A PSEUDO-MARKOV MODEL FOR SERIES OF NEURONAL SPIKE EVENTS

ANDERS EKHOLM *and* JUHANI HYVÄRINEN

From the Finnish Foundation for Alcohol Studies and the Institute of Physiology, University of Helsinki, Helsinki, Finland

ABSTRACT Spike trains of spontaneous neuronal activity in the rabbit brain are submitted to statistical analyses based on the following pseudo-Markov model. The nerve cell is supposed to alternate between a bursting and a resting state. The numbers of consecutive spikes within each state are assumed to be independent integer-valued random variables with discrete probability distributions. Given the state, the interspike intervals are independent real-valued random variables. The two state semi-Markov model is obtained as a special case when the discrete distributions are geometrical. Statistical second-order properties of recorded spike trains are compared with those predicted by the model on the basis of known first-order properties. For that purpose, serial correlation coefficients and intensity functions for spike trains produced by the model are computed. A comparison between observed and predicted results for the spontaneous activity of 17 brain cells yields a good fit in eight cells and discloses some salient features of the statistical structure in the activity of six other cells. By making it feasible to compute theoretical correlograms, the model may advance the understanding of empirical correlograms. The possibilities for integrating this statistical model of spike trains with a model of the mechanism of spike train production are discussed.

INTRODUCTION

The purpose of this paper is to display a statistical model of the single spike train. Statistical models of spike trains are, of course, only a means to the neurophysiological end of constructing and testing models of the mechanisms of spike production (Perkel, Gerstein, and Moore, 1967 *a* and *b*). There seems, however, to be a need for more complicated statistical models than the ones commonly used. The spontaneous spike discharge activity of the neural cell has sometimes been considered to have the statistical characteristics of a Poisson process. This may be the case in some peripheral afferent nerves like the muscle spindle afferents (Buller, Nicholls, and Ström, 1953; Hagiwara, 1954), carotid body afferents (Biscoe and Taylor, 1963), and auditory primary afferents (Kiang, 1965). In most of the studies of the spontaneous unit activity in the central nervous system, others than exponential interspike interval distributions have been reported, particularly distributions resembling the gammatype (Kuffler, Fitzhugh, and Barlow, 1957; Hyvärinen, 1966), normal distributions, and multimodal distributions as well as many unclassified distributions (Rodieck, Kiang, and Gerstein, 1962; Poggio and Viernstein, 1964; Pfeiffer and Kiang, 1965). Therefore, in the central nerve cells all spontaneous activity cannot be considered a simple Poisson process.

It is then natural to assume that the spike trains have the statistical characteristics of a renewal process. Several of the recent models of the mechanisms of spike production end up with this assumption (for example, Gerstein and Mandelbrot, 1964; Stein 1967; for a summary see Moore, Perkel, and Segundo, 1966).

In some recent experimental works there has, however, been found spontaneous spike activity which does not fulfill the criteria of a renewal process (Werner and Mountcastle, 1963; Poggio and Viernstein, 1964). In a study of spontaneous spike trains from 300 cells, Hyvärinen (1966) found that about 30% of the trains did not fit the renewal hypothesis.

The outstanding characteristic of the renewal model is that the serial correlations of all lags are supposed to be zero. We propose a model called the two state pseudo-Markov (PM) model in which this assumption is relieved by assuming two alternating states of activity of the cell. The model makes it possible to compute theoretical correlograms and compare these to empirically found correlograms. We believe that there is a need for models from which the serial correlations of many lags can be computed. Perkel, Gerstein, and Moore (1967 *a*, p. 402) show a considerable interest in the interpretation of empirical correlograms, but their rules of interpretation are derived by simulation experiments (a very laborious way) or from empirical experience (not always reliable and slowly learned). Among the generalizations of the renewal model which Cox and Lewis (1966) survey, the two state semi-Markov model is unique in that the serial correlations of the model for all lags can be easily computed. In fact, we started out from an attempt to fit the two state semi-Markov model to some spike trains but discovered that we needed a model allowing more freedom for the behaviour of the serial correlations. The PM model was developed from this impetus (Ekholm, 1967).

In the latter part of this paper, we perform a statistical analysis based on the PM model of 17 spike trains drawn from the nonrenewal trains of Hyvärinen (1966). Finally, we discuss some possibilities of deriving models of spike production in the mathematical framework given by the PM model.

THE TWO STATE PSEUDO-MARKOV MODEL

Basic Mathematical Concepts

The basis for our mathematical and statistical work is the excellent textbook by Cox and Lewis (1966) with which the reader is supposed to be familiar. The basic concepts are also presented from the neurophysiological point of view in Perkel, Gerstein, and Moore (1967 *a*). We are concerned only with single unit spontaneous activity, a stationary process being assumed.

