Moscou, URSS

Dr D. Blaskovič, Directeur de l'Institut de Virologie, Académie tchécoslovaque des Sciences, Bratislava, Tchécoslovaquie

Dr K. Evang, Directeur général des Services de Santé, Oslo, Norvège

Professeur R. B. Fischer, Department of Biochemistry, University of Edinburgh, Ecosse

Professeur L. Huisman, Département de Génie civil, Institut universitaire de Technologie, Delft, Pays-Bas

Dr J. H. Humphrey, Head, Division of Immunology, National Institute for Medical Research, Londres, Angleterre

Professeur N. K. Jerne, Directeur de l'Institut Paul Ehrlich, Francfort-sur-le-Main, République fédérale d'Allemagne

Professeur J. Lederberg, Department of Genetics, Standford University School of Medicine, Standford, Cal., Etats-Unis d'Amérique

Professeur A. M. Lwoff, Directeur de l'Institut de Recherche scientifique sur le Cancer, Villejuif, France

Professeur O. Maaloe, Institut de Microbiologie, Université de Copenhague, Danemark

Professeur I. Malek, Institut de Microbiologie, Académie tchécoslovaque des Sciences, Prague, Tchécoslovaquie

Professeur M. Meselson, The Biological Laboratories, Harvard University, Cambridge, Mass., Etats-Unis d'Amérique

Dr F. Pasquill, Meteorological Office, Bracknell, Berks., Angleterre

M. H. P. Perry Robinson, Stockholm International Peace Research Institute (SIPRI), Suède

Professeur M. P. Schutzenberger, Faculté des Sciences, Université de Paris, France

Professeur V. W. Sidel, Department of Community Health, Albert Einstein College of Medicine ; Chief, Division of Social Medicine, Montefiore Hospital and Medical Center, New York, Etats-Unis d'Amérique

Dr Berhane Teoume-Lessane, Co-Directeur, Laboratoire et Institut de Recherche impérial central, Addis-Abéba, Ethiopie

M. F. W. J. van Haaren, Chef des Laboratoires du Service municipal des Eaux, Amsterdam, Pays-Bas

Secrétaire

Dr M. Kaplan, Assistant spécial pour les questions scientifiques, Bureau du Directeur général, OMS, Genève, Suisse

8

1. INTRODUCTION

Le 29 janvier 1969, le Secrétaire général de l'Organisation des Nations Unies a prié le Directeur général de l'Organisation mondiale de la Santé de collaborer avec le Groupe d'experts consultants des Nations Unies sur les armes chimiques et bactériologiques (biologiques) à la préparation d'un rapport sur la question. L'OMS était invitée à fournir les renseignements qu'elle jugerait utiles pour l'établissement du rapport des Nations Unies, qui serait communiqué au Comité des dix-huit puissances sur le désarmement,[1] au Conseil de Sécurité et à l'Assemblée générale, si possible avant le 1er juillet 1969, comme l'Assemblée générale l'avait demandé dans sa résolution 2454 A (XXIII) du 20 décembre 1968 (voir annexe 7).

Afin d'aider l'OMS à accomplir la tâche dont elle était chargée, le Directeur général a engagé un certain nombre de consultants. En outre, l'OMS s'est mise en rapport avec la Division des affaires concernant le désarmement de l'ONU (qui a assuré le secrétariat du Groupe d'experts consultants nommés par le Secrétaire général), l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (FAO), le Stockholm International Peace Research Institute (SIPRI), et la Conférence de Pugwash afin d'éviter que les contributions respectives des diverses organisations ne fassent double emploi.

La fabrication et l'utilisation éventuelle d'armes chimiques et bactériologiques et les destructions que ces armes pourraient causer sont depuis plusieurs années une source de préoccupation pour l'OMS. En 1967, la Vingtième Assemblée mondiale de la Santé a adopté, sur la recommandation du Conseil exécutif, une résolution (voir annexe 8) dans laquelle elle accueillait avec satisfaction la résolution 2162 (XXI) de l'Assemblée générale des Nations Unies et invitait tous ses Membres à déployer le maximum d'efforts pour assurer son application. Le Directeur général a donc été heureux d'accéder à la demande de l'ONU et, vers la fin de mai 1969, il a fait tenir au Secrétaire général un rapport intérimaire sur la question. Une partie des renseignements figurant dans ce rapport a été incorporée dans le rapport final du Groupe d'experts consultants des Nations Unies sur les armes chimiques et bactériologiques (biologiques), (ci-après dénommé « Rapport des Nations Unies »),[2] qui a été rendu public le 2 juillet 1969

---

[1] Dénommé « Conférence du Comité sur le désarmement » depuis le 26 août 1969.

[2] Organisation des Nations Unies (1969) *Les armes chimiques et bactériologiques (biologiques) et les effets de leur utilisation éventuelle. Rapport du Secrétaire général,* New York (Publication des Nations Unies, N° de vente : F.69.I.24).

9

et communiqué pour examen au Comité des dix-huit puissances sur le désarmement au cours de l'été de 1969, avant d'être étudié à la vingt-quatrième session de l'Assemblée générale des Nations Unies tenue plus tard dans l'année.

