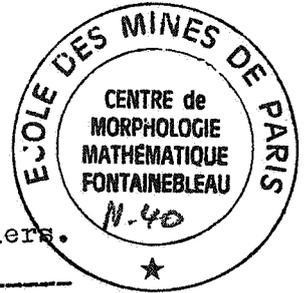


G. Matheron (1960)

- I -

Note géostatistique n° 26



Construction de Modèles Numériques de gisements miniers.

1/ Les méthodes dites de simulation - très utilisées, en particulier, dans le domaine des sciences économiques - consistent à construire un modèle mathématique d'un phénomène naturel. Elles permettent de contrôler la valeur des idées théoriques qui ont servi de base à la construction du modèle - puisque l'on peut examiner dans quelle mesure les propriétés déduites de l'étude mathématique du modèle singent, ou simulent, les propriétés du phénomène réel. Inversement, elles peuvent conduire à la découverte de lois nouvelles du phénomène réel, l'expérience étant infiniment plus facile à réaliser sur un modèle purement mathématique. De même, il serait intéressant pour le géostatisticien de savoir construire des modèles numériques de gisements miniers, c'est-à-dire d'être capable d'aligner des séries de chiffres, présentant les mêmes caractéristiques que celles que l'on pense devoir attribuer à la succession dans l'espace des teneurs d'un gisement réel.

2/ Dans le cas d'un gisement, on considérera la teneur comme une variable régionalisée douée d'une loi de dispersion intrinsèque de Wijsienne. Pour obtenir  $n$  valeurs numériques plausibles  $x_i$ , susceptibles de représenter les teneurs de  $n$  volumes  $v_i$  occupant des positions respectives bien déterminées à l'intérieur du gisement fictif que l'on cherche à construire, la première idée qui vient à l'esprit est de se donner à priori une loi de probabilité - lognormale par exemple - à  $n$  variables  $x_i$  reliées par les  $\frac{n(n-1)}{2}$  covariances de Wijsiennes déduites des distances mutuelles des volumes  $v_i$ . Un tirage au sort simultané selon une telle loi engendrerait un excellent modèle. On est malheureusement arrêté par l'extrême complication des calculs dès que  $n$  prend des valeurs un peu élevées. Un système linéaire de  $n$  équations à  $n$  inconnues conduit au calcul de  $n + 1$  déterminants d'ordre  $n$ , dont chacun contient  $n + 1$  termes, soit, pour  $n = 40$ ,  $8,16 \cdot 10^{47}$  termes - ce qui est beaucoup, même pour une machine électronique.

3/ Les méthodes basées sur les développements en série de Fourier permettent de construire, à bien moindre frais, des modèles de gisement à une seule dimension. Elles ont de plus l'avantage de n'impliquer aucune hypothèse a priori sur la forme mathématique de la loi de distribution statistique des teneurs. La condition de périodisation imposée par de telles méthodes nous conduit à considérer la succession des teneurs le long d'une galerie circulaire infiniment mince.

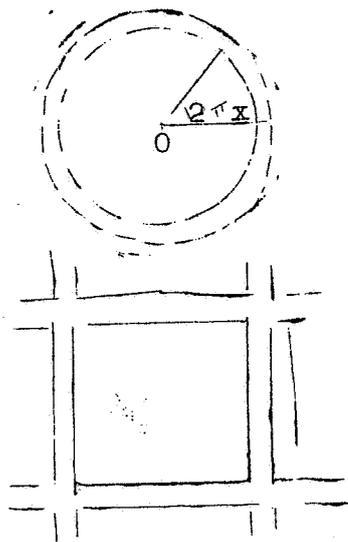


Fig. 1

Une telle galerie (fig.1) est, si l'on veut, l'idéalisation de 4 tronçons de galeries et recoupés ceinturant un panneau carré dans un échantillonnage périphérique. Le schéma de de Wijs en série de Fourier permet alors de construire le modèle cherché avec une relative facilité.

4/ Soit donc une galerie circulaire dont le rayon R nous servira d'unité de longueur ( $R = 1$ ). Désignons par  $2\pi x$  l'angle polaire d'un point M de la galerie. La variable x varie de 0 à 1. On cherchera à déterminer les coefficients  $a_p$  et  $b_p$  d'un développement de Fourier du type :

$$(I) \quad f(x) = m + \sum_{p=1}^{\infty} (a_p \cos 2\pi p x + b_p \sin 2\pi p x)$$

capable de représenter la teneur  $f(x)$  d'une volée de longueur  $2\pi l$  centrée au point M d'angle polaire  $2\pi x$ . Nous avons choisi, comme exemple numérique,

$$l = \frac{1}{40}$$

de manière que la totalité de la galerie soit occupée par 40 volées successives. Le terme m, qui figure dans l'expression (I), n'est autre que la teneur moyenne de la galerie circulaire, et peut être choisie arbitrairement, sous réserve seulement d'être suffisamment élevée pour que  $f(x)$  ne prenne aucune valeur négative. Les coefficients  $a_p$  et  $b_p$ , de leurs côtés, doivent être tels que la covariogramme

$$(2) \quad G(h) = \frac{1}{2} \sum_{h=1}^{\infty} (a_p^2 + b_p^2) \cos 2\pi p h$$

- c'est-à-dire la fonction représentant la covariance des teneurs des volées l distantes de l'angle polaire h - soit du type de Wijsien, c'est-à-dire de la forme<sup>1</sup> :

$$(3) \quad G(h) = 3 \times \sum_{p=1}^{\infty} \frac{\sin^2 \frac{\pi p l}{l}}{\pi^2 p^3 l^2} \cos 2 \pi p h$$

Le coefficient  $3 \times$  est la dispersion absolue et  $l$  est la longueur de la volée rapportée à la longueur de la galerie (dans les exemples qui suivent,  $l = \frac{l}{6}$ ). Nous obtiendrons donc un modèle satisfaisant de la forme (1) à condition de choisir des coefficients  $a_p$  et  $b_p$  vérifiant les relations :

$$(4) \quad a_p^2 + b_p^2 = 6 \times \frac{\sin^2 \frac{\pi p l}{l}}{\pi^2 p^3 l^2}$$

La dispersion absolue pouvant être choisie arbitrairement, nous prendrons  $\frac{1}{6}$ , de manière à faire disparaître le terme  $6 \times$  de la formule (4). Chaque harmonique de Fourier est ainsi déterminé en amplitude, la phase restant arbitraire. Une fonction de la forme

$$(5) \quad f(x) = m + \sum_{p=1}^{\infty} \frac{\sin \frac{\pi p l}{l}}{\pi p^{3/2} l} \cos (2 \pi p x - \phi_p)$$

dans laquelle les phases  $\phi_p$  peuvent être prises arbitrairement, permet de construire un modèle convenable. En pratique, on devra se limiter à un nombre fini de termes de la série (5). La convergence étant relativement lente, il convient cependant de prendre un nombre assez grand de termes. Dans les exemples ci-dessous, on a retenu les 40 premiers termes de la série. Quant aux phases  $\phi_p$ , entièrement arbitraires, elles peuvent être choisies suivant un procédé arbitraire quelconque. Dans toute la suite, nous prendrons  $m = 0$ , puisqu'il suffira d'ajouter une constante à  $f(x)$  pour faire disparaître les tencurs négatives.

5/ Des modèles particuliers, pairs ou impairs, peuvent être obtenus en choisissant les phases  $\phi_p$  de telle manière que  $f(x)$  comporte uniquement soit des termes en sinus, soit des termes en cosinus. Les deux fonctions  $f(x)$  les plus simples que l'on puisse ainsi obtenir sont les deux fonctions  $C(x)$  et  $S(x)$  définies comme suit :

---

1. Cf. Mon "traité de Géostatistique Appliquée", ch. III Paragraphe I7.

$$(6) \quad \begin{cases} C(x) = \sum_{p=1}^{\infty} \frac{\sin \pi p l}{\pi^{3/2} p} \cos 2\pi p x \\ S(x) = \sum_{p=1}^{\infty} \frac{\sin \pi p l}{\pi^{3/2} p} \sin 2\pi p x \end{cases}$$

Pour plus de simplicité, nous avons multiplié  $C(x)$  et  $S(x)$  par  $\pi l = \frac{\pi}{40}$ , de sorte que le coefficient  $C_p$  de la fonction trigonométrique est :

$$(7) \quad C_p = \frac{1}{p^{3/2}} \sin \frac{\pi p}{40}$$

La table suivante donne les valeurs numériques des 40 premiers coefficients  $C_p$

TABLE I

Valeurs des 40 premiers coefficients  $C_p$

$p$	$C_p$	$p$	$C_p$	$p$	$C_p$	$p$	$C_p$
1	0.07845	11	0.02084	21	0.01036	31	0.00376
2	0.05540	12	0.01946	22	0.00957	32	0.00325
3	0.04494	13	0.01819	23	0.00881	33	0.00276
4	0.03863	14	0.01701	24	0.00809	34	0.00229
5	0.03423	15	0.01590	25	0.00739	35	0.00185
6	0.03089	16	0.01486	26	0.00672	36	0.00143
7	0.02847	17	0.01387	27	0.00608	37	0.00104
8	0.02598	18	0.01294	28	0.00546	38	0.00067
9	0.02405	19	0.01204	29	0.00487	39	0.00032
10	0.02236	20	0.01118	30	0.00430	40	0.00000

