MATHÉMATIQUES ET SCIENCES HUMAINES

SIMON RÉGNIER

Études sur le polyèdre des partitions

Mathématiques et sciences humaines, tome 82 (1983), p. 85-111
<http://www.numdam.org/item?id=MSH_1983__82__85_0>

Math. Sci. hum. (21^e année, n°82, *1983*, p.85-111)

ETUDES SUR LE POLYEDRE DES PARTITIONS

Simon REGNIER *

---

* Ce texte rassemble différents travaux, écrits de 1971 à 1975 et correspondant aux textes (11) et (22) de la bibliographie de S. Régnier, pages 11 et 12 de ce numéro de *Mathématiques et Sciences humaines*.

1. INTRODUCTION.

Dans une publication antérieure [1], nous avons introduit, en vue de l'agrégation d'une famille de relations d'équivalence, la représentation de chaque partition x d'un ensemble fini, $E = \underline{n}^*$, par le vecteur X de $R^{E\times E}$ de coordonnées :

$$\forall i,j \in E\times E \qquad X_{ij} = 1 \quad \text{si i et j sont dans la même x-classe}$$
$$= 0 \quad \text{sinon.}$$

$(i,j) \to X_{ij}$, application de $E\times E$ dans $(0,1) \subset R$, est alors la fonction indicatrice** du graphe de la relation d'équivalence associée à x. Les correspondances étant biunivoques, nous convenons de désigner par le même symbole (X, ou Y, ou Z, etc...) une partition, la relation d'équivalence associée, son graphe inclus dans $E\times E$, et la fonction indicatrice de ce graphe, considérée comme un élément de l'espace vectoriel $R^{E\times E}$. Alors $X_{ij} = 1$ s'écrit au choix : $i\ X\ j$, $j \in i\ X$, $(i,j) \in X$, et $X(i,j) = 1$.

De même $\underline{P_n}$ peut désigner à la fois l'ensemble des partitions de E, et l'ensemble des points représentatifs dans $R^{n\times n}$. Les relations quelconques sont représentées par la totalité de l'hypercube de base $(0,1)^{E\times E}$, également noté $(0,1)^{n^2}$. Nous avons donc :

$$\underline{P_n} \subset (0,1)^{n\times n} \subset R^{n\times n}.$$

Dans l'article cité [1], cette représentation nous servait à simplifier et linéariser le problème suivant.

Trouver une équivalence X dont la moyenne des distances à p partitions données s^k pondérées par des poids p_k soit minima, la distance utilisée étant le cardinal de la différence symétrique des graphes.

On montrait que toute relation X, dite "partition centrale", devait maximiser la forme linéaire suivante :

$$L(X) = \Sigma\, t_{ij}\, X_{ij} \qquad \text{où} \qquad t_{ij} = \left(\sum_k p_k\, s^k_{ij} \,/\, \Sigma\, p_k\right) - \frac{1}{2}$$

Nous proposions un algorithme heuristique susceptible de conduire à des maxima locaux.

---

* Toute notation $a = \underline{b}$ signifie ici : b est un cardinal, et a un ensemble de cardinal b. R désigne l'ensemble des réels.

** "Indicatrice" plutôt que "caractéristique", vu l'emploi de ce terme en calcul des probabilités.

L'objet du présent travail est de décider cette question des maxima locaux selon les principes de la programmation linéaire : une partition est optimale absolument si et seulement si elle est meilleure que toutes les partitions adjacentes, sur le polyèdre P_n qui est la fermeture convexe de $\underline{P_n}$.

2. ARETES ET DIAGONALES.

Dans $R^{n\times n}$, la fermeture convexe de l'ensemble des $\underline{P_n}$ points représentatifs est un polyèdre convexe P_n qui présente autant de sommets ; car plus généralement, aucune relation S ne peut être un mélange (combinaison linéaire convexe) d'autres relations $T \in (0,1)^{n\times n}$:

$$S = \Sigma\, P_T\, T \qquad \text{avec} \qquad \Sigma\, P_T = 1 \;\text{ et }\; P_T \geq 0 \text{ pour tout } T$$

car cela entraîne : si $S_{ij} = 1$ $\quad T_{ij} = 1$ ou bien $P_T = 0$

si $S_{ij} = 0$ $\quad T_{ij} = 0$ ou bien $P_T = 0$

donc $P_T > 0$ entraîne $T = S$.

Deux sommets X et Y sont dits *adjacents* si le segment qui les joint est une arête du polyèdre, ce que l'on écrira X A Y. La négation de X A Y s'écrira X D Y, quand X et Y sont différents.

La théorie générale des polyèdres nous propose deux caractérisations des arêtes :

a) $X\,A\,Y \Longleftrightarrow X \neq Y$ et il existe sur $R^{n\times n}$ une forme linéaire L telle que $L(X) = L(Y) > L(Z)$ pour tout sommet Z distinct de X et Y.

b) $X\,D\,Y \Longleftrightarrow$ Il existe des sommets Z_t distincts de Y et des réels $\lambda_t > 0$, tels que :

(1) $\qquad Y - X = \Sigma\, \lambda_t (Z_t - X)$.

La dissymétrie de cette formule n'est qu'apparente, puisqu'on peut imposer $Z_t \neq X$ et qu'en posant $\Lambda = \Sigma\, \lambda_t \quad \beta = \frac{1}{\Lambda} \quad \alpha = 1 - \beta$ et $p_t = \lambda_t / \Lambda$ on peut écrire :

(11) $\qquad Y + (\Lambda - 1)\, X = \Sigma\, \lambda_t\, Z_t$, puis

(111) $\qquad \alpha\, X + \beta\, Y = \Sigma\, p_t\, Z_t$

avec $\alpha + \beta = 1 = \Sigma\, p_t \qquad \alpha, \beta$ et $p_t > 0$.

Le fait que $\lambda - 1$ et α sont > 0 est établi dans 3.2.

A partir de la formule symétrique (111) on peut former réciproquement des coefficients $\lambda_t = p_t / \beta$ ou $\mu_t = p_t / \alpha$ tels que

(2) $X - Y = \Sigma\, \mu_t\, (Z_t - Y).$

Les formules (1) et (2) seront dites "adjointes", il peut exister plusieurs paires de telles formules, et autant de formules (111), avec des partitions Z_t différentes. Cela est particulièrement évident lorsque X et Y n'ont pas les mêmes partitions adjacentes.

Cette relation symétrique signifie que le segment X Y coupe la fermeture convexe des autres sommets. C'est là une définition très intuitive des segments inscrits qui ne sont pas des arêtes, et que l'on peut appeler des diagonales. Dans la relation (1), on peut évidemment imposer aux points Z_t d'être adjacents à X. Alors pour qu'un sommet X maximise une forme linéaire L, il faut et il suffit que $L(X) > L(Z)$ pour tout Z adjacent à X. Puisque alors, pour tout Y non adjacent :

$$L(Y - X) = \Sigma\, \lambda_t\, L(Z_t - X) \leq 0.$$

Pour lever la question des maxima locaux, dans l'algorithme des transferts ou tout autre, il suffit de connaître toutes les partitions adjacentes à une partition donnée.

Nous allons d'abord caractériser les arêtes $\{X,Y \mid X\, A\, Y\}$, puis étudier l'ensemble des adjacentes d'un X donné :

$$A\,[X] = \{Y \mid X\, A\, Y\}.$$

3. CARACTERISATION DES PARTITIONS ADJACENTES.

3.1. L'énoncé du théorème de caractérisation.

Nous allons établir en plusieurs étapes que :

Pour que deux partitions X et Y soient adjacentes, il faut et il suffit :

a) si elles sont comparables, que la moins fine se déduise de l'autre par réunion de deux classes, et

b) si elles ne sont pas comparables, que leurs restrictions aux diverses classes de la partition $X \vee Y$ qui les borne supérieurement, (composantes connexes de la relation "X ou Y") ne diffèrent que sur l'une de ces classes.

a) signifie que X et Y sont en succession immédiate dans le treillis des partitions ordonnées par la relation d'inclusion, équivalente à l'ordre canonique de $R^{n\times n}$: $X_{ij} \leq Y_{ij}$ pour tout i, j .

