

Sur le choix "au hasard" d'une impulsion

par Jörg W. Müller

Bureau International des Poids et Mesures, F-92310 Sèvres

Abstract

A curious difficulty arises when we try to pick out randomly a pulse out of a stochastic sequence of events, requiring that it be surrounded by time intervals of unbiased length. The usual procedure chooses preferably long time intervals, and for a Poisson process such a pulse is then actually preceded by an interval of double average length. This distortion can be explained theoretically and it is well confirmed by experiment. For the time being no simple and reliable method seems to be known which would allow us to avoid the bias completely.

1. Introduction

A première vue, rien ne semble plus facile que de choisir tout à fait au hasard une impulsion qui fait partie d'un processus aléatoire. A la rigueur, on pourrait même se demander comment il serait possible de faire autrement, étant donné que - au moins pour un processus de renouvellement - toute arrivée d'une impulsion correspond à un point de régénération à partir duquel le passé n'a plus d'influence sur le développement futur. Pourquoi, alors, ne devrait-on pas avoir le droit d'utiliser tout simplement la première impulsion qui se présente? Nous allons tenter d'expliquer pour quelle raison, dans l'étude pratique de séquences d'événements, le choix d'une impulsion qui peut être considérée comme vraiment aléatoire pose souvent problème.

C'est dans le contexte de la méthode d'échantillonnage sélectif que ce genre de difficulté s'est déjà manifesté - pas pour la première fois, mais de façon particulièrement nette -, car la répartition des impulsions observées en dépend beaucoup. L'explication que nous avons donnée [1] pour la déformation observée dans la distribution temporelle des impulsions gamma ne semble pas toujours avoir été bien comprise, ce qui n'est pas trop étonnant. En effet, l'observation se trouvait dans un contexte rendu difficile à expliquer par la présence d'un temps mort cumulatif qui, lui aussi, était responsable d'une perturbation importante. Au surplus, on avait constaté que l'endroit d'une intersection avec l'axe du temps, observable dans le cas d'un déclenchement par la première impulsion bêta disponible, ne coïncidait avec la prévision théorique que pour un taux de comptage faible.

Tout cela n'a guère contribué à emporter la conviction générale. Néanmoins, l'essentiel de l'explication était correct, et c'est ce point que nous nous proposons de discuter ici un peu plus en détail dans un contexte simple et clair. Par ailleurs, le problème de l'intersection s'est également éclairci entre temps: le décalage est dû aux impulsions perdues pendant le temps mort cumulatif, mais présentes avant - et donc aussi dans les enregistrements. Cependant, une discussion approfondie doit être reportée à plus tard car elle nécessitera un développement plus large.

## 2. Exposé du problème

Un procédé couramment utilisé pour faire un choix "au hasard" consiste à prendre, à partir d'un moment déterminé arbitrairement dans le temps (et en particulier sans se référer au processus à étudier), la prochaine impulsion qui se présente. Or, si un tel choix est aléatoire dans le temps, il ne l'est pas pour les impulsions!

Désignons par  $f(t)$  la densité des intervalles de temps entre événements consécutifs. Un moment choisi au hasard sur l'axe de temps se place, cependant, de préférence dans un intervalle long, de façon à ce que la densité des intervalles ainsi choisis soit donnée par

$$g(t) = \frac{t \cdot f(t)}{\int_0^{\infty} t \cdot f(t) dt}, \quad (1)$$

où le dénominateur assure la normalisation.

L'illustration la plus simple (et sans doute la plus convaincante) de l'effet produit par (1) se sert d'un simple processus de Poisson, que nous admettrons non perturbé. Dans ce cas, on sait que les intervalles d'ordre  $k$  sont décrits par des densités gammas, c'est-à-dire (pour  $t \geq 0$ )

$$f_k(t) \equiv \{f(t)\}^{*k} = \frac{\rho(\rho t)^{k-1}}{(k-1)!} e^{-\rho t}, \quad k = 1, 2, \dots, \quad (2)$$

où  $\rho$  est le taux de comptage du processus (et avec  $f_1(t) \equiv f(t)$ ).

