ANNALES DE L'I. H. P.

JERZY NEYMAN

Sur la théorie probabiliste des amas de galaxies et la vérification de l'hypothèse de l'expansion de l'univers

Annales de l'I. H. P., tome 14, n° 3 (1954-1955), p. 201-244
<http://www.numdam.org/item?id=AIHP_1955__14_3_201_0>

Sur la théorie probabiliste des amas de galaxies et la vérification de l'hypothèse de l'expansion de l'univers [1]

par

Jerzy NEYMAN,

Laboratoire de Statistique, Université de Californie, Berkeley, Californie.

TABLE DES MATIÈRES.

---

(1) Cette étude a été faite avec l'appui partiel de l'Office of Naval Research, United States Navy.

---

1. Introduction. — Dans un travail récent, entrepris en collaboration avec Elizabeth L. Scott, nous avons développé une théorie générale de la répartition spatiale des galaxies, en partant de l'hypothèse que cette répartition peut être représentée par la combinaison d'une multitude d'amas semblables dont les centres sont répartis au hasard dans l'espace. Cette théorie dépend d'une hypothèse additionnelle, à savoir celle d'un univers statique. Dans la suite, en parlant de cette publication [1], nous l'appellerons l'article I (²). Dans un travail ultérieur [2] (article II), en collaboration avec Elizabeth L. Scott et C. D. Shane, nous avons particularisé la théorie générale de l'article I en adoptant des formes particulières pour représenter les fonctions qui y interviennent. Cette théorie ainsi particularisée a été appliquée aux nombres des images des galaxies identifiées par C. D. Shane et C. A. Wirtanen [3] sur les photographies prises par ces deux astronomes à l'Observatoire Lick. La magnitude limite dans ces plaques est à peu près de 18,3 mag. La région du ciel ainsi étudiée est un rectangle de $93^\circ \times 46^\circ$ situé au voisinage du pôle galactique nord.

La confrontation de la théorie avec les données d'observation présentées dans l'article II est basée sur les valeurs des quasi-corrélations sériales (conception semblable, mais non identique, à celle de corrélations sériales) et l'accord est excellent. Malheureusement, l'estimation de la constante σ, qui caractérise l'espace occupé par un amas de galaxies typique, dépend des propriétés de la répartition de la magni-

---

(²) Les chiffres entre crochets renvoient à la bibliographie à la fin du présent travail.

tude absolue des galaxies. Cette répartition n'étant pas exactement connue, l'article II donne, au lieu d'une estimation unique de σ, toute une série de valeurs de σ, situées dans un très grand intervalle et toutes compatibles avec les quasi-corrélations empiriques.

Dans le troisième article [4] de la même série (article III), E. L. Scott, C. D. Shane et Margaret D. Swanson, ont essayé d'établir une comparaison visuelle entre les conclusions de l'article II et les données de l'Observatoire Lick. Dans ce but, les auteurs ont organisé, effectué et complété une expérience impressionnante de tirages de boules, en conformité parfaite avec tous les détails de la théorie, y compris les hypothèses sur le mécanisme des erreurs commises en comptant les images des galaxies. Le but de l'expérience était de produire « une répartition synthétique » des images des galaxies sur une plaque photographique qui puisse être comparée visuellement avec celle d'une photographie réelle.

Au premier coup d'œil, les deux répartitions, synthétique et réelle, paraissent semblables, mais un examen plus minutieux révèle une différence intéressante : la plaque réelle contient un grand nombre de petites concentrations locales des images des galaxies qui ne sont pas visibles sur la plaque synthétique. Plus tard, l'existence de cette différence a été confirmée par l'énumération des images des galaxies dans des carrés de $10' \times 10'$ sensiblement plus petits que ceux, $1^\circ \times 1^\circ$, qui ont fourni le matériel numérique pour le calcul des quasi-corrélations et pour l'ajustement des valeurs des paramètres. On a trouvé que, pour obtenir un bon accord entre la théorie et les nombres des images des galaxies dans les carrés $10' \times 10'$ il faut admettre une valeur de σ plus petite que celle résultant de l'analyse des carrés $1^\circ \times 1^\circ$.

Ultérieurement le Laboratoire de Statistique de l'Université de Californie a eu l'occasion d'étudier un cliché d'une plus grande magnitude limite. Ce cliché a été pris par M. A. G. Wilson qui se servait du télescope de Schmidt de 48 pouces de l'Observatoire du Mont Palomar. M. Wilson a fait une énumération des images des galaxies dans des petits carrés $5'36'' \times 5'36''$ et a eu l'obligeance de nous communiquer ses résultats. Le Laboratoire de Statistique lui en est bien reconnaissant. La magnitude limite de la plaque de M. Wilson est à peu près 19,6 mag.

Quoique l'analyse des données de M. Wilson soit encore en progrès et que le volume de ses données soit plutôt inférieur à celui de l'Observatoire Lick (1 cliché contre environ 100), il est malgré tout intéressant de constater que les chiffres de M. Wilson indiquent des condensations à petite échelle des images des galaxies, encore plus fortes que celles que l'on a constaté en analysant les carrés $10' \times 10'$ sur les plaques de l'Observatoire Lick. En effet, l'estimation provisoire de σ obtenue en partant de l'énumération de M. Wilson semble être encore plus petite que celle suggérée par les calculs de l'article III.

La théorie implique que la variabilité d'une plaque photographique à une autre doit être très grande et notre expérience, qui résulte de l'analyse des plaques de l'Observatoire Lick confirme cette conclusion. Donc il est bien possible que les différences entre les valeurs progressivement décroissantes de σ obtenues dans les trois cas [c'est-à-dire en partant des énumérations des images des galaxies sur les plaques de l'Observatoire Lick, d'abord dans des carrés $1^\circ \times 1^\circ$, ensuite dans des carrés $10' \times 10'$ et, enfin, dans des carrés $5'36'' \times 5'36''$ (énumération de M. Wilson)] soient entièrement dues au hasard. Cependant il est également possible que cette tendance soit due à une cause plus profonde et, notamment, qu'elle reflète le phénomène de l'expansion de l'univers. Sans entrer dans les détails, on peut dire que l'hypothèse suivant laquelle les amas de galaxies s'éloignent de l'observateur avec des vitesses croissant proportionnellement aux distances, conduit à la conclusion que sur les clichés correspondant à un tel univers en expansion on doit constater un excès d'images des amas lointains (donc, un excès d'amas de petites dimensions angulaires), par rapport à ce qu'on devrait voir dans le cas d'un univers stationnaire. On voit intuitivement que ce phénomène expliquerait les concentrations locales des images des galaxies qu'on constate en analysant les nombres de ces images dans les petits carrés. Il paraît également probable que le même phénomène expliquerait l'intensification de cet effet sur les plaques prises aux plus grandes magnitudes limites. En d'autres termes, il semble probable qu'une généralisation de la théorie de l'article I admettant la possibilité que les amas de galaxies s'éloignent de l'observateur nous mènerait aux conclusions conformes aux tendances remarquées dans les données disponibles.

En conformité avec ces idées, la première partie du présent travail est consacrée à une généralisation de la théorie de l'article I. Cette généralisation admet la possibilité de l'expansion de l'univers. Les formules de la théorie généralisée se réduisent à celles de l'article I lorsque les deux fonctions, que nous introduisons pour caractériser la récession et l'expansion des amas, sont remplacées par zéro.

Une remarque dans l'article II signale une erreur dans la démonstration du résultat principal de l'article I, résultat qui est cependant correct. La démonstration de la formule fondamentale du présent Ouvrage corrige donc l'erreur de la démonstration donnée dans l'article I.

Nous espérons que les calculs numériques basés sur les formules de la première partie du présent travail et la comparaison des résultats avec les données empiriques contribueront à la solution du passionnant problème de savoir si l'univers est en expansion ou non. On sait que la méthode statistique employée jusqu'ici pour traiter ce problème repose sur la relation entre le nombre moyen des images des galaxies par degré carré et la magnitude limite d'une plaque photographique. Malheureusement, la magnitude limite d'une plaque prise avec des instruments donnés dépend des conditions atmosphériques et est difficile à mesurer. C'est peut-être pour cette raison que les résultats d'application de la méthode des moyennes ne sont pas décisifs. La méthode d'étude empirique dérivant des résultats de la première partie du présent Ouvrage est basée sur un autre aspect des observations, à savoir, sur ce que M. Zwicky [5] appelle « morphologie non-dimensionnelle » de la répartition des images des galaxies. Donc, elle fournit une possibilité nouvelle d'utiliser les résultats d'observation déjà accumulés.

La deuxième partie du présent Ouvrage donne certaines formules qui peuvent conduire à une caractérisation de l'image sommaire de la répartition des galaxies dans l'espace. La question principale que ces formules peuvent aider à résoudre est celle du rôle de l'amas. Une des possibilités est la suivante : en général, les amas de galaxies sont des assemblages compacts, séparés les uns des autres par des espaces vides, se prêtant aux études dynamiques comme des systèmes isolés. L'autre possibilité extrême est celle qui consiste à considérer un amas isolé comme une pure fiction; en réalité nous étudions un champ continu de galaxies avec de nombreuses condensations locales, qu'on interprète comme « amas ».

D'après M. Couderc [6], les astronomes contemporains sont disposés à croire que « l'espace est pavé d'amas de galaxies... ». Une image pareille est suggérée par M. Zwicky [7] qui parle des amas de galaxies « remplissant l'espace comme une mousse de bulles de savon ». Ces descriptions pittoresques, et d'autres du même genre, reflètent des impressions qualitatives plutôt que les résultats d'une analyse numérique. Il paraît intéressant de fournir les moyens théoriques d'une telle analyse, avec une possibilité que, en fin de compte, les amas paraîtront comparables à une grande multitude de ballons lâchés un jour de fête.

Pour distinguer entre ces deux possibilités extrêmes ou bien pour en choisir une intermédiaire, il ne suffit pas d'avoir des formules théoriques. Il faut, naturellement, des estimations aussi exactes que possible des divers paramètres qui interviennent dans le modèle général. En nous servant des estimations provisoires de l'article II, nous avons publié, en collaboration avec Élizabeth L. Scott [8], quelques valeurs numériques des probabilités déduites dans la deuxième partie du présent travail. On en trouvera ici une série plus complète. La conclusion générale suggérée par ces chiffres, est que, si la plus grande valeur de σ, indiquée dans l'article II, est correcte, l'idée d'un amas isolé ne peut être qu'un mythe. Si la seconde valeur de σ plus petite que la première, suggérée par l'article III, est correcte, alors un amas isolé joue un rôle important dans l'image générale de la répartition des galaxies dans l'espace. Enfin, par extrapolation, si la valeur de σ est encore plus petite, comme le suggère la plaque de M. Wilson, alors le rôle des amas isolés est prédominant. Une réponse définitive à cette question ne peut être obtenue qu'à la suite d'une analyse plus fine des données empiriques. La théorie de la première partie du présent Ouvrage, admettant la possibilité de l'univers en expansion, peut aider à cette analyse.

PREMIÈRE PARTIE.

RÉPARTITION SIMULTANÉE DES NOMBRES DES IMAGES DES GALAXIES DANS DEUX RÉGIONS D'UN CLICHÉ.

2. Postulats fondamentaux concernant la répartition des galaxies dans l'espace. — Ces postulats sont les mêmes que ceux de l'article I. Nous les reproduisons, à peu près sous la même forme, sauf en ce qui concerne la mention explicite du rôle du temps :

(i) Les galaxies ne se présentent qu'en amas.

(ii) A chaque instant T et à chaque ensemble R mesurable-B dans l'espace (nous ne considérons que les ensembles mesurables-B), il correspond une variable aléatoire $\gamma(R, T)$ représentant le nombre des centres d'amas qui au moment T sont situés en R. Pour T fixe la répartition de $\gamma(R, T)$ ne dépend que de la mesure de R, soit $V(R) = V$. Si $R_1, R_2, \ldots, R_n, \ldots$, est une suite finie ou dénombrable d'ensembles disjoints, alors les variables aléatoires $\gamma(R_1, T)$, $\gamma(R_2, T)$, $\ldots$, $\gamma(R_n, T)$, $\ldots$, correspondant à ces ensembles et au même instant T, sont indépendantes. Si l'univers est stationnaire, alors la répartition de $\gamma(R, T)$ ne dépend pas de T. Dans le cas contraire, elle est une fonction non-constante de T.

Désignons par $G_\gamma(t \mid V) = E[t^{\gamma(R,T)}]$ la fonction génératrice des probabilités de $\gamma(R, T)$. Dans l'article I nous avons signalé que la loi de probabilité de $\gamma(R, T)$ appartient à la classe des lois indéfiniment divisibles de M. Paul Lévy, et que, si la mesure $V(R) = V$ est finie,

$$G_\gamma(t \mid V) = e^{Vh(t)}, \tag{1}$$

avec

$$h(t) = -h_0 + \sum_{k=1}^{\infty} h_k t^k, \tag{2}$$

où $h_0 > 0$ et $h_k \geqq 0$ pour $k = 1, 2, \ldots$, et où

$$\sum_{k=1}^{\infty} h_k = h_0. \tag{3}$$

Réciproquement, toute répartition définie par (1), (2) et (3), avec les h_k non-négatifs, satisfait aux conditions imposées sur $\gamma(R, T)$. On démontre ces faits directement quoiqu'on puisse les considérer comme des conséquences faciles à déduire de certains travaux antérieurs. M. Brockway McMillan a bien voulu attirer mon attention sur le fait qu'on peut les déduire des résultats de M. Norbert Wiener [9] et je lui en suis bien reconnaissant.

(iii) Soit R_1 un ensemble de mesure positive et finie, et soit R_2 un sous-ensemble de R_1. Soit $a_1, a_2, \ldots, a_m$ une combinaison arbitraire de $m \geqq 1$ nombres choisis parmi $1, 2, \ldots, n > m$. Considérons le cas où l'on sait que R_1 contient exactement n centres d'amas, soient $c_1, c_2, \ldots, c_n$.

Nous supposons que la probabilité conditionnelle pour que R_2 contienne les m amas $c_{a_1}, c_{a_2}, \ldots, c_{a_m}$, et pas d'autres, ne dépende pas de la combinaison particulière $a_1, \ldots, a_m$, mais seulement des nombres m, n et, naturellement, des mesures de R_1 et de R_2.

(iv) Le nombre ν des galaxies appartenant à un amas donné est une variable aléatoire non-négative. Cette variable ν est indépendante de toute autre variable aléatoire qui intervient dans le schéma, et sa répartition est la même pour tous les amas.