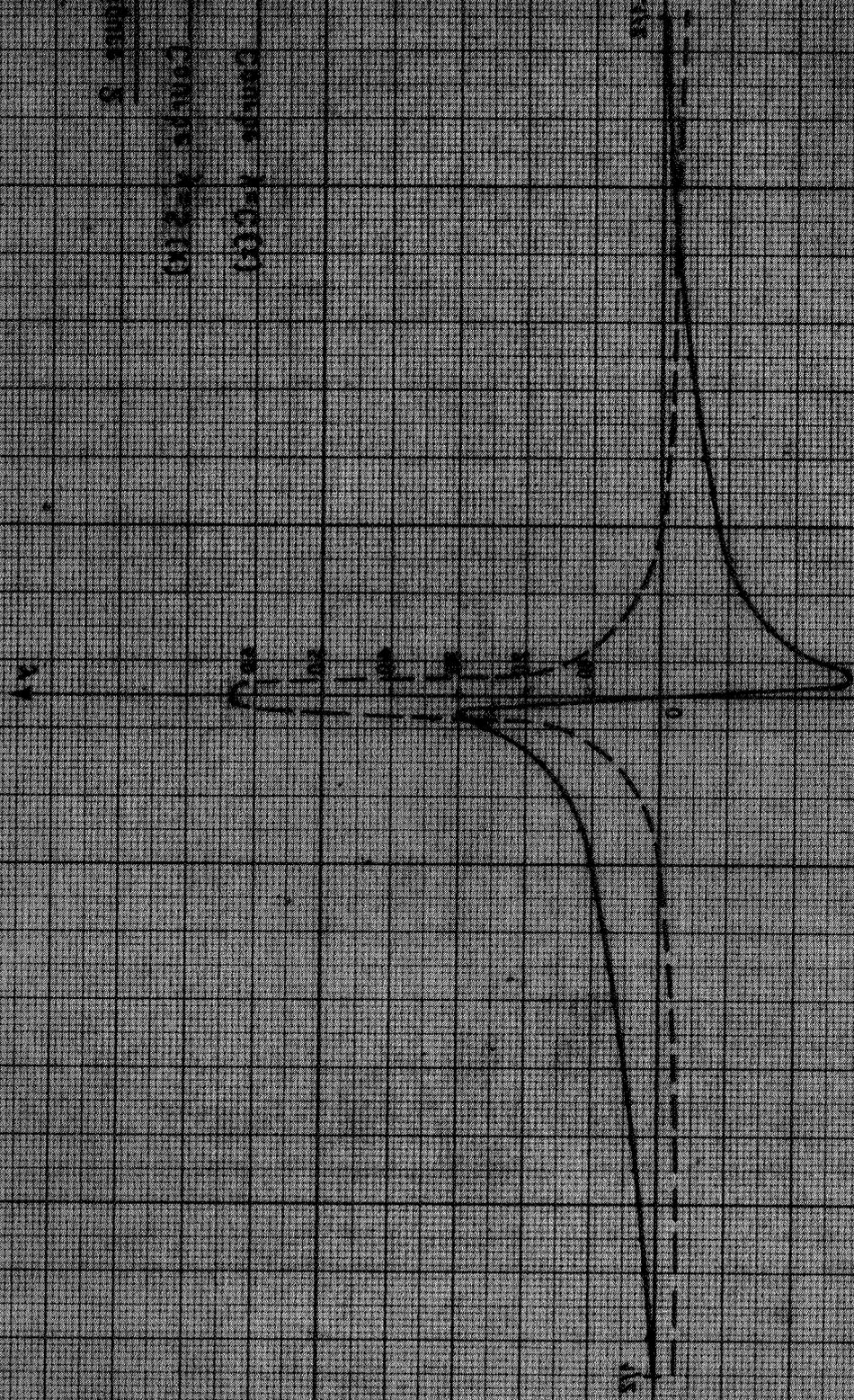
Les valeurs numériques des deux fonctions  $\frac{\pi}{40} C(x) = \sum_{p=1}^{\infty} C_p \cos 2\pi p x$  et  $\frac{\pi}{40} S(x) = \sum_{p=1}^{\infty} C_p \sin 2\pi p x$  ont été calculées pour les valeurs

de  $x$  multiples de  $1/40$ . Les valeurs numériques sont portées sur le table II, tandis que la figure 2 représente graphiquement la variation de ces deux fonctions. Pour ne pas traîner des décimales inutiles, les valeurs numériques ont été multipliées par 100. Pour les valeurs de  $x$  négatives, on a  $C(-x) = C(x)$  et  $S(-x) = -S(x)$ , égalités qui permettent de compléter la table II sur une période entière (de  $x = -\frac{1}{2}$  à  $x = +\frac{1}{2}$ ).

Figure 3

Complex plane

Complex plane



Ces deux fonctions présentent la particularité de localiser aux alentours de l'origine la quasi totalité de leur variabilité. La fonction C (x) représente des teneurs décroissant très rapidement de part et d'autre d'un maximum très accusé. La fonction S (x) représente un seuil relevant brutalement des teneurs dont la tendance générale est à une décroissance régulière. Aucune de ces deux répartitions très particulières n'est susceptible d'être représentée par une loi de distribution (au sens statistique) d'un type simple. Mais toutes deux sont de Wijsiennes. Elles représentent une variation fonctionnelle des teneurs de part et d'autre d'un accident.

TABLE II  
valeurs numériques des fonctions C (x) et S (x)

x	$\frac{100}{40} C(x)$	$\frac{100}{40} S(x)$	x	$\frac{100}{40} C(x)$	$\frac{100}{40} S(x)$
0	+ 62.871	0.000	11/40	- 3.657	+ 4.573
1/40	+ 13.046	+ 24.135	12/40	- 3.928	+ 4.054
2/40	+ 5.898	+ 17.292	13/40	- 4.152	+ 3.462
3/40	+ 2.774	+ 13.458	14/40	- 4.341	+ 2.722
4/40	+ 0.926	+ 11.413	15/40	- 4.470	+ 2.379
5/40	- 0.326	+ 9.933	16/40	- 4.556	+ 1.877
6/40	- 1.353	+ 8.616	17/40	- 4.620	+ 1.400
7/40	- 1.986	+ 7.518	18/40	- 4.650	+ 1.032
8/40	- 2.584	+ 6.760	19/40	- 4.749	+ 0.488
9/40	- 3.041	+ 5.931	20/40	- 4.769	+ 0.000
10/40	- 3.381	+ 5.220			

6/ Plus généralement, des fonctions paires et impaires de la forme :

$$(8) \begin{cases} P(x) = \sum p C_p \cos 2\pi px \\ I(x) = \sum p C_p \sin 2\pi px \end{cases}$$

où  $p$  peut prendre la valeur + 1 ou - 1, qui se déduisent de C (x) et S (x) en affectant un signe arbitraire à chacun des harmoniques, permettent de construire tous les modèles pairs ou impairs possibles.

Nous avons construit deux fonctions paires,  $P_1(x)$  et  $P_2(x)$  et deux fonctions impaires  $I_1(x)$  et  $I_2(x)$  particulières en donnant aux  $\xi_p$  les séquences de signe arbitraires, et à peu près aléatoires présentées sur le tableau suivant :

TABLEAU III

: Signes  $\xi_p$  arbitrairement attribués aux termes de  $P_1, P_2, I_1,$  et  $I_2$  :

: p :	: $P_1$ :	: $P_2$ :	: $I_1$ :	: $I_2$ :	: p :	: $P_1$ :	: $P_2$ :	: $I_1$ :	: $I_2$ :
: 1:	: + :	: - :	: + :	: - :	: 21:	: - :	: - :	: - :	: - :
: 2:	: - :	: - :	: + :	: - :	: 22:	: - :	: + :	: + :	: - :
: 3:	: - :	: - :	: - :	: + :	: 23:	: + :	: + :	: + :	: + :
: 4:	: - :	: + :	: + :	: + :	: 24:	: - :	: - :	: + :	: - :
: 5:	: + :	: - :	: - :	: - :	: 25:	: - :	: - :	: + :	: - :
: 6:	: - :	: + :	: - :	: + :	: 26:	: - :	: - :	: - :	: + :
: 7:	: + :	: - :	: + :	: + :	: 27:	: - :	: - :	: + :	: - :
: 8:	: + :	: - :	: - :	: - :	: 28:	: + :	: + :	: - :	: + :
: 9:	: + :	: + :	: + :	: - :	: 29:	: + :	: + :	: + :	: - :
: 10:	: - :	: + :	: + :	: + :	: 30:	: - :	: - :	: - :	: + :
: 11:	: - :	: + :	: - :	: + :	: 31:	: + :	: + :	: - :	: - :
: 12:	: + :	: - :	: + :	: - :	: 32:	: + :	: + :	: + :	: + :
: 13:	: - :	: + :	: - :	: + :	: 33:	: - :	: - :	: + :	: + :
: 14:	: + :	: - :	: + :	: - :	: 34:	: + :	: + :	: - :	: - :
: 15:	: - :	: - :	: - :	: + :	: 35:	: - :	: - :	: - :	: + :
: 16:	: - :	: - :	: + :	: + :	: 36:	: - :	: - :	: + :	: - :
: 17:	: + :	: + :	: - :	: - :	: 37:	: + :	: + :	: - :	: + :
: 18:	: + :	: - :	: - :	: + :	: 38:	: + :	: + :	: + :	: + :
: 19:	: + :	: - :	: + :	: + :	: 39:	: + :	: + :	: + :	: - :
: 20:	: + :	: + :	: + :	: - :	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:

Il est clair que le tirage au sort du signe de  $C_p$  dans une fonction  $P(x)$  ou  $I(x)$  est un cas particulier du tirage au sort des phases  $\phi_p$  d'un développement du type (5): Les valeurs numériques des deux fonctions paires  $P_1$  et  $P_2$  et des deux fonctions impaires  $I_1$  et  $I_2$  sont représentées sur la table IV.