Pour un processus de Poisson on a donc

$$g(t) = \frac{t \cdot \rho e^{-\rho t}}{\rho \int_0^{\infty} t \cdot e^{-\rho t} dt} = \rho^2 t \cdot e^{-\rho t} = f_2(t). \quad (3)$$

Par conséquent, l'impulsion que l'on croyait choisie au hasard est en réalité précédée d'un intervalle anormalement long. Pour un processus de Poisson, celui-ci est de densité  $f_2(t)$ , au lieu de  $f_1(t)$ , ce qui correspond au double de la distance moyenne entre impulsions (voir Annexe pour quelques détails).

Puisque les intervalles sont indépendants les uns des autres, tous ceux qui précèdent ou suivent ce premier intervalle, qui est décrit par la densité  $g(t)$ , sont "normaux", c'est-à-dire donnés par  $f(t)$ . Par conséquent, pour le  $k$ -ième intervalle précédant l'origine du temps  $t = 0$ , on a la densité ( $x \equiv -t$  et  $g_1(x) \equiv g(x)$ )

$$\begin{aligned} g_k(x) &= g(x) * \{f(x)\}^{*(k-1)} \\ &= g(x) * f_{k-1}(x) = f_{k+1}(x), \end{aligned} \quad (4)$$

dont la forme explicite est donnée par (2).

Il en résulte pour la densité totale

$$\text{- pour } t > 0: F(t) \equiv \sum_{k=1}^{\infty} f_k(t) = \rho \cdot e^{-\rho t} \sum_{k=1}^{\infty} \frac{(\rho t)^{k-1}}{(k-1)!} = \rho, \quad (5)$$

$$\begin{aligned} \text{- pour } t < 0: G(x) &\equiv \sum_{k=1}^{\infty} g_k(x) = \sum_{k=2}^{\infty} f_k(x) \\ (\text{donc } x > 0) & \\ &= \rho - f_1(x) = \rho(1 - e^{-\rho x}). \end{aligned} \quad (6)$$

Ce raisonnement se généralise aisément. Si l'on prend pour point de départ (et origine de temps) l'impulsion numéro  $K \geq 1$ , on trouve pour les densités

$$\begin{aligned} \text{- pour } k < K: g_k(x) &= f_k(x), \text{ mais} \\ \text{- pour } k \geq K: g_k(x) &= f_{K-1}(x) * g_{k-K+1}(x) = f_{k+1}(x). \end{aligned} \quad (7)$$

Il en découle pour la densité totale avant  $t = 0$

$$\begin{aligned} G(x) &= \sum_{k=1}^{\infty} g_k(x) = \sum_{k=1}^{K-1} g_k(x) + \sum_{k=K}^{\infty} g_k(x) \\ &= \sum_{k=1}^{K-1} f_k(x) + \sum_{k=K+1}^{\infty} f_k(x) = \sum_{k=1}^{\infty} f_k(x) - f_K(x) \\ &= \rho - f_K(x) = \rho \left[ 1 - \frac{(\rho x)^{K-1}}{(K-1)!} \cdot e^{-\rho x} \right], \end{aligned} \quad (8)$$

résultat qui se réduit à (6) pour  $K = 1$ . La densité  $F(t)$  pour  $t > 0$ , donnée par (5), reste constante et ne dépend pas de  $K$ .

### 3. Observations expérimentales

P. Bréonce a fait une série d'enregistrements de densités cumulatives (ou somme des fréquences) des temps d'arrivée d'impulsions bêta. On s'est servi dans ce but d'un sélecteur multicanaux, utilisé en mode multiéchelle, sur lequel on a pu accumuler les résultats d'un très grand nombre de cycles. Le schéma de principe est indiqué dans la Fig. 1.