(v) Considérons un système de coordonnées rectangulaires ayant son origine au point où se trouve l'observateur. Désignons par u_1, u_2, u_3 les coordonnées d'un point où se trouve le centre d'un amas particulier au moment T. Désignons aussi par X_1, X_2, X_3 les coordonnées au même moment T d'une galaxie appartenant à cet amas. Nous supposons que X_1, X_2, X_3 sont des variables aléatoires avec une densité de probabilité conditionnelle que nous désignons par $f^*(\eta_1, \eta_2, \eta_3)$, les u_1, u_2, u_3 étant donnés. Nous supposons de plus que cette densité est une fonction continue des arguments $\eta_i = x_i - u_i$, $(i = 1, 2, 3)$, où x_i représente une variable réelle, valeur particulière de X_i. Enfin, nous supposons que, les positions de tous les centres d'amas étant fixés, les triplets des coordonnées de toutes les galaxies sont indépendants les uns des autres et sont aussi indépendants de toutes les autres variables aléatoires du schéma.

Remarque. — Dans les articles antérieurs on a supposé que la fonction f^* ne dépendait que de la distance

$$\eta = \{ \eta_1^2 + \eta_2^2 + \eta_3^2 \}^{\frac{1}{2}}. \tag{4}$$

Cependant cette restriction n'est d'aucune utilité.

Les cinq postulats ci-dessus déterminent la structure générale de la répartition spatiale des galaxies à un moment donné T. Pour obtenir un modèle spécial, il suffit de particulariser les lois de probabilité de $\gamma(R, T)$, de ν et celle de X_1, X_2, X_3. Un des résultats généraux de l'article I est que, tant qu'on n'impose aucune restriction sur la répartition de ν, on ne restreint pas la généralité du modèle en supposant que $\gamma(R, T)$ est une variable de Poisson, de sorte que

$$G_\gamma(t \mid V) = e^{-\lambda V(1-t)}, \tag{5}$$

où λ désigne le nombre moyen des centres d'amas par unité de volume.

Les trois répartitions dont nous avons parlé ne sont pas directement observables. Le problème central de la théorie consiste à établir une relation entre ces répartitions non-observables et la répartition des images des galaxies observable sur un cliché photographique. Pour résoudre ce problème dans le cas de l'univers stationnaire il nous a été nécessaire d'accepter le postulat supplémentaire que voici :

(vi) Étant donné une galaxie de coordonnées x_1, x_2, x_3 fixes et étant donné un nombre m, il existe une probabilité conditionnelle $\theta(\xi, m)$ pour que la magnitude apparente de cette galaxie soit plus petite que m. Cette probabilité $\theta(\xi, m)$ ne dépend que du nombre m et de la distance

$$\xi = \{ x_1^2 + x_2^2 + x_3^2 \}^{\frac{1}{2}} \tag{6}$$

entre la galaxie et l'observateur. Si m est la magnitude limite d'un cliché et si la magnitude apparente d'une galaxie est plus petite que m, l'image de cette galaxie apparaîtra sur le cliché. Dans ce cas, il nous sera commode de dire que la galaxie est « visible » sur le cliché. On postule encore que, si une galaxie est visible, elle est visible dans la direction déterminée par les coordonnées x_1, x_2, x_3. Enfin, on postule que, lorsque les distances ξ_1, ξ_2, ..., ξ_s de s galaxies arbitraires sont déterminées, la « visibilité » d'une quelconque de ces galaxies est indépendante de celles des autres. Cette dernière hypothèse implique que nous négligeons la possibilité des nuages absorbants qui pourraient augmenter à la fois les magnitudes de plusieurs galaxies voisines.

Le résultat principal de l'article I est la fonction génératrice des probabilités $G_{N_1,N_2}(t_1, t_2)$ de deux variables aléatoires N_1 et N_2 définies comme suit.

Soient ω_1 et ω_2 deux angles solides arbitraires, ayant leurs sommets au point où se trouve l'observateur. Supposons qu'on prenne deux clichés aux magnitudes limites m_1 et m_2. Le $i^{\text{ième}}$ cliché est pris de manière qu'il contienne la projection de l'angle solide ω_i. Cette projection de ω_i sur le $i^{\text{ième}}$ cliché sera désignée par le même symbole ω_i. Alors, par définition, la variable aléatoire N_i est le nombre des galaxies « visibles » dans $\omega_i (i = 1, 2)$. La formule publiée dans l'article I, représente la fonction $G_{N_1,N_2}(t_1, t_2)$ comme une fonctionnelle dépendant des fonctions arbitraires $G_\nu(t)$, f^* et $\theta(\xi, m)$. L'article II donne une spécialisation du modèle général obtenue en supposant que la variable ν possède un certain nombre de moments finis et suit peut-être une loi binomiale négative. Aussi, dans l'article II a-t-on admis que les variables η_1, η_2, η_3 sont indépendantes et suivent la même loi de probabilité normale, avec l'espérance mathématique zéro et avec une variance inconnue σ^2. Enfin, on a supposé que la fonction $\theta(\xi, m)$ a la forme indiquée par les résultats des recherches de Hubble et de M. Humason [10], [11]. Premièrement, l'article II admet que la magnitude absolue M d'une galaxie est une variable aléatoire normale avec une espérance mathématique M_0 et une variance σ_M^2. Deuxièmement, on a admis que la magnitude apparente m d'une galaxie, sa magnitude absolue M et sa distance ξ sont liées par la formule

$$m = M - 5 + 5\log_{10}\xi + 5{,}19.10^{-9}\xi, \tag{7}$$

où le dernier terme représente le phénomène de « déplacement vers le rouge » (« red shift ») et la distance ξ est mesurée en parsecs. Le déplacement vers le rouge est traité comme un phénomène établi empiriquement, sans aucune interprétation physique.

Il est clair que, si l'on prend garde de n'appliquer les postulats (i) à (v) qu'aux positions simultanées des galaxies à un moment déterminé T, ces postulats ne contiendront rien de contradictoire à l'hypothèse que l'univers est en expansion. Au contraire, le postulat (vi) dépend de l'hypothèse que l'univers est stationnaire et, si l'on veut construire une théorie pouvant s'appliquer à l'univers en expansion aussi bien qu'à l'univers stationnaire, il doit être remplacé par un postulat plus général. Pour formuler ce postulat nouveau, nous commençons par une analyse des conséquences de l'hypothèse de l'expansion de l'univers sur la « visibilité » des galaxies sur des clichés photographiques. C'est ce que nous faisons dans le chapitre suivant en supposant toujours un espace euclidien et en admettant la mécanique newtonienne.

3. **Postulat (vii) qui établit la relation entre la répartition des galaxies dans un univers en expansion et la répartition des images de ces galaxies sur une plaque photographique.** — Suivant les idées contemporaines sur l'expansion de l'univers, nous postulons que le centre d'un amas arbitraire de galaxies s'éloigne de l'observateur avec une vitesse proportionnelle à sa distance. On a donc

$$\frac{du_i}{d\tau} = u_i(\tau)h_1(\tau) \qquad (i = 1, 2, 3), \tag{8}$$

où τ désigne le temps et $h_1(\tau)$ est un coefficient de proportionnalité non-négatif, qui peut dépendre du temps, mais est le même pour tous les amas de galaxies.

Les formules (8) entraînent

$$u_i(\tau) = u_i(0)\mathrm{H}_1(\tau), \tag{9}$$

où

$$\mathrm{H}_1(\tau) = \exp\left\{\int_0^\tau h_1(x)\,dx\right\} \tag{10}$$

est une fonction non-décroissante de τ. L'origine dont on mesure τ étant arbitraire, nous la placerons dans « le présent », c'est-à-dire au moment où l'on prend une photographie du ciel. Les répartitions considérées dans les postulats (ii), (iii) et (iv) s'appliqueront à cette époque particulière, « le présent ».

En plus du postulat ci-dessus, concernant le phénomène de récession des amas de galaxies, nous adopterons un postulat semblable concernant l'expansion des amas. Soit $x_1(\tau)$, $x_2(\tau)$, $x_3(\tau)$ les coordonnées au temps τ d'une galaxie particulière appartenant à un amas dont le centre possède les coordonnées $v_1(\tau)$, $u_2(\tau)$, $u_3(\tau)$. Nous admettons que ces fonctions de temps sont liées par les équations

$$\frac{d}{d\tau}[x_i(\tau) - u_i(\tau)] = [x_i(\tau) - u_i(\tau)]\,h_2(\tau) \qquad (i = 1, 2, 3), \tag{11}$$

où $h_2(\tau)$ est un coefficient de proportionnalité non-négatif, qui peut dépendre de temps, mais qui est le même pour toutes les galaxies et pour tous les amas. On a

$$\begin{aligned} x_i(\tau) &= u_i(\tau) + [x_i(0) - u_i(0)]\,\mathrm{H}_2(\tau) \\ &= x_i(0)\,\mathrm{H}_2(\tau) + u_i(0)[\mathrm{H}_1(\tau) - \mathrm{H}_2(\tau)], \end{aligned} \tag{12}$$

où $\mathrm{H}_2(\tau)$ est liée à $h_2(\tau)$ par une formule analogue à (10).

On voit que si $h_1(\tau) = h_2(\tau) = 0$, l'univers est stationnaire. Dans le cas contraire, il y a expansion.

Si $h_1(\tau) = h_2(\tau) > 0$, nous dirons que la loi de l'expansion de l'univers et des amas est la même. Si $h_1(\tau) = h_2(\tau) = h = \text{const.}$, nous parlerons d'expansion « uniforme ».

Les formules (12) impliquent que les coordonnées $x_i(\tau)$ d'une galaxie déterminée sont des fonctions certaines du temps τ qui dépendent des coordonnées au temps zéro de cette même galaxie et de celles du centre d'amas correspondant. Il s'ensuit que le caractère aléatoire des coordonnées $x_i(\tau)$ d'une galaxie au temps arbitraire τ se ramène au caractère aléatoire des coordonnées au temps $\tau = 0$ de cette même galaxie et du centre d'amas auquel cette galaxie appartient.

Jusqu'à présent nous n'avons considéré que les coordonnées présentes d'une galaxie (c'est-à-dire celles correspondant à $\tau = 0$) et les coordonnées de cette galaxie à un moment arbitraire τ. Il nous faudra considérer maintenant les coordonnées d'une galaxie que nous appellerons les *coordonnées apparentes* et que nous désignerons par y_1, y_2, y_3. Désignons par $-\tau^\star$ le moment dans le passé où une galaxie déterminée a émis la lumière qui arrive à l'observateur au moment $\tau = 0$. La coordonnée apparente y_i de cette galaxie est définie par $y_i = x_i(-\tau^\star)(i = 1, 2, 3)$. En d'autres termes, les coordonnées apparentes d'une galaxie sont les coordonnées de la position occupée par cette galaxie au moment $-\tau^\star$, où la lumière enregistrée par la plaque photographique a été émise par la galaxie.

La relation entre les coordonnées apparentes et les coordonnées présentes d'une galaxie est bien simple. Soit c la vitesse de la lumière. Le temps nécessaire pour que le signal lumineux émis au point (y_1, y_2, y_3) arrive à l'observateur, soit $\tau^\star$, est donné par

$$\tau^\star = \frac{\{y_1^2 + y_2^2 + y_3^2\}^{\frac{1}{2}}}{c}. \tag{13}$$

Donc, d'après les formules (11),

$$x_i = x_i(0) = \frac{\{y_i - u_i(H_1 - H_2)\}}{H_2} \qquad (i = 1, 2, 3), \tag{14}$$

où, pour abréger les formules,

$$u_i = u_i(0) \qquad \text{et} \qquad H_j = H_j(-\tau^\star) \qquad (j = 1, 2).$$

Pour compléter la généralisation de la théorie de l'article I au cas où l'univers peut être en expansion, remarquons que, quelle que soit la fonction $\theta(\xi, m)$ postulée dans (vi), la distance « présente » ξ qui y intervient doit être remplacée par la distance « apparente », soit

$$\xi^\star = \{y_1^2 + y_2^2 + y_3^2\}^{\frac{1}{2}}. \tag{15}$$

Nous arrivons donc à la formule suivante du postulat (vii) :

(vii) Étant donnée une galaxie qui, au moment où l'on fait une photographie du ciel, possède les coordonnées x_1, x_2, x_3 et, étant donnée la magnitude limite m de la plaque, il existe une probabilité conditionnelle $\theta(\xi^*, m)$ pour que cette galaxie soit visible. Cette probabilité ne dépend que de m et de la distance apparente de la galaxie déterminée par (15), (14) et (13). De plus, si une galaxie est visible, elle l'est dans la direction de sa position apparente, au point (y_1, y_2, y_3). Enfin, étant donné les distances apparentes ξ_1^*, ξ_2^*, ..., ξ_s^* de huit galaxies différentes, les « visibilités » de ces galaxies sont indépendantes.

Désormais, la *distance présente* ξ d'une galaxie définie par la formule (6), n'interviendra plus dans nos raisonnements. Au lieu de ξ nous aurons constamment à considérer la *distance apparente* ξ^* définie par (15). Donc, pour simplifier les formules, l'astérisque utilisé dans la formule (15) sera abandonné et la distance apparente d'une galaxie sera désignée simplement par ξ.

4. Densité de probabilité conditionnelle des coordonnées apparentes d'une galaxie. — Partant de la densité conditionnelle f^* des coordonnées présentes X_1, X_2, X_3 d'une galaxie qui appartient à un amas dont le centre possède les coordonnées présentes u_1, u_2, u_3, calculons la densité conditionnelle, soit $f(y_1, y_2, y_3; u_1, u_2, u_3)$, des coordonnées apparentes Y_1, Y_2, Y_3 de la même galaxie.

