J. SERRA

N-82

NOTE GEOSTATISTIQUE N° 88

THEORIE DES GRAPHES ET MORPHOLOGIE

G. MATHERON

Décembre 1968

N-82

NOTE GEOSTATISTIQUE N° 88

THEORIE DES GRAPHES ET MORPHOLOGIE

Table

NOTE GEOSTATISTIQUE N° 88

THEORIE DES GRAPHES ET MORPHOLOGIE

REMARQUE PRELIMINAIRE

Les deux remarques qui suivent montrent l'intérêt que peut présenter la théorie des graphes pour la morphologie.

a/ Du point de vue expérimental, les techniques de l'analyse des textures, fondées sur un balayage discontinu, remplacent l'objet plan que l'on veut étudier par un réseau discret qu'il est toujours possible de considérer comme un graphe. Cette contrainte expérimentale ne nous oblige en aucune façon à limiter l'analyse théorique à l'élaboration de schémas discrets, plus pauvres que les schémas continus, et, en un certain sens, assez artificiels (la <u>maille</u> est choisie par l'opérateur, et ne constitue pas une caractéristique de l'objet). Mais, inversement, des raisonnements simples empruntés à la théorie des graphes permettent d'établir certains résultats généraux, valables pour tout schéma, discret ou non, et dont nous aurions tort de nous priver.

[On peut faire d'autres remarques, d'ailleurs, qui limitent la supériorité des schémas continus : c'est une supériorité de nature mathématique, qui provient simplement du fait que la structure topologique du plan est plus riche que celle d'un ensemble dénombrable muni (implicitement) de la topologie discrète. Mais, en un sens, cette structure topologique est trop riche pour la réalité physique. Elle incite à passer à des échelles de petitesse où l'objet que l'on étudie n'est plus défini, ou change de nature. Du point de vue probabiliste, d'autre part, décider au départ que l'on n'étudiera d'un objet que les propriétés mettant simultanément en jeu une infinité dénombrable au plus de points d'appui revient très exactement à étudier la fonction aléatoire correspondante à partir de sa loi spatiale, et de la σ-algèbre "maigre" qui lui est associée].

b/ Du point de vue théorique, d'autre part, il est important de distinguer clairement les propriétés <u>métriques</u> d'un objet (une covariance, ou un covariogramme, une traversée, une granulométrie etc...: on ne peut définir ces notions que sur un ensemble muni au moins d'une structure de groupe topologique et d'une mesure invariante par translation; dans un espace euclidien, on utilise à la fois la <u>norme</u> et la <u>mesure de Lebesgue</u>) et ses propriétés <u>topologiques</u> (invariantes pour les déformations continues). Un graphe n'étant défini qu'à un homéomorphisme près, toute propriété déduite de la théorie des graphes sera de nature topologique, et non métrique. Nous n'obtiendrons pas de cette manière toutes les propriétés topologiques de l'objet, mais probablement les plus simples et les plus importantes.

Donnons un exemple. Le nombre des composantes connexes d'un ensemble est une caractéristique topologique de cet ensemble (il s'agit du nombre absolu, et non du nombre spécifique dont la définition fait intervenir la mesure de Lebesgue). Or, une méthode variationnelle permet de rattacher ce nombre au covariogramme, notion typiquement métrique. Le covariogramme, cependant, ne sert que d'intermédiaire, et tout se ramène, en fait, au dénombrement de certaines configurations portant sur quatre points formant un carré ou un rectangle, c'est-à-dire, en définitive, à une caractéristique du graphe obtenu en prenant l'intersection de l'objet et d'un graphe à maille rectangulaire régulière. Comme il s'agit d'une propriété non métrique, la régularité de la maille ne doit jouer aucun rôle, et nous verrons qu'un graphe quelconque fait aussi bien l'affaire.