Pour  $x$  négatif, on prendra  $P(-x) = P(x)$  et  $I(-x) = -I(x)$ .

TABLE IV  
Valeurs numériques des fonctions  $P_1(x)$ ,  $P_2(x)$ ,  $I_1(x)$  et  $I_2(x)$ .

$x$	$P_1(x)$	$P_2(x)$	$I_1(x)$	$I_2(x)$
0	- 3.171	- 0.505	0.000	0.000
1/40	+ 3.159	- 3.879	+ 4.264	+ 1.192
2/40	- 2.082	+ 0.616	- 2.788	+ 2.366
3/40	+ 2.392	- 4.408	+ 2.332	- 0.198
4/40	+ 14.172	+ 9.704	- 6.405	+ 9.047
5/40	+ 2.844	+ 16.582	- 6.781	+ 18.401
6/40	- 0.045	+ 15.871	- 16.050	+ 13.574
7/40	+ 15.340	+ 7.286	- 13.116	+ 9.052
8/40	+ 12.035	+ 3.916	- 9.564	+ 5.252
9/40	+ 12.800	+ 14.935	- 7.505	+ 7.417
10/40	+ 2.325	+ 8.129	- 27.992	+ 8.902
11/40	- 2.800	+ 1.515	- 13.443	+ 9.251
12/40	+ 9.647	- 5.222	- 8.404	+ 5.548
13/40	- 3.280	- 12.710	+ 13.710	+ 18.358
14/40	- 0.789	+ 1.672	+ 8.420	+ 6.612
15/40	- 16.856	- 10.414	- 2.191	+ 1.271
16/40	- 17.998	- 3.036	+ 12.397	- 23.225
17/40	- 23.764	+ 6.654	- 6.972	- 8.632
18/40	- 4.728	- 14.910	- 7.098	+ 4.120
19/40	+ 0.715	- 20.409	+ 4.088	- 9.692
20/40	- 5.207	+ 15.089	0.000	0.000

Enfin, bien que les fonctions ainsi obtenues présentent déjà des caractéristiques très analogues à des séries de teneurs, il convient de s'affranchir des restrictions apportées par la parité  $P_1$  et  $P_2$  et l'imparité de  $I_1$  et  $I_2$ . Nous introduisons les quatre fonctions  $f_1$ ,  $f_2$ ,  $f_3$  et  $f_4$  définies de la manière suivante :

$$(9) \quad \left( \begin{array}{l} f_1(x) = P_1(x) + I_1(x) \\ f_2(x) = P_2(x) + I_1(x) \\ f_3(x) = P_1(x) + I_2(x) \end{array} \right.$$

$$( f_4 (x) = P_2 (x) + I_2 (x) )$$

Ces fonctions ne sont plus liées par aucune condition de symétrie autour de  $x = 0$ .

Ce sont des fonctions du type 5 avec des phases  $\phi_p$  égales à  $\frac{\pi}{4}$ ,  $\frac{3\pi}{4}$ ,  $-\frac{\pi}{4}$ , ou  $-\frac{3\pi}{4}$ .

Elles sont représentées graphiquement sur les figures 3, 4, 5 et 6, et tabulées sur la table V.

TABLE V				
valeurs numériques des 4 fonctions $f_1 (x)$ , $f_2 (x)$ , $f_3 (x)$ et $f_4 (x)$ .				
x	$f_1$	$f_2$	$f_3$	$f_4$
20/40	- 5.207	- 15.089	- 5.207	- 15.089
19/40	- 3.373	- 24.497	+ 10.407	- 10.717
18/40	+ 2.370	- 7.812	- 8.848	- 19.030
17/40	- 16.792	+ 13.626	- 15.132	+ 15.286
16/40	- 30.395	- 15.433	+ 5.227	+ 20.189
15/40	- 14.665	- 8.223	- 18.127	- 11.685
14/40	- 9.209	- 6.748	- 7.401	- 4.940
13/40	- 16.990	- 26.420	- 21.638	- 31.068
12/40	- 18.051	+ 3.182	+ 4.099	- 10.770
11/40	+ 10.643	+ 11.928	- 12.051	- 10.766
10/40	+ 29.717	+ 35.521	- 6.577	- 0.773
9/40	+ 20.305	+ 21.940	+ 5.383	+ 7.018
8/40	+ 21.599	+ 13.480	+ 6.783	- 1.306
7/40	+ 28.456	+ 20.402	+ 6.288	- 1.766
6/40	+ 16.005	+ 31.921	- 13.619	+ 2.297
5/40	+ 9.625	+ 23.363	- 15.557	- 1.819
4/40	+ 20.577	+ 16.109	+ 5.125	+ 0.657
3/40	+ 0.060	- 6.740	+ 2.590	- 4.210
2/40	+ 0.706	+ 3.404	- 4.448	- 1.756
1/40	- 1.105	- 8.143	+ 1.967	- 5.071
			(à suivre)	

