



HAL
open science

Détermination de périodicités dans des phénomènes aléatoires historiques stationnaires

D. Taupin, G. Polian

► **To cite this version:**

D. Taupin, G. Polian. Détermination de périodicités dans des phénomènes aléatoires historiques stationnaires. *Revue de Physique Appliquée*, 1973, 8 (4), pp.351-359. 10.1051/rphysap:0197300804035100 . jpa-00243694

HAL Id: jpa-00243694

<https://hal.science/jpa-00243694v1>

Submitted on 4 Feb 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Classification
Physics Abstracts :
 02.40

DÉTERMINATION DE PÉRIODICITÉS DANS DES PHÉNOMÈNES ALÉATOIRES HISTORIQUES STATIONNAIRES

D. TAUPIN

Laboratoire de Physique des Solides Associé au CNRS
 Bâtiment 510, Université de Paris-Sud, 91405 Orsay, France

G. POLIAN

Centre d'Etude des Faibles Radioactivités
 CNRS, 91190 Gif-sur-Yvette, France

et

Terres Australes et Antarctiques Françaises

(Reçu le 24 janvier 1973, révisé le 2 avril 1973)

Résumé. — Nous nous intéressons ici à la détermination pratique de périodicités dans des phénomènes « historiques », c'est-à-dire aléatoires, non reproductibles, dont les mesures sont limitées dans le temps, et qui en outre comportent des manques dus à diverses causes fortuites. Nous présentons ici, avec critiques, plusieurs méthodes d'étude : périodogramme, spectre de puissance, fonction d'autocorrélation, et enfin une méthode mixte (autocorrélation multipliée par spectre de puissance) qui nous a donné de bons résultats.

Abstract. — We are concerned here with the practical study of periodicities in phenomena which can be qualified as « historical », i. e. stochastic, non-reproducible, whose measurements extend over a limited time, and contain vacancies due to incidental conditions. Several methods are discussed here : periodogram, power spectrum, autocorrelation function, and a mixed method (autocorrelation multiplied by power spectrum) which yielded some good results.

Cette étude a été entreprise à l'occasion d'une recherche de périodicité de phénomènes atmosphériques (concentration en éléments radioactifs, en ozone, mesures météorologiques, etc... [1]).

Les phénomènes étudiés présentent un certain nombre de propriétés communes :

— ils sont entachés d'un bruit important (bruit « statistique » de comptage, fluctuations à longue période, accidents extérieurs affectant plusieurs mesures consécutives, etc...);

— ils sont permanents pendant un temps beaucoup plus grand que la durée de leur étude, sans pour autant être stables (ceci par opposition aux phénomènes éphémères qui sont nuls, ou tendent vers zéro en dehors de l'intervalle étudié);

— ils sont « historiques » et ne peuvent donc être reproduits ni déclenchés à volonté,

— les mesures, effectuées *périodiquement* pendant des temps longs (des années) et dans des lieux parfois inhospitaliers, contiennent des trous de tailles diverses dus à toutes sortes d'incidents (pannes d'appareillage, négligences, incendies, cataclysmes, arrêts périodiques pour maintenance préventive, etc...).

Il s'agit alors, compte tenu des informations recueillies et de celles connues *a priori* (permanence des phénomènes) de déterminer les composantes périodiques (ou pseudo-périodiques en cas de fluctuations d'amplitude) en minimisant autant que possible les conséquences des trous et de la limitation de l'intervalle étudié. Ceci revient à étudier le spectre de fréquences, en décidant, autant que possible, de ce qui est significatif, et de ce qui n'est dû qu'au bruit.

La complication vient de la diversité du bruit, non seulement en ce qui concerne son origine, mais surtout sa nature mathématique :

a) bruit additif d'origine statistique ou accidentelle (le moins gênant en fait) ;

b) bruit multiplicatif :

— trous de mesure, aisément détectés,

— fluctuations d'amplitude qui conduisent à un comportement pathologique des transformées de Fourier,

— fluctuations de phases ou de fréquence, ce qui interdit des mesures étendues sur de très longues durées.

Nous présentons ici, avec critiques, plusieurs méthodes d'étude traditionnelles : le périodogramme, le spectre de puissance, la fonction d'autocorrélation ; ensuite nous présentons une autre méthode (mixte autocorrélation/spectre de puissance) qui nous a donné de bons résultats.

Nous n'avons pas abordé la question sous l'angle théorique de la recherche d'une méthode optimale dans un problème qui finalement se ramène à une analyse spectrale sur données échantillonnées et bruitées. Ce problème n'est pas sans intérêt, mais nous nous adresserons plutôt aux praticiens familiers de méthodes que nous proposons d'améliorer.

1. Terminologie et notations. — Nous employons le terme « périodicité » pour désigner une composante périodique, non sinusoïdale, dont l'harmonique fondamental n'est pas absent.

Ceci nous paraît assez bien coïncider avec la notion intuitive de « phénomène périodique » dans laquelle les harmoniques ne doivent pas être séparés. Il est cependant évident que des difficultés apparaissent si deux ou plusieurs phénomènes distincts (quant à leur origine) ont un rapport de fréquence égal à 2 ou un autre petit entier ; mais nous pouvons considérer que ce dernier problème est le même que celui de décider si un phénomène périodique est un phénomène unique, ou la superposition de deux autres rigoureusement synchrones, ce qui, en l'absence d'autres informations, est insoluble.

D'autre part, nous notons par \bar{y} et $\overline{y^2}$ les valeurs moyenne et quadratique moyenne d'une série de N mesures $\{y_i\}$:

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i ;$$

$$\overline{y^2} = \frac{1}{N} \sum_{i=1}^N y_i^2 .$$

2. Revue de quelques méthodes classiques. —

2.1 PÉRIODOGRAMME. — La référence que nous citons ici [2] est pratiquement introuvable ; en fait, la méthode du « périodogramme » fait partie de la « tradition orale » des géophysiciens et cosmiciens et c'est ainsi que nous en avons eu connaissance. La présente discussion est donc basée sur ce qui se fait usuellement.