Dans b) la partition $X \vee Y$ est par définition la plus fine des partitions moins fines que X et Y à la fois. C'est la fermeture transitive $\overline{S}$ de la relation non transitive "X ou Y", notée $S = X \cup Y$.S est la borne supérieure de X et Y dans le treillis des relations symétriques et dans celui de toutes les relations.

3.2. Partitions intermédiaires et partitions utiles.

Nos démonstrations reposent surtout sur la formule (1)

$$X\ D\ Y \Longleftrightarrow \exists \text{ des } \lambda_t > 0, \text{ des } Z_t \neq Y,\ t \in T, \text{ tels que}$$

$$(1) \qquad Y - X = \sum_{t \in T} \lambda_t\,(Z_t - X).$$

Nous allons montrer d'abord que les partitions Z susceptibles de figurer dans une telle formule vérifient deux propriétés très restrictives :

a) les partitions Z de (1) sont comprises entre les deux relations qui bornent X et Y, nous dirons alors que Z est *intermédiaire*. La borne inférieure est : $X \cap Y = X\,Y$ qui est une relation d'équivalence, l'autre est la relation : $= X \cup Y = X + Y - X\,Y$, qui n'est pas transitive en général.

Nous en déduirons que chaque Z classe $i\,Z$ est incluse dans l'une ou l'autre des deux classes $i\,X$ et $i\,Y$.

b) nous montrerons de plus que si deux classes concourantes $i\,X$ et $i\,Y$ sont non comparables, alors l'inclusion ci-dessus est remplacée par l'égalité :

$$i\,Z = i\,X, \quad \text{ou} \quad i\,Z = i\,Y.$$

Les partitions intermédiaires qui présentent en plus cette propriété seront appelées les *partitions utiles*.

c) enfin, nous montrerons que $\Sigma\,\lambda_t$ dépasse 1, et que par suite, quand X_{ij} diffère de Y_{ij}, la coordonnée Z^t_{ij} des diverses partitions Z ne peut être une constante indépendante de t.

Démonstration.

a) *Partitions intermédiaires.*

Pour tout couple (i,j) de E, on a :

$$Y_{ij} - X_{ij} = \Sigma\,\lambda_t\,(Z^t_{ij} - X_{ij}).$$

Toutes les différences qui interviennent ont le même signe : elles sont ≥ 0 si $X_{ij} = 0$ et ≤ 0 si $X_{ij} = 1$.

Par conséquent, $Y_{ij} = X_{ij}$ entraîne, pour tout t, $Z_{ij} = X_{ij}$.

Comme les X, Y, Z ne prennent que les valeurs 0 et 1, il en résulte que dans tous les cas Z_{ij} est compris entre les deux bornes des nombres X_{ij} et Y_{ij}, d'où :

$$X\,Y \leq Z \leq X \cup Y \quad \text{pour l'ordre partiel sur } R^{n\times n}. \qquad \text{(a)}$$

Ce premier point permet d'établir de plus que toute Z *classe est incluse dans une* X *classe ou une* Y *classe*. Désignons en général par $i\,S$ l'ensemble des i qui entrent avec i dans la relation S : $i\,S = \{j \mid i\,S\,j\}$;
(a) nous donne pour tout i dans E : $i\,Z \subset i(X \cup Y) = i\,X \cup i\,Y$

Maintenant $i\,Z$ ne peut déborder à la fois de $i\,X$ et de $i\,Y$ car si
1) $i\,X\,j$ et non $i\,Y\,j$ (soit $j \in i\,X - i\,Y$)
2) $i\,Y\,k$ et non $i\,X\,k$ (soit $k \in i\,Y - i\,X$)
j et $k \in i\,Z$ est impossible car
$j\,Z\,k$ entraîne $j\,X\,k$ ou $j\,Y\,k$ mais
$j\,X\,k$ et 1) entraîne $i\,X\,k$, contraire à 2), tandis que $j\,Y\,k$ est également impossible.

Partitions utiles.

b) Si $i\,X$ et $i\,Y$ sont non comparables, les deux différences $i\,X - i\,Y$ et $i\,Y - i\,X$ sont non vides. Prenons j dans l'une et k dans l'autre. Alors

(1) $X_{ij} = 1, \; Y_{ij} = 0, \; X_{ik} = 0$ et $Y_{ik} = 1$ si bien que

$X_{jk} = Y_{jk} = 0$ par transitivité, d'où $Z_{ik} = 0$

(1) entraîne en i, j : $-1 = \Sigma\, \lambda_t\, (Z^t_{ij} - 1)$

en i, k : $1 = \Sigma\, \lambda_t\, Z^t_{ik}$

d'où :

(2) $0 = \Sigma\, \lambda_t\, (Z^t_{ij} + Z^t_{ik} - 1)$

mais comme, par transitivité, on a $Z^t_{ij} + Z^t_{ik} - 1 \leq Z^t_{jk} = 0$ aucun terme de cette somme n'est positif. Il faut alors que tous soient nuls.
$i\,Z$ doit contenir soit j soit k. Mais alors, ce raisonnement étant validé pour tout j de $i\,X - i\,Y$, et tout k de $i\,Y - i\,X$, il apparaît que nécessairement, $i\,Z = i\,X$ ou bien $i\,Z = i\,Y$.

c) Comme Y diffère de X, il existe un couple (i,j) tel que :

$$|Y_{ij} - X_{ij}| = 1 = \sum_{Z^t_{ij} \neq X_{ij}} \lambda_t$$

Cette somme partielle étant égale à 1 on en déduit : $\Lambda = \Sigma\, \lambda_t \geq 1$, mais $\Lambda = 1$ est impossible car $\Sigma\, \lambda_t = 1$ et $Y = \Sigma\, \lambda_t\, Z_t$ signifierait que Y est un mélange de partitions distinctes de Y, ce qui est impossible. De ce point résulte que quand X_{ij} diffère de Y_{ij}, Z^t_{ij} ne peut être une constante.

3.3. Deux lemmes.

Dans le treillis des partitions, X et Y ont deux bornes $X \wedge Y = X\,Y = X \cap Y$, et $X \vee Y = \overline{X \cup Y}$ notée $\overline{S}$. Nous appellerons *atomes* les classes de l'intersection X Y, *régions* les classes de $\overline{S}$, *zones* les réunions de régions. Une partie F de E est une zone, si et seulement si F est à la fois une réunion de X-classes et de Y-classes, et aussi de Z-classes pour toutes les partitions Z intermédiaires.

X et Y définissent deux partitions dites "quotients" X' et Y' sur le quotient E' de E par X Y, et des partitions restrictions X_r et Y_r sur toutes les régions $r \in E/\overline{S}$ comme sur n'importe quelle autre partie $r \subset E$. Les théorèmes préalables sont :

THEOREME DU QUOTIENT X et Y sont adjacentes si et seulement si leurs partitions quotients X' et Y' le sont.

THEOREME DES RESTRICTIONS X et Y sont adjacentes si et seulement si leur restrictions sont adjacentes dans une région et identique dans les autres.

Démonstrations.

1) Montrons que $X \, D \, Y \Longleftrightarrow X' \, D \, Y'$.

Dans la formule (2) : $Y - X = \Sigma \, \lambda_t \, (Z_t - X)$ les intermédiaires Z_t sont moins fines que X Y, et définissent sur le quotient $E' = E/X\,Y$ des partitions Z'_t.
Les trois fonctions $(i,j) \to X_{ij}, Y_{ij}$ ou Z_{ij} restent constantes quand i et j parcourent chacun un atome : $i \in i' \in E'$ et $j \in j' \in E'$. Par conséquent (2) entraîne $Y' - X' = \Sigma \, \lambda_t \, (Z'_t - X')$, et réciproquement.

2) Pour le théorème des restrictions, nous montrerons d'abord que s'il existe une partition de E en deux zones r_1 et r_2, telles que X et Y aient sur l'une et l'autre des restrictions distinctes, alors X et Y sont non adjacentes. Puis inversement, que ssi X et Y ont des restrictions adjacentes dans une région r, et identiques dans toutes les autres, elles sont adjacentes.