Les valeurs particulières y_1, y_2, y_3 des Y_1, Y_2, Y_3 sont liées aux valeurs particulières x_1, x_2, x_3 de X_1, X_2, X_3 par les formules (14). Suivant la méthode bien connue, pour obtenir $f(y_1, y_2, y_3; u_1, u_2, u_3)$, on substitue dans f^* à la place des x_i leurs expressions en fonctions des y_i et l'on multiplie le résultat par la valeur absolue du jacobien de la transformation. A cause de la forme spéciale de la fonction f^*, on a

(16) $$f \equiv f(y_1, y_2, y_3; u_1, u_2, u_3) = f^*(\eta_1, \eta_2, \eta_3)\,|\,J\,|,$$

où

(17) $$\eta_i = x_i - u_i = \frac{y_i - H_1 u_i}{H_2}$$

et

(18) $$J = \frac{\partial(x_1, x_2, x_3)}{\partial(y_1, y_2, y_3)} = \frac{1}{H_2^3}\left\{1 + \frac{H_2'}{cH_2}\xi + \frac{H_1' H_2 - H_1 H_2'}{cH_2}\sum_{i=1}^{3} u_i y_i\right\}.$$

Dans les formules (17) et (18) l'argument des fonctions H_j et de leurs dérivées H'_j ($j = 1, 2$) est égal à $-\frac{\xi}{c}$, où ξ est définie par (15).

Les formules (16), (17) et (18) peuvent être utiles dans le cas où quelque théorie physique suggérerait les formes spéciales des fonctions $H_1(\tau)$ et $H_2(\tau)$. En l'absence de telles théories, ou devra probablement s'appuyer sur les résultats empiriques de Hubble et de M. Humasson, à savoir que, pour les valeurs de τ entre zéro et, environ 500 millions d'années dans le passé, la valeur de $h_1(\tau)$ est à peu près constante. Donc, si l'on veut confronter la théorie avec les observations, il faudra probablement poser $h_1(\tau) = h^* =$ const. de Hubble et adopter, en outre, une hypothèse convenable concernant la fonction $H_2(\tau)$. L'hypothèse la plus simple est que

$$H_1(\tau) \equiv H_2(\tau).$$

Il existe une autre hypothèse simplificatrice qn'il faudra probablement adopter lorsqu'on travaille empiriquement. Elle concerne la densité de probabilité f^*. Au lieu d'admettre, comme on a fait plus haut, que cette fonction dépend de trois arguments η_1, η_2, η_3, on pourrait se limiter probablement au cas où il n'existe qu'un seul argument, soit

$$\eta = (\eta_1^2 + \eta_2^2 + \eta_3^2)^{\frac{1}{2}},$$

la distance entre la galaxie et le centre de l'amas auquel elle appartient.

En admettant ces trois hypothèses simplificatrices, on tire des formules (14)

$$x_i = y_i e^{g\xi} \qquad (i = 1, 2, 3), \tag{19}$$

où, pour abréger, on a posé $g = \frac{h^*}{c}$. La formule (18) se réduit à

$$J = e^{3g\xi}(1 + g\xi). \tag{20}$$

Dans ce qui suit on aura besoin de la formule représentant la probabilité conditionnelle, soit $p(\omega \mid u)$, pour qu'une galaxie déterminée appartenant à un amas dont le centre a la position présente $u = (u_1, u_2, u_3)$, soit visible dans un domaine ω d'une plaque photographique à magnitude limite m. En désignant par la même lettre ω l'angle solide déterminé par le domaine ω, on trouve aisément

$$p(\omega \mid u) = \iiint_{\omega} \theta(\xi, m) f \, dy_1 \, dy_2 \, dy_3, \tag{21}$$

ou, avec l'hypothèse $h_1(\tau) \equiv h_2(\tau) \equiv h^\star$,

$$(22) \qquad p(\omega \mid u) = \iiint_\omega \theta(\xi, m) f^\star(y_1 e^{g\xi} - u_1, y_2 e^{g\xi} - u_2, y_3 e^{g\xi} - u_3) \times e^{3g\xi}(1 + g\xi)\, dy_1, dy_2\, dy_3.$$

5. **Formule fondamentale.** — Dans ce chapitre nous déduirons une formule, appelée formule fondamentale, qui donne la fonction génératrice simultanée de deux variables aléatoires entières N_1 et N_2, se rattachant à deux ensembles arbitraires ω_1 et ω_2 dans l'espace. La définition de ces variables N_1 et N_2 est un peu artificielle et, au premier coup d'œil, son intérêt peut ne pas apparaître clairement. Cette définition a été inventée pour généraliser trois problèmes qui n'ont rien en commun en apparence. Néanmoins, au moyen de quelques spécialisations simples, la formule fondamentale donne les réponses aux questions posées par tous ces problèmes. L'un des problèmes en question est traité dans le paragraphe 7, les deux autres dans le paragraphe 15.

Soient ω_1 et ω_2 deux ensembles tout à fait arbitraires mesurables dans l'espace, disjoints ou non. Considérons un amas de galaxies C et désignons par $\nu(C)$ le nombre des galaxies qui lui appartiennent. Nous allons nous intéresser à celles des galaxies de l'amas C dont les positions *apparentes* appartiennent soit à ω_1 soit à ω_2. Ultérieurement, nous aurons aussi à nous intéresser à celles des galaxies de l'amas C dont les *positions présentes* appartiennent à ω_1 ou à ω_2. Il est clair que la réponse à cette dernière question peut être obtenue de celle qui répond à la première, en substituant $h_1(\tau) \equiv h_2(\tau) \equiv 0$. Pour cette raison nous nous occuperons des positions apparentes des galaxies.

Pour $i = 1, 2$, désignons par $\nu_i(C)$ le nombre de galaxies de l'amas C, dont les positions apparentes sont dans ω_i. Nous dirons que ces galaxies sont dans ω_i « en apparence » ou qu'elles « paraissent » être dans ω_i.

Imaginons maintenant que pour chaque galaxie de l'amas C qui paraît être dans ω_i, donc telle que sa position apparente $(y_1, y_2, y_3) \in \omega_i$, on fasse une épreuve au hasard avec une probabilité de « succès » $\theta_i^\star(y_1, y_2, y_3) \equiv \theta_i^\star$ qui est une fonction mesurable de y_1, y_2, y_3. Si pour une galaxie donnée cette épreuve donne le succès, alors cette galaxie sera appelée une « galaxie choisie dans ω_i » ($i = 1, 2$). Postulons encore que pour toute galaxie de l'amas C dont la position apparente appartient à la partie commune de ω_1 et ω_2, donc telle que $(y_1, y_2, y_3) \in \omega_1\omega_2$, il existe une fonction mesurable $\rho(y_1, y_2, y_3) \equiv \rho$ de ses coordonnées apparentes, telle que le produit $\theta_1^\star \theta_2^\star \rho$ représente la probabilité composée pour que cette galaxie soit une galaxie choisie simultanément dans ω_1 et dans ω_2. On voit qu'à partir de $\theta_1^\star$, $\theta_2^\star$ et ρ on peut calculer la probabilité pour qu'une galaxie soit une galaxie choisie dans ω_1 mais pas dans ω_2, etc.

Postulons enfin que, lorsque les positions apparentes de plusieurs galaxies de l'amas C sont déterminées, les épreuves correspondantes à ces galaxies sont indépendantes.

Désignons par $n_i(\mathrm{C})$ le nombre de galaxies de l'amas C qui sont les galaxies choisies dans ω_i, pour $i = 1, 2$. Soit encore $\mathrm{N}(r_1, r_2)$ une fonction arbitraire mais déterminée de deux arguments r_1 et r_2, définie pour les valeurs entières non-négatives de ces arguments, et capable elle-même de ne prendre que des valeurs entières non-négatives. Enfin, soit $\mathrm{N}_i = \mathrm{N}[n_i(\mathrm{C}), \nu(\mathrm{C})]$ pour $i = 1, 2$.

Avec ces conventions préalables, les deux variables N_1 et N_2 sont définies par les égalités

$$\mathrm{N}_i = \sum_c \mathrm{N}_i(\mathrm{C}) = \sum_c \mathrm{N}[n_i(\mathrm{C}), \nu(\mathrm{C})] \qquad (i = 1, 2), \tag{23}$$

où la sommation s'étend à tous les amas dans l'espace.

Le problème central du présent chapitre est de déduire la fonction génératrice simultanée de deux variables N_1, N_2 ainsi définies. Nous commençons par trois remarques préliminaires.

Remarquons d'abord que les postulats adoptés impliquent que, lorsque les positions présentes de plusieurs centres d'amas, soient $\mathrm{C}_1, \mathrm{C}_2, \ldots,$ sont déterminées, les triplets de variables aléatoires $[n_1(\mathrm{C}_k), n_2(\mathrm{C}_k), \nu(\mathrm{C}_k)]$ pour $n = 1, 2, \ldots,$ sont indépendants. Il en résulte que les couples de variables aléatoires $\{\mathrm{N}_1(\mathrm{C}_k), \mathrm{N}_2(\mathrm{C}_k)\}$, correspondants à différentes valeurs de k, sont aussi indépendants.

La deuxième remarque évidente est que, si l'on connaît la fonction génératrice simultanée des trois variables $n_1(\mathrm{C}_k)$, $n_2(\mathrm{C}_k)$ et $\nu(\mathrm{C}_k)$, soit $\mathrm{G}_{n_1(\mathrm{C}), n_2(\mathrm{C}), \nu}(t_1, t_2, t_3)$, la fonction génératrice simultanée du couple $\{\mathrm{N}_1(\mathrm{C}_k), \mathrm{N}_2(\mathrm{C}_k)\}$ pourra être obtenue par une simple opération que nous appellerons l'opération de réduction. Pour effectuer cette opération on n'a qu'à développer $\mathrm{G}_{n_1(\mathrm{C}), n_2(\mathrm{C}), \nu}(t_1, t_2, t_3)$ en série de puissances de t_1, t_2, t_3 et à y remplacer tout produit $t_3^{r_1} t_3^{r_2} t_3^{r_3}$ par le produit $t_1^{\mathrm{N}(r_1, r_3)} t_2^{\mathrm{N}(r_2, r_3)}$. Il sera commode d'avoir un symbole spécial pour désigner cette opération de réduction et nous écrirons

$$G_{N_1(C),\,N_2(C)}(t_1,\ t_2 \equiv {}^{\star}G_{n_1(C),\,n_2(C),\,\nu}(t_1,\ t_2,\ t_3) \tag{24}$$

La troisième remarque préliminaire consiste à observer que les variables N_1 et N_2 définies par (23) peuvent être considérées comme des limites de certaines autres variables, soient $M_1(A)$ et $M_2(A)$, définies pour tout $A > 0$. Désignons par $R(A)$ un cube défini par les trois inégalités $|x_j| < A$, $(j = 1, 2, 3)$. Alors, par définition, $M_i(A)$ est la somme des nombres $N_i(C)$ qui correspondent aux amas C dont les centres ont leurs positions présentes dans le cube $R(A)$. Évidemment

$$N_i = \lim_{A\to\infty} M_i(A) \qquad (i = 1,\ 2). \tag{25}$$

Nous allons suivre maintenant les lignes générales des raisonnements de l'article I. Soit $s > 1$ un entier qui est un cube parfait. Divisons le cube $R(A)$ en s cubes égaux R_{sj} $(j = 1, 2, \ldots, s)$ et distribuons les points aux frontières de façon que les R_{sj} soient disjoints. Désignons par $\Delta_s^3 = \dfrac{8A^3}{s}$ le volume de chacun des R_{sj}. Désignons aussi par M_{isj} la contribution à $M_i(A)$ des amas dont les centres ont leurs positions présentes dans R_{sj} pour $i = 1, 2$. Alors, pour tout s,

$$M_i(A) = \sum_{j=1}^{s} M_{isj} \qquad (i = 1,2). \tag{26}$$

Remarquons que les couples $\{M_{1sj},\ M_{2sj}\}$, correspondants aux valeurs de $j = 1, 2, \ldots, s$, sont indépendantes et que, par conséquent,

$$G_{M_1(A),\,M_2(A)}(t_1,\ t_2) = \prod_{j=1}^{s} G_{M_{1sj},\,M_{2sj}}(t_1,\ t_2). \tag{27}$$

Pour calculer (27) il nous faut une expression pour la fonction génératrice simultanée des variables M_{1sj} et M_{2sj} qui correspondent au $j^{\text{ième}}$ cube R_{sj}. Pour simplifier les formules, écrivons γ_j au lieu de $\gamma(R_{sj},\ 0)$; γ_j est donc le nombre des centres d'amas qui au moment où l'on prend la photographie du ciel se trouvent dans R_{sj}. On a évidemment

$$\begin{aligned} (28)\quad G_{M_{1sj},\,M_{2sj}}(t_1,\ t_2) &= E\{t_1^{M_{1sj}}\,t_2^{M_{2sj}}\} \\ &= P\{\gamma_j = 0\} + \sum_{d=1}^{\infty} P\{\gamma_j = d\}\, E\{t_1^{M_{1sj}} t_2^{M_{2sj}} \mid \gamma_j = d\}, \end{aligned}$$

où l'espérance mathématique dans le second membre de la deuxième ligne est une espérance mathématique conditionnelle, étant donné que γ_j est égale à d.