On peut aussi s'interroger sur la stationnarité. Ce n'est pas une propriété topologique, puisque les homeomorphismes ne la conservent pas. Elle ne fait intervenir, en principe, que la structure de groupe (invariance de la loi spatiale pour les translations), mais les applications les plus intéressantes de la théorie ergodique supposent l'existence d'une mesure invariante par translation (mesure de Lebesgue). C'est donc, en fait, une propriété métrique. Il y a lieu, pourtant, d'introduire aussi une notion plus faible, purement topologique, de la stationnarité. Les écailles d'une peau de serpent n'ont pas la même dimension moyenne, mais conservent les mêmes caractéristiques morphologiques sur tout le corps de l'animal. On peut

On peut aussi bien décrire une partition aléatoire du type écailles de serpent comme un graphe aléatoire : la stationnarité topologique doit exprimer, en un sens précis, que le graphe aléatoire présente les mêmes caractéristiques probabilités lorsqu'on le regarde à partir d'un quelconque de ses sommets. [On prendra garde que certaines propriétés non métriques, mais algébriques (liées à la structure d'espace vectoriel) de la partition aléatoire, comme, par exemple, la convexité des écailles de serpent, ne se conservent pas par homeomorphisme et ne constituent une propriété du graphe aléatoire stationnaire, bien que l'on puisse, à leur sujet aussi, envisager une certaine forme de stationnarité].

GRAPHE PLANAIRE

Nous appellerons <u>arète</u> un sous-ensemble du plan homéomorphe au segment ouvert (o,1) : c'est un arc de courbe continu, simple (sans point multiple), fini ou non. Si l'on compactifie le plan par la donnée du point à l'infini ω, toute arète possède alors deux extrémités [images de o et 1 par le prolongement de l'homéomorphisme au segment fermé (o,1)], distinctes ou non et pouvant éventuellement coïncider avec le point à l'infini ω. Ces extrémités sont les <u>sommets</u> adjacents à cette arète. Si les deux sommets adjacents à une arète sont confondus, on dit que l'arète est une <u>boucle</u>. On appelle <u>graphe</u> un ensemble (A, S) constitué d'une famille A d'arètes et de la famille S des sommets des arètes $a \in A$. Le graphe (A, S) est <u>planaire</u> si deux arètes quelconques dans A sont disjointes ou confondues (ou encore si tout point adhérent à deux arètes disjointes est un sommet) - On appelle <u>face</u> toute composante connexe de l'ensemble obtenu en soustrayant du plan les arètes de A et les sommets de S. Si F est la famille des faces du graphe (A, S), les familles A, S et F constituent une partition du plan. Une arète est adjacente à une face si elle contient un point adhérent à cette face. On dit qu'une face est d'ordre n si elle admet n arètes adjacentes distinctes exactement, qu'un sommet est d'ordre n s'il est adjacent à n arètes distinctes exactement. On dit que le graphe est fini si $\omega \notin S$ et si la famille A, et donc aussi les familles S et F sont finies; qu'il est <u>localement fini</u> si $\omega \notin S$ et si tout ensemble borné ne rencontre qu'un nombre fini d'arètes ; les sommets constituent alors un ensemble discret, tout sommet est d'ordre fini, toute face est un ensemble ouvert (mais n'est pas nécessairement d'ordre fini). Pour qu'un graphe localement fini soit fini, il faut et il suffit que l'ensemble S soit borné. Nous dirons qu'un graphe est d'<u>ordre fini</u> s'il est localement fini et si toute face est d'ordre fini. Sauf mention expresse du contraire, nous ne considérerons que des graphes localement finis (mais pas nécessairement d'ordre fini).

Un graphe (A', S', F') est un sous-graphe du graphe (A,S,F) si $A' \subset A$. On a alors $S' \subset S$, et toute face $f' \in F'$ est réunion de faces $f \in F$, d'arêtes $a \in A \cap \complement A'$ et de sommets $s \in S \cap \complement S'$ - On appelle <u>composante connexe</u> d'un graphe (A,S,F) tout sous-graphe (A',S',F') de (A,S,F) tel que la réunion des arêtes $a' \in A'$ et de leurs sommets $s' \in S'$ soit une composante connexe de l'ensemble $S \cup \bigcup_{a \in A} a$.