Première partie.

$E = r_1 + r_2$ donc $E \times E = r_1 \times r_1 + r_2 \times r_2 + r_1 \times r_2 + r_2 \times r_1$.

Nous pouvons noter X_1 et Y_1 les deux restrictions à r_1 ;
X_2 et Y_2 les deux restrictions à r_2.

Dans la décomposition de $E \times E$ indiquée ci-dessus, le vecteur (X_{ij}) $i,j \in E \times E$ peut s'écrire $(X_1, X_2, 0, 0)$, car pour tout couple (i,j) pris dans $r_1 \times r_2$ ou $r_2 \times r_1$, la coordonnée X_{ij} est nulle du fait que aucune X-classe ne coupe plusieurs zones ou régions.

De même $Y = (Y_1, Y_2, 0, 0)$.

Les vecteurs $U = (X_1, Y_2, 0, 0)$ et $V = (X_2, Y_1, 0, 0)$ représentent deux partitions intermédiaires

U coïncide avec X sur r_1 et avec Y sur r_2,
V coïncide avec Y sur r_1 et avec X sur r_2.

On a : $X + Y = U + V$ ce qui prouve que X et Y ne sont pas adjacentes d'après la formule b)(11) du 2.

Nota. Cette première partie présente une réciproque intéressante :
Si $Y - X = p\,(U - X) + q(V - X)$, avec U et $V \neq X$ et Y, p et $q \neq 0$ alors $p = q = 1$ et il existe une partition en deux zones r_1 et r_2 telles que :

sur r_1, $U = X$ et $V = Y$
sur r_2, $U = Y$ et $V = X$

En effet, écrivons : $Y + (p + q - 1)\,X = p\,U + p\,V$.

S'il existe i et j tels que $Y_{ij} = 1$, $V_{ij} = 0$, nécessairement $U_{ij} = 1$, $p = 1$ et $X_{ij} = 0$; s'il existe i et j tels que $V_{ij} = 1$ et $Y_{ij} = 0$, on a $(p + q - 1)\,X_{ij} = p\,U_{ij} + q$; $q > 0$ entraîne $X_{ij} = 1$ d'où $p = 1$ et $U_{ij} = 0$.

On montre de même $q = 1$, et donc $Y + X = U + V$.

Cela dit, comme en général $i\,U \subset i\,X$ ou $i\,Y$, ici par symétrie $i\,U = i\,X$ ou $i\,Y$ (X et Y étant intermédiaires pour U et V). Il nous reste à poser $r_1 = \{i \mid i\,V = i\,X \neq i\,Y\}$ $r_2 = \{i \mid i\,V = i\,Y\}$ pour construire les deux zones annoncées. (Il peut exister d'autres couples r_1, r_2).

Deuxième partie.

Si X et Y sont adjacentes dans une région r, et coïncident ailleurs, X A Y.
Montrons que $X\,D\,Y \Longleftrightarrow X_r\,D\,Y_r$.

En effet, dans les régions où X et Y coïncident, les intermédiaires Z coïncident également.

Si $Y - X = \Sigma\,\lambda_t\,(Z_t - X)$, $\lambda_t > 0$ et $Z_t \neq Y$

$Z_t \neq Y$ doit s'appliquer isolément aux restrictions de Z et Y à la région r.

Ce qui assure que $X_r\,D\,Y_r$ d'après (1). Inversement, $X_r\,D\,Y_r$ entraîne D X Y.

3.4. Le cas des partitions comparables.

Nous pouvons désormais simplifier l'étude en nous restreignant à la région où

X et Y diffèrent, puis à son quotient par X Y. Nous sommes alors ramené au cas où X Y = 0 et X v Y = 1. Nous dirons alors que X et Y forment une grille connexe (grille si X Y = 0 seulement).

Si X et Y sont comparables, par exemple $X \leq Y$, cela nous impose X Y = X = 0 et Y = 1.

Mais quand X = 0 il est facile de distinguer l'ensemble des arêtes issues de X. Prenons la représentation dans R^m avec $m = \frac{n(n-1)}{2}$. Dans le cube $(0,1)^m$ (représentant toutes les relations symétriques et réflexives) l'origine 0 est l'extrêmité de m arêtes parallèles aux axes, au même titre que n'importe quel autre sommet, mais de plus ces arêtes appartiennent au polyèdre P_n, parce que l'autre extrêmité est un point avec une seule coordonnée non nulle, c'est-à-dire une partition en n-1 classes, où une seule paire d'objets est réunie. Maintenant, il est connu que si un point extrêmal d'un convexe appartient à un autre convexe Q inclus, il est extrêmal sur Q. Pour la même raison les arêtes du cube sont des arêtes de P_n. On pourrait voir aussi que X = 0 et Y à n-1 classes n'ont pas d'intermédiaires. Ce point résulte aussi du fait que les deux partitions considérées n'ont pas d'intermédiaires.

Réciproquement, comme le cube $(0,1)^m$ est tout entier inclus dans le cône convexe engendré par les m arêtes, qui est le cône des $Y \geq 0$, P_n est à fortiori inclus dans ce cône, ce qui veut dire qu'aucune autre partition n'est adjacente à X.

3.5. Le cas des partitions non comparables.

Pour en terminer avec l'énoncé de 3.1., il nous reste à montrer que si deux partitions non comparables X et Y ont des restrictions différentes dans une seule région, alors elles sont adjacentes. Pour cela nous considérerons la plus petite zone où X et Y supposées non adjacentes, ont des restrictions distinctes, et nous montrerons qu'elle comporte au moins deux régions. On peut convenir que cette zone est E tout entier. (On pourrait convenir aussi que X Y = 0 en traitant les atomes comme des éléments). Il nous suffira alors de démontrer simplement que l'existence de *partitions utiles* implique que E se partage en deux zones non vides.

3.6. La partition de comparaison locale associée à deux partitions.

De deux partitions quelconques X et Y nous déduisons une partition de E en trois classes :

$E_1 = \{i \mid i X \subseteq i Y\}$ $\qquad E_2 = \{i \mid i Y \subset i X\}$ (inclusion stricte)

$E_0 = E - E_1 - E_2$

Nous l'appellerons partition de comparaison locale L(X, Y).

E_1 est la réunion de X-classes qui sont des atomes : $i\,X = i\,X\,Y$.

E_2 celle des Y-classes qui sont des atomes mais non des X-classes :

$$i\,Y = i\,X\,Y \neq i\,X.$$

Par suite les complémentaires $E_0 + E_2$ et $E_0 + E_1$ sont également l'un une réunion de X-classes, l'autre de Y-classes. Pour $i \in E_0$, sa X-classe ne peut déborder E_0 que dans E_2 ; sa Y-classe que dans E_1.

Pour mieux décrire la situation, on peut encore introduire les deux subdivisions :

$E_1 = E_3 + E_5$ et $E_2 = E_4 + E_6$

$E_3 = \{i \in E_1 \mid i\,X \subseteq i\,Y \subset E_1\}$

$E_5 = \{i \in E_1 \mid i\,Y \text{ coupe } E_0\}$

et de même sur E_2 :

$E_4 = \{i \in E_2 \mid i\,Y \subset i\,X \subset E_2\}$

$E_6 = \{i \in E_2 \mid i\,X \text{ coupe } E_0\}$

E_3 est une zone, réunion de X-classes et aussi de Y-classes.

E_4 également. E_3 contient d'ailleurs la zone : $E_7 = \{i \mid i\,X = i\,Y\}$ où chaque atome est une région.

La situation la plus générale est résumée par l'exemple ci-contre. Quand E constitue une seule région, seuls E_0, E_5 et E_6 sont non vides.

3.7. Fin de la démonstration.

Quand X et Y sont non adjacentes (et non comparables), il existe des partitions utiles Z liées par une relation du type (1).