Admettons que γ_j ait une valeur déterminée $d > 0$. Désignons par π_d une combinaison arbitraire des positions de d centres d'amas situés dans R_{sj}. Soit aussi F_d la fonction de probabilité simultanée de $3\,d$ coordonnées de ces d points, conditionnée par la donnée $\gamma_j = d$. Enfin soit $E\{t_1^{M_{1sj}} t_2^{M_{2sj}} \mid (\gamma_j = d), \pi_d\}$ l'espérance mathématique conditionnelle du produit indiqué, étant donné que $\gamma_j = d$ et que les d centres d'amas dans R_{sj} ont les positions π_d. On aura alors

$$(29)\qquad E\{t_1^{M_{1sj}} t_2^{M_{2sj}} \mid \gamma_j = d\} = \int_{R_{sj}} E\{t_1^{M_{1sj}} t_2^{M_{2sj}} \mid (\gamma_j = d)\, \pi_d\}\, dF_d.$$

Désignons par $C_1, C_2, \ldots, C_d$ les d centres d'amas situés dans R_{sj} et supposons que leurs positions π_d soient déterminées. Sous cette condition, les contributions $N_1(C_k)$ et $N_2(C_k)$, respectivement, à M_{1sj} et M_{2sj} sont des couples aléatoires indépendants les uns des autres. Donc

$$(30)\qquad E\{t_1^{M_{1sj}} t_2^{M_{2sj}} \mid (\gamma_j = d), \pi_d\} = \prod_{k=j}^{d} G_{N_1(C_k), N_2(C_k)}(t_1, t_2).$$

D'après la formule (24), on a

$$(31)\qquad G_{N_1(C_k), N_2(C_k)}(t_1, t_2) = {}^{\star}G_{n_1(C_k), n_2(C_k), \nu}(t_1, t_2, t_3).$$

A cause de la continuité postulée de la fonction $f^\star$, le second membre de (31) est une fonction continue des coordonnées, soient u_{1k}, u_{2k}, u_{3k}, du centre C_k. Donc le cube R_{sj}, ou sa frontière, doit contenir deux points, soient C' et C'' où la fonction (31) atteint respectivement son minimum et son maximum. Les formules (30) et (29) impliquent alors

$$(32)\qquad \{{}^{\star}G_{n_1(C'), n_2(C'), \nu}(t_1, t_2, t_3)\}^d \leqq E\{t_1^{M_{1sj}} t_2^{M_{2sj}} \mid \gamma_j = d\} \leqq \{{}^{\star}G_{n_1(C''), n_2(C''), \nu}(t_1, t_2, t_3)\}^d$$

et la formule (28) entraîne

$$(33)\qquad G_\gamma\{{}^{\star}G_{n_1(C'), n_2(C'), \nu}(t_1, t_2, t_3) \mid \Delta_s^3\} \leqq G_{M_{1sj}, M_{2sj}}(t_1, t_2) \leqq G_\gamma\{{}^{\star}G_{n_1(C''), n_2(C''), \nu}(t_1, t_2, t_3) \mid \Delta_s^3\}.$$

A cause de la continuité de toutes les fonctions qui interviennent, la formule (33) implique que le cube R_{sj}, ou sa frontière, doit contenir un point Γ_j, tel que

$$(34)\qquad \begin{aligned} G_{M_{1sj},\,M_{2sj}}(t_1,\ t_2) &= G_\gamma\{\,{}^\star G_{n_1(\Gamma_j),\,n_2(\Gamma_j),\,\nu}(t_1,\ t_2,\ t_3)\mid \Delta_s^3\,\} \\ &= \exp\{\,\Delta_s^3\, h[\,{}^\star G_{n_1(\Gamma_j),\,n_2(\Gamma_j),\,\nu}(t_1,\ t_2,\ t_3)]\,\}, \end{aligned}$$

en vertu de la formule (1). En substituant ce résultat dans la formule (27) et en passant à la limite lorsque s augmente indéfiniment, on obtient

$$(35)\qquad \log G_{M_1(A),\,M_2(A)}(t_1,\ t_2) = \iiint_{R(A)} h({}^\star G_{n_1(\Gamma),\,n_2(\Gamma),\,\nu}(t_1,\ t_2,\ t_3)\,du_1\,du_2\,du_3,$$

où Γ est un point arbitraire ayant les coordonnées u_1, u_2, u_3. Un autre passage à la limite, avec $A \to \infty$, donne

$$(36)\qquad \log G_{N_1,\,N_2}(t_1,\ t_2) = \iiint_{-\infty}^{+\infty} h[\,{}^\star G_{n_1(\Gamma),\,n_2(\Gamma),\,\nu}(t_1,\ t_2,\ t_3)]\,du_1\,du_2\,du_3.$$

A cause des propriétés de la fonction(2), si l'intégrale (36) diverge, elle divergera vers $-\infty$. Dans ce cas, la probabilité pour que N_1 ou N_2 possèdent une valeur infinie est égale à l'unité.

Il nous reste à calculer la fonction génératrice simultanée des trois variables aléatoires $n_1(\Gamma)$, $n_2(\Gamma)$ et $\nu(\Gamma)$, où Γ est un centre d'amas à coordonnées présentes u_1, u_2, u_3 que nous supposerons fixées. u_1, u_2, u_3 étant données, désignons par p_1, p_2, p_3 respectivement les probabilités conditionnelles, pour qu'une galaxie de l'amas de centre Γ soit une galaxie choisie dans ω_1 mais pas dans ω_2, pour qu'elle soit une galaxie choisie dans ω_2 mais pas dans ω_1 et pour qu'elle soit une galaxie choisie ω_1 et dans ω_2 simultanément. Avec une notation évidente, on a

$$(37)\qquad \begin{aligned} p_1 = p_1(u_1,\ u_2,\ u_3) = &\iiint_{\omega_1-\omega_1\omega_2} f\theta_1^\star\,dy_1\,dy_2\,dy_3 \\ &+ \iiint_{\omega_1\omega_2} f\theta_1^\star(1-\rho\theta_2^\star)\,dy_1\,dy_2\,dy_3, \end{aligned}$$

$$(38)\qquad \begin{aligned} p_2 = p_2(u_1,\ u_2,\ u_3) = &\iiint_{\omega_2-\omega_1\omega_2} f\theta_2^\star\,dy_1\,dy_2\,dy_3 \\ &+ \iiint_{\omega_1\omega_2} f\theta_2^\star(1-\rho\theta_1^\star)\,dy_1\,dy_2\,dy_3, \end{aligned}$$

$$(39)\qquad p_3 = p_3(u_1,\ u_2,\ u_3) = \iiint_{\omega_1\omega_2} f\theta_1^\star\,\theta_2^\star\,\rho\,dy_1\,dy_2\,dy_3,$$

où f est définie par la formule (16). Au moyen des probabilités p_1, p_2, p_3, un raisonnement élémentaire donne

$$(40)\qquad G_{n_1(\Gamma),\,n_2(\Gamma),\,\nu}(t_1,\ t_2,\ t_3) = G_\nu\{[1-p_1(1-t_1)-p_2(1-t_2)-p_3(1-t_1\,t_2)]\,t_3\}.$$

En substituant ce résultat dans la formule (36), on obtient

$$(41)\quad \log G_{N_1, N_2}(t_1, t_2) = \iiint_{-\infty}^{+\infty} h\{{}^{\star}G_{\nu}[(1-p_1(1-t_1)-p_2(1-t_2)-p_3(1-t_1t_2))t_3]\}\,du_1\,du_2\,du_3$$

qui est la formule fondamentale cherchée. Dans le paragraphe suivant nous lui donnons une forme (44) équivalente mais plus simple.

6. Répartition simultanée des nombres de galaxies comprises « en apparence » dans deux régions arbitraires. — Dans ce paragraphe nous considérons la première spécialisation de la formule fondamentale (41). Cette spécialisation est obtenue au moyen d'une définition des fonctions $N(r_1, r_2)$, $\theta_1^{\star}$, $\theta_2^{\star}$ et ρ telle que les symboles N_1 et N_2 définis par (23) représentent les nombres totaux des galaxies dont les positions apparentes tombent dans des régions arbitraires fixes ω_1 et ω_2 dans l'espace. Il est évident que, pour obtenir ce résultat, on n'a qu'à poser

$$N(r_1, r_2) \equiv r_1 \quad , \text{et} \qquad \theta_1^{\star} \equiv \theta_2^{\star} \equiv \rho \equiv 1.$$

L'opération de réduction correspondante donne

$$(42)\qquad \begin{aligned} &{}^{\star}G_{\nu}\{[1-p_1(1-t_1)-p_2(1-t_2)-p_3(1-t_1t_2)]t_3\} \\ &\equiv G_{\nu}\{1-p_1(1-t_1)-p_2(1-t_2)-p_3(1-t_1t_2)\}, \end{aligned}$$

et il en résulte

$$(43)\quad \log G_{N_1, N_2}(t_1, t_2) = \iiint_{-\infty}^{+\infty} h\{G_{\nu}[1-p_1(1-t_1)-p_2(1-t_2)-p_3(1-t_1t_2)]\}\,du_1\,du_2\,du_3.$$

Cette formule est valable indépendamment des fonctions $h_1(\tau)$ et $h_2(\tau)$ qui caractérisent le phénomène de l'expansion de l'univers. Donc elle est aussi valable lorsqu'on substitue $h_1(\tau) = h_2(\tau) \equiv 0$, cas dans lequel les symboles N_1 et N_2 représentent les nombres de galaxies dont les positions présentes tombent respectivement dans ω_1 et ω_2. En d'autres termes avec les substitutions indiquées, la formule (43) représente la répartition présente des galaxies dans des régions arbitraires ω_1 et ω_2, la plus générale impliquée par nos postulats.

En répétant les raisonnements de l'article I, il est aisé de démontrer que, si l'on n'impose aucune restriction sur la fonction $G_{\nu}(t)$, on ne restreindra pas la généralité de la répartition en supposant que les variables $\gamma(\mathrm{R}, 0)$ suivent une loi de Poisson, de façon que $h(t) \equiv -\lambda(1-t)$, où λ représente le nombre moyen de centres d'amas par unité de volume. Dans ce qui suit nous supposerons toujours que $h(t)$ a cette forme particulière. En utilisant cette forme, la formule fondamentale (41) peut être écrite plus simplement

$$(44)\quad \log \mathrm{G}_{\mathrm{N}_1,\mathrm{N}_2}(t_1, t_2) = -\lambda \iiint_{+\infty}^{+\infty} \{1 - {}^{\star}\mathrm{G}_{\nu}[1 - p_1(1-t_1) - p_2(1-t_2) - p_3(1-t_1 t_2)t_3]\}\, du_1\, du_2\, du_3.$$

Nous utiliserons cette formule dans les paragraphes 7 et 15.

7. Répartition simultanée des images des galaxies visibles sur des clichés photographiques pris dans deux angles solides arbitraires. — Dans ce chapitre, nous considérons une deuxième spécialisation de la formule fondamentale que nous prenons sous sa forme simplifiée (44) équivalente à (41). Comme dans le cas précédent, cette spécialisation de la formule fondamentale est réalisée par un choix convenable des fonctions $\mathrm{N}(r_1, r_2)$, $\theta_1^\star$, $\theta_2^\star$ et ρ. Aussi imposons-nous une restriction sur les régions ω_1 et ω_2. Notre but est d'identifier les variables aléatoires (23) avec les nombres des images des galaxies dans deux régions arbitraires sur la même plaque ou sur deux plaques photographiques différentes. Posons donc que ω_i représente un angle solide arbitraire, ayant son sommet à l'observateur, et dans lequel on prend une plaque photographique à magnitude limite m_i pour $i = 1, 2$. L'intersection de l'angle solide ω_i avec la plaque correspondante sera désignée par la même lettre ω_i. Admettons que $m_1 \geqq m_2$ de sorte que, si une galaxie a sa position apparente dans la partie commune $\omega_1 \omega_2$ de deux angles solides et si elle est suffisamment brillante pour être visible sur la première plaque, elle le sera aussi sur la seconde. S'il se trouve que $m_1 = m_2$, on est dans le cas d'une seule plaque photographique.

Pour que le N_i de la formule (23) soit égal au nombre des images des galaxies visibles dans la région ω_i sur la $i^{\text{ième}}$ plaque, il suffit de poser $\mathrm{N}(r_1, r_2) \equiv r_1$, d'identifier $\theta_i^\star$ à $\theta(\xi, m_i)$ et d'admettre $\rho\theta_1^\star \equiv 1$. On obtient aisément

$$(45)\quad \log \mathrm{G}_{\mathrm{N}_1,\mathrm{N}_2}(t_1, t_2) = -\lambda \iiint_{-\infty}^{+\infty} \{1 - \mathrm{G}_{\nu}[1 - p_1(1-t_1) - p_2(1-t_2) - p_3(1-t_1 t_2)]\}\, du_1\, du_2\, du_3,$$

où p_1, p_2, p_3 sont représentées par les intégrales ayant la forme générale (21).

La formule (45) représente une généralisation directe du résultat principal de l'article I. En effet, la formule (45) ci-dessus et la formule (41) de l'article I ont la même apparence. Ce qu'il y a de nouveau dans le résultat que nous venons d'établir c'est que la même formule, notamment la formule (45), caractérise la répartition simultanée des nombres N_1 et N_2 des images des galaxies, que l'univers soit en expansion ou non, excepté que les probabilités p_1, p_2 et p_3, qui y interviennent, doivent être calculées en partant de la formule (21) où la densité f, donnée dans (16), dépend des hypothèses sur l'état de l'univers. En particulier, si l'on pose

$$h_1(\tau) \equiv h_2(\tau) \equiv 0,$$

donc si l'on admet que l'univers est stationnaire, toutes les formules déduites ici coïncident avec les formules correspondantes de l'article I.

La formule (45) représente le résultat principal de la première partie du présent travail. Cette formule concerne deux variables aléatoires N_1 et N_2. Mais il est évident que la généralisation au cas d'un nombre quelconque de variables définies d'une manière semblable est immédiate. La même remarque s'applique à la formule fondamentale (44). En d'autres termes, la formule (45) et ses généralisations déterminent le processus stochastique général qui représente la répartition des images des galaxies sur les plaques photographiques et qui satisfait aux postulats (i)-(v) et (vii).

8. **Possibilité d'une vérification empirique de l'hypothèse de l'expansion de l'univers.** — Lorsque le présent chapitre a été abordé j'ai pensé à l'idée de M. Borel [12] suivant laquelle le problème central de la statistique mathématique consiste à inventer un système de tirages de boules des urnes, tel que le résultat le plus probable de ces tirages coïncide avec les faits d'observation. Les postulats (i)-(v) et (vii) du présent Ouvrage correspondent au système de tirages des boules de M. Borel et la formule (45) détermine les probabilités de tous les résultats possibles de tels tirages. Mais les postulats adoptés ne précisent pas entièrement le contenu de toutes les « urnes ». En effet, ces postulats ne caractérisent que la structure générale de la repartition des galaxies dans l'espace au moyen de trois fonctions dont la nature n'a pas été

précisée. La première de ces fonctions est la fonction génératrice $G_\nu(t)$ du nombre ν des galaxies par amas. La deuxième fonction est la densité de probabilité conditionnelle $f^\star$ des coordonnées présentes d'une galaxie, étant donnée la position du centre d'amas auquel cette galaxie appartient. Enfin la troisième fonction est la probabilité $\theta(\xi, m)$ pour que la magnitude apparente d'une galaxie à la distance apparente ξ soit plus petite qu'un nombre donné m.

Dans ces circonstances, avant d'aborder les probabilités numériques des résultats possibles d'observation, il est nécessaire de spécialiser le modèle général de la répartition des galaxies dans l'espace en substituant aux fonctions arbitraires $G_\nu(t)$, $f^\star$ et $\theta(\xi, m)$ certaines fonctions complètement précisées. Naturellement, cela peut être accompli d'une infinité de manières. Tout d'abord on peut penser à utiliser certaines recherches antérieures, si elle existent, et à employer les fonctions particulières suggérées par leurs résultats. C'est précisément ce qu'on a fait dans l'article II par rapport à la fonction $\theta(\xi, m)$: la forme adoptée était celle suggérée par les recherches de Hubble et Humason. Lorsque les recherches antérieures manquent de résultats suffisamment explicites pour pouvoir être utilisés de cette manière, on est forcé d'appliquer une autre méthode. Cette méthode consiste en un appel à l'intuition pour deviner les propriétés générales des fonctions inconnues, telles que $G_\nu(t)$ et $f^\star$, et en un choix de formules d'interpolation, avec un nombre limité des paramètres ajustables, pour représenter approximativement ces fonctions. L'interpolation consiste en un choix des valeurs des paramètres tel que les résultats des calculs théoriques s'accordent autant que possible avec les données empiriques.