Ne disposant que d'un temps relativement court pour établir le rapport qu'elle devait adresser aux Nations Unies, l'OMS n'a pas été en mesure, avant mai, d'étudier de façon aussi étendue et approfondie qu'ils le méritaient les problèmes que pose la guerre chimique et biologique du point de vue de la santé et des sciences biomédicales. Pour cette raison, et en application de la résolution WHA22.58 (annexe 9) adoptée par la Vingt-Deuxième Assemblée mondiale de la Santé en juillet 1969, une étude complémentaire a été entreprise afin de revoir et d'augmenter diverses sections du rapport intérimaire.

2. COMPARAISON DES RAPPORTS DE L'OMS ET DES NATIONS UNIES ET DE LEURS CONCLUSIONS

Le rapport des Nations Unies donne une vue d'ensemble du problème : aspects militaires, maladies des plantes et des animaux, écologie, considérations économiques et de sécurité, et incidences pour la santé humaine. Il a été intentionnellement rédigé dans un style général accessible à des lecteurs de formation très diverse ; aussi n'y trouve-t-on pas de développements très techniques, ni d'analyse détaillée des problèmes que pose la guerre chimique et bactériologique du point de vue de la santé publique et de la médecine.

En revanche, le présent rapport de l'OMS traite la question d'un point de vue plus spécialisé et présente des estimations quantitatives ; il s'adresse plus particulièrement aux autorités de la santé publique et aux médecins. Le rapport de l'OMS et celui des Nations Unies sont donc complémentaires. Sur le plan technique, les conclusions auxquelles ils aboutissent sont essentiellement les mêmes, encore qu'il y ait inévitablement entre eux, quant à l'importance accordée à tel ou tel aspect particulier, ainsi que dans l'appréciation des effets éventuels des armes chimiques et biologiques sur la santé publique, certaines divergences qui traduisent les différences d'optique et d'orientation technique des deux groupes qui ont rédigé les rapports. Le présent document devrait donc apporter aux Etats Membres de l'OMS les renseignements techniques précis dont ils ont besoin pour mieux apprécier les effets qu'aurait sur la santé publique l'utilisation éventuelle d'armes chimiques et biologiques.

Les principales conclusions qui se dégagent de l'analyse faite par l'OMS sont les suivantes :

1. *Les armes chimiques et biologiques sont une menace réelle pour les populations civiles. En effet, elles ne se prêtent généralement pas à un emploi sélectif. En outre, aux concentrations très élevées auxquelles elles seront probablement utilisées dans les opérations militaires, elles pourraient causer indirectement des ravages étendus parmi la population civile, non seulement dans la zone visée, mais aussi jusqu'à des distances considérables dans les secteurs sous le vent.*

2. *L'emploi massif — et même, dans le cas de certains agents, l'emploi restreint — d'armes chimiques et biologiques pourrait provoquer une morbidité capable de déborder les services de santé existants.*

3. *L'emploi massif d'armes chimiques et biologiques pourrait également causer dans le milieu naturel de l'homme des altérations durables, tout à fait imprévisibles.*

4. *Il est extrêmement difficile de déterminer et de prévoir à l'avance les effets possibles des armes chimiques et biologiques, car ils dépendent de l'interaction de facteurs complexes et extrêmement variables, d'ordre météorologique, physiologique, épidémiologique, écologique, etc.*

5. *Des systèmes d'armes très perfectionnés seraient nécessaires pour que l'emploi d'agents chimiques et biologiques contre de grands objectifs civils présente un réel intérêt militaire, mais, dans certaines circonstances et avec certains agents, des opérations isolées ou de sabotage menées avec des moyens plus simples pourraient être efficaces contre de tels objectifs.*

Ces conclusions sont en harmonie avec celles du Groupe d'experts consultants des Nations Unies sur les armes chimiques et bactériologiques (biologiques) ainsi qu'avec le vœu exprimé dans la préface du rapport des Nations Unies par le Secrétaire général U Thant, qui espère que d'autres décisions interviendront pour faire face à la menace que constitue l'existence de ces armes.

3. BUT ET CONTENU DU RAPPORT DE L'OMS

Les auteurs du présent rapport se sont efforcés de déterminer quels seraient les effets de l'utilisation éventuelle d'armes chimiques et biologiques sur la santé des populations civiles de pays ayant atteint différents niveaux de développement économique et social, ainsi que les conséquences qui pourraient en résulter pour l'OMS et ses Etats Membres. Cette analyse laisse de côté tous les aspects purement militaires du problème, sauf l'éventualité d'une attaque dirigée intentionnellement contre des populations civiles. Les aspects militaires de la guerre chimique et biologique sont étudiés dans le rapport des Nations Unies et dans un rapport que le SIPRI est en train de rédiger. On trouvera en outre dans les annexes des estimations qualitatives et quantitatives concernant les effets de certains agents chimiques et biologiques employés dans des conditions hypothétiques bien définies.