D'autre part, le terme « périodogramme » est malheureusement ambigu car il est aussi utilisé [3] pour désigner ce que nous appelons ici « spectre de puissance », lorsque celui-ci est tracé ou calculé en fonction de la période (ou l'inverse de la fréquence).

2.1.1 Description générale. — Le principe consiste à étudier successivement toutes les périodicités p possibles qui soient un multiple de l'intervalle de temps h qui s'écoule (sauf incident) entre deux mesures consécutives du phénomène :

$$p = nh ; \quad (1)$$

pour cela, on « replie » sur lui-même l'intervalle de temps étudié et on calcule la moyenne de toutes les mesures y_k distantes de nh . On obtient une série de n moyennes $\{m_i\}$ qui représenteront bien l'aspect d'une « période » (si le phénomène comporte une composante de période nh) et qui seront sensiblement égales entre elles si la composante de période p est inexistante ou négligeable.

$$m_i = \frac{1}{N_i} \sum_j y_{i+nj}, \quad i = 1, n-1 ; \quad (2)$$

où la somme est étendue à tous les y_{i+nj} effectivement mesurés, pour une valeur donnée de i et où N_i est leur nombre, en général de l'ordre de N/n .

En fait, sauf dans le cas où p est connu à l'avance, il est fastidieux de faire un tracé de la variation des $\{m_i\}$ pour chaque valeur de p ; aussi l'amplitude de la composante de période p est elle usuellement repérée en calculant, pour chaque valeur de n , la dispersion des $\{m_i\}$, en remarquant que cette dispersion croît avec l'« intensité de la périodicité p » (nous employons à dessein cette expression ambiguë car elle recouvre le plus grave défaut de principe de cette méthode, ce que nous préférons discuter *in fine*) :

$$\sigma_n^2 = \frac{1}{n} \sum_{i=1}^n (m_i - \bar{y})^2 . \quad (3)$$

2.1.2 Améliorations. — L'expression ci-dessus contient malheureusement une simplification abusive : au lieu de \bar{y} , il faudrait mettre la moyenne des $\{m_i\}$; cette simplification était peut-être justifiable lorsque les calculs étaient faits à la main, mais certainement pas lorsqu'on dispose d'ordinateurs. Un deuxième défaut vient du fait que la dispersion statistique des $\{m_i\}$ est *a priori* d'autant plus faible que ces $\{m_i\}$ sont obtenus en moyennant un plus grand nombre de mesures ; or le nombre de mesures intervenant dans chaque m_i est de l'ordre de N/n , où N est le nombre total (constant) de mesures. En particulier, dans le cas où ces mesures ne seraient qu'un bruit incohérent, on trouverait :

$$\sigma_{n,\text{incoh.}}^2 \approx \frac{n}{N} (\overline{y^2} - \bar{y}^2) ; \quad (4)$$

c'est pourquoi nous avons préféré corriger l'expression (3) et la normaliser en tenant compte de (4) de façon à repérer l'intensité de la composante de période $p = nh$ par l'expression :

$$q_n = \frac{\sigma_n^2}{\sigma_{n,\text{incoh.}}^2} = \frac{N}{n} \cdot \frac{\frac{1}{n} \sum_{i=1}^n m_i^2 - \left(\frac{1}{n} \sum_{i=1}^n m_i\right)^2}{\overline{y^2} - \bar{y}^2} . \quad (5)$$

Il ne suffit cependant pas de calculer une fonction exprimant la présence de périodicités, il faut aussi faire un calcul d'erreur pour déterminer ce qui est significatif et ce qui ne l'est pas. Ce calcul d'erreur sera fait de la manière la plus pessimiste possible. Pour

cela nous ne considérerons pas seulement l'erreur de mesure due à l'appareillage, mais nous supposerons que la mesure est affectée d'un bruit aléatoire additif incohérent de dispersion égale à la dispersion générale des mesures effectuées.

L'erreur sur q_n est alors donnée par l'expression classique :

$$(\Delta q_n)^2 = \sum_{i=1}^n \left(\frac{\partial q_n}{\partial m_i} \right)^2 (\Delta m_i)^2 \quad (6)$$

soit :

$$(\Delta q_n)^2 = \frac{4N}{n^2} \cdot \frac{(\Delta m_i)^2}{y^2 - \bar{y}^2} \quad (7)$$

Or, l'incertitude sur les $\{m_i\}$ peut être facilement évaluée :

$$(\Delta m_i)^2 = \frac{n}{N} (\Delta y)^2 = \frac{n}{N} (y^2 - \bar{y}^2) \quad (8)$$

d'où :

$$\Delta q_n = 2 \sqrt{\frac{q_n}{n}} \quad (9)$$

Avec cette méthode l'élimination des effets des «trous» dans les mesures et de la limitation de l'intervalle de temps étudié est facile : il suffit, dans l'éq. (2), de ne pas compter les mesures absentes et de diminuer en conséquence N_i ; ceci revient à admettre que la mesure absente est égale à la moyenne des autres mesures situées au même endroit dans la période de longueur p présupposée. On remarquera que la valeur admise par défaut pour les mesures manquantes n'est donc pas nulle comme dans les transformations de Fourier effectuées sans précaution, et qu'elle est variable avec n ; ceci présente à notre avis l'avantage de ne risquer de favoriser aucune périodicité par rapport aux autres.

Un ennui plus gênant se produit lorsque le nombre de trous est tel que dans l'éq. (2) l'un des N_i devient nul (cas fréquent surtout lorsque n devient grand); nous y avons remédié d'une manière assez grossière mais simple en déterminant une valeur approchée du m_i incalculable, par interpolation linéaire entre ses voisins immédiats. On aurait pu aussi ignorer le m_i incalculable et prendre les formules précédentes avec n plus petit, mais cela revient encore à attribuer au point manquant une valeur égale à la moyenne des autres, ce qui est finalement moins plausible que de prendre la moyenne des voisins. Heureusement ce cas est assez rare et marginal, et l'essentiel est alors surtout d'éviter des divisions par zéro, mal tolérées par les ordinateurs.