Cette deuxième méthode a été employée, elle aussi, dans l'article II. Pour la représentation approchée de la fonction $G_\nu(t)$ on a postulé que la variable $\nu - 1$ suit une répartition binomiale négative

$$G_\nu(t) = [1 + \beta(1 - t)]^{-\alpha}, \tag{46}$$

où $\alpha > 0$ et $\beta > 0$ sont deux paramètres ajustables. Aussi, pour obtenir une approximation de $f^\star$ a-t-on adopté la densité normale de probabilité

$$f^\star = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^3 e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{3}(x_i - u_i)^2}, \tag{47}$$

avec un seul paramètre ajustable, σ.

Malheureusement, même avec ce choix destiné à faciliter les calculs, on a trouvé une formule (45) tellement compliquée que le calcul numérique des probabilités qu'elle détermine est impossible. Par conséquent, la comparaison entre le modèle théorique et les observations n'a pas pu être effectuée sur la base de ce qui semblait le plus probable, d'après le modèle (comme on aurait voulu le faire en suivant la suggestion de M. Borel). Toute autre méthode moderne, ayant une propriété optimum intelligible, nous est apparue inapplicable et il nous a fallu nous résigner au seul procédé pratiquement possible, sans justification théorique.

Le procédé de comparaison entre la théorie et l'observation qui paraissait techniquement possible est l'ancien procédé des moments. Pour l'appliquer il faut admettre qu'au moins les deux premiers moments de ν existent. En prenant les dérivées partielles de (45) par rapport à t_1 et à t_2 et en y substituant $t_1 = t_2 = 1$, on obtient à tour de rôle les espérances mathématiques de N_1 et de N_2

$$E(N_1) = \lambda\nu_1 R_{100}, \qquad E(N_2) = \lambda\nu_1 R_{010}, \tag{48}$$

les variances de ces deux variables,

$$\sigma^2_{N_1} = E(N_1) + \lambda(\nu_2 - \nu_1)R_{200}, \qquad \sigma^2_{N_2} = E(N_2) + \lambda(\nu_2 + \nu_1)R_{020} \tag{49}$$

et leur covariance,

$$\sigma_{N_1 N_2} = \lambda(\nu_2 - \nu_1)R_{110} + \lambda\nu_1 R_{001}, \tag{50}$$

où en général ν_k désigne le moment de ν d'ordre k et

$$R_{rst} = \iiint_{-\infty}^{+\infty} (p_1 + p_3)^r (p_2 + p_3)^s p_3^t \, du_1 \, du_2 \, du_3. \tag{51}$$

En examinant ces formules, on voit que l'expression Q définie par

$$Q = \frac{R_{110}}{\sqrt{R_{200} R_{020}}} = \frac{\sigma_{N_1 N_2} - \lambda\nu_1 R_{001}}{\sqrt{[\sigma^2_{N_1} - E(N_1)][\sigma^2_{N_2} - E(N_2)]}} \tag{52}$$

possède des propriétés intéressantes. En effet, on voit que la valeur de Q ne dépend pas de λ ni des propriétés de la variable aléatoire ν. Par contre, Q dépend de la structure interne des amas caractérisée par la fonction f^*, de la probabilité $\theta(\xi, m)$ et de ce que l'univers est en expansion ou non. Grâce à la structure particulière de la formule (52), l'expression Q a été appelée la quasi-corrélation entre les nombres N_1

et N_2. On voit que Q ne peut être négatif. En appliquant l'inégalité de Schwartz, on trouve que $Q \leqq 1$. On trouve aussi que pour que $Q = 1$ il est nécessaire et suffisant que les angles solides ω_1 et ω_2 coïncident et que $m_1 = m_2$.

Citons deux cas particuliers qui ont un intérêt spécial du point de vue des applications. Dans le premier cas, ω_1 et ω_2 sont deux carrés sur la même plaque photographique, égaux et disjoints, ayant la même orientation et avec une distance angulaire entre leurs centres, soit 2β. Dans ce cas la partie commune $\omega_1 \omega_2$ est vide et $p_3 \equiv R_{001} \equiv 0$. La quasi-corrélation calculée sous ces conditions sera désignée par $Q_1(\beta)$.

$$Q_1(\beta) = \frac{\sigma_{N_1 N_2}}{\sigma^2_{N_1} - E(N_1)} = \frac{R_{110}}{R_{200}}. \tag{53}$$

En calculant les moyennes des nombres des images des galaxies dans des carrés ω de mêmes dimensions que ω_1 et ω_2, en calculant les moyennes des deuxièmes puissances de ces nombres et, enfin, les moyennes des produits de ces nombres qui correspondent aux couples des carrés séparés par la distance 2β, on parvient à estimer le numérateur et le dénominateur de la formule (53), donc on parvient à estimer la quasi-corrélation $Q_1(\beta)$. Désignons par $Q_1^*(\beta)$ l'estimation ainsi obtenue. En changeant β on obtient une suite d'estimations, $Q_1^*(\beta_1)$, $Q_1^*(\beta_2)$, ..., ou, comme on dit, on obtient les quasi-corrélations sériales. De plus, les quasi-corrélations sériales peuvent être calculées pour des carrés ω de différentes grandeurs et en utilisant des plaques photographiques ayant différentes magnitudes limites.

Les valeurs empiriques des quasi-corrélations ainsi obtenues peuvent être comparées avec les valeurs théoriques calculées en partant des formes spéciales choisies des fonctions f^* et $\theta(\xi, m)$ et en calculant les intégrales R_{110} et R_{200}. C'est précisément ce qu'on a fait dans l'article II après avoir choisi (47) pour représenter f^*. Le calcul a été répété plusieurs fois en partant de valeurs différentes de σ et en cherchant à approcher autant que possible la suite des quasi-corrélations empiriques calculées d'avance pour les carrés $1^\circ \times 1^\circ$ sur les plaques de l'Observatoire Lick. Pour chaque système plausible de valeurs des paramètres, intervenant dans la formule de $\theta(\xi, m)$, on a pu trouver une valeur de σ qui donne des quasi-corrélations théoriques très proches de leurs valeurs empiriques. A présent le même travail se poursuit pour les carrés $10' \times 10'$. Tous les calculs postulent que l'univers est stationnaire, donc que

$$h_1(\tau) \equiv h_2(\tau) \equiv 0.$$

Les résultats qui ont déjà été obtenus indiquent que les valeurs de σ ajustées aux chiffres empiriques concernant les carrés $1^\circ \times 1^\circ$ paraissent être trop grandes pour traiter de la même manière les carrés $10' \times 10'$.

S'il se trouve impossible de déterminer une seule valeur de σ et un seul système de valeurs des paramètres dans $\theta(\xi, m)$ qui produisent une harmonie satisfaisante entre les quasi-corrélations sériales théoriques et empiriques calculées pour les carrés ω de grandeurs différentes, et pour plusieurs collections de plaques photographiques ayant des magnitudes limites différentes, alors il faudrait probablement conclure ou bien que les formules interpolatrices adoptées pour approcher les fonctions f^* et $\theta(\xi, m)$ ne sont pas suffisamment élastiques, ou bien qu'il y a quelque faute dans le modèle de l'univers stationnaire. Ultérieurement, si l'on trouve que les difficultés à obtenir une harmonie entre la théorie et les matériaux empiriques disponibles disparaissent ou diminuent considérablement dès que les intégrales R_{110} et R_{200} sont calculées en partant de l'hypothèse d'un univers en expansion, alors ce résultat pourrait être interprété comme une indication en faveur de cette hypothèse. Il est à regretter que les calculs numériques indiqués soient très compliqués et que pour les compléter il nous faudra encore assez de temps.

Le second cas particulier de la formule (52) qui devrait conduire à des résultats intéressants est celui où les angles solides ω_1 et ω_2 coïncident et où les magnitudes limites des deux photographies sont différentes, $m_1 > m_2$. Dans ce cas $\omega_1\omega_2 = \omega_1 = \omega_2 = \omega$, et l'on voit que $p_2 \equiv 0$, et $R_{010} = R_{001}$. En se rapportant à (48), la formule (52) peut être écrite sous la forme suivante :

$$(54) \qquad Q_2 = \frac{R_{110}}{\sqrt{R_{200}R_{020}}} = \frac{\sigma_{N_1N_2} - E(N_2)}{\sqrt{[\sigma^2_{N_1} - E(N_1)][\sigma^2_{N_2} - E(N_2)]}}.$$

Ici encore le second membre se prête à l'estimation au moyen des données empiriques. Pour arriver à une telle estimation il faudrait se servir de deux collections de plaques à deux magnitudes limites assez différentes, prises de la même région du ciel. Également, ce qui paraît même plus difficile, il faudrait que les images des galaxies sur ces plaques soient énumérées dans les mêmes systèmes des carrés ω. La partie du centre dans (54) dépend des mêmes intégrales R_{110}, R_{200} et R_{020} que dans le premier cas. Ces intégrales peuvent être calculées en partant des fonctions f^*, $\theta(\xi, m)$ et des valeurs choisies des paramètres. Donc, ces quasi-corrélations du deuxième genre (54), calculées pour des carrés ω de grandeur variable et pour un nombre de combinaisons des magnitudes limites des plaques, offrent une avenue nouvelle pour juger si les postulats adoptés ici correspondent à la réalité.

Malheureusement, les comparaisons diverses entre la théorie et les données empiriques signalées plus haut se compliquent par la présence des erreurs d'observation. Ce problème a été traité dans l'article II mais la solution obtenue n'est pas tout à fait satisfaisante.

DEUXIÈME PARTIE.

LE PROBLÈME DE L'INTERPÉNÉTRATION DES AMAS DE GALAXIES.

9. **Remarques préalables.** — La théorie présentée dans la première partie du présent Ouvrage s'occupe de la relation entre les propriétés postulées de la répartition spatiale des galaxies d'une part et les propriétés de la répartition des images des galaxies sur les clichés photographiques de l'autre. Évidemment, cette relation dépend de ce que l'univers est stationnaire ou non et les formules déduites reflètent cette dépendance.

Les problèmes traités dans la seconde partie de ce travail ont un caractère tout à fait différent. Il est vrai que la théorie traitée dans la deuxième partie est basée sur les mêmes postulats que celle de la première partie. Mais tandis que dans la première partie tous les six postulats, (i) à (v) et (vii), nous étaient nécessaires, dans la deuxième partie nous n'utiliserons que les cinq premiers. Les problèmes que nous aborderons ici ne se rattachent pas à la question de ce qu'on devrait observer sur les plaques photographiques mais, plutôt, à celle de ce qu'on pourrait voir dans l'espace s'il nous était possible de faire un voyage avec des vitesses beaucoup plus grandes que celle de la lumière.

En des termes plus précis, nous nous intéressons à la question de savoir si les amas de galaxies dans leurs positions présentes peuvent ou non être considérés comme des systèmes séparés. L'autre possibilité est qu'en général les amas de galaxies s'interpénètrent mutuellement à un tel degré que la conception d'un amas isolé ne correspond à rien qui existe dans l'espace.

C'est parce que nous ne nous intéressons qu'aux *positions présentes* des amas et des galaxies, que nos raisonnements dans la deuxième partie de l'Ouvrage sont indépendants du postulat (vii) ainsi que de l'hypothèse de l'expansion de l'univers. D'autre part, si l'on arrive à la conclusion que le modèle de l'univers considéré représente la réalité avec une précision satisfaisante et si l'on cherche à obtenir une image de ce qui se passe dans l'espace en calculant les valeurs numériques des formules déduites ci-dessous, alors les constantes diverses dont ces formules dépendent devront être déterminées au moyen de la théorie de la première partie de cet Ouvrage, donc en utilisant le postulat (vi) en combinaison avec une hypothèse convenable concernant l'état de l'univers, son état stationnaire ou son expansion.

Dans ce qui suit nous postulerons que la fonction f^* déterminant la structure interne des amas ne dépend que de la distance η entre les positions présentes d'une galaxie et du centre d'amas correspondant. Comme tous les raisonnements qui suivent ne dépendent que de la fonction f^*, l'astérisque qui distingue cette fonction de la densité de probabilité f des *coordonnées apparentes* de la galaxie sera inutile. Donc, le symbole $f(\eta)$ désignera désormais la densité de probabilité conditionnelle des coordonnées présentes X_1, X_2, X_3 de la galaxie, étant donné que le centre d'amas correspondant a les coordonnées fixées u_1, u_2, u_3.

Pour caractériser le degré d'interpénétration mutuelle des amas de galaxies, nous indroduisons plusieurs variables aléatoires et nous déduisons leurs lois de probabilité.

10. Problème de pénétration d'un amas déterminé, par des galaxies appartenant à l'amas le plus voisin. — Considérons un amas particulier, soit A_1, que nous appellerons l'amas choisi, et plaçons à son centre l'origine d'un système orthogonal des coordonnées. Désignons par $\nu_1 \geqq 1$ le nombre des galaxies qui appartiennent à A_1. Nous traiterons ν_1 comme une variable aléatoire satisfaisant aux postulats de la première partie de cet Ouvrage. Numérotons les galaxies de l'amas A_1

dans l'ordre de leur distance à l'origine et commençons par la galaxie la plus éloignée. Désignons par η_k la distance de la $k^{\text{ième}}$ galaxie de façon que

$$\eta_1 \geqq \eta_2 \geqq \ldots \geqq \eta_{\nu_1}.$$

Les distances η_k ($k = 1, 2, \ldots, \nu_1$) seront traitées comme des variables aléatoires satisfaisant aux postulats de la première partie du présent article.

Pour tout nombre positif x, désignons par $S(x)$ une sphère de rayon x ayant son centre au centre de l'amas choisi, c'est-à-dire, à l'origine du système des coordonnées. Soit également $S_k = S(\eta_k)$ la sphère de rayon aléatoire η_k.