4. DÉFINITIONS PRATIQUES DES ARMES CHIMIQUES ET BIOLOGIQUES AUX FINS DU PRÉSENT RAPPORT

Les agents chimiques de guerre comprennent toutes les substances employées en raison de leurs effets toxiques sur l'homme, les animaux et les plantes.[1]

Les agents biologiques sont ceux dont les effets sont fonction de leur aptitude à se multiplier dans l'organisme attaqué, et qui sont destinés à être utilisés en cas de guerre pour provoquer la mort ou la maladie chez l'homme, les animaux ou les plantes.[2]

Un agent létal est conçu pour provoquer la mort lorsque l'homme y est exposé à des concentrations aisément réalisables dans les opérations militaires.[3]

Les agents incapacitants sont destinés à provoquer temporairement soit la maladie soit une incapacité mentale ou physique dont la durée dépasse de beaucoup la période d'exposition.[4]

---

[1] Cette définition exclut les substances chimiques actuellement employées à des fins militaires — explosifs, fumigènes et substances incendiaires (napalm, magnésium et phosphore blanc, par exemple) — dont l'action principale est de provoquer des destructions physiques, le feu, une privation temporaire d'air ou une réduction de la visibilité.

[2] Cette définition exclut par conséquent les toxines élaborées par certains microbes (toxine botulique et entérotoxine staphylococcique, par exemple) qui sont formées à l'extérieur et non à l'intérieur de l'organisme attaqué. Dans certaines études sur les armes chimiques et biologiques, ces toxines sont classées comme agents biologiques du seul fait que la technologie de leur production s'apparente à celle des agents biologiques et non à celles des agents chimiques.

[3] A plus faibles doses, ces agents peuvent causer une incapacité étendue et durable ; c'est principalement une action de ce genre qui est attendue de certains d'entre eux dans les engagements militaires.

[4] Il est impossible de tracer une ligne de démarcation précise entre agents létaux et agents incapacitants car ces derniers peuvent provoquer la mort ou l'incapacité permanente dans certaines circonstances (par exemple chez les personnes affaiblies par la malnutrition ou une maladie préexistante, chez les très jeunes enfants et les personnes âgées ou en cas d'exposition aux concentrations exceptionnellement élevées qui se rencontreront dans des espaces clos ou à proximité immédiate d'armes chimiques et biologiques en action). Pour des raisons analogues, on ne peut pas tracer de ligne de démarcation précise entre les agents neutralisants et les autres agents chimiques antipersonnels ; dans la guerre, les agents neutralisants peuvent en outre être employés en association avec d'autres armes (explosives, à fragmentation, etc.) pour accroître leur efficacité létale ; dans la répression des émeutes, ils sont utilisés au contraire pour éviter qu'il y ait des affrontements violents et des victimes.

Les agents neutralisants (ou incapacitants de courte durée) provoquent rapidement une incapacité qui ne se prolonge guère au-delà de la période d'exposition.

Par *victimes*, on entend les morts et les personnes réduites à l'incapacité.

5. SÉLECTION D'AGENTS CHIMIQUES ET BIOLOGIQUES PROPRES À SERVIR DE MODÈLES POUR DES ÉVALUATIONS QUALITATIVES ET QUANTITATIVES

De nombreux agents chimiques et biologiques sont potentiellement utilisables pour la guerre. Aux fins du présent rapport, on a fait, en se fondant sur les données déjà publiées et sur diverses considérations théoriques, un choix des agents qui seraient le plus probablement utilisés pour leurs effets létaux, incapacitants ou neutralisants.

Les agents chimiques et biologiques décrits dans les annexes 1 et 2 sont ceux qu'il a semblé bon de retenir pour illustrer différents usages et effets possibles des armes chimiques et biologiques : exposition à des aérosols ; contamination de l'eau par sabotage ; action immédiate ou à retardement ; maladies infectieuses pouvant ou non se propager largement ; introduction de vecteurs de maladies. En dehors de descriptions qualitatives de tous ces agents, on trouvera dans la suite une évaluation quantitative des effets de l'utilisation éventuelle de certains types d'agents représentatifs.

6. MODE D'ESTIMATION DU NOMBRE DES VICTIMES

L'action et la toxicité des agents de guerre chimique considérés ici sont assez bien connus. Pour beaucoup d'entre eux, il existe des estimations des doses létales qui sont généralement admises, mais il ne faut pas perdre de vue que ces estimations reposent souvent sur des hypothèses concernant la sensibilité relative de différentes espèces animales, ni que la sensibilité varie d'un individu à l'autre. Pour la description des symptômes cliniques, le pronostic et le choix des méthodes générales de traitement et de prévention, s'il en existe, on peut se référer à la littérature médicale courante.

Les agents dont l'utilisation éventuelle comme armes biologiques est étudiée dans les pages qui suivent ont des effets cliniques bien connus ; pour certains d'entre eux, on sait aussi quelles sont à peu près les doses infectantes par inhalation ou ingestion d'après des communications publiées à la suite d'accidents de laboratoire ou d'études sur des volontaires humains. En revanche, on ne possède guère de données sur la sensibilité de l'homme aux aérosols artificiels d'agents infectieux, ni sur les conséquences de l'exposition à des doses très fortes. Le présent rapport ne traite que d'agents connus, dont l'action infectante ou létale varie dans des limites très larges, mais il n'est pas inconcevable que d'autres agents puissent être mis au point. En dépit des incertitudes qui en découlent, il a paru utile d'essayer de déterminer quelle pourrait être la gamme des effets immédiats des agents considérés.