2.1.3 Critique. — Le périodogramme est une méthode d'analyse facile à mettre en œuvre, même sans ordinateur, et jouit de ce fait d'une certaine faveur auprès des géophysiciens; en outre nos améliorations permettent de réduire les conséquences de mesures incomplètes. Malheureusement elle présente de très graves défauts de principe qui nous ont conduits

à l'abandonner, du moins comme moyen d'analyse qualitative.

En effet, elle détecte, outre la période fondamentale, les multiples de cette période (avec des intensités apparentes ne décroissant que lentement); un exemple saisissant en est la figure 1, qui représente le périodogramme d'une sinusoïde de période 30 jours, mesurée pendant 1 an, dans laquelle on trouve des périodicités parasites de 60 et 90 jours. Cet inconvénient peut aussi se prévoir de manière théorique en étudiant la signification de l'éq. (3) en termes d'analyse de Fourier. Posons :

$$y_i = \sum_q a_q \exp(+2\pi i q) \quad (10)$$

où q prend un certain nombre de valeurs de module compris entre $1/N$ et $1/2$ (seul domaine significatif), ainsi que la valeur 0. Alors :

$$m_i = \frac{1}{N_i} \sum_k \sum_q a_q \exp(j2\pi q(i+nk)); \quad (11)$$

remarquons que seuls les termes pour lesquels qn est à peu près entier (à $1/N_i$ près) donneront une contribution non négligeable, de sorte que :

$$m_i \simeq \sum_r a_{r/n} \exp\left(j2\pi \frac{ri}{n}\right), \quad r \sim \text{entier}; \quad (13)$$

en remarquant que \bar{y} est à peu près égal au terme a_0 , l'éq. (3) devient :

$$\sigma_n^2 \simeq \sum_{i=1}^n \sum_{r \neq 0} \sum_{s \neq 0} a_{r/n} a_{s/n} \exp\left(j2\pi \frac{(r+s)i}{n}\right) \quad (14)$$

où seuls les termes pour lesquels $r+s=0$ donnent une contribution importante, de sorte que :

$$\sigma_n^2 \simeq \sum_r |a_{r/n}|^2 \quad (15)$$

σ_n^2 est donc la somme des intensités des divers harmoniques (fondamental compris) de la composante de fréquence $1/nh$; par suite de la nature de l'échantillonnage, cette somme est limitée à l'harmonique $n/2$ (période 2 h). Du fait de la limitation de l'intervalle de temps étudié, s'y ajoutent des contributions de périodicités voisines (phénomène classique d'élargissement spectral lorsque le temps de mesure n'est pas très grand par rapport à la période).

La périodicité $p = nh$, si elle existe, sera donc aussi détectée, en tant qu'harmonique lors de la recherche d'une éventuelle périodicité $p' = 2nh$, et ainsi de suite.

D'autre part, il est également évident que les harmoniques de p seront eux aussi détectés séparément, ce qui, combiné avec la détection des périodes multiples donne rapidement un diagramme inextricable; on trouvera figure 2 le périodogramme d'une fonction créneau dissymétrique où l'on trouve les périodes $p, 2p, 3p, p/2, 3/2p, 5/2p, p/3, 2/3p, 4/3p$, etc...

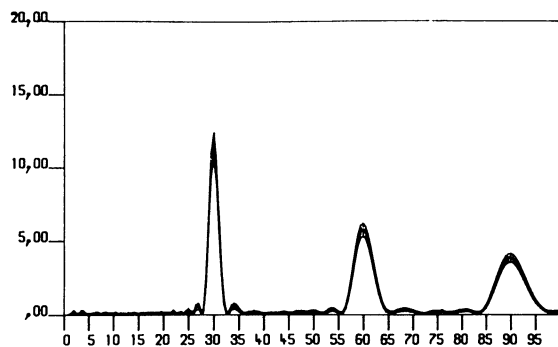


FIG. 1. — Périodogramme d'une sinusoïde (période 30 jours, durée 1 an).

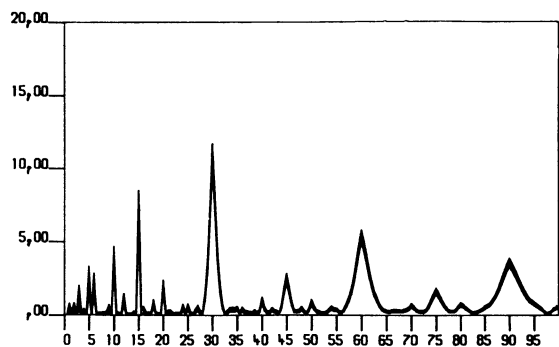


FIG. 2. — Périodogramme d'une fonction créneau périodique (période 30 jours, durée 1 an).

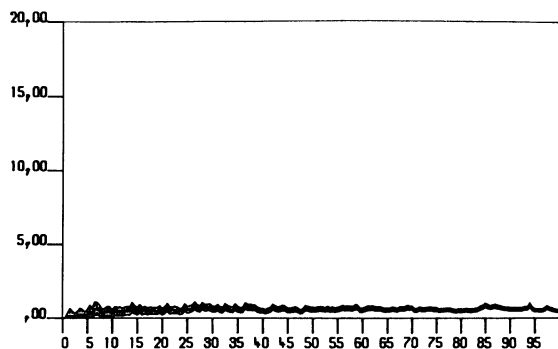


FIG. 3. — Périodogramme de mesures réelles (Radon atmosphérique, Ile Crozet, durée 1 an ; périodicité attendue : environ 28 jours).

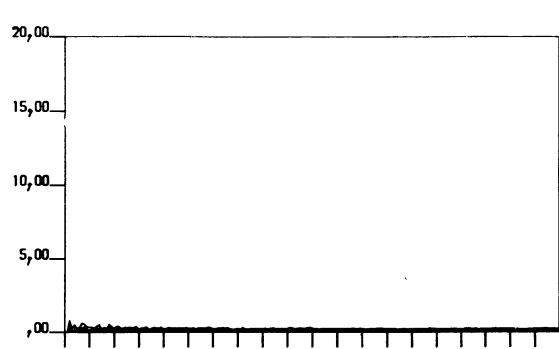


FIG. 4. — Périodogramme d'un générateur aléatoire de même répartition que la figure 3 (durée 1 an).