Figure 4

Courbe  $y = f(x)$

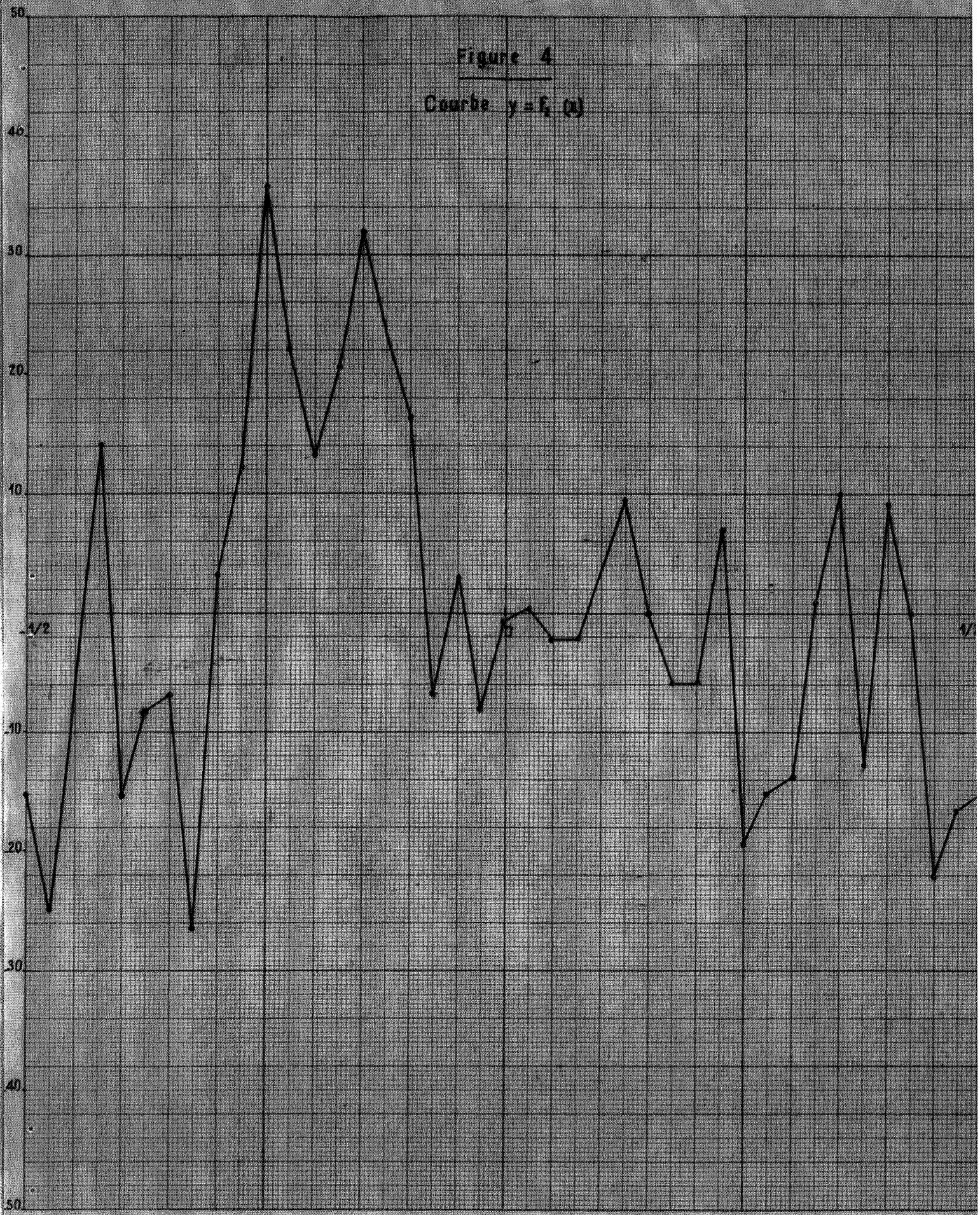


Figure 5  
Course  $y=f(x)$

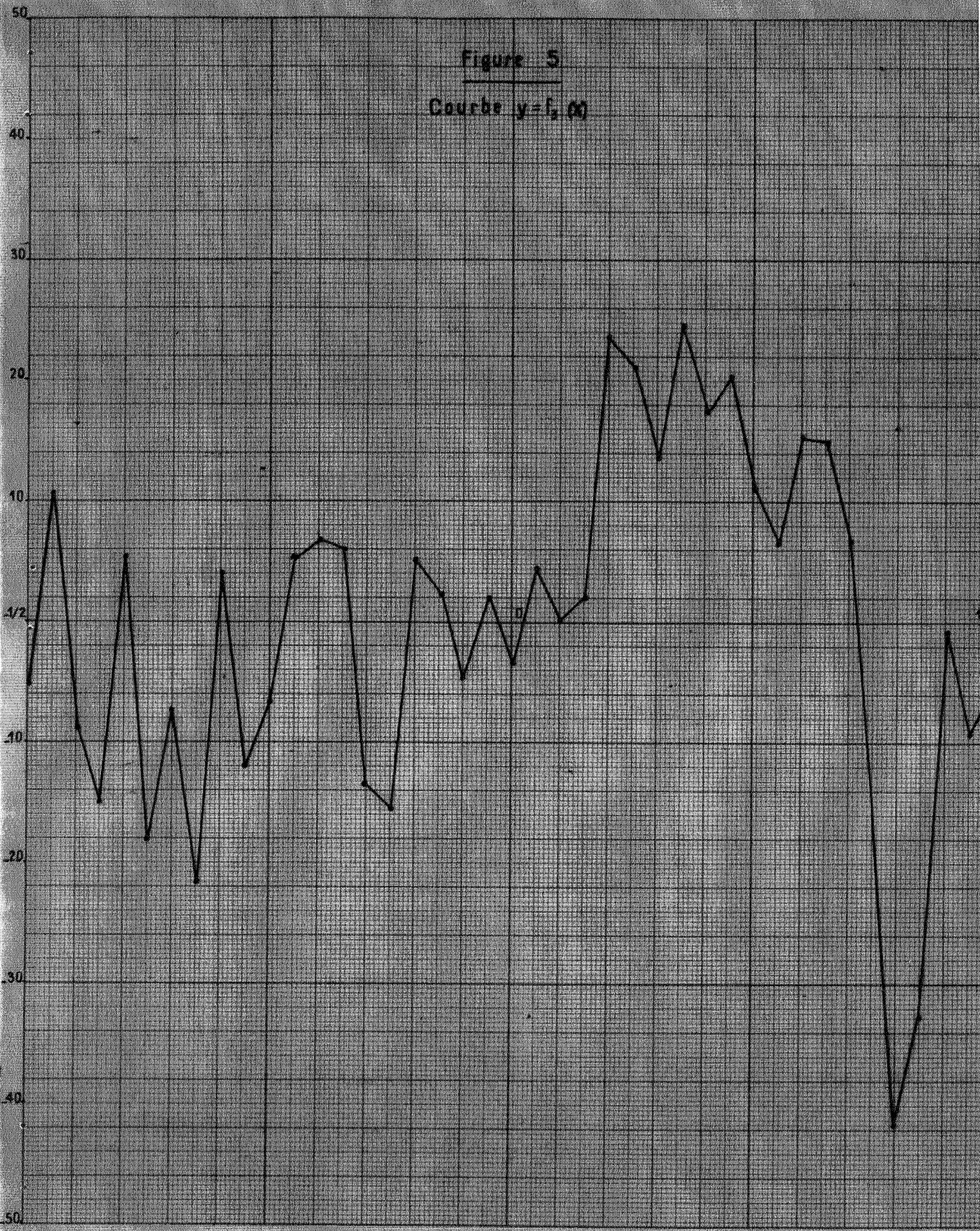
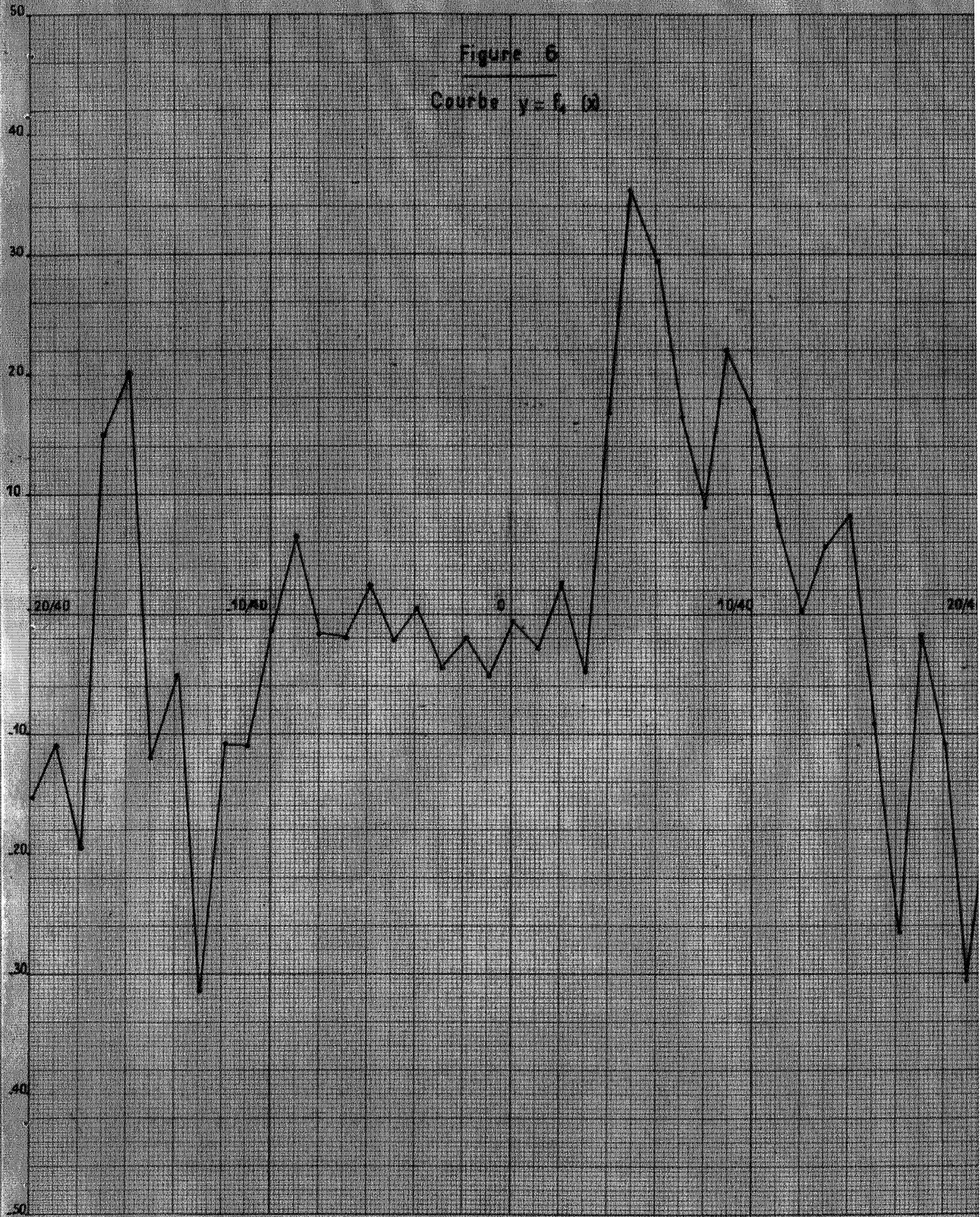


Figure 6  
Courbe  $y = f(x)$



suite Table V				
x	f <sub>1</sub>	f <sub>2</sub>	f <sub>3</sub>	f <sub>4</sub>
0	- 3.171	- 0.505	- 3.171	- 0.505
1/40	+ 7.423	+ 0.385	+ 4.351	- 2.687
2/40	- 4.870	- 2.172	+ 0.284	+ 2.982
3/40	+ 4.724	- 2.176	+ 2.194	- 4.606
4/40	+ 7.767	+ 3.299	+ 23.219	+ 18.751
5/40	- 3.937	+ 9.801	+ 21.245	+ 34.983
6/40	- 16.095	- 0.172	+ 13.529	+ 29.445
7/40	+ 2.224	- 5.830	+ 24.392	+ 16.338
8/40	+ 2.471	- 5.648	+ 17.287	+ 9.168
9/40	+ 5.295	+ 6.930	+ 20.217	+ 21.852
10/40	- 24.967	- 19.263	+ 11.227	+ 17.031
11/40	- 16.243	- 14.958	+ 6.451	+ 7.736
12/40	+ 1.243	- 13.626	+ 15.195	+ 0.326
13/40	+ 10.430	+ 1.000	+ 15.078	+ 5.648
14/40	+ 7.631	+ 10.092	+ 5.823	+ 8.284
15/40	- 19.047	- 12.605	- 15.585	- 9.143
16/40	- 5.601	+ 9.361	- 41.223	- 26.261
17/40	- 30.766	- 0.318	- 32.396	- 1.978
18/40	- 11.826	- 22.008	- 0.608	- 10.790
19/40	+ 4.803	- 16.321	- 8.967	- 30.101

Pour  $x = + \frac{20}{40}$ , on retrouve les mêmes valeurs que pour  $x = -\frac{20}{40}$ . Ces quatre fonctions varient avec assez d'irrégularité pour qu'il ne soit pas possible d'interpoler graphiquement entre les points connus - d'où le tracé en dent de scie des courbes des figures 3, 4, 5 et 6, et cependant leur variation présente un certain degré de pseudo-continuité. Ce sont bien là des caractères de variable régionalisée de Wijsienne.

7/ Nous pouvons maintenant nous demander ce que donnerait une étude statistique ordinaire des teneurs représentées par les quatre fonctions ainsi construites. Ces quatre fonctions ayant, par construction, même valeur moyenne  $m = 0$  et même loi de dispersion intrinsèque (en particulier, même variance) il nous suffit d'étudier la loi de distribution statistique de la population obtenue en mélangeant les 160 valeurs calculées

pour les quatre fonctions. L'histogramme de cette population est présenté sur la table VI.