Les conclusions générales de la présente étude n'auraient probablement guère été différentes s'il avait été possible d'avoir accès aux renseignements sur les agents chimiques et biologiques qui sont tenus secrets ; peut-être aurait-on pu toutefois déterminer de façon plus réaliste leurs possibilités d'utilisation dans telles ou telles circonstances déterminées.

Les estimations de victimes figurant aux tableaux 8 et 10 de l'annexe 3 sont fondées sur des effectifs hypothétiques de groupes de population exposés à l'action de tel ou tel agent chimique ou biologique. Les hypothèses formulées et les variables prises en considération sont indiquées dans l'annexe 3. Ces hypothèses concernant le mode de lancement, la dissémination et la persistance des agents ; les conditions météorologiques ; les concentrations et doses réelles ; les doses infectantes pour l'homme, les taux d'atteinte et de létalité ; enfin les possibilités de chimiothérapie.

Dans l'analyse détaillée d'une attaque par les agents de la tularémie et de la peste pulmonaire et par le nouveau toxique VX, on a tenu compte des effets qu'auraient l'existence et l'utilisation de certains équipements sanitaires sur les résultats de cette attaque (annexe 4). L'annexe 5 décrit les effets de la contamination par sabotage des eaux de consommation au moyen de la toxine botulique, du bacille de la typhoïde et du LSD.

Les estimations de victimes présentées dans l'annexe 3 valent pour une attaque surprise effectuée sur une zone urbaine dépourvue de moyens spéciaux de protection contre les agents chimiques et biologiques, et dans laquelle les bâtiments sont normalement aérés. On verra que, même si la population de la zone atteinte n'est pas très dense, les effets d'une attaque effectuée par un seul aéronef poseraient des problèmes de santé d'une ampleur sans précédent.

Il est certain que l'étendue et l'efficacité d'une attaque (ainsi que son coût et ses possibilités de réalisation) dépendent du niveau atteint par la technologie militaire. Examiner cet aspect de la question de façon tant soit peu approfondie n'est pas l'objet du présent rapport. Aussi a-t-on choisi d'étudier le cas d'une attaque limitée lancée contre une population civile dans des conditions propres à favoriser au maximum l'efficacité des armes employées. Cet exemple devrait permettre d'illustrer convenablement l'ampleur des dangers qui résulteraient de l'utilisation d'armes chimiques et biologiques. Des résultats analogues pourraient être obtenus par des

14

attaques de plus grande envergure au cas où les circonstances ne se prêteraient pas au mode d'attaque envisagé ici.

7. EFFETS À LONG TERME

Un certain nombre de considérations générales relatives aux effets éventuels à long terme des agents chimiques et biologiques sont ici de rigueur. En premier lieu, les connaissances actuelles sont insuffisantes pour qu'il soit possible de faire des prédictions sûres. Dans bien des cas, on ne peut faire plus que de mentionnner diverses possibilités appelant des études plus approfondies. Au-delà, il se pose des problèmes d'évaluation qui, s'ils comportent encore un élément technique important, supposent aussi des jugements de valeur qui débordent manifestement le cadre du présent rapport. Par exemple, on peut considérer que, par rapport aux effets directs d'une attaque chimique ou biologique létale, un risque limité d'effets nocifs à long terme ne présente qu'une importance secondaire. Au contraire, les effets à long terme de l'utilisation militaire d'agents n'ayant pas d'effet létal direct peuvent être jugés plus importants que leurs effets immédiats. A cet égard, il ne faut pas perdre de vue qu'il peut être risqué de s'appuyer sur l'expérience acquise dans un contexte non militaire, au sujet de l'action d'organismes et de composés chimiques pathogènes déjà présents dans le milieu, pour essayer de prévoir les effets de ces mêmes agents dans les conditions totalement différentes de la guerre et aux doses généralement plus fortes qui seraient utilisées à des fins militaires.

Les effets sanitaires à long terme que pourrait avoir l'utilisation d'armes chimiques et biologiques sont les suivants : 1) maladies chroniques causées par l'exposition aux agent chimiques et biologiques (voir les descriptions des annexes 1 et 2) ; 2) effets à retardement sur les personnes directement exposées aux agents chimiques et biologiques ; 3) création de nouveaux foyers de maladies infectieuses ; et 4) effets liés à des modifications écologiques.

Effets à retardement sur les personnes directement exposées

On s'inquiète beaucoup depuis quelque temps des effets nocifs à retardement que peut avoir sur l'homme l'exposition aux agents infectieux ou chimiques déjà présents dans le milieu. Les plus inquiétants de ces effets sont les suivants :

a) *Cancérogénèse*. Parmi les agents étiologiques présumés du cancer humain figurent à la fois des virus et des substances chimiques. On ignore actuellement si l'infection par l'un des virus qui pourraient éventuellement être utilisés à des fins militaires risque d'être cancérogène pour l'homme.