Un graphe est connexe, en particulier, si la réunion de ses arêtes et de ses sommets est un ensemble connexe.

RELATION D'EULER POUR LES GRAPHES FINIS

Soient a, s, f et ν respectivement les nombres d'arêtes, de sommets, de faces, et de composantes connexes d'un graphe fini (A,S,F) - On a :

(1) $$\boxed{s + f = 1 + a + \nu}$$

soient, en effet, $L_1, \ldots L_a$ les arêtes de A, et les familles $A_o = \emptyset \subset A_1 \subset \ldots \subset A_a = A$ $\quad (A_n = \{L_1, \ldots L_n\},\ o < n < a)$

Pour le sous-graphe A_o, on a $s_o = a_o = \nu_o = 0$ et $f_o = 1$, et (1) est vérifiée. Si (1) est vérifiée pour A_{n-1}, montrons qu'il en est encore ainsi pour A_n. Si les deux sommets de L_n sont distincts de ceux des arêtes de A_{n-1}, on a :

$$s_n = s_{n-1} + 2, \quad f_n = f_{n-1}, \quad a_n = a_{n-1} + 1, \quad \nu_n = \nu_{n-1} + 1$$

Si l'un des sommets de L_n figure parmi ceux du graphe A_{n-1}, mais non l'autre (arète pendante), on a

$$s_n = s_{n-1} + 1 \quad , \; f_n = f_{n-1} \quad , \quad a_n = a_{n-1} + 1 , \quad \nu_n = \nu_{n-1}$$

Si les deux sommets s et s' de L_n figurent déjà parmi ceux du graphe A_{n-1}, deux cas sont possibles : ou bien s et s' appartiennent à une même composante connexe de A_{n-1}, et on a :

$$s_n = s_{n-1} \quad , \quad f_n = f_{n-1} + 1 , \quad a_n = a_{n-1} + 1 , \quad \nu_n = \nu_{n-1}$$

ou bien s et s' appartiennent à deux composantes distinctes du graphe A_{n-1}, et on a :

$$s_n = s_{n-1} \quad , \quad f_n = f_{n-1} , \quad a_n = a_{n-1} + 1 , \quad \nu_n = \nu_{n-1} - 1$$

Par suite si A_{n-1} vérifie (1), A_n vérifie encore cette relation.

DUALITE

Toute arète admet un ou deux sommets adjacents. On déduit de la relation d'Euler que toute arète est adjacente à une face au moins et à deux faces au plus. Si le graphe est fini, en effet, la suppression de cette arète diminue d'une unité au plus le nombre des faces. Cette propriété, qui semble intuitive, reste vraie pour un graphe localement fini (mais non pour un graphe quelconque).