Enfin, l'aspect du périodogramme dépend fortement de la fréquence des mesures ou de leur regroupement.

Cependant, malgré ses graves inconvénients, la méthode du périodogramme reste utile une fois que la périodicité a été déterminée (par d'autres méthodes, ou connue d'avance) ; en effet le simple tracé des $\{m_i\}$ fournit directement l'allure de la composante de périodicité donnée dans le phénomène étudié.

On trouvera figure 3 le périodogramme de mesures réelles ⁽¹⁾ et figure 4 celui d'un générateur aléatoire incohérent de même répartition (on remplace chaque mesure existante par une quelconque, tirée au hasard, avec probabilité uniforme, dans l'ensemble des N mesures réelles).

2.2 TRANSFORMÉE DE FOURIER. SPECTRE DE PUISSANCE.

2.2.1 *Description générale.* — Les transformées de Fourier et les spectres de puissance (le spectre de puissance est le carré du module de la transformée de Fourier) sont généralement calculés et représentés comme des fonctions de la fréquence ; lorsqu'on les calcule par points, ceux-ci correspondent à des valeurs équidistantes de la fréquence ; ceci présente une certaine importance lorsque l'on veut effectuer ensuite la transformation inverse, c'est-à-dire une intégration, auquel cas un problème de métrique pourrait se poser.

Mais lorsque l'on désire seulement étudier cette transformée de Fourier il est intéressant et pratique de la calculer pour des valeurs de la période p , variant en progression arithmétique et multiples de l'intervalle élémentaire h entre les mesures de y .

Il est alors facile de voir que la transformée de Fourier $Y(p)$ peut être calculée à partir de l'éq. (2) :

$$Y(p = nh) = h \sum_{i=1}^n m_i \exp\left(-j 2 \pi \frac{i}{n}\right). \quad (16)$$

2.2.2 *Améliorations.* — L'expression ci-dessus a l'inconvénient de donner des résultats dépendants de l'ordre de grandeur des mesures ce qui peut être gênant lorsqu'on veut tracer ou comparer des phénomènes de nature différente ; c'est pourquoi nous avons préféré calculer la fonction sans dimension :

$$C(p) = \frac{\sqrt{2}}{n} \cdot \frac{\sum_{i=1}^n m_i \exp\left(-j 2 \pi \frac{i}{n}\right)}{\sqrt{(y^2 - \bar{y}^2)}} \quad (17)$$

et son module :

$$|C(p)| = \sqrt{2} \frac{\left[\left(\sum_{i=1}^n m_i \cos 2 \pi \frac{i}{n} \right)^2 + \left(\sum_{i=1}^n m_i \sin 2 \pi \frac{i}{n} \right)^2 \right]}{(y^2 - \bar{y}^2)}. \quad (18)$$

⁽¹⁾ Radon atmosphérique, Ile Crozet, TAAF, 1968-1969.

Ces expressions ne sont en fait (aux corrections près) que des coefficients de corrélation classiques ; elles ont donc l'avantage d'être sans dimensions et de module inférieur à 1.

Par rapport à l'équation ci-dessus (19)

$$C'(p) = \frac{\sqrt{2}}{N} \frac{\sum_{i=1}^N y_i \exp\left(-j 2 \pi \frac{t_i}{nh}\right)}{\sqrt{(\bar{y}^2 - \bar{y}^2)}} \quad (19)$$

où le numérateur est l'expression traditionnelle (t_i est l'instant où la mesure y_i a été faite), elles ont l'avantage de combler les trous de mesure par des valeurs beaucoup plus probables que les valeurs nulles qui sont implicitement admises dans (19), au coefficient N près ; cette dernière hypothèse parasite peut toutefois être aussi éliminée de l'éq. (19) moyennant un terme correctif, lequel tient compte du fait que la moyenne des diverses exponentielles complexes n'est pas nulle ; on obtient alors :

$$C''(p) = \frac{\sqrt{2}}{N} \frac{\sum_{i=1}^N y_i \exp\left(-j 2 \pi \frac{t_i}{nh}\right) - \bar{y} \sum_{i=1}^N \exp\left(-j 2 \pi \frac{t_i}{nh}\right)}{\sqrt{(\bar{y}^2 - \bar{y}^2)}} \quad (20)$$

qui est ainsi l'expression exacte du coefficient de corrélation (complexe) des fonctions y_i et

$$\exp(-j 2 \pi (t_i/nh)) ;$$

la deuxième partie du numérateur ne s'y annule que lorsque l'intervalle étudié contient un nombre entier de périodes et est sans trous ; en toute rigueur il faudrait aussi corriger le coefficient $\sqrt{2}$ qui représente la variance des sin ou cos, mais cet effet est négligeable.

Pratiquement, nous utilisons soit (18) soit (20).

Ici encore il est en pratique intéressant d'estimer ce qui, dans une transformée de Fourier de phénomènes entachés d'un bruit important, est significatif et ce qui ne l'est pas.

Comme précédemment nous ferons le calcul d'erreur le plus pessimiste en prenant pour erreur sur chaque mesure la dispersion de l'ensemble de ces mesures, ce qui revient à faire ce calcul d'erreur dans le cas limite où toutes les variations seraient dues à un bruit non significatif. Donc :

$$(\Delta C(p))^2 = \sum_{k=1}^n \left| \frac{\partial C_p}{\partial y_k} \right|^2 (\Delta y_k)^2 \quad (21)$$

où

$$(\Delta y_k)^2 = \bar{y}^2 - \bar{y}^2 \quad (22)$$

ce qui donne l'expression très simple :

$$|\Delta C(p)| = \sqrt{\frac{2}{N}} \quad (23)$$

Pratiquement, on s'intéresse surtout à la quantité $|C(p)|^2$ c'est-à-dire à ce qu'on appelle généralement le « spectre de puissance ».