Pour chacune de ces partitions utiles Z et chaque i de E_0, $i\,Z = i\,X$ ou $i\,Z = i\,Y$. Ce qui définit une nouvelle partition de E_0 en deux catégories :

$$A_Z = \{i \mid i\,Z = i\,X\}$$

$$B_Z = \{i \mid i\,Z = i\,Y\}$$

Quand $i\,Z\,j$, i et j sont de même catégorie. A_Z est donc une réunion de X-classes (qui sont des Z-classes) et B_Z une réunion de Y-classes. Par suite, $E'_Z = A_Z + E_1$ et $E''_Z = B_Z + E_2$ sont deux zones, puisque le complémentaire d'une réunion de X-classes est une autre réunion de X-classes. Il nous reste à prouver que, pour au moins une partition utile Z, E'_Z et E''_Z sont non vides.

Si $E_0 = \emptyset$, E_1 ni E_2 ne peuvent l'être, sinon X et Y seraient comparables *.

Si E_1 et E_2 sont vides, $E = E_0$, $E'_Z = A_Z$, $E''_Z = B_Z$, mais A_Z et B_Z ne sont jamais vides, puisque les partitions utiles sont distinctes de X et Y.

Enfin, si $E_1 = \emptyset \neq E_2$, A_Z ne peut être constamment vide en raison de la propriété plus forte suivante :

Pour toute classe T de partitions utiles liées à X et Y par une formule (1), les deux familles A_Z et B_Z ont toutes deux des intersections vides et des réunions égales à E_0. Autrement dit, tout i de E_0 appartient à certains ensembles A_Z (avec $i\,Z = i\,X$) et certains ensembles B_Z (avec $i\,Z = i\,Y$).

En effet E_7 étant inclus dans E_2, il est exclu que $i\,X = i\,Y$; il existe un objet j tel que $|X_{ij} - Y_{ij}| = 1$. On sait alors, selon c) de 3.2. que Z_{ij} atteint les deux valeurs 0 et 1. Cela veut dire que certaines classes $i\,Z$ incluent j, d'autres non, et prouve que, quand Z parcourt T, $i\,Z$ atteint les deux valeurs possibles $i\,X$ et $i\,Y$. Q.E.D.

4. PROPORTIONS DES ARETES ET DES DIAGONALES

4.1. Notations.

Le polyèdre P_n comporte B_n sommets. Le nombre de Bell, B_n ou B(n), est connu par de nombreuses propriétés, notamment la récurrence :

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k \qquad (10) \;**$$

qui traduit la possibilité de distinguer un élément parmi n + 1 et l'effectif de sa classe : m = n + 1 - k.

---

* Nota : le cas particulier, E_0 vide, mérite quelque attention : on a au choix cinq caractérisations :
1 - L'existence de la partition en deux zones E_1 et E_2, où etc...
2 - Une X-classe et une Y-classe sont disjointes ou comparables.
3 - Toute X Y classe est soit une X-classe, soit une Y-classe.
4 - Les deux classes d'un même élément x sont comparables.
5 - La relation "X ou Y" est transitive.
Nous dirons alors que X et Y sont *localement comparables*. (Deux telles partitions ont été appelées associables par Ore, N.D.L.R.).

** L'auteur a trouvé utile de numéroter les formules de ce paragraphe à partir de (10). NDLR.

Il y a $S_n = B_n\left(B_{n-1}\right) / 2$ segments qui joignent une paire de sommets, dont A_n seront des arêtes, les autres des diagonales. Il est intéressant de calculer A_n / S_n car on peut démontrer que cette proportion α_n tend vers 1 quand n tend vers l'infini.

On a montré au paragraphe 3 que deux partitions sont adjacentes si et seulement si leurs restrictions ne diffèrent que dans une seule région de leur borne supérieure, où en notant X et Y ces restrictions sur R :

$\overline{X \cup Y} = 1_R^*$, et si X et Y sont comparables, la plus fine admet deux classes dans R. (11)

Pour R fixé, d'effectif r, le nombre des restrictions possibles (égales) à E - R est B(n - r). Soit alors C(r) celui des paires de restrictions à R. Il ne dépend que de l'effectif r, par conséquent :

$$A_n = \sum_0^n \binom{n}{r} B(n-r)\ C(r) \qquad (12)$$

C(r) est le nombre des paires {X,Y} de partitions de R, vérifiant (11). Il est plus facile de calculer par récurrence :

D(r) = nombre de paires telles que $\overline{X \cup Y} = 1_R$ (11bis)

où figurent en trop des paires comparables non adjacentes.
Le nombre de paires comparables est B_{r-1} , puisque la moins fine, soit Y, vaut $\overline{X \cup Y} = 1$, l'autre X, est quelconque mais différente de Y. Parmi ces partitions X, les seules adjacentes à Y sont en nombre $2^{r-1} - 1$, nombre de partitions de R en 2 classes non vides. Ainsi :

$$C(r) = D(r) - B(r) + 2^{r-1}. \qquad (13)$$

4.2. <u>Récurrence</u>.
Maintenant D(r) peut se calculer par une récurrence sur r analogue à (10).

Considérons toutes les S_n paires de {X,Y} de partitions de E, et distinguons un objet $a \in E$ et sa région :

$R = a\ \overline{X \cup Y}$, classe de $\overline{X \cup Y}$ incluant a.

On peut dénombrer ces paires, pour R fixé d'effectif r = n - t. Les restrictions à R doivent être une paire vérifiant (11bis), ou bien être confondues et égales à 1_R. Celles à E - R sont une paire quelconque, en nombre S_t, ou bien confondues en nombre B_t. Mais les restrictions ne peuvent être partout confondues. Le nombre annoncé est donc :

$$D_r(S_t + B_t) + S_t.$$

---

* 1_R désigne, rappelons-le, la partition en 1 seule classe ; et le mot région n'importe quelle classe de la borne supérieure des 2 partitions.

En distinguant maintenant tous les R possibles ($a \in R \subset E$) nous avons : $\binom{n-1}{t}$ parties R d'effectifs $r = n-t$, incluant a, et :

$$S_n = \sum_{\substack{t=0 \\ r=n-t}}^{n-1} \binom{n-1}{t} \left[D_r(S_t + B_t) + S_t\right] . \qquad (14)$$

Le premier terme, en $t=0$, vaut D_n : $S_0 = 0$, mais $B_0 = 1$ car l'ensemble vide comporte une seule partition. (14) nous permet donc de calculer D_n si D_r est connu pour $r = 1, 2, \ldots, n-1$.

4.3. Algorithme et résultats numérique.

Pour ce calcul progressif de D_n, A_n et α_n, l'organisation la plus naturelle sur machine est la suivante. Les formules utilisées sont, dans la boucle de calcul entre n et $n' = n+1$:

$$S_n = \sum_{\substack{t=0 \\ (t=n-r)}}^{n-1} \binom{n-1}{t} \left[D_r(S_t + B_t) + S_t\right] \qquad (14) \quad \text{(qui donne } D_n\text{)}$$

$$C_r = D_r - B_r + 2^{r-1} \qquad (13) \quad \text{(qui donne } C_n\text{)}$$

$$A_n = \sum_0^n \binom{n}{r} B_{n-r}\, C_r \qquad (12) \quad \text{(qui donne } A_n\text{)}$$

$$S_n = B_n \frac{B_{n-1}}{2} , \quad \text{et} \quad B_{n+1} = \sum_0^n \binom{n}{r} B_r \qquad (10\text{bis})$$

et enfin :

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1} . \qquad (15) \text{ pour la matrice de Pascal des coefficient binomiaux.}$$

Puisque l'argument n de ces derniers, progresse de 1 entre (13) et (12), nous pouvons ne mettre en mémoire que la ligne :

$K(r) = \binom{n-1}{r}$, à condition d'insérer sa mise à jour, par (15), entre (14) et (12). L'ordre des calculs le long de la boucle sera alors :

(10), puis (14), (13) et (15), puis (12) et

et $n = n+1$, retour à (10)...,

Tel est le principe du programme FORTRAN "ARETE" (tableau 2), où les valeurs initiales correspondant à $n=2$, sont facilement calculées à la main.