On possède en revanche quelques données, tirées surtout d'expériences sur l'animal, concernant la cancérogénicité de certaines catégories de substances chimiques. De nombreux agents alcoylants, par exemple, se sont révélés cancérogènes. Or, certains composés utilisables à des fins militaires, comme l'ypérite et le CS, sont des alcoylants. Comme on le verra à l'annexe 1, il y a lieu de penser que la fréquence du cancer des voies respiratoires a été nettement supérieure à la normale parmi les anciens combattants qui avaient été exposés à l'ypérite au cours de la première guerre mondiale ; un accroissement considérable de la fréquence des cas de ce type de cancer a d'autre part été signalé chez les travailleurs employés à la fabrication d'ypérite pendant la deuxième guerre mondiale.

b) *Tératogénèse*. Certaines substances chimiques et certains agents infectieux peuvent avoir de graves répercussions sur le développement du fœtus humain. L'action tératogène de la thalidomide et celle du virus de la rubéole sont particulièrement bien connues. On ignore si des agents susceptibles d'être utilisés pour la guerre chimique ou biologique auraient des effets tératogènes aux doses auxquelles risqueraient d'être soumises les femmes enceintes faisant partie de populations civiles directement visées ou accidentellement exposées. A cet égard, il est à noter que le Gouvernement des Etats-Unis a récemment limité l'utilisation de l'acide trichloro-2,4,5 phénoxyacétique, substance phytocide qui a été largement utilisée à des fins civiles aussi bien que militaires, car diverses expériences ont montré que l'absorption par voie buccale de doses relativement élevées de ce composé est tératogène pour la souris et le rat.

c) *Mutagénèse*. Jusqu'à ces dernières années, on n'a guère prêté d'attention aux altérations du génome humain qui pourraient résulter des maladies infectieuses ou de l'action de substances chimiques présentes dans le milieu. On sait que plusieurs substances chimiques provoquent de telles altérations chez l'animal d'expérience et dans les cultures de cellules humaines. L'inoculation de certains virus cause de nombreuses ruptures chromosomiques chez l'homme, mais on ignore si ce nouveau caractère est transmissible. Dans le cas de la rubéole, tout au moins, on peut dire que le dommage génétique n'est pas massif, encore qu'on ne puisse pas exclure la possibilité d'un accroissement moins important, mais néanmoins notable, de la fréquence des mutations.

Nouveaux foyers de maladies infectieuses

Comme on le verra à l'annexe 2, la guerre biologique risque d'entraîner directement la formation de nouveaux foyers de maladies infectieuses, soit dans les populations humaines, soit dans les populations d'animaux inférieurs, y compris les arthropodes vecteurs. Cette possibilité a été examinée dans le rapport des Nations Unies :

« L'une des conséquences possibles d'une attaque bactériologique (biologique) peut être la création artificielle de foyers d'infection multiples et répartis de manière dense, à partir desquels, si les conditions écologiques sont favorables, des foyers naturels peuvent se développer là où il n'en a jamais existé, ou dans des lieux où l'on était parvenu à les éliminer par des mesures de protection efficaces. »

Effets liés à des modifications écologiques

De nouveaux foyers de maladie pourraient également s'établir à la suite de modifications écologiques résultant de l'emploi d'agents biologiques infectieux pour l'homme et les animaux. Cette possibilité a également été examinée dans le rapport des Nations Unies :

« ..., l'emploi à grande échelle d'une arme bactériologique (biologique) pourrait réduire la population des animaux sauvages réceptifs à un niveau inférieur à celui qui lui permet de se perpétuer. L'élimination d'une espèce animale ou d'un groupe d'espèces animales, dans une région, créerait un vide dans la communauté écologique, et ce vide risquerait de gravement perturber l'équilibre de cette dernière, ou d'être comblé par une autre espèce présentant plus de danger pour l'homme parce qu'elle serait porteuse d'une zoonose, acquise soit par un processus naturel, soit à la suite de l'attaque bactériologique (biologique). Ainsi s'établirait un nouveau foyer naturel de maladie. »

En ce qui concerne les armes chimiques anti-personnel, on peut dire en tout cas que la dissémination massive de plusieurs agents chimiques pendant la première guerre mondiale n'a apparemment pas causé à long terme de dommages écologiques importants en Europe.