2.2.3 Critique. — Ainsi déterminée la transformée de Fourier est un moyen d'étude puissant, beaucoup plus significatif que le périodogramme, et dont les propriétés mathématiques sont bien connues.

Elle présente néanmoins deux défauts :

1° Les raies sont inévitablement convoluées par la fonction de forme de l'intervalle de mesure ;

2° Lorsque les phénomènes étudiés ne sont pas sinusoidaux, on enregistre aussi leurs harmoniques qu'il est difficile (à moins de faire des études détaillées de phase) de distinguer des phénomènes indépendants (Fig. 5 et 6).

La figure 7 représente un exemple de transformée de Fourier de mesures réelles (les mêmes que dans la figure 3) et la figure 8 celle d'un générateur aléatoire de même répartition.

2.3 LA FONCTION D'AUTOCORRÉLATION. — 2.3.1 Description générale. — Sa définition, très simple, est bien connue :

$$F(\tau) = \int_{-\infty}^{+\infty} y(t) y(t + \tau) dt \quad (24)$$

soit, pour N mesures discrètes et régulièrement espacées de h :

$$F(n) = \sum_{i=1}^{N-n} y_i y_{i+n} h \quad (25)$$

Lorsque les phénomènes à étudier sont de nature ou de grandeur différente, il est intéressant de normaliser cette fonction qui devient alors un « coefficient d'autocorrélation » :

$$A(n) = \frac{\sum_{i=1}^{N-n} y_i y_{i+n} - \bar{y}^2}{(N-n)(\bar{y}^2 - \bar{y}^2)} = \frac{\sum_{i=1}^{N-n} (y_i - \bar{y})(y_{i+n} - \bar{y})}{(N-n)(\bar{y}^2 - \bar{y}^2)} \quad (26)$$

Les termes $(N-n)$ proviennent du fait que le nombre de mesures est fini et qu'il n'existe que $N-n$ couples (y_i, y_{i+n}) .

2.3.2 Améliorations. — Lorsqu'il manque des mesures il est nécessaire d'aménager l'expression (26) afin de ne pas introduire de valeurs parasites nulles. Pour cela, on remplace au numérateur les \bar{y} par des moyennes calculées sur les seuls u_i ou y_{i+n} intervenant effectivement dans la sommation, soit :

$$A(n) = \frac{\sum_{i=1}^{v_n} y_i y(t_i + nh) - \frac{1}{v_n} \sum_{i=1}^{v_n} y_i \sum_{i=1}^{v_n} y(t_i + nh)}{v_n(\bar{y}^2 - \bar{y}^2)} \quad (27)$$

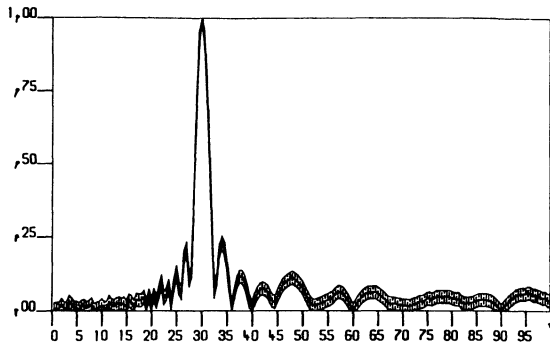


FIG. 5. — Spectre de puissance d'une sinusoïde (période 30 jours, durée 1 an).

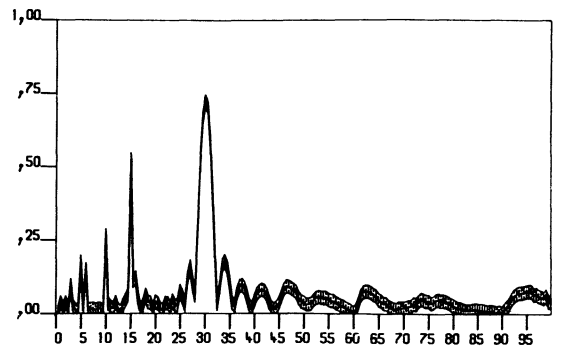


FIG. 6. — Spectre de puissance d'une fonction créneau périodique (période 30 jours, durée 1 an).

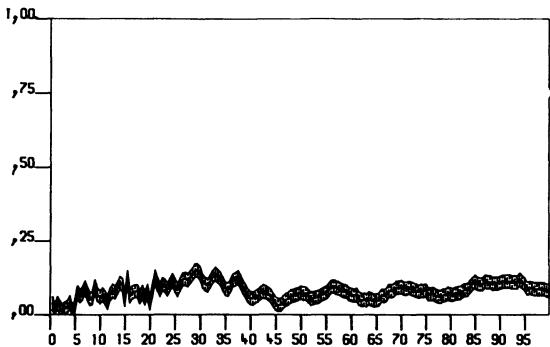


FIG. 7. — Spectre de puissance de mesures réelles (Radon atmosphérique, Ile Crozet, durée 1 an ; périodicité attendue : environ 28 jours).

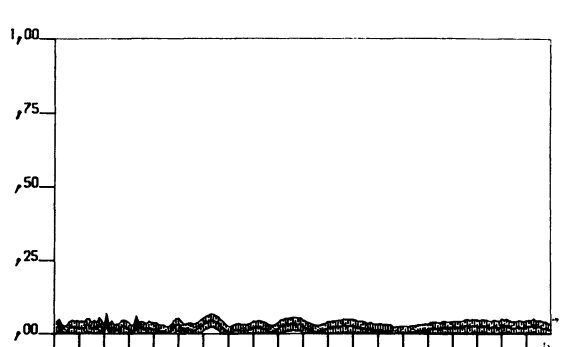


FIG. 8. — Spectre de puissance d'un générateur aléatoire de même répartition que la figure 7 (durée 1 an).