Considérons maintenant l'ensemble de tous les amas dans l'espace, autres que l'amas choisi A_1, et désignons par A_2 l'amas particulier de cet ensemble dont le centre est le plus proche de celui de l'amas A_1. Cet amas A_2 sera appelé l'amas le plus proche de l'amas choisi A_1. Désignons par ξ la variable aléatoire qui représente la distance entre le centre de l'amas A_2 de celui de l'amas A_1. Désignons aussi par ν_2 le nombre des galaxies appartenant à A_2. Ce nombre sera traité comme une variable aléatoire satisfaisant aux postulats de la présente théorie.

S'il se trouve qu'une galaxie de l'amas A_2 a sa position présente à l'intérieur de la sphère S_k, nous dirons que cette galaxie pénètre l'amas A_1 à la profondeur k. Une des caractéristiques du degré de l'interpénétration mutuelle des amas de galaxies que nous considérons ici est le nombre aléatoire $\mu_k (k = 1, 2, \ldots)$ des galaxies de l'amas A_2 le plus proche qui pénètrent l'amas choisi jusqu'à la profondeur k.

Dans ce qui suit nous déduisons la formule de la fonction génératrice de μ_k sous deux hypothèses différentes. La première de ces hypothèses est que l'amas choisi A_1 contient un nombre déterminé $N \geqq k$ de galaxies, donc que la variable aléatoire $\nu_1 = N$. La seconde hypothèse est plus générale, à savoir que $\nu_1 \geqq N \geqq k$. En particulier, nous nous intéresserons à la probabilité pour que $\mu_1 = 0$, donc que l'amas choisi A_1 soit isolé de l'amas A_2 le plus proche. Il sera également intéressant de calculer l'espérance mathématique de μ_k, ou le nombre moyen de galaxies de l'amas le plus proche qui pénètrent l'amas choisi jusqu'à la profondeur k.

11. Problème de pénétration d'un amas déterminé, par des galaxies d'un amas quelconque. — Les problèmes qui se rattachent à la variable aléatoire μ_k sont les problèmes de pénétration de l'amas choisi par des galaxies de l'amas le plus proche. Mais il est évident que l'amas choisi A_1 peut être pénétré par des galaxies des amas plus éloignés, A_3, A_4, ... et la fréquence relative de ce phénomène caracterise, elle aussi, l'image globale de la répartition des galaxies dans l'espace. Donc, il est intéressant d'étudier la variable aléatoire τ_k définie pour $k = 1, 2, \ldots$ comme le nombre de tels amas, autre que l'amas choisi A_1, qui possèdent au moins une galaxie pénétrant l'amas choisi A_1 jusqu'à la profondeur k. Dans ce qui suit nous déduirons la fonction génératrice de τ_k sous les deux hypothèses mentionnées dans le paragraphe 10, à savoir que le nombre ν_1 de galaxies de l'amas A_1 est égal à un nombre donné $N \geqq k$ et que ce nombre ν_1 est au moins égal à N.

En examinant les divers détails possibles du phénomène d'interpénétration mutuelle des amas de galaxies, on s'aperçoit que, du point de vue de la question de savoir si les amas peuvent être considérés comme des systèmes dynamiques isolés, l'importance de toutes les pénétrations n'est pas la même. Supposons, par exemple, qu'un amas géant d'un millier de galaxies contienne un seul ou même plusieurs autres petits amas tout à fait encastrés dans son intérieur. Du point de vue d'un système dynamique isolé, ces petits amas encastrés pourraient, probablement, être ignorés. Au plus, pourrait-on les considérer comme des irrégularités structurelles du grand amas. Par conséquent, en étudiant les variables aléatoires μ_k et τ_k, il peut être intéressant de se limiter aux cas où l'amas pénétrant l'amas choisi a un certain minimum, soit N_2, de galaxies.

Il est aisé de voir que ce nouveau problème est un cas particulier du premier. En effet, si l'on obtient la fonction génératrice de μ_k ou celle de τ_k en n'imposant aucune restriction sur la variable ν_2 représentant le nombre de galaxies d'un amas pénétrant l'amas choisi, alors le résultat analogue correspondant à la restriction $\nu_2 \geqq N_2$ pourra être obtenu immédiatement au moyen de deux substitutions. En premier lieu, le nombre moyen λ de centres d'amas par unité de volume, doit être remplacé par $\lambda P\{\nu \geqq N_2\}$ qui est le nombre moyen de centres de tels amas contenant au moins N_2 galaxies. En second lieu, la fonction génératrice absolue $G_\nu(t)$ de la variable ν_2 doit être remplacée par la fonction génératrice conditionnelle $G_\nu(t \mid \nu \geqq N_2)$, étant donné que $\nu_2 \geqq N_2$. On voit donc que la restriction aux amas pénétrant l'amas choisi qui possèdent au moins un nombre déterminé N_2 de galaxies n'exige aucune formule générale nouvelle. D'autre part, si l'on veut obtenir des valeurs numériques (peut-être celles correspondant à la spécialisation de l'article II), la condition $\nu_2 \geqq N_2$ devra être introduite comme nous l'avons indiqué.

12. Problème des amas encastrés dans un amas choisi. — Considérons le problème signalé plus haut des irrégularités dans la structure d'un amas. Si toutes les galaxies d'un amas A' (autre que A_1) pénètrent dans l'amas choisi A_1 jusqu'à la profondeur k, nous dirons que l'amas A' est encastré dans A_1 à la profondeur k. Définissons la troisième variable aléatoire capable de caractériser la structure générale de la répartition des galaxies dans l'espace. Nous la désignons par ε_k et la définissons comme le nombre d'amas A' encastrés dans A_1 à la profondeur k. En désignant par ν_2 le nombre des galaxies de A', nous déduirons la fonction génératrice de ε_k sous l'hypothèse que ν_2 ne surpasse pas un nombre donné N_2.

13. Quelques formules utiles. — Les solutions des problèmes énumérés tout à l'heure exigent une formule donnant l'espérance mathématique d'une fonction $\varphi(\eta_k)$. Dans ce paragraphe nous déduisons cette formule dans deux hypothèses différentes. D'abord nous supposons que le nombre ν_1 de galaxies dans l'amas choisi A_1 a une valeur fixe $\nu_1 = N \geqq k$. Ensuite nous supposons que $\nu_1 \geqq N \geqq k$.

Comme auparavant, la lettre η désignera la variable aléatoire égale à la distance d'une galaxie de l'amas A_1 à son centre. On sait bien que la densité de probabilité de η est représentée par la formule

$$p_\eta(x) = 4\pi x^2 f(x). \tag{55}$$

Désignons par $F_\eta(x)$ la fonction de répartition correspondante

$$F_\eta(x) = P\{\eta \leqq x\} = \int_0^x p_\eta(t)\,dt. \tag{56}$$

Étant donné que l'amas choisi contient exactement $\nu_1 = N \geqq k$ galaxies, où N est un nombre certain, la densité de probabilité conditionnelle de η_k sera donnée par la formule bien connue,

$$(57)\quad p_{\eta_k}(y \mid \nu_1 = \mathrm{N}) = \frac{\mathrm{N}\,!}{(k-1)\,!\,(\mathrm{N}-k)\,!}\,\mathrm{F}_{\eta}^{\mathrm{N}-k}(y)\,[1-\mathrm{F}_{\eta}(y)]^{k-1}\,p_{\eta}(y).$$

Pour obtenir la densité de probabilité conditionnelle de η_k, étant donné que $\nu_1 \geqq \mathrm{N}$, il suffit de substituer dans (57) n à N, de multiplier le résultat par la probabilité conditionnelle pour que $\nu_1 = n$, étant donnée $\nu_1 \geqq \mathrm{N}$ et de sommer pour $n = \mathrm{N}, \mathrm{N}+1, \ldots$. De cette façon on obtient

$$(58)\quad p_{\eta_k}(y \mid \nu_1 \geqq \mathrm{N}) = \frac{[1-\mathrm{F}_{\eta}(y)]^{k-1}p_{\eta}(y)}{(k-1)\,!\,\mathrm{P}\{\nu_1 \geqq \mathrm{N}\}}\sum_{n=\mathrm{N}}^{\infty}\frac{n\,!}{(n-k)\,!}\,\mathrm{F}_{\eta}^{n-k}(y)\,\mathrm{P}\{\nu_1 = n\}.$$

Désignons maintenant par $\varphi(y)$ une fonction mesurable de l'argument non-négatif y. En substituant η_k dans $\varphi(y)$ au lieu de y, on obtient une variable aléatoire $\varphi(\eta_k)$. Si $\nu_1 = \mathrm{N}$, l'espérance mathématique conditionnelle de $\varphi(\eta_k)$ s'obtient en multipliant $\varphi(y)$ par (57) et en intégrant par rapport à y entre zéro et l'infini. Si l'on se sert de (58) au lieu de (57), la même méthode donne l'espérance mathématique conditionnelle de $\varphi(\eta_k)$, étant donné que $\nu_1 \geqq \mathrm{N} \geqq k$.

Dans les deux cas, il est commode d'introduire une variable d'intégration v liée à y par la formule

$$(59)\qquad \mathrm{F}_{\eta}[y(v)] \equiv v \qquad (0 \leqq v < 1).$$

De cette manière, on obtient

$$(60)\quad \mathrm{E}[\varphi(\eta_k) \mid \nu_1 = \mathrm{N}] = \frac{\mathrm{N}\,!}{(k-1)\,!\,(\mathrm{N}-k)\,!}\int_0^1 \varphi[y(v)]\,v^{\mathrm{N}-k}(1-v)^{k-1}\,dv$$

et

$$(61)\quad \mathrm{E}[\varphi(\eta_k) \mid \nu_1 \geqq \mathrm{N}] = \frac{1}{(k-1)\,!\,\mathrm{P}\{\nu_1 \geqq \mathrm{N}\}}\int_0^1 \varphi[y(v)]\,(1-v)^{k-1} \times \sum_{n=\mathrm{N}}^{\infty}\frac{n\,!}{(n-k)\,!}\,v^{n-k}\,\mathrm{P}\{\nu_1 = n\}\,dv.$$

Remarquons que si $\mathrm{N} = k$, la formule (61) se réduit à une forme simple

$$(62)\quad \mathrm{E}[\varphi(\eta_k) \mid \nu_1 \geqq k] = \frac{1}{(k-1)\,!\,\mathrm{P}\{\nu_1 \geqq k\}}\int_0^1 \varphi[y(v)]\,(1-v)^{k-1}\,\frac{d^k\mathrm{G}_{\nu_1}(v)}{dv^k}\,dv.$$

Dans le cas général, le calcul de la formule (61) est assez compliqué si le nombre $N-k$ est grand. D'autre part, si l'on adopte la spécialisation (46) de la fonction $G_\nu(t)$, particulièrement si l'on y substitue la valeur $\alpha = 1$ suggérée dans l'article II, la formule (61) peut être simplifiée. En effet, avec ces hypothèses,

$$(63)\qquad P\{\nu = n\} = \frac{1}{\beta+1}\left(\frac{\beta}{\beta+1}\right)^{n-1} \qquad (n=1,2,\ldots)$$

et

$$(64)\qquad P\{\nu \geqq N\} = \left(\frac{\beta}{\beta+1}\right)^{N-1}.$$

En substituant (63) dans la somme de (61), on voit que cette somme peut être écrite sous la forme

$$(65)\qquad \frac{1}{\beta+1}\left(\frac{\beta}{\beta+1}\right)^{k-1}\frac{d^k}{dx^k}\sum_{n=N}^{\infty} x^n = k!\left(\frac{\beta}{\beta+1}\right)^{N-1}\sum_{m=0}^{k}\frac{N!}{m!\,(N-m)!}\,\beta^{k-m}\,\frac{v^{N-m}}{(1+\beta-\beta v)^{k+1-m}},$$

où, pour simplifier, $x = \frac{\beta v}{\beta+1}\cdot$ Donc, la formule (61) se réduit à

$$(66)\qquad E[\varphi(\eta_k)\mid \nu_1 \geqq N] = k\int_0^1 \varphi[y(v)]\,(1-v)^{k-1}\sum_{m=0}^{k}\frac{N!}{m!\,(N-m)!}\,\beta^{k-m}\,\frac{v^{N-m}}{(1+\beta-\beta v)^{k+1-m}}\,dv.$$

Si le nombre k n'est pas très grand, (66) se prête aux calculs numériques.

14. Nombre des galaxies du plus proche amas qui pénètrent dans un amas déterminé. — Dans ce paragraphe nous nous occupons des fonctions génératrices $G_{\mu_k}(t \mid \nu_1 = N)$ et $G_{\mu_k}(t \mid \nu_1 \geqq N)$ du nombre aléatoire μ_k des galaxies de l'amas A_2, le plus proche de A_1, qui pénètrent dans A_1 jusqu'à la profondeur k. Plus précisément, la première de ces fonctions génératrices correspond à l'hypothèse $\nu_1 = N \geqq k$ et la seconde à l'hypothèse $\nu_1 \geqq N \geqq k$. Quelle que soit l'hypothèse H qu'on admette en calculant la génératrice de μ_k, on peut écrire

$$(67)\qquad G_{\mu_k}(t \mid H) = E(t^{\mu_k} \mid H)$$

et après, en appliquant la formule bien connue qui lie l'espérance mathématique absolue à l'espérance mathématique conditionnelle,

$$G_{\mu_k}(t \mid H) = E\{ E(t^{\mu_k} \mid \eta_k) \mid H \}. \tag{68}$$

Pour calculer $E(t^{\mu_k} \mid \eta_k)$ nous appliquons la même méthode et écrivons

$$E(t^{\mu_k} \mid \eta_k) = E\{ E(t^{\mu_k} \mid \eta_k, \xi, \nu_2) \mid \eta_k \}. \tag{69}$$

Si les variables η_k, ξ, ν_2 ont des valeurs fixes arbitraires, μ_k sera une variable binomiale ($=$ nombre des « succès » dans ν_2 épreuves indépendantes, avec la même probabilité de succès), et sa fonction génératrice s'écrira

$$E(t^{\mu_k} \mid \eta_k, \xi, \nu_2) = [1-(1-t)P(\xi, \eta_k)]^{\nu_2}, \tag{70}$$

ou le symbole $P(\xi, \eta_k)$ représente la probabilité conditionnelle pour qu'une galaxie se trouve dans la sphère $S(\eta_k)$, étant donnée la valeur de η_k et étant donné que la galaxie en question appartient à un amas dont le centre se trouve à la distance donnée ξ de l'origine. Pour des raisons de symétrie, on peut écrire, pour tout η positif,

$$P(\xi, \eta) = \iiint_{S(\eta)} f\left\{ [(x-\xi)^2 + y^2 + z^2]^{\frac{1}{2}} \right\} dx\, dy\, dz \tag{71}$$

et, si l'on adopte la spécialisation (47) et si l'on pose $\sigma = 1$,

$$P(\xi, \eta) = \frac{1}{\sqrt{2\pi}} \int_{\xi-\eta}^{\xi+\eta} e^{-\frac{t^2}{2}} dt - \frac{1}{\xi\sqrt{2\pi}} \left(e^{-\frac{(\xi-\eta)^2}{2}} - e^{-\frac{(\xi+\eta)^2}{2}} \right). \tag{72}$$