TABLE VI			
Histogramme des 160 teneurs de la table V			
classe	Fréquence	classe	Fréquence
- 45 à - 40	1	0 à 5	23
- 40 à - 35	0	5 à 10	20
- 35 à - 30	5	10 à 15	9
- 30 à - 25	2	15 à 20	9
- 25 à - 20	4	20 à 25	12
- 20 à - 15	16	25 à 30	3
- 15 à - 10	12	30 à 35	2
- 10 à - 5	17	35 à 40	1
- 5 à 0	24		

De l'histogramme de la table VI, nous déduisons les valeurs numériques suivantes de la moyenne  $m$  et de la variance  $\sigma^2$

$$( m = - 022$$

$$\left. \begin{array}{l} \\ \end{array} \right\} \sigma^2 = 225$$

La moyenne  $m = - 022$  n'est pratiquement pas différente de 0. D'après la manière dont les fonctions  $f(x)$  ont été construites, elle devrait être rigoureusement nulle.

Cette valeur légèrement négative peut être imputée au fait que l'on a regroupé les 160 teneurs en classes de taille 5, sur l'histogramme de la table VI. Ce groupage doit entraîner aussi une erreur sur la variance. Un calcul plus précis - utilisant des classes de taille 1 - donne en effet la valeur :

$$\sigma^2 = 212$$

Nous allons comparer cette valeur numérique  $\sigma^2 = 212$  avec la variance théorique  $\sigma_t^2$  que l'on observerait, si l'on connaissait les valeurs des  $f(x)$  en tous les points du segment 0 - 1, et non plus seulement en quarante points équidistants.  $\sigma^2 = 212$  est la variance des teneurs des 40 volées en quoi on a découpé la galerie ;  $\sigma_t^2$  sera la variance du nombre infini de volées de même taille existant virtuellement dans la

galerie, et se chevauchant les uns les autres.

D'après les formules (8) et (9), cette variance théorique a pour valeur :

$$\sigma_t^2 = \frac{10^4}{2} \sum C_p^2 + C_p^2 = \sum \frac{10^4}{p^3} \sin^2 \frac{\pi p}{40}$$

de l'égalité

$$\sum \frac{\sin^2 \frac{\pi p}{40}}{p^3} = \frac{3}{2} - \log_e 2 + \log_e 1 + 0 \quad (1^2)$$

On tire, en négligeant les termes en  $1^2$  et d'ordre supérieur, et en faisant  $l = \frac{1}{40}$  :

$$(10) \quad \sigma_t^2 = \frac{\pi^2 10^4}{1600} \left( \frac{3}{2} - \log_e 2 + \log 40 \right) = 206$$

Cette valeur théorique  $\sigma_t^2 = 206$  est en excellent accord avec la valeur observée, 212.

On notera sur la formule (10), que la dispersion absolue des modèles  $f(x)$  est :

$$(11) \quad 3 \sigma = \frac{\pi^2 10^4}{1600} = 61,5$$

Il reste à examiner si l'histogramme de la table VI peut être représenté par une loi de distribution statistique de forme mathématique connue. La symétrie approximative de cet histogramme nous incite à essayer la loi normale, ou loi de Gauss. Nous avons comparé, au moyen d'un test  $X^2$ , la distribution observée avec la loi normale de moyenne  $m = 0$  et de variance  $\sigma_t^2 = 206$ . L'accord est excellent<sup>1</sup>.

---

1 - On a trouvé  $X^2 = 8,86$  pour 9 degrés de liberté, soit  $P(X) = 0,45$ , ce qui est excellent. Cependant, les 160 teneurs - de par leur procédé de construction - ne sont pas indépendantes les uns des autres, ce qui enlève toute valeur quantitative au test classique du  $X^2$ . Néanmoins, en tant que mesure de l'écart quadratique moyen entre la distribution observée et la loi normale, le  $X^2$  conserve une valeur qualitative ; il est bien évident que l'accord reste d'autant meilleur que le  $X^2$  est plus petit.

Bien qu'il ne nous soit pas permis de nous exprimer quantitativement, il nous paraît certain que l'accord est ici très bon.

Ainsi donc le procédé de construction utilisé - qui consiste à tirer au sort indépendamment les unes des autres les phases d'un développement de Fourier de type (5) - engendre une distribution statistique gaussienne. Ce point doit sans doute pouvoir se démontrer théoriquement. Nous avons donc réussi à construire un modèle de gisement à une dimension dont les teneurs obéissent à la fois à une loi de dispersion absolue de Wijsienne et à une loi de distribution statistique gaussienne. De ces deux caractéristiques, seule la première apparaît comme fondamentale. Si, au lieu de tirer les phases au sort indépendamment les unes des autres, nous nous les donnions à priori selon une certaine loi - on en a un exemple avec les fonctions  $C(x)$  et  $S(x)$  de la formule (6) - ou même si nous les tirions au sort selon une loi admettant des intercorrelations, la distribution ne serait sans doute plus normale, tandis que la loi de dispersion intrinsèque resterait de Wijsienne.

8/ Il est aisé, à partir des modèles précédents - normaux et de Wijsiens - d'obtenir des modèles dont la loi de distribution statistique soit lognormale, tandis que la loi de dispersion intrinsèque reste de Wijsienne. Il suffit d'admettre que les fonctions  $f(x)$  représentent, non plus les teneurs elles-mêmes, mais leurs logarithmes. Pour obtenir une échelle de variation rappelant celle d'un gisement à basse teneur, nous avons divisé les valeurs numériques de  $f(x)$  par 40 et posé :

$$(I2) \quad \log_{10} y = \frac{1}{40} f(x)$$

D'où quatre nouvelles fonctions  $y_1, y_2, y_3$  et  $y_4$ , dont les valeurs numériques sont données table VII, et dont la variation est représentée graphiquement sur les figures 7, 8, 9 et 10. L'allure de ces fonctions - à la fois erratiques et pseudo-continues, et présentant en outre cette dissymétrie entre les valeurs élevées et basses, si caractéristique des distributions lognormales - évoque parfaitement la succession des teneurs de 40 volées. Le long de la périphérie d'un panneau carré, dans un gisement à basse teneur.

Les modèles obtenus semblent donc satisfaisants. Ils correspondent, d'après (II) et (I2) à une dispersion absolue logarithmique (exprimée, suivant l'usage, en log népériens) égale à :

$$(I3) \quad 3K = \frac{5,3}{1600} \times 61,5 = 0,204 \quad \text{ou } K = \frac{6,8}{100}$$

Nous allons maintenant examiner ce qu'aurait donné l'étude

Figure 7

Courbe  $y=y_1(x)$  représentant la variation de la teneur  
log normale et de Vienne d'une volée de longueur  $\frac{2\pi}{40}$   
le long d'une circonférence de rayon 1