Nous obtenons :

Pour n =	0	1	2	3	4	5	6	7	8
B_n =	1	1	2	5	15	52	203	877	4140
S_n =	0	0	1	10	105	1326	20503	384126	8567730
D_n =	0	0	1	7	62	741	11327	213102	4805940
C_n =	0	0	1	6	55	705	11156	212289	4801930
A_n =	0	0	1	9	91	1150	17861	333858	7403280
$S_n - A_n$ =	0	0	0	1	14	176	3642	50268	1164450
$1-\alpha_n = \delta_n$			0	0.1	0.3333	0.13273	.12885	.130863	.135911

La proportion de diagonales, $\delta_n = 1-\alpha_n$ est seule présentée, et nous avons calculé cette proportion jusqu'à $n=42$ (voir tableau 1).

5. DISTANCES ET DIAMETRES.

Le diamètre d'un polyèdre est classiquement défini comme la plus grande distance qui puisse séparer deux sommets.

Si l'on prenait la distance Euclidienne dans $R^{n(n-1)/2}$, le diamètre de P_n serait évidemment celui de l'hypersphère et de l'hypercube circonscrit : $\sqrt{\frac{n(n-1)}{2}}$, car les partitions extrêmales "0" et "1" sont diamétralement opposées sur cette sphère. Ce sont d'ailleurs les seules : la négation d'une relation transitive R ne peut l'être également que si R = $\emptyset$ ou E×E. Mais, en général, la distance suivante est jugée plus intéressante. On appelle chemin toute suite d'arêtes séquentiellement adjacentes. Leur nombre est la longueur du chemin. Deux sommets X et Y sont les extrêmités de nombreux chemins dont certains sont les plus courts possibles. Cette longueur minima définit la distance D(X,Y).

Nous allons voir que, sur P_n, le diamètre : $D = \sup_{X,Y} D(X,Y) = 3$, et préciser les distances pour toutes les paires X,Y.

Ce résultat simple semble intuitivement lié à la forte proportion d'arêtes. Considérons d'abord les deux partitions extrêmales du treillis 0 et 1.

$D[0,1] = 3$ si $n \geq 4$, pour la raison suivante.

D'abord une distance 2 suppose l'existence d'une partition A adjacente à la fois à 0 et 1. A doit donc réunir une seule paire a, b et comporter deux classes, ab et c. Donc E = {a,b,c} et n = 3 (de même D = 1 pour n = 2). Ensuite, pour $n \geq 4$, tout chemin allant de 0 à 1 débute par une partition A qui réunit une seule paire d'objets (a,b) un "atome" du treillis et se termine, avant "1", par une partition B en deux classes, une "dichotomie". Mais A et B peuvent être adjacentes ; il suffit que B sépare a de b. Ce cas de figure fournit tous les chemins de longueur 3 allant de 0 à 1, et prouve que $D(0,1) \leq 3$.

Soit ensuite X une partition en x classes, $0 < x < n$. Alors $D(1,X) = \inf(x-1,2)$. C'est évident pour $x < 3$. Sinon, soit ab une paire d'objets réunis par X. Il existe une dichotomie B qui les sépare. X et B sont visiblement adjacents avec E entier pour région caractéristique : $X \cup B$ est un graphe connexe. Puisque B est adjacente à 1, $D(1,X) = 2$. Par contre $D(0,X)$ atteint 3. Cette distance est 1 si $x = n-1$, 2 si Y comporte au plus deux classes d'effectif > 1. Soient en effet a et b deux représentants des deux classes prépondérantes {a,b} définit un atome A adjacent à X comme à 0 : X et A ne diffèrent que dans l'union des deux classes aX et bX. Partout ailleurs les restrictions sont celles de 0. D = 2 devient impossible si X comporte plus de deux classes d'effectifs > 1. Il faudrait qu'un atome A, où une seule paire (a,b) est réunie, soit adjacent à X. Mais A et X vont différer dans plusieurs régions, d'abord $aX \cup bX$, puis toutes les autres X-classes d'effectifs > 1. Ainsi $D(0,X)$ est ≥ 3. Cela dit nous pouvons construire des chemins de longueur 3. Il suffit d'exhiber des dichotomies B adjacentes à X, puisque $D(0,B) = 2$. Or il suffit que B sépare une des paires réunies par X pour que $B \cup X$ soit connexe, ce qui entraîne l'adjacence. Ainsi $D(0,X) = 3$.

Enfin prenons deux partitions X et Y toutes deux $\neq 0$, alors $D(X,Y) = 2$ pour une raison analogue. Il existe une dichotomie B adjacente à X et à Y, car X réunit au moins une paire a,b , Y au moins une paire c,d et B peut séparer a de b, et c de d. Ce cas, le plus général, est le plus vite analysé. Après cet inventaire complet, nous voyons que : le diamètre de P_n, pour $n \geq 4$ est toujours égal à 3, et que de plus, les distances $D(X,Y)$ sont :

- 0 si X = Y
- 1 pour les paires adjacantes
- 3 pour les paires {0,X} ssi X comporte plus de 2 classes d'effectifs > 1 , ou si X = 1 et n > 3
- 2 dans tous les autres cas.

Les diamètres de P_4, P_3, P_2, P_1 sont évidemment 3, 2, 1, 0 atteints seulement par la paire extrêmale {0,1}.

On voit ainsi que le rayon du graphe d'adjacence est égal à deux, avec pour "centres" les partitions qui comportent au moins deux classes, dont au plus deux réunissent plusieurs objets. On les appellera des bicliques.

n	$1 - \alpha_n = \delta_n$	$B(n + 1)$
3	.1	15
4	.133333	52
5	.13273	203
6	.128859	977
7	.130863	4140
8	.135911	21147
9	.140051	115975
10	.141542	678570
11	.140294	0.421360E+07
12	.136866	0.276444E+08
13	.131918	0.190899E+09
14	.126006	0.138296E+10
15	.119559	0.104801E+11
16	.1129	0.828648E+11
17	.106246	0.682076E+12
18	0.997614E-01	0.583273E+13
19	0.935413E-01	0.517241E+14
20	0.876606E-01	0.474869E+15
21	0.821539E-01	0.450670E+16
22	0.770226E-01	0.441519E+17
23	0.722860E-01	0.445958E+18
24	0.679174E-01	0.463858E+19
25	0.639101E-01	0.496311E+20
26	0.602398E-01	0.545716E+21
27	0.568845E-01	0.616051E+22
28	0.538028E-01	0.713396E+23
29	0.509914E-01	0.846746E+24
30	0.484191E-01	0.102933E+26
31	0.460641E-01	0.128064E+27
32	0.439052E-01	0.162959E+28
33	0.419031E-01	0.211950E+29
34	0.400717E-01	0.281599E+30
35	0.383821E-01	0.381970E+31
36	0.368201E-01	0.528681E+32
37	0.353726E-01	0.746286E+33
38	0.340284E-01	0.107388E+35
39	0.327753E-01	0.157450E+36
40	0.316113E-01	0.235114E+37
41	0.305192E-01	0.357423E+38
42	0.294969E-01	0.552947E+39

Tableau 1. Proportion δ_n de diagonales et le nombre $B(n+1)$ de partitions.

```
ARETE

DIMENSIØN B(150),C(150),D(150),K(150),S(150)
REAL K
READ(5,*)M
B(1)=1
B(2)=2
C(2)=1
D(2)=1
S(2)=1
K(1)=2
K(2)=1
N=3
N1=2
U=5
10 B(N)=U                          Forme S(n) selon (10 bis)
U=U*(U-1)/2
S(N)=U
DØ 14 I=1,N1
14 U=U-K(I)*(D(N-I)*(S(I)+B(I))+ S(I))   D(n) selon (14)
D(N)=U
C(N)=D(N)-B(N)+2.**N1              C(n) selon (13)
I=N
K(I)=1
15 I=I-1
K(I)=K(I)+K(I-1)                   K(I) = (N I), selon (15)
IF(I-2) 16,16,15
16 K(1)=N
U=1+N +B(N)                        Forme U = B(n + 1) selon (10)
A=C(N)
DØ 12 I=2,N1                       et A = A(n) selon (12)
U=U+K(I)*B(I)
12 A=A+K(I)*B(N-I)+C(I)
A=1.-A/S(N)     = δ(n)
WRITE(6,*) N,A ,U                  édite n, δ(n), B(n + 1)
N1=N
N=N+1                              Pour n < m, retour à 10 bis)
IF(N-M) 10,10,99
99 STØP
END
```

Tableau 2. Programme Fortran "ARETE".