Cependant, de nouveaux foyers de maladie humaine pourraient également s'établir par suite de l'emploi d'agents phytocides. En endommageant gravement la flore de vastes territoires, on risque de créer des conditions favorables à l'établissement de nouveaux vecteurs ou à la formation de réservoirs de maladies infectieuses. Un exemple du processus par lequel des atteintes à la vie végétale peuvent créer de nouveaux dangers pour la santé humaine est donné dans le rapport des Nations Unies :

« Lorsqu'une forêt en état d'équilibre écologique est détruite par des abattages, elle se constitue en une forêt secondaire qui contient beaucoup moins d'essences, de plantes et d'espèces d'animaux, mais beaucoup plus de sujets des espèces qui survivent. Si la forêt secondaire est remplacée par des pâturages, cette évolution est encore plus marquée. Si une ou plusieurs des espèces animales dont la population augmente souffrent chroniquement d'une infection dangereuse pour l'homme (zoonose), le risque de contamination de l'homme en est considérablement accru. On en a eu la preuve dans le cas du typhus acarien en Asie du Sud-Est, où l'espèce de rat qui entretient l'infection et les acariens vecteurs sont beaucoup plus abondants dans la forêt secondaire, et encore plus dans les pâturages, de sorte que le risque représenté pour l'homme par le typhus acarien augmente lorsque la formation végétale passe de la forêt aux pâturages. »

Enfin, toute diminution importante de la qualité ou de la quantité des ressources alimentaires risquerait, à long terme, d'avoir de graves conséquences pour la santé humaine. Cette diminution pourrait résulter directement de l'utilisation d'agents de destruction des cultures ou se produire indirectement à la suite de modifications écologiques qui seraient la conséquence de la guerre chimique ou biologique.

8. RÉSUMÉ

A. *Considérations qualitatives*

Les agents chimiques et biologiques susceptibles d'être utilisés à des fins militaires sont décrits dans les annexes 1 et 2 ; les hypothèses formulées à ce sujet et divers facteurs d'ordre général sont examinés dans les pages qui précèdent ainsi que dans les annexes 3, 4 et 5. Les effets à long terme dont il est question à la section 7, et les conséquences psychosociales de l'utilisation d'armes chimiques et biologiques (annexe 6) sont également importants, quoique plus difficiles à évaluer.

Etant donné la rapidité d'action des *agents chimiques* létaux (voir annexe 1), il ne serait pas possible de réduire notablement la mortalité en appliquant un traitement spécifique. Les masques et abris contre les gaz pourraient offrir une protection, mais seulement si la population était bien préparée à les utiliser et hautement disciplinée, conditions qui ne sont pas remplies aujourd'hui dans la plupart des pays ; la mise en place de tels moyens de protection poserait d'autre part de sérieux problèmes d'ordre économique et psychosocial.

Les principales caractéristiques des *armes biologiques* (voir annexe 2) qui pourraient être éventuellement exploitées pour la guerre sont les suivantes :

a) Grande diversité des agents biologiques et combinaisons d'agents biologiques utilisables à des fins militaires.

b) Possibilité de manipuler à des fins de guerre biologique des souches de micro-organismes actuellement présentes dans le milieu, afin de produire des germes antigéniquement différents ou résistants aux antibiotiques (tularémie, peste, charbon, grippe) contre lesquels les moyens de prophylaxie ou de traitement ordinaires seraient inopérants.[1]

c) Impossibilité de prévoir les effets directs. Il n'est pas exclu en effet qu'une attaque biologique destinée à produire une action létale massive se révèle relativement inefficace, alors qu'une attaque visant seulement à produire des effets incapacitants pourrait, contre toute attente, tuer une fraction importante de la population visée. D'autre part, certains agents (charbon, coccidioïdomycose) pourraient persister pendant de longues périodes sous forme de spores résistantes que les vents, à la longue, transporteraient à de très grandes distances.

d) Impossibilité de prévoir certains effets secondaires : risques de contagion et d'épidémies, par exemple. Des épidémies risqueraient en outre de se produire accidentellement en cas de « fuites » de souches virulentes cultivées en laboratoire.

e) Si les agents biologiques sont faciles à produire, des systèmes complexes de production et d'acheminement jusqu'au but seraient nécessaires pour qu'une attaque ait quelque chance de succès, sauf peut-être s'il s'agissait seulement de désorganiser la vie de la collectivité par une action limitée de sabotage (visant à provoquer une épidémie de variole, par exemple).

Des caractéristiques susmentionnées, *a*) et *b*) seraient favorables à l'attaquant, tandis que *c*) et *d*) limiteraient la valeur militaire des armes biologiques.

B. *Estimations quantitatives* (tableaux 8, 9 et 10, annexe 3)

1. On s'est efforcé d'apprécier ce que seraient les effets primaires d'attaques aériennes de faible envergure sur des villes de 0,5 à 5 millions d'habitants dans des pays industrialisés et des pays en voie de développement. On a supposé que l'attaque serait effectuée par un bombardier ou une petite escadrille, qui disperseraient des agents chimiques ou biologiques déterminés sur un front de 2 km perpendiculaire à la direction du vent. Sur la base des hypothèses retenues on est parvenu aux conclusions ci-après :

a) De tous les agents de guerre chimiques connus, seuls les gaz neurotoxiques, et peut-être aussi la toxine botulique, seraient capables de faire autant de victimes que les agents biologiques.

b) Dans des conditions atmosphériques favorables à l'attaquant, le déversement efficace de 4 tonnes de sarin sur une agglomération urbaine

[1] Les vaccinations de masse n'auraient qu'une valeur protectrice très douteuse étant donné la multiplicité des agents et des souches susceptibles d'être employés, sans parler des réactions adverses (réactions immunologiques secondaires) qu'elles pourraient produire.