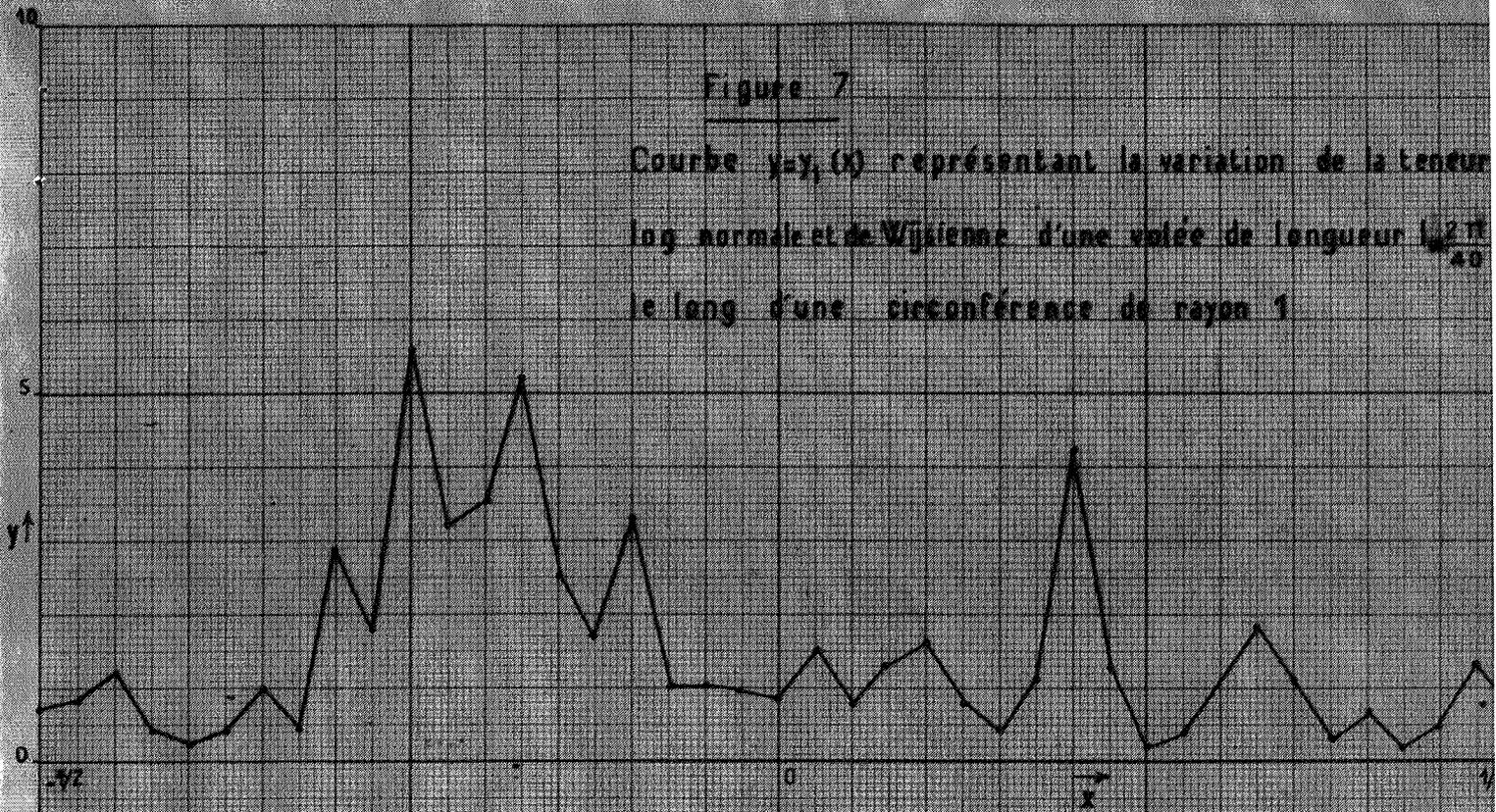


Figure 8

Courbe  $y=y_2(x)$

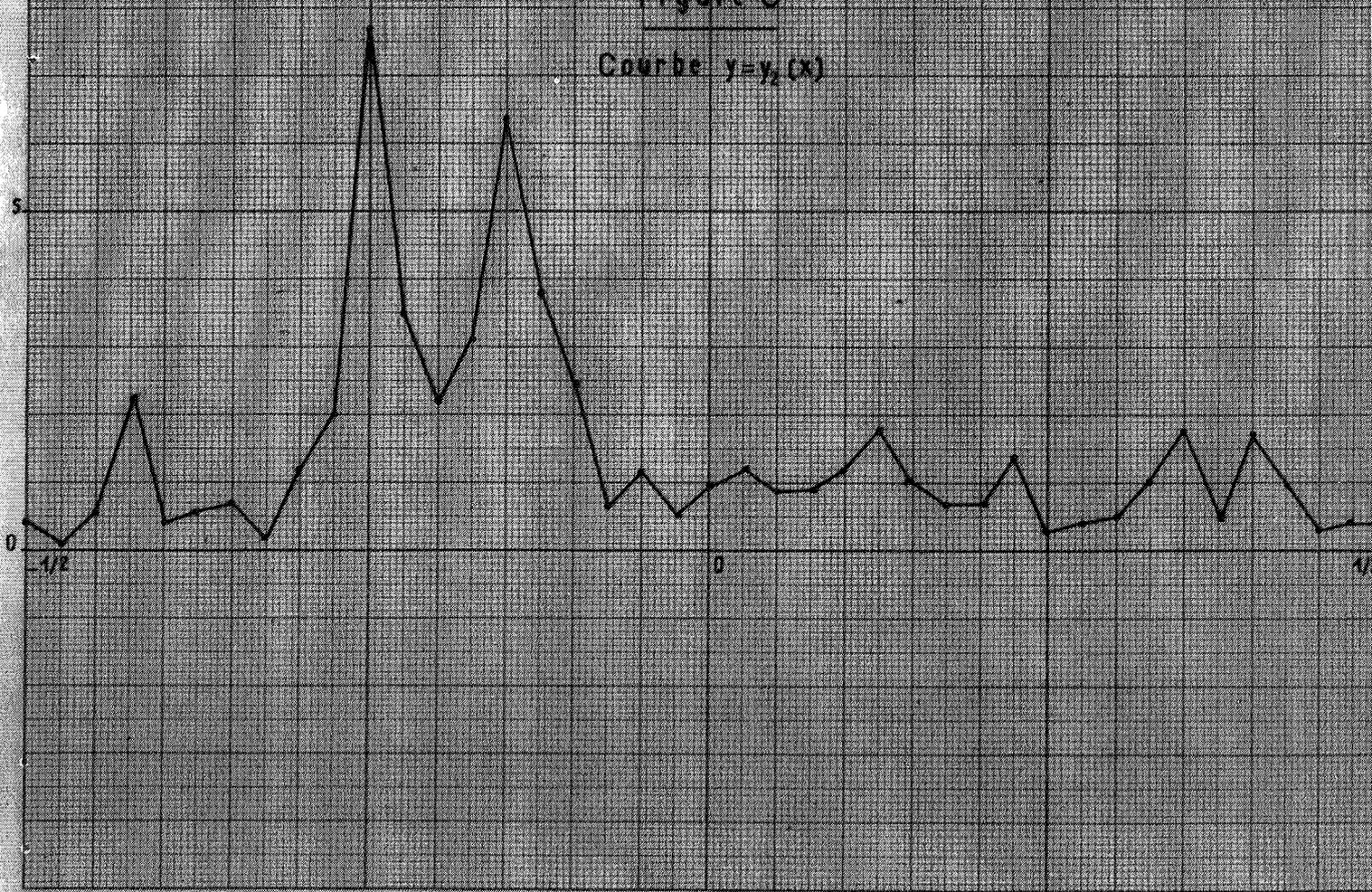


Figure 9

Courbe  $y = y_1(x)$

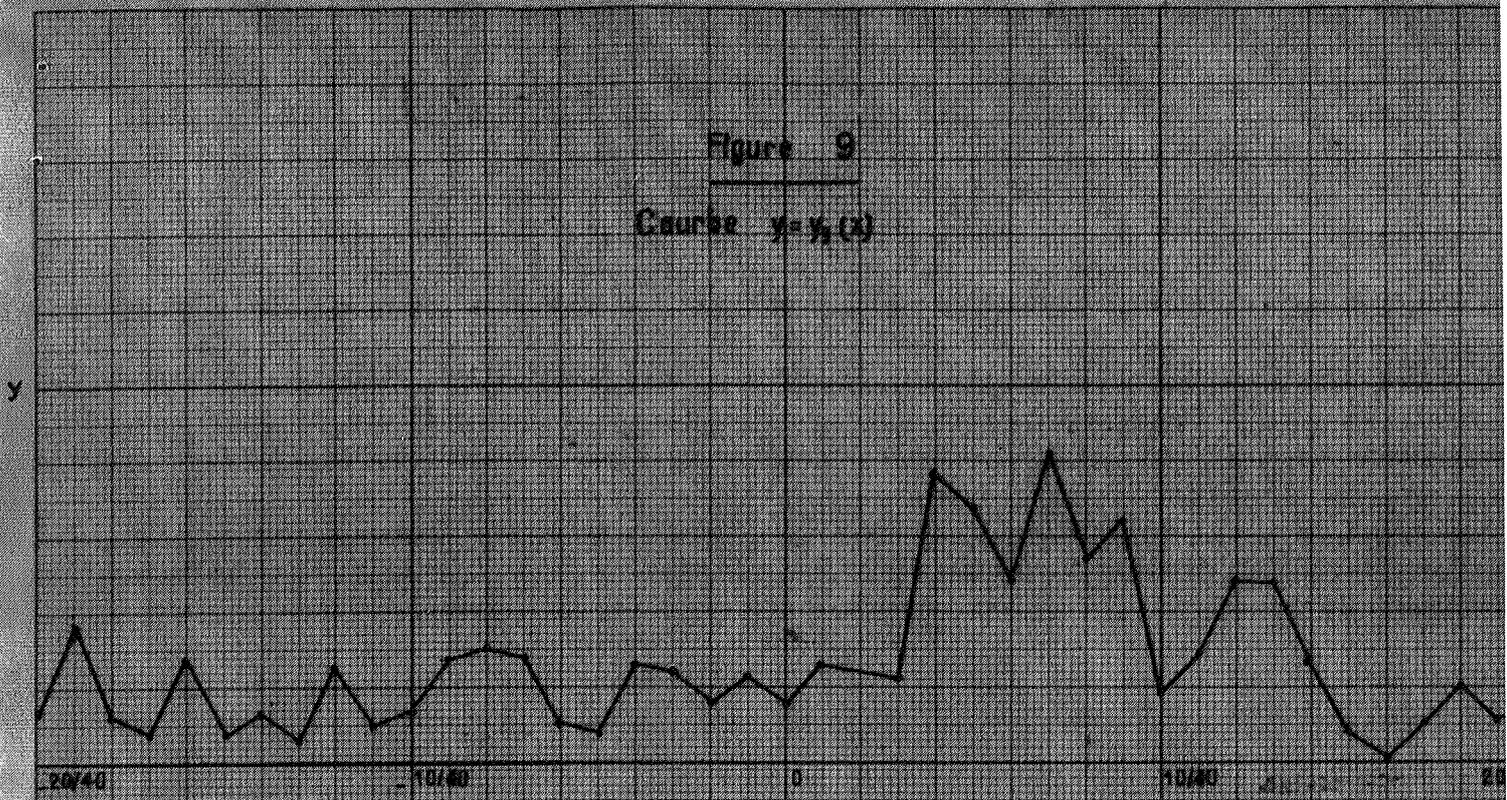
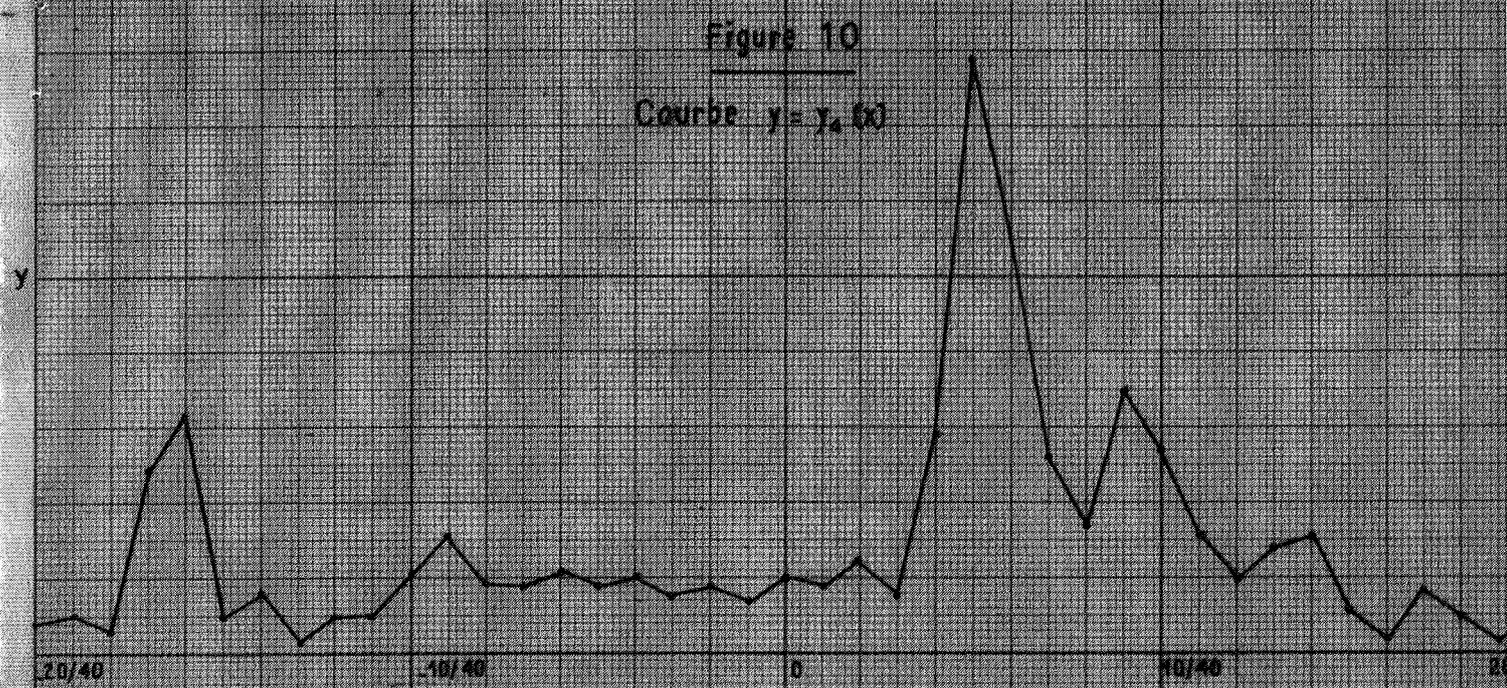