où v_n est le nombre de couples $(y_i, y(t_i + nh))$ existants et où $y(t_i + nh)$ est la mesure faite avec un décalage de nh par rapport à y_i ; si aucun trou n'existe entre t_i et $t_i + nh$ on a évidemment :

$$y_{i+nj} \equiv y(t_i + nh). \tag{28}$$

On peut, comme précédemment, faire un calcul de l'erreur sur $A(n)$:

$$|\Delta A(n)|^2 = \sum_{k=1}^N \left| \frac{\partial A(n)}{\partial y_k} \right|^2 \Delta y_k^2 \tag{29}$$

où d'après (27) :

$$\begin{aligned} \frac{\partial A_n}{\partial y_n} &= \frac{1}{v_n(\bar{y}^2 - \bar{y}^2)} \times \\ &\times \left[y(t_k + nh) + y(t_k - nh) - \frac{1}{v} \sum_i y_i - \frac{1}{v} \sum y(t_i - nh) \right] \end{aligned} \tag{30}$$

d'où :

$$\begin{aligned} \sum \left| \frac{\partial A_n}{\partial y_k} \right|^2 &= \\ &= \frac{2 \sum y(t_i + nh) y(t_i - nh) + 2 v_n \bar{y}^2 - 4 v_n \bar{y}^2}{v_n^2 (\bar{y}^2 - \bar{y}^2)}. \end{aligned} \tag{31}$$

Si maintenant on se rappelle l'hypothèse de base de ce calcul d'erreur, à savoir que les mesures sont supposées totalement erratiques il est alors légitime de considérer que le premier terme du numérateur est égal à :

$$2 v_n \bar{y}^2 \tag{32}$$

ce qui compte tenu de (22), donne

$$|\Delta A_n| = \sqrt{\frac{2}{v_n}} \neq \frac{2}{N - n}. \tag{33}$$

2.3.3 Critique. — La fonction d'autocorrélation présente l'avantage de ne pas être fortement perturbée par la présence d'harmonique dans la fonction à étudier, mais elle présente le très grave inconvénient de donner les périodes multiples du phénomène fondamental (Fig. 9 et 10).

La figure 11 représente la fonction d'autocorrélation des mêmes mesures réelles que sur les figures 3 et 7 ; la figure 12 représente, pour comparaison, la fonction d'autocorrélation d'un générateur aléatoire de même répartition.

On remarquera ainsi que la transformée de Fourier (spectre de puissance) et la fonction d'autocorrélation présentant des avantages complémentaires, l'une ne donnant pas les périodes multiples, l'autre étant insensible aux harmoniques non significatifs, c'est ce

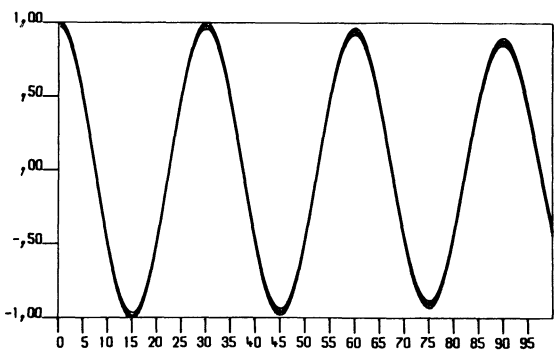


FIG. 9. — Coefficient d'autocorrélation d'une sinusoïde (période 30 jours, durée 1 an).

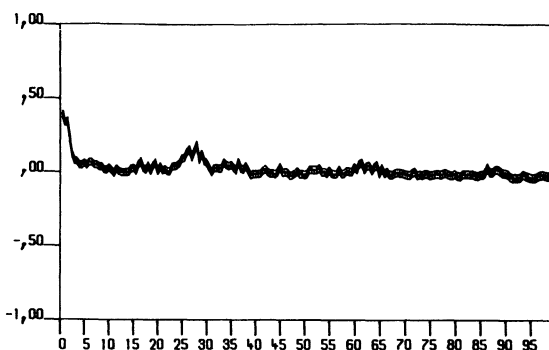


FIG. 11. — Coefficient d'autocorrélation de données réelles (Radon atmosphérique, Ile Crozet, durée 1 an ; périodicité attendue 28 jours).

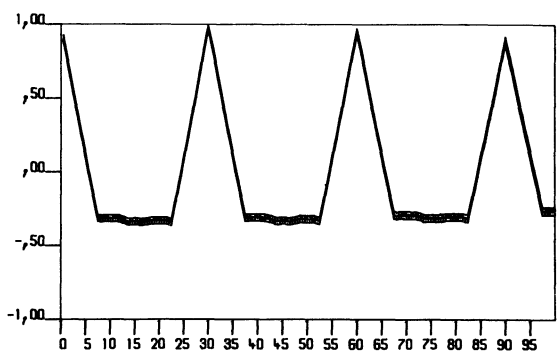


FIG. 10. — Coefficient d'autocorrélation d'une fonction créneau périodique (période 30 jours, durée 1 an).

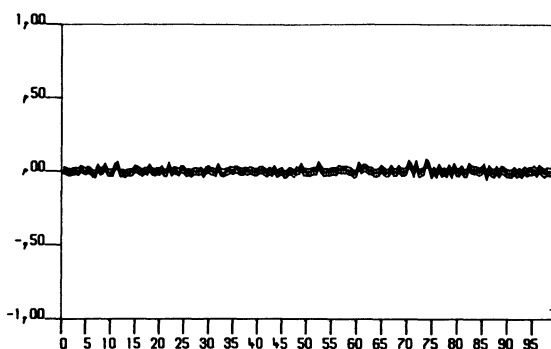


FIG. 12. — Coefficient d'autocorrélation d'un générateur aléatoire de même répartition que la figure 11 (durée 1 an).

qui nous a amenés à imaginer une méthode mixte que nous présentons ci-dessous.

3. Méthode mixte fonction d'autocorrélation/spectre de puissance. — Elle est basée sur le fait que les avantages de la transformée de Fourier (spectre de puissance) et de la fonction d'autocorrélation sont complémentaires ; elle consiste à calculer ces deux fonctions en fonction de la même variable $p = nh$ et à en faire le produit algébrique

$$\Phi(p) = A(p) |C(p)|^2. \tag{34}$$

Ainsi seules les périodicités fondamentales seront restituées, les harmoniques étant annulés par la fonction d'autocorrélation, les périodicités multiples étant au contraire annulées par le spectre de puissance.