(ce qui nécessiterait des projectiles d'un poids total de 15 à 20 tonnes) causerait des dizaines de milliers de morts (sur une superficie d'environ 2 km²). Même si les conditions étaient défavorables, les morts pourraient se compter par milliers. Au cas où, pour une telle attaque, on utiliserait 4 tonnes de VX, le nombre des victimes ne serait pas sensiblement plus grand si les conditions météorologiques étaient défavorables à l'attaquant, mais si elles étaient favorables, cette attaque de portée limitée affecterait une zone d'environ 6 km² et pourrait causer de 50 000 à 180 000 morts.

c) L'utilisation de 4 tonnes d'une toxine botulique bien stabilisée ou de VX en aérosol à particules fines (5 μ de diamètre) — préparations qui restent à mettre au point — pourrait causer des centaines de milliers de morts en raison de la grande dispersabilité de ces agents (12 km² pour la toxine botulique, 40 km² pour l'aérosol de VX monodispersé). Dans ce cas, le poids total des projectiles nécessaires pourrait être deux ou trois fois plus grand que dans celui des agents mentionnés au paragraphe *b*).

d) Dans le cas d'un agent biologique comme la bactéridie du charbon, l'attaque d'une ville, même par un seul appareil de bombardement déversant 50 kg de l'agent desséché sous la forme d'un aérosol approprié, affecterait une superficie très supérieure à 20 km² et causerait des dizaines, voire des centaines de milliers de morts. Une attaque analogue au moyen d'un agent biologique plus labile pourrait affecter des superficies de 1 km² à plus de 20 km², selon l'agent utilisé, et faire des dizaines ou des centaines de milliers de victimes ainsi que de nombreux milliers de morts.

2. La contamination des eaux de consommation par une action limitée au moyen du bacille de la fièvre typhoïde, de LSD ou d'une toxine botulique stable, pourrait provoquer une désorganisation et une mortalité considérables dans une grande ville (annexe 5), et affecterait des dizaines de milliers d'individus.

3. Une action de sabotage ou une attaque ouverte qui provoqueraient la propagation secondaire d'épidémies de fièvre jaune, de peste pulmonaire, de variole ou de grippe pourrait, dans certaines conditions, se solder en fin de compte par de nombreux millions de malades et de morts (voir annexe 2).

4. Les estimations de victimes qui sont présentées ici correspondent aux effets que pourrait produire une attaque de très faible envergure n'ayant qu'un objectif très limité. Un certain nombre de pays sont déjà parfaitement en mesure de lancer une telle attaque, et il est possible que leur nombre aille en augmentant. Dans le cas d'attaques de plus grande envergure menées avec des moyens très perfectionnés — attaques que les grandes puissances militaires pourraient lancer sans trop de difficultés — l'ampleur des destructions causées par les armes chimiques et biologiques serait évidemment beaucoup plus grande.

9. ORIENTATIONS OFFERTES À L'ORGANISATION MONDIALE DE LA SANTÉ ET À SES ÉTATS MEMBRES

Aux termes de l'article 2 *d*) de sa Constitution, l'OMS doit « fournir l'assistance technique appropriée et, dans les cas d'urgence, l'aide nécessaire, à la requête des gouvernements ou sur leur acceptation ». L'utilisation d'armes chimiques et biologiques aurait sans nul doute pour effet de créer des problèmes médico-sanitaires d'une ampleur et d'une urgence extrêmes — mortalité et morbidité massives, épidémies, etc. — à la solution desquels l'OMS pourrait être appelée à participer. On trouvera à l'annexe 4 une estimation de l'ampleur des problèmes de santé publique que poserait une attaque minimale menée au moyen de quelques agents représentatifs (tularémie, peste, VX). Cette analyse des effets d'une attaque limitée, dont les résultats sont confirmés par les autres analyses présentées ailleurs dans le présent rapport ainsi que dans le rapport des Nations Unies, montre l'immensité des efforts qu'il faudrait déployer, essentiellement en pure perte, pour mettre en œuvre les mesures complexes que suppose une défense contre tel ou tel agent. D'autre part, comme l'indique l'annexe 6, de telles mesures risqueraient de renforcer les craintes d'autres pays qui se sentiraient menacés d'anéantissement. Les sentiments de peur réciproque qui en résulteraient entre les nations pourraient favoriser à leur tour une prolifération des armes chimiques et biologiques et une accélération de la course aux armements, ce qui augmenterait dans d'énormes proportions les risques de libération accidentelle ou délibérée d'agents chimiques et biologiques.