Figure 10

Courbe  $y = y_2(x)$



statistique, entreprise par les procédés habituels, des populations de teneurs figurant sur la table VII, et comparer les résultats ainsi obtenus avec les prévisions théoriques déduites du schéma de de Wijs en série de Fourier. Indiquons tout de suite que les écarts observés seront imputables à deux causes très différentes :

- tout d'abord, l'usage habituel de représenter les histogrammes au moyen de classes logarithmiques de raison  $10^{\frac{1}{10}}$  introduit quelques erreurs sur l'estimation des médianes et des variances.

TABLE VII

Valeurs numériques des quatre fonctions de Wijsiennes et lognormales  $y_1, y_2, y_3$  et  $y_4$

x	$y_1$	$y_2$	$y_3$	$y_4$	x	$y_1$	$y_2$	$y_3$	$y_4$
- 20/40	0.741	0.420	0.741	0.420	0	0.834	0.971	0.834	0.971
- 19/40	0.824	0.244	1.820	0.541	1/40	1.531	1.248	1.283	0.858
- 18/40	1.146	0.638	0.601	0.335	2/40	0.755	0.873	1.177	1.189
- 17/40	0.381	2.193	0.419	2.410	3/40	1.312	0.883	1.133	0.767
- 16/40	0.174	0.412	1.352	3.199	4/40	1.563	1.208	3.802	2.844
- 15/40	0.430	0.624	0.352	0.505	5/40	0.798	1.758	3.396	7.925
- 14/40	0.948	0.678	0.653	0.753	6/40	0.396	0.991	2.444	5.445
- 13/40	0.376	0.218	0.288	0.167	7/40	1.138	0.716	4.065	2.559
- 12/40	2.825	1.200	1.265	0.538	8/40	4.150	0.723	2.704	1.675
- 11/40	1.845	1.986	0.501	0.538	9/40	1.279	1.406	3.199	3.516
- 10/40	5.534	7.727	0.685	0.957	10/40	0.298	0.330	1.910	2.661
- 9/40	3.214	3.532	1.365	1.496	11/40	0.393	0.423	1.449	1.560
- 8/40	3.468	2.173	1.476	0.927	12/40	1.074	0.456	2.399	1.019
- 7/40	5.141	0.157	1.436	0.904	13/40	1.824	1.059	2.383	1.384
- 6/40	2.512	6.281	0.457	1.140	14/40	1.553	1.787	1.321	1.611
- 5/40	1.742	3.837	0.408	0.902	15/40	0.334	0.484	0.408	0.592
- 4/40	3.266	2.529	1.343	1.037	16/40	0.725	1.714	0.093	0.221
- 3/40	1.002	0.679	1.162	0.785	17/40	0.170	0.982	0.552	0.893
- 2/40	1.043	1.216	0.775	0.904	18/40	0.507	0.282	0.966	0.538
- 1/40	0.938	0.627	1.120	0.747	19/40	1.319	0.391	0.597	0.177

Ces erreurs, dont nous apprécierons l'ordre de grandeur proviennent simplement du procédé de calcul adopté. Elles n'ont pas de

signification théorique.

- En deuxième lieu, les populations observées représentent les teneurs de 40 volées obtenues en découpant la galerie en 40 tronçons successifs, et correspondent à un mode de découpage particulier. Les prévisions théoriques concernent au contraire la population infinie de toutes les volées existant virtuellement dans la galerie et se chevauchant. Il est très important, pour la pratique, de se faire une idée de l'ordre de grandeur des écarts imputables à l'adoption d'un mode de découpage discontinu. De tels écarts, observés à propos de gisements réels, pourront, en effet, être considérés comme non significatifs.

9/ La table VIII représente les histogrammes des quatre populations  $y_1, y_2, y_3, y_4$ , avec les classes logarithmiques habituelles. Les classes sont désignées par leur borne inférieure. La classe marquée 0.158, par exemple, contient toutes les données comprises entre 0.158 et 0.199. A partir d'un histogramme, on a calculé les paramètres lognormaux habituels - médiane ou moyenne géométrique  $\bar{y}$ , variance  $\sigma^2$  (qui est, suivant l'usage, la variance des logarithmes népériens des teneurs) des populations  $y_1, y_2, y_3, y_4$  séparément, puis de la population totale obtenue par le mélange des quatre précédentes. Les valeurs numériques de ces paramètres sont indiquées sur la table IX. Pour la population totale, on a indiqué d'abord les paramètres déduits de l'histogramme et des classes logarithmiques, puis les valeurs plus exactes déduites de la moyenne et de la variance des fonctions  $f_1(x), f_2(x), f_3(x)$  et  $f_4(x)$ . Deux conclusions se dégagent :

- Tout d'abord, les erreurs, introduites par le procédé de calcul à partir des classes logarithmiques, sur les variances comme sur les médianes, apparaissent comme négligeables, comme on le voit en comparant les valeurs exactes et déduites de l'histogramme des paramètres de la population totale.

- Seul donc le mode de découpage de la galerie en volée discontinues est responsable des écarts entre les valeurs trouvées et les précisions théoriques. Si les fluctuations des quatre médianes des  $y$  restent d'amplitude faible (6% au maximum), les variances présentent des écarts plus importants (18 % au maximum). Par effet de compensation, la variance de la population totale (0.700) est en bon accord avec la valeur théorique (0.683).

TABLE VIII

Histogramme des 4 populations  $y_1, y_2, y_3, y_4$ , et de leur total.

Classe	Fréquence $y_1$	Fréquence $y_2$	Fréquence $y_3$	Fréquence $y_4$	TOTAL
0.0794	0	0	I	0	I
0.100	0	0	o	0	0
0.126	0	0	o	0	0
0.158	2	0	0	2	4
0.199	0	2	0	I	3
0.25I	I	I	I	0	3
0.316	5	2	I	I	9
0.398	I	5	4	I	II
0.50I	I	2	4	6	I3
0.63I	3	5	4	4	I6
0.794	5	5	2	8	20
I.00	5	5	4	4	I8
I.26	6	I	9	3	I9
I.58	3	4	2	2	II
I.99	0	2	3	I	6
2.5I	2	I	I	3	7
3.16	3	3	3	2	II
3.98	I	0	I	0	2
5.0I	2	I	0	I	4
6.3I	0	I	0	I	2
TOTAL	40	40	40	40	I60