Pratiquement, on calcule d'abord la fonction d'autocorrélation ensuite l'application des propriétés classiques du produit de convolution et de la transformation de Fourier permettent de calculer le spectre de puissance par transformation de Fourier de la fonction d'autocorrélation (cf. éq. (20)) :

$$|C(p = nh)|^2 = \frac{\sqrt{2} \sum_{j=1}^M A(j) \cos 2\pi \frac{j}{M} - \frac{1}{M} \sum_{j=1}^M A(j) \sum_{j=1}^M \cos 2\pi \frac{j}{M}}{\sqrt{\frac{1}{M} \sum_{j=1}^M A(j)^2 - \left(\frac{1}{M} \sum_{j=1}^M A(j)\right)^2}}. \tag{35}$$

Comme pour les autres méthodes, il est possible de faire un calcul d'erreur pour déterminer le seuil au-dessus duquel la fonction Φ devient significative ; en partant de l'éq. (34), en maximisant $A(p)$ et $C(p)$ à 1, et en utilisant les éq. (23) et (33) on trouve :

$$\Delta\Phi(p = nh) = \sqrt{\frac{8}{N} + \frac{2}{N-n}} \tag{36}$$

soit :

$$\Delta\Phi \# \frac{3}{\sqrt{N}} \tag{37}$$

lorsque $n < N/2$, seul domaine significatif.

Cette méthode présente, outre ses caractéristiques fondamentales relatives aux harmoniques et périodes multiples, plusieurs avantages :

1° un gain de temps puisque la fonction d'autocorrélation est utilisée pour calculer la transformée de Fourier ;

2° une élimination simple de la plupart des trous : ceux-ci sont déjà en partie éliminés en calculant la fonction d'autocorrélation, ce qui conduit à ce que certains v_n (cf. éq. (2)) soient inférieurs à la valeur théorique $N - n$; quant aux valeurs de n pour lesquelles $A(n)$ est incalculable parce que v_n est nul, elles sont en majorité concentrées dans les régions où n est grand, c'est-à-dire dans les régions de faible intérêt.

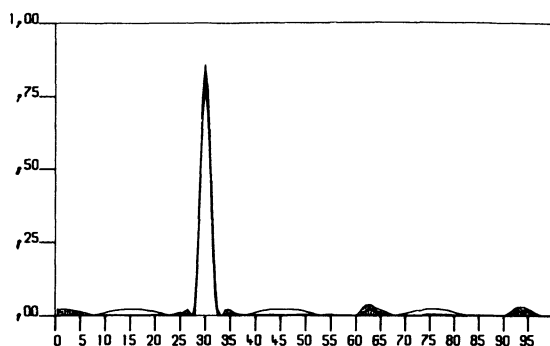


FIG. 13. — Spectre de puissance après apodisation d'une sinusoïde (période 30 jours ; durée 1 an).

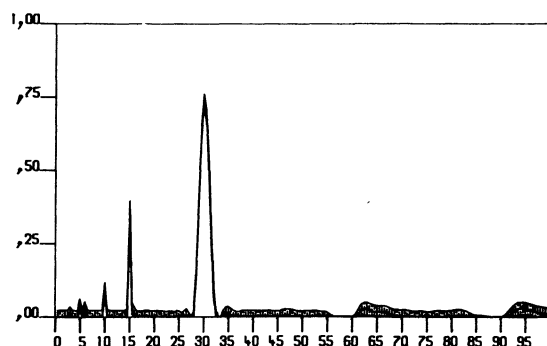


FIG. 14. — Spectre de puissance après apodisation d'une fonction créneau périodique (période 30 jours ; durée 1 an).

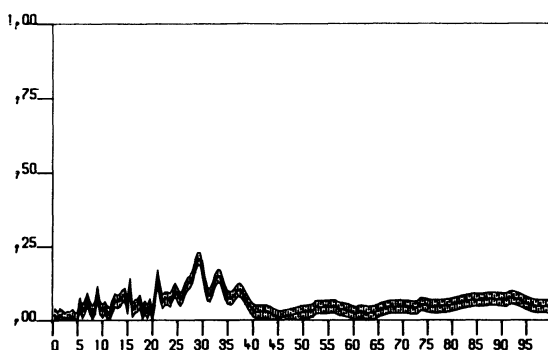


FIG. 15. — Spectre de puissance de mesures réelles (Radon atmosphérique, Ile Crozet, durée 1 an ; périodicité attendue 28 jours).

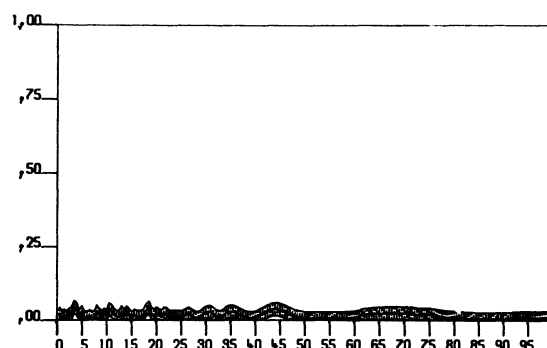


FIG. 16. — Spectre de puissance d'un générateur aléatoire de même répartition que la figure 15 (durée 1 an).

La perturbation apportée dans le calcul de la transformée de Fourier par les artifices décrits dans le paragraphe 1. B sera donc sans grande importance.

3° L'application des procédés d'apodisation qui éliminent les effets de bords y est particulièrement commode. En effet ces procédés consistent à modifier la fonction originale (qui est le produit de la fonction intéressante par une fonction de forme rectangulaire, égale à 1 pendant l'intervalle observé et à zéro ailleurs), en la multipliant par une nouvelle fonction de forme s'annulant moins brutalement aux extrémités de l'intervalle de mesure [4]. Si l'on applique ce procédé aux mesures originales on perturbe les informations apportées près des extrémités de l'intervalle de mesure, bien que celles-ci soient tout aussi valables que celles situées au centre ; au contraire, l'application de cet artifice à la fonction d'autocorrélation affecte préférentiellement les régions les moins significatives.