6. FACES ET FACETTES.

6.1. Introduction et notations.

On appelle *face* d'un polyèdre $P \subset R^d$ chaque partie F de P où une forme linéaire se trouve maximisée.

La forme $L \equiv 0$ donne $F = P$. Sinon la dimension de F est au plus $d-1$. Les $(d-1)$ faces s'appellent des *facettes*. Toute face de dimension $n-k$ est intersection de k facettes (de plusieurs façons en général) et toute intersection de faces (ou de facettes) est une face * .

Nous sommes alors tentés de chercher toutes les facettes de P_n pour connaître toutes les faces. Ces deux objectifs se révèlent hors d'atteinte car l'inventaire des facettes est trop complexe. 6.2. et 6.3.1. donnent un inventaire partiel des facettes. Dans 6.3.3. sera caractérisée une classe de faces particulières : les tronçons. On donne enfin dans 6.4. un "théorème de restriction" qui caractérise les facettes de P_n en fonction du support des vecteurs normaux associés.

Notations.

Soit $E = \{1, 2, \dots, n\}$ noté $[n]$. A toute forme linéaire $X \to LX$ définie sur $R^{\frac{n(n-1)}{2}}$ nous associons le nombre $\mathrm{MaxL} = \mathrm{MaxLX}$ pour $X \in P_n$ et l'ensemble $F = \mathrm{XamL} = \{X \in P_n \mid LX = \mathrm{MaxL}\}$. De la sorte, chaque face F est le Xam d'une forme L (et de toutes les formes hL pour $h > 0$).

L'équation $LX = \mathrm{MaxL}$ est appelée *équation de la face* F. Le vecteur ligne L, qui suffit à définir F s'appelle un *vecteur normal* à F.

Nous notons (e_{ij}) les éléments de la base canonique $R^{\frac{n(n-1)}{2}}$ (e_{ij} correspond à la relation réduite à la paire (i,j)), et nous notons (e^*_{ij}) les éléments de la base canonique du dual. Ainsi une forme linéaire L s'écrira $L = \Sigma\, l_{ij}\, e^*_{ij}$ et l'on a, pour $X \in R^{\frac{n(n-1)}{2}}$: $LX = L(X) = \Sigma\, l_{ij}\, X_{ij}$.

6.2. Facettes qui incluent l'origine 0.

On pose $m = \frac{n(n-1)}{2}$. Si $0 \in F$, on a $\mathrm{MaxL} = 0$, donc pour tout (ij), $l_{ij} \leq 0$. Montrons alors que F est une facette ssi un seul l_{ij} est < 0. Montrons

---

* On considère souvent la partie vide $\emptyset$, intersection de toutes les faces comme une face de dimension -1. Nous serons amenés à le convenir ici aussi, bien qu'aucune forme linéaire ne soit maximisée sur cette face.

d'abord que, par exemple, $X_{ij} = 0$ est l'équation d'une facette F_{ij} : c'est une face parce que $0 = \text{MaxL}$ pour $LX = -X_{ij}$, et elle est de dimension $m-1$ parce qu'elle inclut toutes les partitions atomiques (en $n-1$ classes) définies par les $m-1$ paries (k,l) autres que (i,j) et que ces partitions sont linéairement indépendantes. Inversement, si plusieurs l_{ij} sont <0, la face $F = \text{XamL}$ est l'intersection des facettes F_{ij} associées. F ne peut être une facette parce que plus généralement si $C = A \cap B$, les fermetures affines vérifient $\overline{C} \subseteq \overline{A} \cap \overline{B}$. La fermeture affine de F ne peut donc pas être un hyperplan. Ces facettes F_{ij} sont en nombre m, comme leurs vecteurs normaux, $-e_{ij}$. Pour les autres facettes on a $\text{MaxL} > 0$. Leur structure est plus compliquée.

6.3. Facettes qui n'incluent pas l'origine.

6.3.1. Facettes associées aux contraintes de transitivité.

Soit $L(X) = x_{ik} + x_{kj} - x_{ij}$, i, j et k distincts. Montrons que l'équation $x_{ik} + x_{kj} - x_{ij} = \text{MaxL} = 1$ définit une facette à laquelle on peut associer la contrainte de transitivité $x_{ik} + x_{kj} - x_{ij} \leq 1$.

Pour montrer que F = XamL est une facette, on vérifie d'abord que F contient les équivalences suivantes, notées chacune par leurs classes d'équivalence :

$e_{ik} = (i,k)$

$e_{jk} = (j,k)$

(i,j,k)

(i,k,l), (j,k,l) et (i,j,k,l), pour tout $l \notin \{i,j,k\}$

(i,k), (l,p), pour tout $l \notin \{i,j,k\}$, $p \notin \{i,j,k,l\}$

à l'aide desquelles on peut exprimer linéairement tout vecteur de la base canonique de $R^{\frac{n(n-1)}{2}}$.

En effet, $e_{ik} \in F$, $e_{jk} \in F$ et l'on a dans le vectoriel $R^{\frac{n(n-1)}{2}}$

$e_{ij} = (i,j,k) - e_{ik} - e_{jk}$

$e_{il} + e_{jl} + e_{kl} = (i,j,k,l) - (i,j,k)$, pour tout $l \notin \{i,j,k\}$

$e_{jl} + e_{kl} = (j,k,l) - (j,k)$

et e_{il} s'obtient en retranchant l'une de l'autre les deux égalités précédentes.

Enfin :

$e_{lp} = \{(i,k),(l,p)\} - (i,k)$, pour tout $l \notin \{i,j,k\}$ et pour tout $p \notin \{i,j,k,l\}$.

6.3.2. Autres facettes de la même espèce.

Plus généralement, nous avons pu démontrer que l'équation

$$L(X) = - \sum_{2 \leq i \leq j \leq n} x_{ij} + (k-1) \sum_{2 \leq i \leq n} x_{1i}$$

définit, pour chaque valeur de $k > 2$, une facette de P_n. Etablissons d'abord deux lemmes.

Lemme 1. Soit $[n] = \{1, 2, \ldots, n\}$, $n > 3$. Soit k un entier > 2, différent de n et de $n-1$. L'ensemble $\mathcal{K}_k$ des cliques à k sommets pris dans [n] forme un système générateur de $R^{\frac{n(n-1)}{2}}$.

Démonstration. Pour chaque sous-ensemble A (resp.B, resp.C,...) inclus dans [n], nous désignons par $A^{[2]}$ (resp. $B^{[2]}$, resp. $C^{[2]}$, ...) la relation constituée par toutes les paires d'éléments de A. Les valeurs de k exclues par le théorème le sont pour des raisons de dimensionalité évidentes. Pour les autres valeurs, il suffit de montrer que l'on peut exprimer chaque élément e_{ij} de la base canonique de $R^{\frac{n(n-1)}{2}}$ à l'aide des fonctions indicatrices (notées $A^{[2]}$, $B^{[2]}$,...) des relations correspondant aux cliques considérées. Fixons i et j et partitionnons $\mathcal{K}_k$ en 4 ensembles $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$:

$$\mathcal{A} = \{A^{[2]} \ / \ i \in A, j \in A\} \qquad \mathcal{C} = \{C^{[2]} \ / \ i \notin C, j \in C\}$$

$$\mathcal{B} = \{B^{[2]} \ / \ i \in B, j \notin B\} \qquad \mathcal{D} = \{D^{[2]} \ / \ i \notin D, j \notin D\}$$

L'expression cherchée est de la forme :

$$\binom{n}{k-2} e_{ij} = \sum_{\mathcal{A}} A^{[2]} + \alpha \left(\sum_{\mathcal{B}} B^{[2]} + \sum_{\mathcal{C}} C^{[2]} \right) + \beta \sum_{\mathcal{D}} D^{[2]}$$

où α et β doivent satisfaire aux équations :

$$0 = \binom{n-3}{k-3} + \alpha \binom{n-2}{k-3} \qquad \text{et} \qquad 0 = \binom{n-2}{k-2} + \beta \binom{n-2}{k-2}$$

Lemme 2. Supposons que le cardinal n soit supérieur à 2 et soit k un entier compris entre 2 et n-3. L'ensemble des cliques à k ou k+1 sommets pris dans [n] et incluant un élément $a \in [n]$, fixé, forme un système générateur de $R^{\frac{n(n-1)}{2}}$.