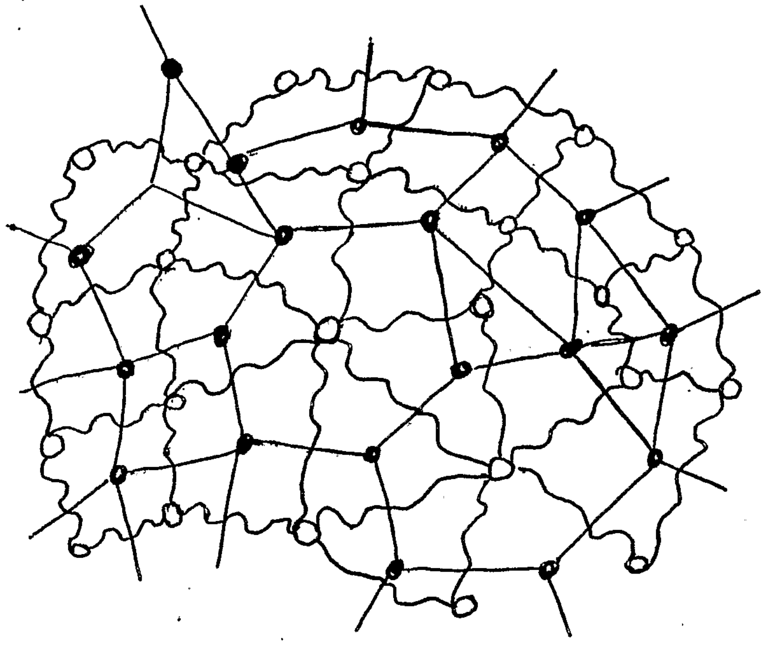
En effet, soit L une arète d'un graphe localement fini (A,S,F), et x et y deux points de L. L étant homéomorphe à (o,1), on peut définir le "segment" fermé de L limité par x et y (homéomorphe à un intervalle fermé contenu dans (o,1). Ce segment est compact. Comme le graphe est localement fini, on peut trouver un ouvert V borné connexe contenant ce segment et ne rencontrant aucun sommet ni aucune arète autre que L, et tel que le segment ouvert $L \cap V$ soit connexe. L'ouvert $V - L$ comporte deux composantes connexes V_1 et V_2, disjointes des arètes et des sommets du graphe. Chacune d'elles est donc contenue dans une face, soit $V_1 \subset f_1$ et $V_2 \subset f_2$ (avec éventuellement $f_1 = f_2$). Tout point de L compris entre x et y est donc adhérent aux faces f_1 et f_2 et à celles-là seulement. Le point x étant fixé, donc aussi f_1 et f_2, on peut choisir y quelconque sur L, et par suite :

> Dans un graphe localement fini, toute arète est adjaçente à une ou à deux faces. Elle est adjacente à une face f (rencontre la frontière de cette face) si et seulement si elle est contenue dans la frontière de cette face f. Ses sommets appartiennent alors aussi à la frontière de cette face f.

Une arète L qui n'admet qu'une extrémité est une <u>boucle.</u> Elle est adjacente à deux faces distinctes. Trois circonstances sont possibles : ou bien L (et son unique sommet) constituent une composante connexe du graphe ; ou bien de l'unique sommet de L partent des arètes toutes extérieures (ou toutes intérieures) à la boucle L, et le graphe admet des composantes intérieu- (extérieures) à la boucle L et disjointes de L; ou bien enfin ce sommet est relié à la fois à des arètes intérieures et à des arètes extérieures à L.

De même une arète L adjacente à une seule face f admet deux sommets distincts. Elle peut, (avec ses deux sommets) constituer une composante du graphe, et ses deux sommets sont alors d'ordre 1. Elle peut admettre un sommet d'ordre 1 et l'autre d'ordre supérieur à 1 (c'est une arète pendante). Ou bien enfin ses deux sommets sont d'ordre>1 : c'est alors une <u>arète critique</u> (on ne peut pas l'enlever sans diminuer d'une unité le nombre des composantes connexes du graphe).

A tout graphe connexe (A,S,F,) on fait correspondre un graphe (A',S',F') <u>dual</u> construit de la manière suivante : on choisit un point s' dans chaque face $f \in F$. A toute arète $a \in A$, on fait correspondre une arète a' rencontrant a en un seul point et joignant les points s'_1 et s'_2 choisis dans les deux faces adjacentes à a (a' est donc une boucle si $f_1 = f_2$ et si par suite $s'_1 = s'_2$). C'est une transformation réciproque, qui échange faces et sommets <u>en conservant leur ordre de multiplicité</u>. En particulier, le dual d'un graphe d'ordre fini est d'ordre fini.