Toutefois, dans l'état actuel des besoins et des ressources, il serait possible de prendre certaines mesures qui viendraient à l'appui des activités d'hygiène et de prévention courantes, sans donner naissance aux craintes dont il vient d'être question ; ces mesures comprennent notamment le renforcement des moyens de détection rapide de la pollution et des moyens de diagnostic des maladies transmissibles, ce qui aurait incontestablement des conséquences favorables pour les services de santé et les services de laboratoire en général ; l'amélioration des moyens médicaux à mettre en œuvre en cas de catastrophe naturelle, notamment celle des méthodes de décontamination ; la généralisation des dispositifs de sécurité installés dans les bâtiments (ventilation filtrée) et dans les réseaux de distribution d'eau (voir annexe 5). Ces mesures exerceraient peut-être un certain effet de dissuasion sur les groupes irresponsables et pourraient réduire notablement les pertes résultant d'une attaque de très faible envergure ou de la propagation des effets d'une attaque dirigée contre un pays voisin, mais elles n'offriraient guère de protection à un pays victime d'une attaque résolue et de grande ampleur.

Tant que des recherches chimiques et biologiques seront menées à des fins purement militaires, certains pays jugeront nécessaire de continuer à vouloir se donner les moyens de détecter les agents chimiques et biologiques et de protéger leur population contre ces agents. Or, ces efforts de recherche pourraient faire découvrir des agents encore plus dévastateurs que ceux que l'on connaît actuellement. Etant donné que les agents déjà connus ont un pouvoir de destruction énorme s'ils sont utilisés dans des conditions favorables et que de nouvelles armes encore plus dangereuses risquent d'être mises au point, il est impératif de trouver rapidement des moyens d'obtenir que nul ne se sente plus obligé de faire dans ce domaine des recherches à buts militaires.

En dernière analyse, il paraît donc évident que l'intérêt bien compris de tous les Etats Membres et de l'humanité en général est de donner rapidement effet aux résolutions de l'Assemblée générale des Nations Unies et de l'Assemblée mondiale de la Santé sur la guerre chimique et biologique (annexes 7 et 8), et de prendre toutes les mesures supplémentaires qui pourront se révéler nécessaires pour proscrire la mise au point et l'utilisation en toutes circonstances d'agents chimiques et biologiques comme armes de guerre.

Enfin, il est possible que l'ONU demande à l'OMS de l'aider à juger du bien-fondé d'allégations relatives à l'emploi d'armes chimiques et biologiques dans les conflits entre nations et qu'elle sollicite son concours pour l'action qu'elle mène en faveur de la limitation des armes chimiques et biologiques et en faveur du désarmement. Grâce aux moyens techniques dont elle dispose,[1] l'OMS pourrait grandement contribuer à la solution de beaucoup des problèmes qui se posent à cet égard et qui sont actuellement étudiés par l'Organisation des Nations Unies.

REMERCIEMENTS

Le Groupe tient à remercier les nombreux membres du personnel de l'OMS dont le concours lui a été extrêmement précieux pour la préparation du présent rapport. Il remercie également le Stockholm International Peace Research Institute, la Conférence de Pugwash et l'Organisation météorologique mondiale de leur collaboration et des nombreux renseignements de grand intérêt que ces organisations lui ont communiqués.

[1] On peut mentionner par exemple la collecte de données épidémiologiques sur les maladies transmissibles, à laquelle l'OMS procède depuis de longues années par l'intermédiaire de ses banques de sérums et de ses programmes de surveillance portant sur des maladies déterminées. Ces données sont extrêmement utiles pour déterminer les changements qui interviennent, ou qui pourraient intervenir, dans le tableau des maladies transmissibles, ainsi que pour approfondir la connaissance des maladies qui sévissent déjà dans les collectivités. En dehors de son intérêt épidémiologique général, l'extension de cette activité pourrait être très utile pour aider, le cas échéant, à déterminer le bien-fondé d'allégations concernant l'utilisation d'armes biologiques.

Table des matières

Tome VII

Marcel-Paul Schützenberger

ŒUVRES COMPLÈTES

éditées par Jean Berstel, Alain Lascoux et Dominique Perrin

Les treize tomes de cette édition contiennent l'ensemble des œuvres de Marcel-Paul Schützenberger qui ont fait l'objet d'une publication dans une revue scientifique ou un livre. Ses travaux couvrent une période de plus de 50 ans, depuis sa première note aux Comptes Rendus en 1943 jusqu'à son dernier article, paru en 1997.
Les publications sont présentées dans l'ordre chronologique. Chaque tome est précédé d'une courte introduction qui essaie d'éclairer certains des travaux, tant pour leur intérêt scientifique intrinsèque que pour l'écho qu'ils ont rencontré et les développements qu'ils ont suscités.

Tome 7 : 1969 – 1970

Une longue collaboration débute entre Eilenberg et Schützenberger. L'un des premiers produits de cette coopération est l'article « Rational sets in commutative monoids » sur les parties rationnelles dans les monoïdes commutatifs. Cette coopération se poursuivra pendant toute la préparation du livre « Automata, Languages and Machines ».
L'article « A combinatorial problem in the theory of free monoids » est écrit avec André Lentin. Il contient un résultat très important de la combinatoire des mots.
La monographie « Théorie géométrique des polynômes eulériens » est une illustration de la méthode bijective, appelée ici géométrique, et appliquée de manière systématique dans ce texte. Les résultats seront par la suite généralisés par divers auteurs et repris dans de nombreux ouvrages dont le volume 3 du livre de Knuth et dans un chapitre du volume de Lothaire.