Démonstration. Nous allons montrer que pour tout $A = \{a_1, a_2, \ldots, a_k\} \subset [n]$, $A \not\ni a$, la clique $A^{[2]}$ est engendrée linéairement par la famille des cliques considérées. Comme les cliques à k sommets incluant a appartiennent déjà à cette famille, le lemme 2 résultera alors du lemme 1.

Posons $A_i = A - a_i \cup \{a\}$, $1 \le i \le k$ et $A_o = A \cup \{a\}$.

L'équation $A^{[2]} = \lambda A_o^{[2]} + \mu \sum_{1 \le i \le k} A_i^{[2]}$ équivaut aux équations

$$A^{[2]}(a,a_i) = 0 = \lambda + \mu(k-1), \quad 1 \le i \le k \ ;$$

$$\text{et} \quad A^{[2]}(a_i,a_j) = 1 = \lambda + \mu(k-2), \quad 1 \le i < j \le k.$$

qui ont une solution évidente.

Nous venons maintenant à l'équation générale de la nouvelle classe de facettes que nous avons mises en évidence.

Théorème 2.1. Soit k un entier vérifiant $2 \le k \le n-3$. Alors

$$L(X) = - \sum_{2 \le i < j \le n} x_{ij} + (k-1) \sum_{2 \le i \le n} x_{1i}$$

est l'équation d'une facette de P_n.

Démonstration. Il est clair que L(X) est maximisée par des cliques contenant 1. Soit donc X une clique contenant 1 et de taille h. On a :

$$2L(X) = -(h-1)(h-2) + 2(k-1)(h-1) = \varphi(h).$$

On a $\varphi(h) = \varphi(k+1)$ ce qui suffit à établir que le maximum du trinôme $\varphi(h)$ est atteint exactement pour $h = k$ et $h = k+1$.

6.3.3. Tronçons.

Sur l'ensemble des relations symétriques chaque forme $L(X) = \Sigma\, l_{ij} X_{ij}$ est maximisée par toutes les relations intermédiaires entre :

$$A_{ij} \iff l_{ij} > 0 \quad \text{et} \quad B_{ij} \iff l_{ij} \ge 0 .$$

Si certaines de ces relations sont transitives, elles constituent nécessairement la face de P_n : F = XamL.

Pour cela, il faut et il suffit, puisque $A \subset B$, que la fermeture transitive $\overline{A}$ soit elle aussi incluse dans B. Désignant par $B_{n,k}$ le nombre de relations d'équivalences sur [n] comprenant k classe non-vides, nous voyons que le nombre de ces faces particulières, appelées des *tronçons* est égal à

$\sum_{1 \leq k \leq n} B_{n,k} \, 2^{\frac{k(k-1)}{2}}$. Nous allons préciser la dimension de ces faces.

Définitions. Soit A une équivalence sur [n]. Une relation R sur [n] est dite *divisible* par A si son indicatrice R(x,y) ne dépend que des classes de x et de y dans A. L'ensemble $\mathcal{R}_A$ des relations divisibles par A est en correspondance biunivoque avec l'ensemble quotient $\mathcal{R}_A/A$ formé des relations dans l'ensemble [n]/A. Si R et S appartiennent à $\mathcal{R}_A$, $R \wedge S(x,y)$ et $R \vee S(x,y)$ ne dépendent eux aussi que des classes de x et de y dans A. Ainsi l'ensemble des relations divisibles par A et plus fines qu'une relation donnée B forme un sous-treillis de l'ensemble des relations dans [n]. Ce sous-treillis possède un plus grand élément qui est la réunion de tous ses éléments et que nous noterons B_A. B_A n'est pas nécessairement une équivalence. On a précisément :

$(x,y) \in B_A \Longleftrightarrow (z \in xA \;\text{ et }\; t \in yA \Rightarrow (z,t) \in B)$.

Théorème 2.2. Soit A une équivalence sur [n], B une relation sur [n] moinsfine que A. L'ensemble T_A^B des partitions intermédiaires entre A et B forme une face de P_n. La dimension de cette face est égale à $\mathrm{Card}(B_A/A)$.

Démonstration. Notons d'abord que si R est une équivalence moins fine que A, alors nécessairement R est divisible par A. En effet soient x,y,z et t des éléments de [n] vérifiant $z \in xA$ et $t \in yA$. On a :

$$R(z,t) \geq R(z,x)\, R(x,y)\, R(y,t) \geq R(x,y)$$

et de même :

$$R(x,y) \geq R(z,t),$$

soit en définitive :

$$R(z,t) = R(x,y)$$

qui exprime la divisibilité de R par A. Ainsi les partitions intermédiaires entre A et B coïncident avec les partitions intermédiaires entre A et B_A.

Considérons maintenant la forme linéaire

$$L_{A,B} = \sum_{(i,j) \in A} e^*_{ij} - \sum_{(k,l) \in [n]^{[2]} - B_A} e^*_{kl}$$

Clairement, cette forme n'est maximisée que par les partitions intermédiaires entre A et B, ce qui démontre la première assertion du théorème. La seconde assertion résulte de ce que les $|B_A/A|$ partitions obtenues à partir de A en réunissant deux classes réunies dans B_A, pour tous les choix possibles de ces deux classes, sont linéairement indépendantes et engendrent avec A, T_A^B linéairement.

On peut préciser davantage dans le cas où B est la relation pleine. Les partitions intermédiaires entre A et B sont alors les partitions moins fines que A et l'on a de plus :

<u>Théorème 2.3</u>. Soit A une partition de [n] en h classe non vides, Φ_h l'ensembles des faces de P_h, $\mathcal{F}_A^n$ l'ensemble des face de P_n formées de partitions moins fines que A. On a l'équivalence :

$$F = \{B_i\} \in \mathcal{F}_A^n \iff \varphi = \{B_i/A\} \in \Phi_n$$

où B_i/A désigne la relation d'équivalence induite dans [n]/A par B_i.

Démonstration. Soient $\alpha_1, \alpha_2, \ldots, \alpha_h$ les classes de A, F une face de P_n formée de partitions moins fines que A, L une forme linéaire associée à F :

$$L = \Sigma \lambda_{kl} e_{kl}^* .$$

Définissons L' par

$$L' = \sum_{1 \leq i \leq h-1} \sum_{i < j \leq h} \lambda_{ij} e_{ij}^*$$

où

$$\lambda_{ij} = \sum_{k \in \alpha_i} \sum_{l \in \alpha_j} \lambda_{kl} .$$

Soit B une partition de [n] moins fine que R. On a

$$L'(B/A) = L(B) - \mu$$

où μ est une constante : $\mu = \sum_i \sum_{k \in \alpha_i} \sum_{l \in \alpha_i, l > k} \lambda_{kl}$.

Ainsi L' est maximisée sur P_h par les partitions de la forme B/A, pour B maximisant L, c'est-à-dire pour $B \in F$, ce qui établit l'implication du théorème dans le premier sens.

Pour établir l'implication réciproque, considérons une face φ de P_h et soit L' une forme linéaire associée à φ.

$$L' = \Sigma \lambda'_{ij} e_{ij}^* .$$

Associons à L' la forme linéaire L définie par :

$$L = \sum_{1 \le i < j \le h} X_{ij} \sum_{k \in \alpha_i,\, l \in \alpha_j} e^*_{kl} + \upsilon \sum_i \sum_{k \in \alpha_i} \sum_{l \in \alpha_i,\, l > k} e^*_{kl}$$

où υ est une très grande constante positive, ce qui implique que L ne peut être maximisée que par les partitions moins fines que R. Pour une telle partition B_i, on a :

$$L(B_i) = L'(B_i/R) + \upsilon'$$

où υ' est une constante.