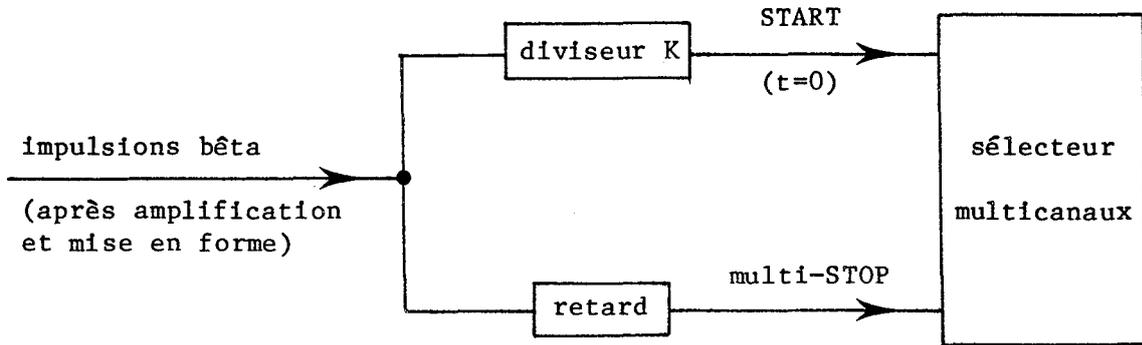


Fig. 1 - Principe du dispositif expérimental utilisé pour les mesures présentées dans la figure 2.

Pour toutes ces mesures le taux de comptage, provenant d'une source radioactive de  $^{60}\text{Co}$ , était d'environ  $35\,000\text{ s}^{-1}$ , et pour la durée d'un canal sur le sélecteur on a choisi une valeur de  $1,5\ \mu\text{s}$ . Pour simplifier la comparaison des courbes superposées dans la Fig. 2, le nombre moyen d'impulsions enregistrées par canal a été ajusté pour le domaine  $t > 0$  sensiblement à la même valeur.

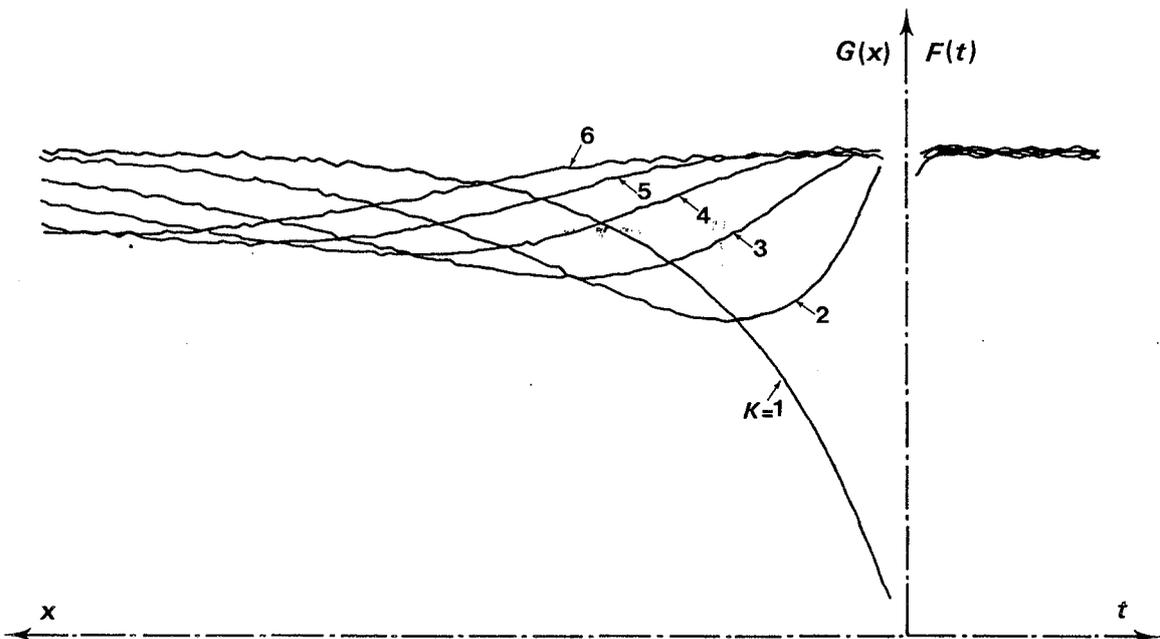


Fig. 2 - Répartition observée des temps d'arrivée des impulsions au voisinage d'une impulsion choisie au hasard.  $K$  est le rang de l'impulsion qui définit l'origine du temps.