On trouvera dans les figures 13, 14, 15, 16 les spectres de puissance correspondant aux mêmes données que sur les figures 5 à 8, après apodisation (correction d'effets de bords).

La fonction Φ pour les mêmes données est représentée sur les figures 17 à 20.

4. **Quelques remarques pratiques.** — Il est particulièrement intéressant de comparer ce que les diverses méthodes fournissent à partir des données réelles : alors que le périodogramme (Fig. 3) ne donne aucune

indication notable quant à la périodicité du phénomène, la plus grande intensité du spectre de puissance se trouve vers 28 jours (Fig. 7) ; on trouve également un pic très notable à cette valeur sur la figure 11 (autocorrélation) mais les résultats les plus probants sont ceux de la figure 19, obtenus par la présente méthode. Ceci est d'ailleurs en accord avec des prévisions faites par ailleurs (activité solaire).

Les diverses méthodes présentées ici ont été programmées au centre de calcul de l'Université de Paris-Sud (2). Le temps de calcul est de quelques minutes pour traiter un ensemble de 5 000 mesures, par 2 ou 3 des méthodes ci-dessus.

Afin d'éviter des résultats parasites liés à des trous périodiques toujours possibles (par exemple des pannes déclenchées préférentiellement par des opérations de maintenance périodiques) nous avons préféré effectuer, conjointement à chaque traitement, un passage partant de données aléatoires de même répartition globale, même durée, mêmes trous, que les données réelles (Fig. 4, 8, 12, 16 et 20) ; ceci est pratiquement obtenu en remplaçant chaque mesure existante par une autre tirée au hasard parmi toutes les autres. Ceci est effectué automatiquement par nos programmes.

Nous avons également fait des essais en remplaçant

(2) Ordinateur Univac 1108.

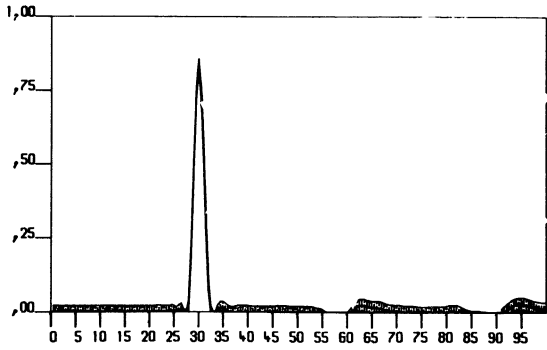


FIG. 17. — Fonction Φ d'une sinusoïde (période 30 jours, durée 1 an).

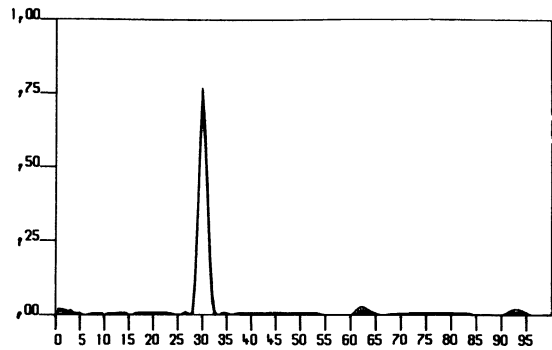


FIG. 18. — Fonction Φ d'une fonction créneau périodique (période 30 jours, durée 1 an).

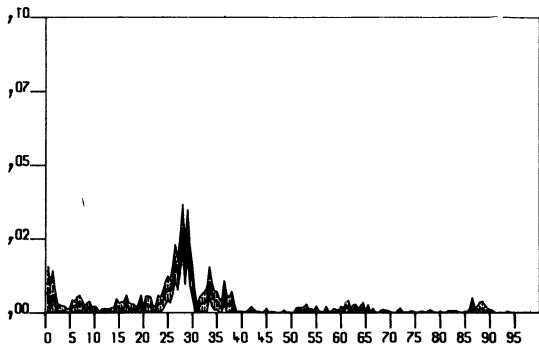


FIG. 19. — Fonction Φ de données réelles (Radon atmosphérique, Ile Crozet, durée 1 an, périodicité attendue 28 jours).

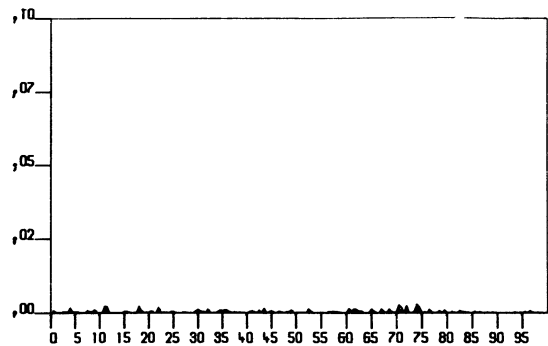


FIG. 20. — Fonction Φ d'un générateur aléatoire de même répartition que la figure 19 (durée 1 an).

chaque mesure existante par une valeur constante non nulle, afin de vérifier qu'aucune mesure absente n'était implicitement remplacée par un zéro : tous les résultats obtenus sont rigoureusement nuls. Ceci vérifie

donc que la fonction de forme de l'intervalle (« troué ») de mesure n'intervient jamais seule dans nos diagrammes, mais seulement pour convoluer des mesures non blanches.

Bibliographie

- [1] LAMBERT, G., POLIAN, G., TAUPIN, D., *J. Geophys. Res.* 75 (1970) 2341.
- [2] CHARAKHTCHYAN, A. N., TULINOV, V. F., CHARAKHTCHYAN, A. N., International Conference on Cosmic Rays, I. G. Y. Prog., Sect 7.4, Moscou (1961), NSA-TT-F 127, 274-280.
- [3] ANDERSON, T. W., *The Statistical Analysis of Time Series* (Wiley & sons, N. Y.) 1971 107.
- [4] JACQUINOT, P., ROIZEN-DOSSIER, B., *Progress in Opt.* 3 (1964) 62.