Revenons à la formule (70). En y substituant $\nu_2 = n$, en multipliant le résultat par $P\{\nu_2 = n\}$ et en sommant, on obtient

$$E(t^{\mu_k} \mid \eta_k, \xi) = G_{\nu_2}\{1-(1-t)P(\xi, \eta_k)\}. \tag{73}$$

Dans les formules précédentes ξ désigne la distance à l'origine du centre de l'amas A_2 le plus proche. Cette définition implique que, pour tout $x \geqq 0$, la fonction de répartition de ξ, $F_\xi(x)$, est égale à la différence entre l'unité et la probabilité pour que la sphère $S(x)$ ne contienne aucun centre d'amas. On a donc

$$F_\xi(x) = 1 - e^{-\frac{4\pi\lambda x^3}{3}} \qquad (x \geqq 0) \tag{74}$$

et la densité de probabilité correspondante s'écrit

$$p_\xi(x) = 4\pi\lambda x^2 e^{-\frac{4\pi\lambda x^3}{3}}. \tag{75}$$

En substituant dans (73) x à la place de ξ, en multipliant le résultat par (75) en intégrant pour x de zéro à l'infini, on obtient

$$(76) \qquad E(t^{\mu_k} \mid \eta_k) = 4\pi\lambda \int_0^\infty x^2 G_{\nu_2}\{1-(1-t)P(\xi, \eta_k)\} e^{-\frac{4\pi\lambda x^3}{3}} dx$$

ou encore, en introduisant la fonction $x(u)$ définie par

$$(77) \qquad F_\xi[x(u)] \equiv u \qquad (0 \leqq u < 1),$$

$$(78) \qquad E(t^{\mu_k} \mid \eta_k) = \int_0^1 G_{\nu_2}\{1-(1-t)P[x(u), \eta_k]\} du,$$

où

$$(79) \qquad x(u) = \sqrt[3]{\frac{-3\log(1-u)}{4\pi\lambda}}.$$

Pour compléter la solution du problème, il suffit d'appliquer les formules (60) et (61) et de poser que la fonction $\varphi(\eta_k)$ a la forme (78). Donc

$$(80) \qquad G_{\mu_k}(t \mid \nu_1 = N) = \frac{N!}{(k-1)!(N-k)!} \times \int_0^1 \left\{ v^{N-k}(1-v)^{k-1} \times \int_0^1 G_{\nu_2}\{1-(1-t)P[x(u), y(v)]\} du \right\} dv$$

et

$$(81) \qquad G_{\mu_k}(t \mid \nu_1 \geqq N) = \frac{1}{(k-1)!\,P\{\nu_1 \geqq N\}} \times \int_0^1 \left\{ (1-v)^{k-1} \sum_{n=N}^{\infty} \frac{n!}{(n-k)!} v^{n-k} P\{\nu_1 = n\} \times \int_0^1 G_{\nu_2}\{1-(1-t)P[x(u), y(v)]\} du \right\} dv.$$

15. Nombre des amas qui pénètrent l'amas déterminé et nombre des amas encastrés. — La variable aléatoire τ_k, définie dans le paragraphe 11, représente le nombre des amas, autres que l'amas A_1, qui possèdent au moins une galaxie pénétrant l'amas choisi A_1 jusqu'à la profondeur k. Nous appliquons la méthode du paragraphe 14 pour calculer la fonction génératrice de μ_k sous une hypothèse H, soit que $\nu_1 = N \geqq k$, soit que $\nu_1 \geqq N \geqq k$.

Comme plus haut, nous écrirons

$$G_{\tau_k}(t \mid H) = E\{E(t^{\tau_k} \mid \eta_k) \mid H\} \tag{82}$$

et le problème se réduit au calcul de la fonction génératrice de la variable $\tau(\eta)$, qui est le nombre des amas possédant au moins une galaxie située dans la sphère $S(\eta)$ de rayon fixe $\eta \geqq 0$. Une fois la fonction génératrice de $\tau(\eta)$ calculée, le résultat cherché s'obtient en appliquant les formules (60) ou (61), exactement comme dans le chapitre précédent.

Les mêmes remarques s'appliquent au problème de la fonction génératrice de la variable ε_k. On commence par définir la variable $\varepsilon(\eta)$ qui représente le nombre d'amas dont toutes les galaxies sont contenues dans la sphère $S(\eta)$ de rayon fixe $\eta \geqq 0$. Ensuite on substitue $\eta = \eta_k$ et l'on applique les formules (60) et (61).

Les fonctions génératrices des variables $\tau(\eta)$ et $\varepsilon(\eta)$ s'obtiennent au moyen de deux particularisations à savoir de la formule fondamentale (44) le choix des définitions de $N(r_1, r_2)$, ω_1, ω_2 et θ_i^* telles que $N_2 \equiv 0$ et que N_1 coïncide respectivement avec $\tau(\eta)$ ou avec $\varepsilon(\eta)$.

Conformément à ce que nous avons signalé dans les paragraphes 10 et 11, pour augmenter la généralité des résultats, il est avantageux d'admettre que le nombre aléatoire des galaxies de l'amas choisi peut suivre une loi de probabilité différente de celle du nombre aléatoire des galaxies des amas qui le pénètrent. Donc, le nombre aléatoire des galaxies de l'amas choisi A_1 sera désigné par ν_1 et celui de tout autre amas A_2 pénétrant A_1 par ν_2.

Pour obtenir $G_{\tau(\eta)}(t)$ de la formule fondamentale (44), posons $\theta^* \equiv 1$, substituons $S(\eta)$ à ω_1 et l'ensemble vide à ω_2. Alors $p_2 = p_3 = 0$ et

$$p_1 = P(\xi, \eta), \tag{83}$$

où $\xi^2 = u_1^2 + u_2^2 + u_3^2$. Posons aussi, $N(r_1, r_2) \equiv 0$ pour $r_1 = 0$ et $N(r_1, r_2) \equiv 1$ pour $r_1 > 0$. On voit immédiatement qu'alors le N_1 de la formule (23) coïncide avec $\tau(\eta)$ et que $N_2 \equiv 0$.

La substitution indiquée des valeurs de p_1, p_2 et p_3 dans l'argument de G_ν dans (44) et l'opération de réduction correspondant à la définition de $N(r_1, r_2)$ adoptée, donnent

$$^{\star}G_\nu\{[1-(1-t_1)P(\xi, \eta)]t_3\} = t_1 + (1-t_1)G_\nu[1-P(\xi, \eta)]. \tag{84}$$

Comme il s'agit des amas pénétrant l'amas choisi, la variable ν doit être identifiée avec ν_2. En substituant (84) dans (44), on obtient

$$\log G_{\tau(\eta)}(t) = -\lambda(1-t)\,I(\eta), \tag{85}$$

où

$$I(\eta) = \iiint_{-\infty}^{+\infty} \{1 - G_{\nu_2}[1 - P(\xi, \eta)]\}\, du_1\, du_2\, du_3 \tag{86}$$

$$= 4\pi \int_0^{\infty} \xi^2 \{1 - G_{\nu_2}[1 - P(\xi, \eta)]\}\, d\xi.$$

La fonction génératrice de τ_k s'obtient maintenant par une simple application de la formule (60) ou (61) et l'on a

$$G_{\tau_k}(t \mid \nu_1 = N) = \frac{N!}{(k-1)!\,(N-k)!} \int_0^1 v^{N-k}(1-v)^{k-1}\, e^{-\lambda(1-t)\,I[y(v)]}\, dv \tag{87}$$

et

$$G_{\tau_k}(t \mid \nu_1 \geqq N) = \frac{1}{(k-1)!\,P\{\nu_1 \geqq N\}} \int_0^1 (1-v)^{k-1} \tag{88}$$

$$\times\, e^{-\lambda(1-t)\,I[y(v)]} \sum_{n=N}^{\infty} \frac{n!}{(n-k)!}\, v^{n-k}\, P\{\nu_1 = n\}\, dv,$$

où la fonction $y(v)$ est définie par (59).

La fonction génératrice de ε_k s'obtient par la même méthode. On commence par calculer $G_{\varepsilon(\eta)}(t)$, au moyen de la formule fondamentale (44) et en y posant $\omega_1 = S(\eta)$, $\omega_2 =$ ensemble vide, $\theta_1^\star \equiv 1$, qui conduit à $p_1 = P(\xi, \eta)$, $p_2 = p_3 = 0$ et $N_2 \equiv 0$. Adoptons maintenant une spécialisation nouvelle de la fonction $N(r_1, r_2)$. Posons notamment $N(r_1, r_2) \equiv 1$ pour $r_1 = r_2 > 0$ et $N(r_1, r_2) = 0$ pour toutes les autres valeurs des deux arguments. On voit bien qu'avec cette définition de $N(r_1\, r_2)$, la variable N_1 définie par (23) coïncide avec $\varepsilon(\eta)$. Après avoir substitué $p_1 = P(\xi, \eta)$ et $p_2 = p_3 = 0$ dans (44), l'opération de réduction donne

$$^\star G_\nu\{[1 - (1-t_1)\,P(\xi, \eta)]\, t_3\} = 1 - (1-t_1)\,\{G_\nu[P(\xi, \eta)] - P\{\nu = 0\}\} \tag{89}$$

et l'on a

$$\log G_{\varepsilon(\eta)}(t) = -\lambda(1-t)\,J(\eta), \tag{90}$$

où

$$J(\eta) = 4\pi \int_0^{\infty} \xi^2 \{G_{\nu_2}[P(\xi, \eta)] - G_{\nu_2}(0)\}\, d\xi. \tag{91}$$

Par conséquent, en appliquant les formules (60) et (61), on obtient

$$(92)\qquad G_{\varepsilon_k}(t \mid \nu_1 = N) = \frac{N!}{(k-1)!\,(N-k)!}\int_0^1 v^{N-k}(1-v)^{k-1}\, e^{-\lambda(1-t)\,J[y(v)]}\, dv$$

et

$$(93)\qquad G_{\varepsilon_k}(t \mid \nu_1 \geqq N) = \frac{1}{(k-1)!\,P\{\nu_1 \geqq N\}}\int_0^1 (1-v)^{k-1} \times e^{-\lambda(1-t)\,J[y(v)]}\sum_{n=N}^{\infty}\frac{n!}{(n-k)!}\, v^{n-k}\, P\{\nu_1 = n\}\, dv.$$

Les formules (87), (88), (92) et (93) représentent la solution générale des problèmes des amas pénétrant l'amas choisi et des amas encastrés.

16. Spécialisation des formules générales. — Dans le présent paragraphe nous suivons les méthodes de l'article II et spécialisons les formules générales obtenues plus haut. Plus précisément, nous adoptons les formules (46) avec $\alpha = 1$ et (47) avec $\sigma = 1$, et calculons les probabilités pour que chacune des variables μ_k, τ_k et ε_k soit égale à zéro, ainsi que les espérances mathématiques de ces variables. Tous ces calculs sont basés sur l'hypothèse que $\nu_1 \geqq N_1$ et $\nu_2 \geqq N_2$, où N_1 et N_2 sont des nombres arbitraires $N_1 \geqq k$ et $N_2 \geqq 1$.

Remarquons que, pour tenir compte de l'hypothèse $\nu_1 \geqq N$ en admettant (46), il suffit de conduire les calculs non pas à partir des formules (81), (88) et (93) qui concernent le cas général où la répartition de ν_1 est quelconque, mais à partir des formules analogues qu'il est aisé de déduire en employant la formule (66) au lieu de (61). Dans ces calculs il sera commode d'introduire le symbole $\Phi(v, N_1, k)$ pour désigner l'expression

$$(94)\qquad \Phi(v, N_1, k) = k\,\frac{v^{N_1-k}(1-v)^{k-1}}{1+\beta-\beta v}\sum_{m=0}^{k}\frac{N_1!}{m!\,(N_1-m)!}\left(\frac{\beta v}{1+\beta-\beta v}\right)^{k-m}$$

qui y intervient constamment. L'hypothèse que $\nu_2 \geqq N_2$ influence les calculs de deux manières différentes. Premièrement, au lieu de λ, qui représente le nombre moyen de centres d'amas par unité de volume, il nous faut écrire $\lambda(N_2)$, le nombre moyen de centres de tels amas qui contiennent au moins N_2 galaxies,

$$(95)\qquad \lambda(N_2) = \lambda\, P\{\nu_2 \geqq N_2\} = \lambda\left(\frac{\beta}{\beta+1}\right)^{N_2-1}.$$

Donc, sous les hypothèses (46) et $\nu_2 \geqq N_2$, la formule (79) qui définit la fonction $x(u)$ doit être remplacée par

$$x(u) = \sqrt[3]{\frac{-3\log(1-u)}{4\pi\lambda}\left(\frac{\beta+1}{\beta}\right)^{N_2-1}}. \tag{96}$$

Deuxièmement, les hypothèses (46) et $\nu_2 \geqq N_2$ entrent dans nos calculs en attribuant une forme précise à la fonction génératrice de la variable aléatoire ν_2,

$$G_{\nu_2}(t \mid \nu_2 \geqq N_2) = \frac{t^{N_2}}{1+\beta-\beta t}. \tag{97}$$

Notons encore les deux conséquences de l'hypothèse $\sigma = 1$. La première consiste dans le fait que si l'on a une estimation de σ exprimée en parsecs, la valeur de λ à utiliser dans nos calculs devra être égale à la moyenne du nombre des centres d'amas par unité de volume égale à $\frac{4\pi\sigma^3}{3}$. La deuxième conséquence de la même hypothèse $\sigma = 1$ est le caractère particulier de la fonction $y(v)$ définie par (59). En tenant compte de (56), (55) et (47) on voit qu'avec $\sigma = 1$ l'équation (59) se réduit à

$$\sqrt{\frac{2}{\pi}}\int_0^{y(v)} x^2 e^{-\frac{x^2}{2}}\,dx = v. \tag{98}$$

Donc, pour tout v compris entre zéro et un, $y(v)$ représente un nombre tel que la probabilité pour que la variable aléatoire χ avec trois degrés de liberté soit plus grande que $y(v)$ est égale à $1-v$. On voit comment utiliser cette propriété dans les calculs numériques, en se servant des tables de la loi de probabilité de χ.