Il est facile de constater qu'en général l'accord des répartitions mesurées avec les courbes calculées à l'aide de (8) est excellent. Les quelques écarts, visibles surtout aux alentours de  $t = 0$ , proviennent de l'inévitable distorsion que la série d'impulsions a subi dans la chaîne électronique (amplificateurs). Cette perturbation est bien connue de nos études antérieures; elle correspond approximativement à l'effet produit par un temps mort d'une durée d'environ  $1,2 \mu\text{s}$ .

#### 4. Compléments et conclusion

On se demandera maintenant sans doute si - et dans l'affirmative comment - le problème que nous venons de discuter est lié à ce qui est parfois appelé le "paradoxe de temps d'attente" pour un processus de Poisson [2]. Il s'agit du fait assez surprenant que le temps qu'il faut attendre pour observer le prochain événement est exactement le même si l'on commence l'enquête avec l'arrivée d'une impulsion ou à n'importe quel moment. Or, si l'on admet que ce moment tombe dans un intervalle "normal", décrit par la densité  $f(t)$  et dont la valeur moyenne est donc  $m_1 = 1/\rho$ , on devrait penser que le temps d'attente, avec début choisi au hasard à l'intérieur de cet intervalle, serait plus court, et sans doute de valeur moyenne  $m_1/2$ .

La contradiction peut se résoudre au moins de deux manières, dont aucune n'est triviale. Dans la première, on souligne d'abord que l'hypothèse selon laquelle on a affaire à un intervalle "normal" est fautive: la densité de l'intervalle en question n'est pas  $f(t)$ , mais  $g(t) = f(t)*f(t) = f_2(t)$ , comme on l'a vu dans (3). Cet intervalle a une durée moyenne de  $2m_1$ . Le temps d'attente moyen, qui en est en effet la moitié, nous ramène donc à la valeur attendue (et observée) de  $1/\rho$ . Ceci établit le lien étroit qui existe entre ce "paradoxe" et le problème qui nous préoccupe ici.

Le deuxième dénouement, plus souvent proposé, utilise quelques notions élémentaires de la théorie de renouvellement [3]. Dans un processus, dont les intervalles séparant deux événements consécutifs ont la densité  $f(t)$ , on peut envisager plusieurs façons de le commencer. En particulier, si l'origine de temps coïncide avec l'arrivée d'un événement, on parle d'un processus ordinaire. Par contre, si cette origine est choisie au hasard, on a affaire à un processus stationnaire et on montre que, dans ce cas, le premier intervalle (entre l'origine de temps et la première arrivée) est donné par

$$f_1(t) = \frac{1 - \int_0^t f(x) dx}{m_1} . \quad (9a)$$

En termes de transformées de Laplace, ceci correspond à

$$\tilde{f}_1(s) = \frac{1 - \tilde{f}(s)}{m_1 \cdot s} . \quad (9b)$$

Pour le cas particulier d'un processus de Poisson, c'est-à-dire pour

$$f(t) = \rho \cdot e^{-\rho t} \quad \text{ou} \quad \tilde{f}(s) = \frac{\rho}{\rho + s} ,$$

on trouve aisément avec (9) qu'on a, en effet,

$$f_1(t) = f(t) , \quad (10)$$

puisque  $m_1 = 1/\rho$ . Ce résultat "explique" formellement le paradoxe du temps d'attente. On peut montrer que le processus de Poisson (à densité exponentielle pour les intervalles) est le seul cas où (10) est valable.

Comme nous l'avons vu précédemment, on doit conclure de (8) que la distribution des impulsions qui arrivent avant celle qui marque l'origine de temps, donnée par  $G(x)$ , est toujours plus ou moins distordue. Cette déformation est assez spectaculaire pour  $K = 1$ , c'est-à-dire si l'on choisit le premier événement disponible. Elle diminue en augmentant  $K$ , mais en principe la distorsion ne disparaît jamais complètement.