- Dualité des deux graphes -

RELATIONS COMPLEMENTAIRES

Soit un graphe fini, connexe, sans boucle et sans arête pendante ni arête critique (toute arête est adjacente à deux sommets distincts et deux faces distinctes). Soit s,(s_n) le nombre des sommets (d'ordre n), f,(f_n) le nombre des faces (d'ordre n), a le nombre des arêtes. On a alors :

(2) $$2a = \sum_n n\, s_n = \sum_n n\, f_n$$

(lorsque l'on fait la somme des nombres d'arêtes adjacentes à chacun des sommets, ou à chacune des faces, chaque arête est comptée deux fois exactement).

Désignons par n_F (n_s) l'ordre de multiplicité moyen des faces (des sommets), soit :

$$n_F = \frac{1}{f} \sum n f_n$$

$$n_s = \frac{1}{s} \sum n s_n$$

La relation d'Euler

$$f + s = a + 2$$

donne ici :

$$\frac{2a}{n_F} + \frac{2a}{n_s} = a + 2$$

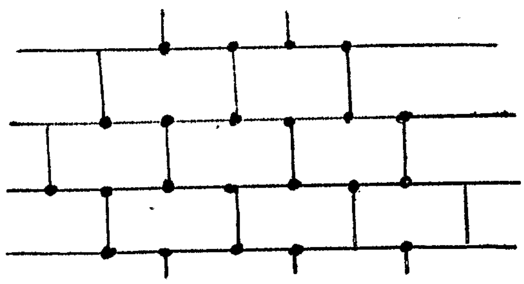
c'est-à-dire

(3) $$\frac{1}{n_F} + \frac{1}{n_s} = \frac{1}{a} + \frac{1}{2}$$

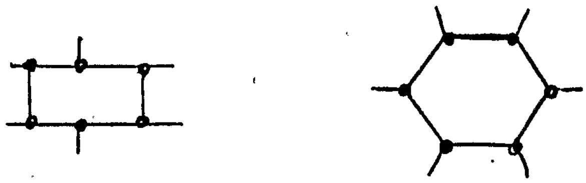
Pour un graphe infini (a grand), on aura (approximativement)

(4) $$\frac{1}{n_s} + \frac{1}{n_F} = \frac{1}{2}$$

Si, en moyenne, les faces sont des rectangles, les sommets sont en moyenne d'ordre 4 - Si les faces sont, en moyenne, des hexagones, les sommets sont, en moyenne, d'ordre 3, etc... (On prendra garde que sur le graphe dessiné ci-joint les faces sont des hexagones et non des rectangles, puisque chacune d'elles est adjacente à 6 arètes :



les figures :



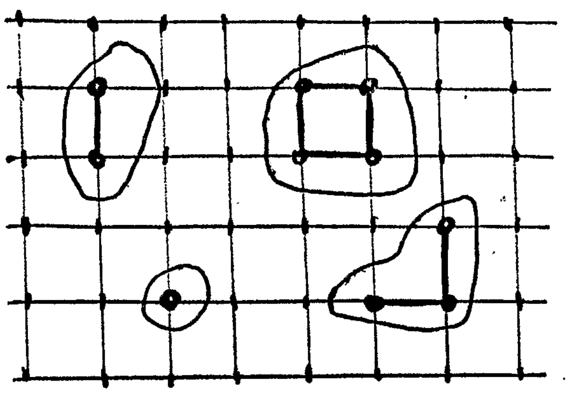
sont équivalentes pour les homéomorphismes).

APPLICATIONS DE LA RELATION D'EULER.

1/ Dénombrement des composantes connexes d'un ensemble

Dans les définitions données plus haut, nous avons exclu les sommets isolés. Tout sommet isolé augmente s et ν d'une unité, sans modifier a et f. La relation d'Euler reste donc valable s'il y a des sommets isolés.