La probabilité pour que μ_k soit égal à zéro s'obtient en substituant $t=0$ dans l'expression de la génératrice correspondante. En effectuant les substitutions indiquées, on obtient

$$\begin{aligned} &P\{\mu_k = 0 \mid (\nu_1 \geqq N_1)(\nu_2 \geqq N_2)\} \\ &\quad = \int_0^1 \left\{ \Phi(v, N_1, k) \int_0^1 \frac{(1-P[x(u), y(v)])^{N_2}}{1+\beta P[x(u), y(v)]}\,du \right\} dv. \end{aligned} \tag{99}$$

Pour obtenir l'espérance mathématique de μ_k, on calcule la dérivé par rapport à t de sa fonction génératrice et l'on y substitue $t=1$. Le résultat est

$$\begin{aligned} &E\{\mu_k \mid (\nu_1 \geqq N_1)(\nu_2 \geqq N_2)\} \\ &\quad = (N_2+\beta)\int_0^1 \left\{ \Phi(v, N_1, k) \int_0^1 P[x(u), y(v)]\,du \right\} dv. \end{aligned} \tag{100}$$

Pour obtenir les résultats analogues concernant τ_k, commençons par spécialiser le $I(\eta)$ de la formule (86). En adoptant (46) et en supposant $\nu_2 \geqq N_2$, on obtient

$$(101) \qquad I(\eta, N_2) = 4\pi \int_0^\infty \xi^2 \frac{1 + \beta P(\xi, \eta) - [1 - P(\xi, \eta)]^{N_2}}{1 + P(\xi, \eta)} d\xi.$$

Des calculs simples donnent

$$(102) \quad P\{\tau_k = 0 \mid (\nu_1 \geqq N_1)(\nu_2 \geqq N_2)\} = \int_0^1 \Phi(v, N_1, k)\, e^{-\lambda(N_2) I[y(v), N_2]}\, dv.$$

De la même façon

$$(103) \quad E\{\tau_k \mid (\nu_1 \geqq N_1)(\nu_2 \geqq N_2)\} = \left(\frac{\beta}{\beta+1}\right)^{N_2-1} \lambda \int_0^1 \Phi(v, N_1, k)\, I[y(v), N_2]\, dv.$$

En appliquant la même méthode à la variable ε_k, on commence par spécialiser la fonction $J(\eta)$. Pour indiquer la dépendance du nombre N_2, nous la désignons par $J(\eta, N_2)$ et l'on a

$$(104) \qquad J(\eta, N_2) = 4\pi \int_0^\infty \xi^2 \frac{[P(\xi, \eta)]^{N_2}}{1 + \beta - \beta P(\xi, \eta)} d\xi.$$

Alors

$$(105) \quad P\{\varepsilon_k = 0 \mid (\nu_1 \geqq N_1)(\nu_2 \geqq N_2)\} = \int_0^1 \Phi(v, N_1, k)\, e^{-\lambda(N_2) J[y(v), N_2]}\, dv$$

et

$$(106) \quad E\{\varepsilon_k \mid (\nu_1 \geqq N_1)(\nu_2 \geqq N_2)\} = \left(\frac{\beta}{\beta+1}\right)^{N_2-1} \lambda \int_0^1 \Phi(v, N_1, k)\, J[y(v), N_2]\, dv.$$

Dans le paragraphe suivant nous donnons quelques valeurs numériques des formules que nous venons de calculer.

17. Le degré d'interpénétration des amas de galaxies, à la lumière des résultats empiriques actuellement disponibles. — Malgré l'apparente simplicité des formules du paragraphe 16, les calculs numériques correspondants sont extrêmement pénibles. Je suis très reconnaissant à ma collègue, M^lle Elizabeth L. Scott pour avoir bien voulu analyser les fonctions sous les signes somme et inventer des méthodes d'intégration numérique convenables. Les calculs laborieux, dont les résultats sont présentés ci-dessous, ont été effectués par M^me J. Lovasich et M^me M. Vasilevskis sous la direction de M^lle Scott. Je les en remercie bien cordialement.

Les chiffres calculés se rattachent aux trois variables aléatoires μ_k, τ_k

et ε_k et sont donnés dans trois tableaux correspondants. Chaque tableau donne pour quelques valeurs de k, les probabilités pour que la variable en question soit égale à zéro et l'espérance mathématique de cette variable. En interprétant les tableaux, il faut se rappeler que les chiffres qu'ils contiennent dépendent : *a*. de la spécialisation du modèle général de répartition des galaxies adoptée dans l'article II, et *b*. des estimations des constantes qui interviennent dans le modèle spécialisé, obtenues de l'analyse des données de M. Shane et M. Wirtanen et publiées dans l'article II et puis dans l'article III. On se rappellera également que ces estimations dépendent de l'hypothèse que l'univers est stationnaire. Donc les chiffres des tableaux I, II et III ne représentent rien de définitif mais, par contre, on doit s'attendre à ce qu'une révision soit nécessaire.

Voici les valeurs des constantes dont on s'est servi :

$\lambda \equiv$ l'espérance mathématique du nombre des centres d'amas par parsec$^3 = 1{,}63.10^{-19}$;

$\alpha \equiv$ l'exposant dans la fonction génératrice binomiale négative adoptée pour interpoler la génératrice de $\nu - 1$, où ν représente le nombre des galaxies par amas $= 1$;

$\beta \equiv$ le deuxième paramètre dans la même fonction génératrice binomiale négative, dont l'estimation est égale à $91{,}5$;

$m_1 - M_0 \equiv$ la différence entre la magnitude limite m_1 des plaques de M. Shane et M. Wirtanen et la moyenne M_0 des magnitudes absolues des galaxies $= 32{,}5$ mag;

$\sigma_M \equiv$ l'écart type des magnitudes absolues des galaxies $= 1{,}00$ mag;

$\sigma \equiv$ l'écart type de la densité normale (47) à trois variables adoptée pour caractériser la structure interne des amas.

Les valeurs indiquées de λ, α, β, $m_1 - M_0$ et de σ_M ont été utilisées pour obtenir les deux colonnes des tables que nous donnons ci-dessous. Quant à σ, nous lui avons attribué deux valeurs différentes. La première des colonnes des trois tables est basée sur la valeur $\sigma = 5.10^5$ parsecs qu'on a choisi entre les valeurs plausibles indiquées dans l'article II. Comme nous l'avons signalé dans l'introduction au présent Ouvrage, les résultats ultérieurs, publiés dans l'article III, suggèrent que cette valeur de σ est trop grande. Par conséquent, la deuxième colonne de tables I, II et III a été calculée en supposant que $\sigma = 2{,}5.10^5$.

TABLE I.

Probabilités et espérances mathématiques concernant les nombres aléatoires μ_k des galaxies de l'amas le plus proche qui pénètrent l'amas choisi jusqu'à la profondeur k.

Probabilité ou espérance mathématique.	Valeur de σ adoptée dans les calculs. 5,0.10^5 parsecs.	2,5.10^5 parsecs.
$P\{\mu_1 = 0 \mid (\nu_1 \geqq 1)(\nu_2 \geqq 1)\}$	0,0106	0,237
$E\{\mu_1 \mid (\nu_1 \geqq 1)(\nu_2 \geqq 1)\}$	69,9	27,0
$P\{\mu_1 = 0 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 100)\}$	0,00330	0,426
$E\{\mu_1 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 100)\}$	120,0	31,0
$P\{\mu_5 = 0 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 100)\}$	0,0142	0,546
$E\{\mu_5 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 100)\}$	84,5	17,9

TABLE II.

Probabilités et espérances mathématiques concernant les nombres aléatoires τ_k d'amas, dont les galaxies pénètrent l'amas choisi jusqu'à la profondeur k.

Probabilité ou espérance mathématique.	Valeur de σ adoptée dans les calculs. 5,0.10^5 parsecs.	2,5.10^5 parsecs.
$P\{\tau_1 = 0 \mid (\nu_1 \geqq 1)(\nu_2 \geqq 1)\}$	0,00172	0,204
$E\{\tau_1 \mid (\nu_1 \geqq 1)(\nu_2 \geqq 1)\}$	14,0	1,8
$P\{\tau_1 = 0 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 100)\}$	0,0017	0,40
$E\{\tau_1 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 100)\}$	7,8	0,98
$P\{\tau_5 = 0 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 100)\}$	0,0084	0,54
$E\{\tau_5 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 100)\}$	4,9	0,62

TABLE III.

Probabilité et espérance mathématique concernant le nombre aléatoire ε_1 d'amas encastrés dans l'amas choisi.

Probabilité ou espérance mathématique.	Valeur de σ adoptée dans les calculs. 5,0.10^5 parsecs.	2,5.10^5 parsecs.
$P\{\varepsilon_1 = 0 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 1)\}$	0,737	0,958
$E\{\varepsilon_1 \mid (\nu_1 \geqq 100)(\nu_2 \geqq 1)\}$	0,324	0,040

On voit que, si le modèle considéré de répartition des galaxies correspond à la réalité et si la valeur de $\sigma = 5.10^5$ parsecs est à peu près exacte, un amas de galaxies tout à fait isolé des autres sera une exception rare. En effet, dans ce cas, la conception d'un amas n'apparaît que comme un élément d'une structure interpolatrice de la répartition des galaxies.

D'autre part, si la valeur 5.10^5 est trop grande pour représenter σ et la valeur $2,5.10^5$ parsecs plus proche de la réalité, les chiffres de la deuxième colonne des trois tables indiquent une situation différente. En effet, la première ligne de la table II montre que, dans ce cas, à peu près 20 % de tous les amas sont des amas isolés. Si l'on se limite aux amas d'au moins 100 galaxies, cette proportion monte jusqu'à 40 %. On voit donc qu'avec la valeur $2,5.10^5$ parsecs de σ un amas isolé doit être un phénomène assez fréquent.

Enfin, si l'on trouve, comme le suggère l'analyse préliminaire de la plaque de M. Wilson, que la valeur $2,5.10^5$ parsecs est encore trop grande pour représenter σ, et qu'on devrait en adopter une autre beaucoup plus petite, alors, il faudrait conclure par extrapolation, que la répartition des galaxies dans l'espace se réduit à la répartition des amas isolés les uns des autres.

Ces résultats indiquent que le problème suivant à traiter est celui de trouver des estimations plus exactes des paramètres qui interviennent dans le modèle considéré de la répartition spatiale des galaxies. On peut espérer que l'analyse des données empiriques déjà accumulées par les divers observateurs, au moyen de formules basées sur l'hypothèse de l'univers stationnaire et puis au moyen de celles basées sur l'hypothèse contraire, aboutira à de telles estimations. Si une telle analyse conduit à la détermination de fonctions $G_\nu(t)$, f^* et $\theta(\xi, m)$ et d'estimations des paramètres qui y interviennent telles que les conséquences de la théorie s'accordent bien avec les résultats d'énumération des galaxies dans des carrés ω petits ou grands et sur des plaques aux magnitudes limites variables, alors de telles estimations des paramètres pourront être substituées dans les formules que nous venons de déduire pour obtenir une caractérisation de la structure de l'univers plus proche de la réalité.

Avant de conclure, remarquons que la théorie basée sur l'hypothèse de l'expansion de l'univers devrait être développée en tenant compte de la théorie de relativité, ce que nous ne pouvions pas faire. M. McVittie a eu l'obligeance de s'occuper de quelques formules de la première partie de cet ouvrage et ses résultats vont paraître prochainement [13].

BIBLIOGRAPHIE.

[1] J. Neyman et E. L. Scott, *A theory of the spatial distribution of galaxies* (*Astroph. J.*, t. 116, 1952, p. 144-163).

[2] J. Neyman, E. L. Scott et C. D. Shane, *On the spatial distribution of galaxies; a specific model* (*Astroph. J.*, t. 117, 1953, p. 92-133).

[3] C. D. Shane et C. A Wirtanen, *The surface distribution of the extragalactic nebulæ* (*Astron. J.*, t. 59, 1954, p. 285-312).

[4] E. L. Scott, C. D. Shane et M. D. Swanson, *Comparison of the synthetic and actual distribution of galaxies on a photographic plate* (*Astroph. J.*, t. 119, 1954, p. 91-112).

[5] F. Zwicky, *Neue Methoden der kosmologischen Forschung* (*Helv. Phys. Acta*, t. 26, 1953, p. 241-254).

[6] Paul Couderc, *The expansion of the universe*, p. 231, McMillan, New York, 1952.

[7] F. Zwicky, *Dispersion in the large scale distribution of galaxies* (*Publ. Astron. Soc. of the Pacific*, t. 64, 1952, p. 247-255).

[8] J. Neyman et E. L. Scott, *Frequency of separation and of interlocking of clusters of galaxies* (*Proc. National Acad. Sc.*, t. 39, 1953, p. 737-743).

[9] Norbert Wiener, *The homogeneous chaos* (*Amer. J. Math.*, t. 60, 1938, p. 897-936).

[10] E. Hubble, *The luminosity function of nebulæ*. I (*Astroph. J.*, *t.* 84, 1936, p. 158-179).

[11] E. Hubble, *The luminosity function of nebulæ*, II (*Astroph. J.*, t. 84, 1936, p. 270-295).

[12] Émile Borel, *Éléments de la théorie des probabilités*, Hermann, Paris, 1924.

[13] G. C. McVittie, *Relativity and the statistical treatment of the distribution of galaxies*, à paraître dans l'*Astronomical Journal.*

[14] J. Neyman, E. L. Scott et C. D. Shane, *The index of clumpiness of the distribution of images of galaxies*, (*Astrophysical Journal*, *Supplément*, t. 1954, p. 269-293).

[15] J. Neyman et E. L. Scott, *Spatial distribution of galaxies-analysis of the theory of fluctuations* (*Proc. National Acad. Sc.*, t. 40, 1954, p. 873-881).

[16] J. Neyman et E. L. Scott, *On the inapplicability of the theory of fluctuations to galaxies*, (*Astronomical Journal*, t. 60, 1955, p. 33-38).