Pour certaines mesures ce fait peut poser un problème sérieux, étant donné que le subterfuge de choisir  $K \gg 1$  n'est pas rigoureux et implique au moins une perte de temps. Pour le moment, aucune solution idéale ne paraît connue, bien qu'on puisse trouver des approximations qui peuvent être suffisantes. Il nous semble que toute approche qui se sert d'une manière ou d'une autre d'un point pris "au hasard" dans le temps pour trouver, à partir de lui, un "candidat" approprié, doit présenter le même type de difficulté.

Par contre, une méthode permettant de choisir directement un événement (et non un intervalle), par exemple à l'aide d'une coïncidence fortuite avec une série indépendante de signaux, pourrait donner lieu à une solution utilisable en pratique. Quelques essais dans cette direction ont déjà été faits et ont donné des résultats encourageants.

La rencontre avec un problème qui semble à la fois élémentaire et profond, comme dans notre cas le choix "au hasard" d'une impulsion, est toujours un peu inquiétant, car on se sent à proximité d'un piège. Ou faut-il plutôt conclure que tout sujet est encore plein de petites curiosités si l'on y regarde de plus près?

## ANNEXE

Détermination du nouvel intervalle moyen

Pour une densité d'intervalle  $f(t)$  et des moments  $E\{t^r\} = m_r$ , la densité de l'intervalle qui précède l'origine du temps est donnée, d'après (1), par

$$g(t) = t \cdot f(t) / m_1 . \quad (A1)$$

Que peut-on déduire de cette relation générale pour les moments de  $g(t)$ , et en particulier pour le nouvel intervalle moyen  $\bar{t} = M_1$ ?

Il est commode de se servir des transformées intégrales. Ainsi, pour les transformées de Laplace on connaît la relation

$$\mathcal{L}\{t \cdot f(t); s\} = - \frac{d\mathcal{L}\{f(t); s\}}{ds} ,$$

d'où

$$\tilde{g}(s) \equiv \mathcal{L}\{g(t); s\} = - \frac{1}{m_1} \cdot \frac{d\tilde{f}(s)}{ds} . \quad (A2)$$

A l'aide du développement valable pour toute densité

$$\begin{aligned} \tilde{f}(s) = E\{e^{st}\} &= 1 - s E\{t\} + \frac{s^2}{2!} E\{t^2\} - \frac{s^3}{3!} E\{t^3\} \pm \dots \\ &= 1 - s m_1 + \frac{s^2}{2} m_2 - \frac{s^3}{6} m_3 \pm \dots , \end{aligned} \quad (A3)$$

donc aussi

$$\frac{d\tilde{f}(s)}{ds} = - m_1 + s m_2 - \frac{1}{2} s^2 m_3 \pm \dots ,$$

on arrive, avec (A2), à l'expression générale

$$\begin{aligned} \tilde{g}(s) &= - \frac{1}{m_1} (-m_1 + s m_2 - \frac{1}{2} s^2 m_3 \pm \dots) \\ &= 1 - s \frac{m_2}{m_1} + \frac{s^2}{2} \frac{m_3}{m_1} \mp \dots . \end{aligned} \quad (A4)$$

Les moments de  $g(t)$  sont donc donnés par

$$M_r = m_r / m_1 , \quad (A5)$$

d'où, en particulier,  $\bar{t} = M_1 = m_2 / m_1$  .

Ainsi, par exemple, pour une densité exponentielle  $f(t) = \rho \cdot \exp(-\rho t)$ , aux premiers moments  $m_1 = 1/\rho$  et  $m_2 = 2/\rho^2$ , on a

$$M_1 = 2 m_1 ,$$

donc un redoublement de l'espace moyen entre les impulsions. Ceci est en accord avec l'observation faite plus haut que  $g(t) = f_2(t)$ , compte tenu de l'additivité des premiers moments dans les opérations de convolution.

### Références

- [1] J.W. Müller: "Comment choisir les cycles d'enregistrement dans l'échantillonnage sélectif?", BIPM WPN-221 (1981),  
in Recueil de Travaux du BIPM, Vol. 8, 1981-1982
- [2] W. Feller: "An Introduction to Probability Theory and its Applications, Volume II" (Wiley, New York, 1966 ou 1971<sup>2</sup>),  
chapitre I,4
- [3] D.R. Cox: "Renewal Theory" (Methuen, London, 1962)

(Mars 1985)