Soit alors E un ensemble dont nous voulons dénombrer les composantes connexes. Partons du graphe défini par un réseau rectangulaire (à maille assez petite pour que toute composante de E contienne un noeud au moins de ce réseau). Considérons le sous-graphe G déduit du précédent en ne conservant que les sommets $s \in E$ et les arètes joignant deux sommets dans E



Soit a', s', f' les nombres d'arètes, de sommets et de faces de G', et ν' le nombre de composantes connexes. On a :

(1) $\nu' = s' + f' - a' - 1$

s' est le nombre N(+) des configurations +

f' est le nombre $N\left(\begin{smallmatrix}+&+\\+&+\end{smallmatrix}\right) + 1$ <u>s'il n'y a pas d'enclaves</u>

a' est le nombre $N\left(\begin{smallmatrix}+\\+\end{smallmatrix}\right) + N(+\,+)$

Or

$$s' = N\left(\begin{smallmatrix}+&+\\+&+\end{smallmatrix}\right) + N\left(\begin{smallmatrix}+&+\\+&o\end{smallmatrix}\right) + N\left(\begin{smallmatrix}+&o\\+&+\end{smallmatrix}\right) + N\left(\begin{smallmatrix}+&o\\+&o\end{smallmatrix}\right) +$$
$$+ N\left(\begin{smallmatrix}o&+\\+&+\end{smallmatrix}\right) + N\left(\begin{smallmatrix}o&+\\+&o\end{smallmatrix}\right) + N\left(\begin{smallmatrix}o&o\\+&+\end{smallmatrix}\right) + N\left(\begin{smallmatrix}o&o\\+&o\end{smallmatrix}\right)$$

$$a' = N\left(\begin{smallmatrix}+&+\\+&+\end{smallmatrix}\right) + N\left(\begin{smallmatrix}+&+\\+&o\end{smallmatrix}\right) + N\left(\begin{smallmatrix}+&o\\+&+\end{smallmatrix}\right) + N\left(\begin{smallmatrix}+&o\\+&o\end{smallmatrix}\right)$$
$$+ N\left(\begin{smallmatrix}+&+\\+&+\end{smallmatrix}\right) + N\left(\begin{smallmatrix}+&o\\+&+\end{smallmatrix}\right) + N\left(\begin{smallmatrix}o&+\\+&+\end{smallmatrix}\right) + N\left(\begin{smallmatrix}o&o\\+&+\end{smallmatrix}\right)$$

On trouve donc

(5) $\nu' = N\left(\begin{smallmatrix}o&\\+&o\end{smallmatrix}\right) - N\left(\begin{smallmatrix}+&o\\+&+\end{smallmatrix}\right)$

S'il y a des enclaves, f' est le nombre $N\left(\begin{smallmatrix}+&+\\+&+\end{smallmatrix}\right) + 1$ augmenté du nombre des enclaves, et la formule (5) donne alors la différence entre le nombre des composantes de E et celui des enclaves.(ν' est l'évaluation faible, celle que l'on obtient en n'autorisant pas les liaisons diagonales).

Il est clair que ce résultat s'applique à tout réseau 4×4 (sommets et faces d'ordre 4) et non aux seuls réseaux rectangulaires réguliers.

On peut aussi utiliser des réseaux (6 × 3), (3 × 6), etc. Prenons par exemple un réseau (6 × 3) (réseau triangulaire). En l'absence d'enclaves, on a