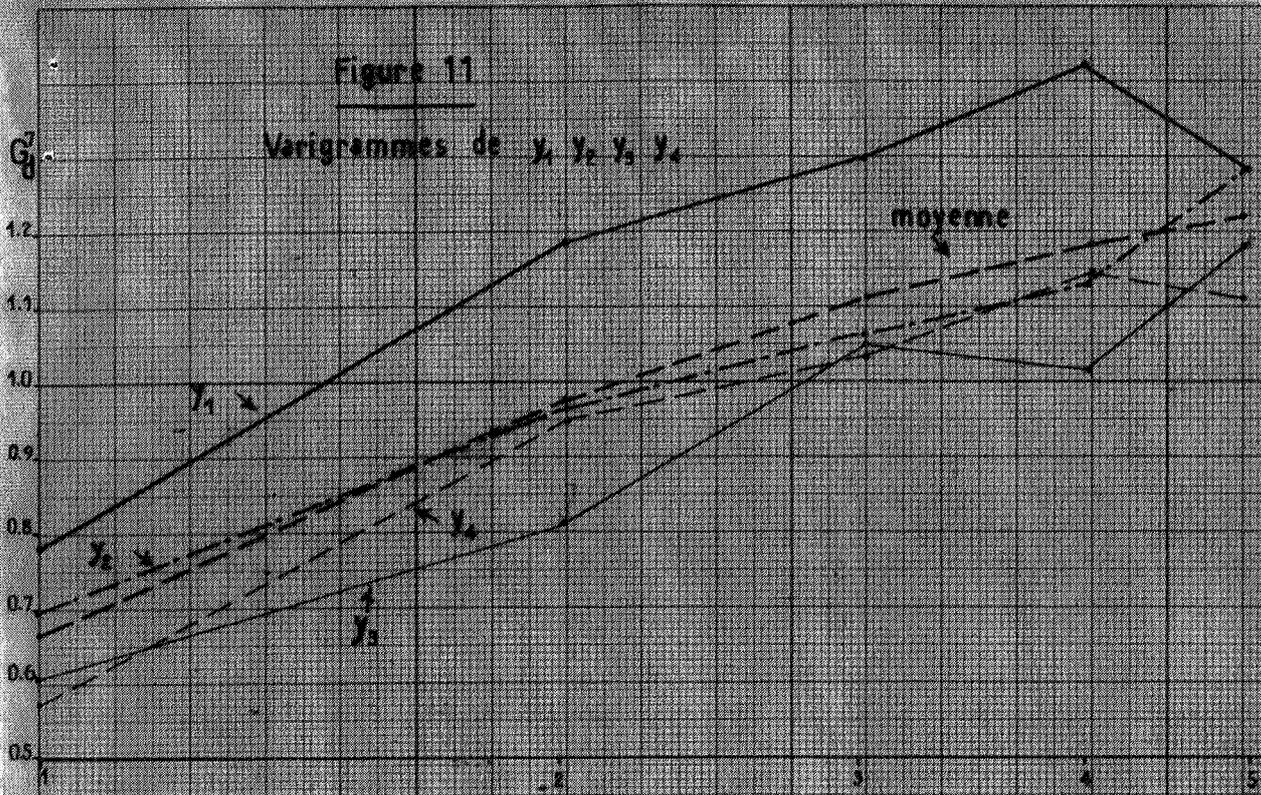
IO/ De la même façon, il convient de comparer les variogrammes des quatre populations disponibles, calculés par les procédés habituels, avec le variogramme théorique déduit de  $G(h)$  du schéma de de Wijs en série de Fourier.

La table X présente les valeurs numériques des 5 premiers points des variogrammes des 4 populations et de leurs moyennes.

Les quantités tabulées sont les variances  $\sigma_d^2$  du logarithme népérien de la différence des teneurs des volées situées à la distance angulaire  $2\sqrt{d}$  l'une de l'autre. La figure 11 donne la représen-

Figure 11

Varigrammes de  $y_1, y_2, y_3, y_4$



tation graphique de ces variogrammes. L'amplitude des fluctuations est importante. Or, pour chacune des populations, si l'on connaissait, non pas seulement 40 valeurs de y en 40 points discontinus, mais la courbe totale représentant la variation de y en fonction de x, les quatre variogrammes seraient, par construction, identiques entre eux et avec covariogramme théorique. Les fluctuations observées sont entièrement imputables au mode de découpage discontinu. Si l'on en rencontre de semblables lors de l'étude de gisements réels, on est en droit de les considérer comme non significatives. Dans les études pratiques, il sera donc souvent difficile de mettre en évidence une véritable régionalisation de la dispersion absolue, celle-ci risquant d'être marquée par les fluctuations aléatoires du type observé ici.

L'équation théorique du variogramme, de son côté, est :

$$\frac{\sigma^2}{d} = 2 \left( \sigma^2 - 3 \times \int \frac{\sin^2 \angle p_1}{r^3 p^2} \cos 2 \angle pd \right)$$

A partir d'une distance angulaire d supérieure à  $2\hat{r}$ , cette expression ne diffère pas beaucoup de :

$$(I4) \quad \frac{\sigma^2}{d} = 2 \left( \sigma^2 - 3 \times \log_0 2 \cdot (\sin \angle d) \right)$$

Cette formule (I4) a servi à calculer les valeurs théoriques du variogramme qui figurent en dernière colonne de la table X : elle sont en bon accord avec les valeurs moyennes des 4 variogrammes à partir de  $d = 3$ , c'est-à-dire à partir du moment où la formule simplifiée (I4) devient vraiment valable.

#### 10/ Influence de la longueur de la volée

A partir des 4 séries de teneurs représentées sur la table VIII, et supposées représenter l'évolution de la minéralisation dans un gisement réel, il nous a paru intéressant d'examiner ce qui se passerait dans un tel gisement pour des volées deux fois plus longues ( $l = \frac{l}{20}$ ) et quatre fois plus longues ( $l = \frac{l}{10}$ ). Compte tenu de la périodicité, et en prenant des volées chevauchantes, on obtien 160 valeurs pour chacune de ces deux populations. Les deux histogrammes sont donnés table XI.

Ces histogrammes permettent d'examiner s'il y a permanence de la lognormalité, et comment évoluent les variances lorsque la taille l de la volée augmente.

TABLE IX		
Paramètres lognormaux des $y_1, y_2, y_3, y_4$ .		
	Médiane	Variance
Population des $y_1$	1.06	0.800
" $y_2$	0.99	0.725
" $y_3$	0.99	0.622
" $y_4$	0.99	0.705
" Totale (histogramme)	1.01	0.700
Population totale (valeurs exactes)	0.99	0.702
Population théorique	1.00	0.683

TABLE X							
Variogrammes de $y_1, y_2, y_3, y_4$							
d	$y_1$	$y_2$	$y_3$	$y_4$	Variogramme		
Distance unité : $\frac{21}{40}$	$y_1$	$y_2$	$y_3$	$y_4$	moyen	théorique	
1	0.777	0.694	0.602	0.568	0.660	(0.609)	
2	1.178	0.962	0.809	0.942	0.972	(0.892)	
3	1.288	1.057	1.049	1.030	1.106	1.054	
4	1.418	1.130	1.023	1.135	1.176	1.169	
5	1.280	1.284	1.192	1.110	1.216	1.256	

TABLE XI		
Histogrammes des volées de longueurs $l = \frac{I}{20}$ et $l = \frac{I}{10}$		
Classe	Fréquence $l = \frac{I}{20}$	Fréquence $l = \frac{I}{10}$
0.25I	5	0
0.3I6	7	2
0.398	II	7
0.50I	8	II
0.63I	20	I7
0.794	28	37
I.	24	3I
I.26	I4	IO
I.58	9	I2
I.99	7	6
2.5I	I2	8
3.I6	6	I5
3.98	5	4
5.0I	3	0
6.3I	I	0

A/ Un simple examen des histogrammes montre qu'apparemment il n'y a pas permanence de la lognormalité. Le test  $X^2$  donne des résultats désastreux : il est vrai qu'ici, plus encore que pour la population primitive des  $y$ , le test  $X^2$  n'a plus aucune signification quantitative. Les volées chevauchantes ne peuvent absolument pas être regardées comme indépendantes. Mais, d'un simple point de vue qualitatif, le résultat demeure. Très caractéristique est, d'ailleurs, l'allure bimodale des deux histogrammes. A côté du mode principal situé dans la classe 0.794-1 apparaît un mode secondaire. Encore peu marqué et situé dans la classe 25I-3.I6 pour les volées  $l = \frac{I}{20}$ , ce mode secondaire prend une ampleur très forte pour les volées  $\frac{I}{10}$ , et se situe dans la classe 398-5.0I. Si ce dernier histogramme était d'origine expérimentale, et si son mode de construction ne nous garantissait pas à priori l'homogénéité de la population, on conclurait à coup sûr à une hétérogénéité, et on chercherait à séparer deux populations l'une riche et l'autre pauvre : mais ici nous sommes certains, à priori, que la population est unique. Cet exemple doit nous inciter à

une grande prudence lorsqu'il s'agira de diagnostiquer une hétérogénéité à partir d'une distribution expérimentale. On doit penser que toute séparation de population non justifiée par un critère géologique solide est très hasardeuse. Enfin, la permanence de la lognormalité, qui n'est pas respectée ici, devient un phénomène un peu énigmatique.

B/ Examinons maintenant l'évolution de la variance. Le tableau ci-après donne les valeurs numériques de la médiane et de la variance des volées de longueur  $\frac{l}{40}$ ,  $\frac{l}{20}$  et  $\frac{l}{10}$

	$l = \frac{l}{40}$	$l = \frac{l}{20}$	$l = \frac{l}{10}$
Médiane $\checkmark$	0.99	1.10	1.19
Variance $\sigma^2$	0.702	0.526	0.375

Pour  $l = \frac{l}{40}$ , on a vu que la valeur de la variance est en bon accord avec la formule théorique. Le passage à des volées de longueur double doit faire décroître la variance de  $3 \times \log_e 2 = 0.204 \times 0.693 = 0.141$ . On observe en fait :  $0.702 - 0.526 = 0.176$  et  $0.526 - 0.375 = 0.151$  : le désaccord ne porte que sur la deuxième décimale.

Il est plus suggestif de faire la comparaison sous la forme suivante :

Population	Variance observée	Variance théorique
$l = \frac{l}{40}$	0.702	0.683
$l = \frac{l}{20}$	0.526	0.542
$l = \frac{l}{10}$	0.375	0.401

On peut conclure que la loi d'évolution de la variance en fonction de la taille de l'échantillon est à peu près conforme aux prévisions théoriques.