Ainsi les partitions maximisant L sont exactement celles induites sur [n] par les partitions de P_h maximisant L', c'est-à-dire par les partitions formant φ. Q.E.D.

6.4. Théorème de restriction.

Appelons *support* d'une forme linéaire $L = \Sigma\, \lambda_{ij}\, e^*_{ij}$ l'ensemble des indices i (et j) pour lesquels on a $\lambda_{ij} \neq 0$ pour au moins un j (un i). On a le théorème suivant :

Théorème. Soit $L = \Sigma\, \lambda_{ij}\, e^*_{ij}$ une forme linéaire dont le support est l'ensemble $[m] = \{1,2,\dots,m\}$, $m \geq 2$.

On a l'équivalence :

L est l'équation d'une facette de $P_m \Leftrightarrow$ L est l'équation d'une facette de P_n pour tout $n > m$.

Démonstration. Supposons que L soit l'équation d'une facette de P_{m+1}. Alors l'ensemble $\mathcal{A} = \{A_1,\dots,A_q\}$ des partitions de [m+1] pour lesquelles L est maximum génère $R^{\frac{m(m+1)}{2}}$ linéairement. Il en résulte que l'ensemble $\mathcal{B} = \{B_1,\dots,B_q\}$ des restrictions des éléments de $\mathcal{A}$ à l'ensemble [m] génère linéairement $R^{\frac{m(m+1)}{2}}$. En effet toute relation de la forme $e_{ij} = \Sigma\, \lambda^k_{ij}\, A_k$ où A_k désigne la fonction indicatrice du graphe de la relation A_k et i et j appartiennent à [m], implique la relation analogue pour les restrictions

$$e_{ij} = \Sigma\ \lambda_{ij}^{k}\ B_k\ .$$

Cela suffit à établir que $\mathcal{B}$ est une facette de P_m.

Supposons maintenant que L soit l'équation d'une facette de P_m. Soit $\mathcal{A} = \{A_1, A_2, \ldots, A_q\}$ l'ensemble des sommets de cette facette. Soit pour $1 \leq i \leq q$ $\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in_i}$ les classes d'équivalence de A_i (comme sous-ensembles de [m]), $a_{i1}, a_{i2}, \ldots, a_{in_i}$ les relations d'équivalence correspondantes. L'ensemble des a_{ij} génère aussi $R^{\frac{m(m-1)}{2}}$ car s'il n'en était pas ainsi, les A_i qui sont des combinaisons linéaires des a_{ij} :

$$A_i = \sum_{1 \leq j \leq n_i} a_{ij}\ , \quad 1 \leq i \leq q$$ ne pourraient pas le générer.

Soit $\mathcal{B}$ l'ensemble des partitions de [m+1] dont les restrictions à [m] coïncident chacune avec l'un des $A_i \in \mathcal{A}$. $\mathcal{B}$ contient $\mathcal{A}$ et, pour $1 \leq i \leq q$ et $1 \leq k \leq n_i$, la partition notée A_i^k dont les classes notées α_{ij}^k sont définies par $\alpha_{ij}^k = \alpha_{ij}$ pour $j \neq k$, et $\alpha_{ik}^k = \alpha_{ik} \cup \{m+1\}$.

Montrons que la réunion des A_i et des A_i^k génère linéairement $R^{\frac{m(m+1)}{2}}$. Il suffit pour cela de montrer que l'on peut exprimer les $e_{i,m+1}$, $1 \leq i \leq m$, linéairement en fonction de cet ensemble. On a d'abord pour $1 \leq i \leq q$, $1 \leq j \leq n_i$

$$\sum_{l \in \alpha_{ij}} e_{l,m+1} = A_i^j - A_i\ .$$

Comme les $a_{i,j}$ génèrent linéairement $R^{\frac{m(m-1)}{2}}$, il résulte d'un lemme auxiliaire (démontré ci-dessous) que les $\alpha_{i,j}$ génèrent linéairement R^m. Ainsi, désignant par $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m)$ la base canonique de R^m, on a :

$$\varepsilon_k = \Sigma\ \lambda_{ij}^k\ \alpha_{ij} \quad 1 \leq k \leq n\ .$$

Définissons l'application $\tau\ :\ R^n \times R^n \to R^{\frac{n(n-1)}{2}}$ par les conditions

$$\tau\ (\varepsilon_i, \varepsilon_j) = \tau\ (\varepsilon_j, \varepsilon_i) = e_{ij},\ 1 \leq i < j \leq n\ ; \quad \tau \text{ est bilinéaire.}$$

On a : $e_{k,m+1} = \tau(\varepsilon_k, \varepsilon_{m+1}) = \Sigma\, \lambda^k_{ij}\ \tau(\alpha_{ij}, \varepsilon_{m+1})\ ,\ 1 \leq k \leq n$

$$e_{k,m+1} = \Sigma\, \lambda^k_{ij} \sum_{l \in \alpha_{ij}} \tau(\varepsilon_l, \varepsilon_{m+1}) = \Sigma\, \lambda^k_{ij} \sum_{l \in \alpha_{ij}} e_{l,m+1}$$

$$e_{k,m+1} = \Sigma\, \lambda^k_{ij}\ (A^j_i - A_i)$$

qui est la relation linéaire cherchée. La démonstration s'achève par récurrence.

Enoncé et démonstration du lemme auxiliaire.

Soit $E = [n] = \{1,2,\ldots,n\}$. Les C_i, $i = 1,2,\ldots m$, éléments de $\{0,1\}^n \subset R^n$ indiquent des parties de $[n]$. Les γ_i, $i = 1,2,\ldots,m$, éléments de $\{0,1\}^{\frac{n(n-1)}{2}} \subset R^{\frac{n(n-1)}{2}}$, indiquent les équivalences associées :

$\gamma_i(k,l) = 1 \iff C_i(k) = 1$ et $C_i(l) = 1$.

Lemme. Pour $n > 2$, (γ_i) générateur du vectoriel $R^{\frac{n(n-1)}{2}}$ implique (C_i) générateur du vectoriel R^n.

Démonstration. Soit (e_i) la base canonique de R^n, (e_{ij}) la base canonique de $R^{\frac{n(n-1)}{2}}$. On a, si (γ_i) est générateur :

$$e_{12} = \sum_1^m \lambda^t\ \gamma_t = u = \Sigma\, e_{ij}\ u_{ij}$$

où seul u_{12} est différent de zéro (et égal à 1).

Considérons $\mathcal{V}_i = \sum_j u_{ij} = 0$ si $i \neq 1$ et $i \neq 2$
$\qquad = 1$ si $i = 1$ ou $i = 2$.

On a : $\mathcal{V}_i = \sum_1^m \lambda^t \left(\sum_j \gamma_t\ (i,j)\right) = \sum_1^m \lambda^t\ (|t| - 1)\ C_t(i)$

où la dernière expression désigne naturellement la restriction de la somme

$$\sum_1^m \lambda^t\ (|C_t| - 1)$$

aux indices t pour lesquels i appartient à la partie de fonction indicatrice C_t.

Donc : $\mathcal{V} = \sum_{1}^{m} \lambda^t (|C_t| - 1)\ C_t = e_1 + e_2$.

Ainsi les C_t engendrent les paires $e_i + e_j = f_{ij}$. Mais les paires f_{ij}

engendrent tous les singletons car :

$e_1 + e_2 = f_{12}$

$e_1 + e_3 = f_{13}$

$e_2 + e_3 = f_{23}$

$f_{12} + f_{13} = 2e_1 + f_{23}$

permettent d'engendrer e_1. Q.E.D.

REFERENCE

[1] "Sur quelques aspects mathématiques de problèmes de classification automatique", I. aspect algébrique", 27cm, 17p., in *I.C.C. Bulletin*, Rome, 1965, vol.4.
repris dans *Mathématiques et Sciences humaines*, n°82, 1983, pp.13-29.