$$s' = N\,(+)$$

$$f' = N\,(\,+^{+}+) + N\,({}^{+}+^{+}) + 1$$

$$a' = N\,(+^{+}) + N\,({}^{+}+) + N\,(+\ +)$$

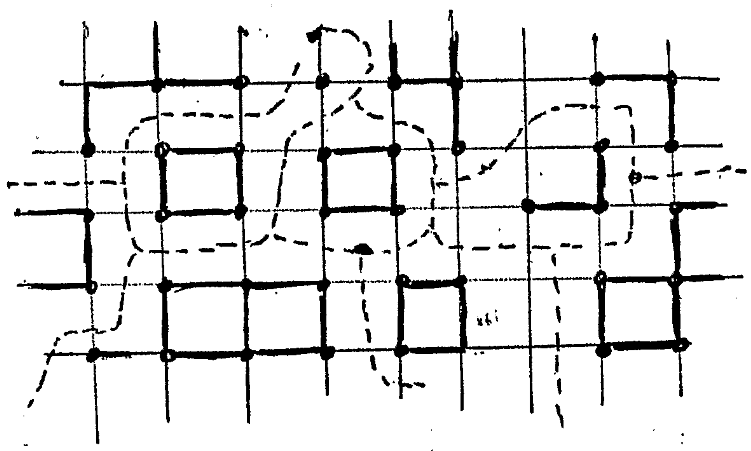
On trouve (entre autres combinaisons possibles):

(5') $$\nu' = N\,(+^{o}o) - N\,(o^{+}+) + N\,(+^{o}+) - N\,({}^{+}o^{+})$$

(noter que $N\,(+^{o}+)$ et $N\,({}^{+}o^{+})$ peuvent différer notablement en cas de dissymétrie orientée).

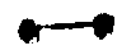
2/ <u>Partitions aléatoires (à composantes connexes)</u>

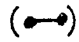
On partira du même graphe rectangulaire (par exemple) que ci-dessus, et on conservera tous les sommets, mais uniquement les arètes qui ne rencontrent pas de ligne de discontinuité (restent contenues dans une même composante). On fera le même décompte que ci-dessus, <u>toujours dans l'hypothèse où il n' y a pas d'enclaves</u> (composantes simplement connexes):



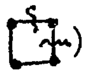
$s' = N(\bullet)$

$f' = N($  $) + 1$

$a' = N($  $) + N($  $)$

et $\nu' = N(\bullet) + N($  $) - N($  $) - N($  $)$

On trouve, en figurant par  la présence d'une interruption entre deux sommets, et par • • l'absence de spécification (c'est-à-dire présence ou absence de coupure indifférente) :

(6) $\nu' = N($  $) - N($  $) - N($  $) - N($  $)$

$= N($  $) + N($  $) - N($  $)$

(les deux derniers termes doivent, en général, être assez faibles)

S'il y a <u>des enclaves</u>, ν' représente le nombre des composantes diminué du nombre des enclaves. Dans une autre optique, la partition aléatoire elle-même peut être considérée comme un graphe, dont les arètes sont les lignes de discontinuité et les sommets les points triples, quadruples, etc... de ces lignes. Soit $(A,S;F)$ ce graphe. Le nombre f des faces coïncide avec le nombre ν' déterminé en (6). Si ce graphe est connexe ($\nu = 1$), la relation d'Euler donne

$$a - s = f - 2$$

Comme f est déterminé, on obtient la différence $a - s$. Si, de plus, une observation directe montre que tous les sommets sont d'ordre n (par exemple d'ordre 3, et ce cas sera relativement fréquent dans les applications), (2) fournit la relation supplémentaire

$$n\ s = 2\ a$$

d'où

$$s = \frac{2(f-2)}{n-2} \quad , \quad a = \frac{n(f-2)}{n-2}$$

CALCUL THEORIQUE DU NOMBRE DES COMPOSANTES CONNEXES

On remarque d'abord que (6) se met sous la forme :

$$\nu' = N\ (\square) - N\ (\vert) - N\ (\bullet\!\!-\!\!\bullet) + N\ (\bullet)$$

Si a et b sont les côtés du rectangle $\square$, posons

$$T\ (a,b) = P\ (\square)$$

On aura donc

$$(7) \qquad \nu' = \lim_{\substack{a \to +0 \\ b \to +0}} \frac{[T(a,b) - T(a,o) - T(o,b) + T(o,o)]}{a\ b}$$

Nous désignerons cette "dérivée seconde à droite" par T''_o : $\nu' = T''_o$