For the times between successive spikes we write $X_1, X_2, \cdots$. These are random variables, and for the observations we use the corresponding small letters $x_1, x_2, \cdots$. A spike train can be described formally in two different ways. One can study the sequence of intervals $\{X_k; k = 1, 2, \cdots\}$ or the process of counts $\{N(t); t > 0\}$, where $N(t)$ denotes the number of impulses up to and including the time point t. It is shown by Cox and Lewis (1966, p. 66) that the second-order properties of each aspect afford genuinely new information about the spike train not readily derived from the other aspect.

To depict the sequence of intervals, we shall use the probability density function (pdf), denoted $f(x)$, of the marginal distribution of the X_i's with mean (μ) and variance (σ^2) and the serial correlation coefficients $\rho(k)$, $k = 1, 2, \cdots$, where k denotes the order of the lag.

The second-order properties of the process of counts $\{N(t); t > 0\}$ will be described by the intensity function $h(t)$ defined as

$$h(t) = \lim_{\Delta t \to 0} \frac{\text{Prob}\,\{\text{event in } (t, t + \Delta t) \mid \text{event at } 0\}}{\Delta t}. \tag{1}$$

Perkel, Gerstein, and Moore (1967 *a*, p. 404) call this function in the general case the renewal density, which is a most unfortunate use of names. We shall compare the estimated intensity function with the intensity functions computed under the hypothesis advanced here and under the renewal hypothesis. We shall call only the latter of these two the renewal density.

The Poisson process is characterized by the fact that the X_i's are independent and identically distributed with $f(x) = \theta \exp(-\theta x)$. The serial correlations are zero for all lags and the intensity function is constantly equal to θ. If the Poisson process is an adequate statistical model of a spike train, the information of the data is aptly summarized by the estimate of the parameter θ.

In a renewal process the X_i's are still independent and identically distributed but with any pdf $f(x)$. The serial correlations are all zero and the intensity function is not constant but uniquely determined by the pdf. If the renewal model is adequate, the data are summarized by the estimate of the pdf $f(x)$. The estimate of the continuous function $f(x)$ can either be in the form of a histogram or in the form of estimates of the parameters of an analytically defined distribution.

In the PM model the X_i's are still substantially independent, but are not formally independent. This is so because we assume that the cell alternates between a "bursting" and a "resting" state. In the bursting state the times between the spikes follow a distribution with pdf $f_1(x)$ (mean μ_1, and variance σ_1^2) and in the resting state they follow another distribution with pdf $f_2(x)$, (μ_2, σ_2^2). The common marginal distribution of the X_i's is then a mixture of these two distributions. The fundamental assumption is that the number of consecutive spikes in one state before the cell shifts to the opposite state is an integer-valued random variable, independent of the number of spikes in the other state. The number of consecutive spikes in the bursting state follows the discrete probability distributuion $\{p_1(k); k = 1, 2, \cdots\}$ and the number of consecutive spikes in the resting state follows the distribution $\{p_2(k); k = 1, 2, \cdots\}$. Conditional upon the state of the cell, the times between spikes are independent. Under these assumptions the serial correlations are not zero and they can be computed from a knowledge of the discrete distributions $\{p_\nu(k)\}$, and the means and variances μ_ν, σ_ν^2, $\nu = 1, 2$. The intensity function can be obtained from a knowledge of the distributions $\{p_\nu(k)\}$ and the pdfs, $f_\nu(x)$, $\nu = 1, 2$. If the PM model is adequate, the data are summarized by the estimates of the two discrete distributions $\{p_\nu(k)\}$, $\nu = 1, 2$ and the two pdfs, $f_\nu(x)$. Again, the estimates can be either in the form of histograms or in the form of estimated parameters of analytically defined distributions.

The PM model is reduced to the two state semi-Markov model with two types of intervals (Cox and Lewis, 1966, p. 194) if the distributions $\{p_\nu(k)\}$, $\nu = 1, 2$ are geometrical, that is for $k = 1, 2, \cdots$,

$$p_\nu(k) = (1 - \alpha_\nu)\alpha_\nu^{k-1}, \qquad \nu = 1, 2, \tag{2}$$

where $0 < \alpha_\nu < 1$. It is thus clear that the PM model also includes, as a special case, the renewal model.

Description of the PM Process

To make quite clear how the PM process is built up, we shall indicate how a realization could be constructed by simulation. We shall use the expression "spike of state 1" (or 2). Actually, of course, all spikes are indistinguishable and a spike of state 1 means that the interval following this spike is sampled from the distribution $f_1(x)$. A sequence of spikes of like kind will be called a run.

Throughout this paper we will assume that the origin of the time axis coincides with a spike. Let us start the realization in state 1, that is, the spike at the origin is of state 1. The serial number of that spike is 0. We fix the length of the first run of spikes of state 1 (denoted by Y_{11}) by choosing a positive integer in accordance with the discrete probability distribution $\{p_1(k)\}$. Let the result be $Y_{11} = k$. We then sample, independently of each other, k observations from the interval distribution $f_1(x)$. The spike numbers $0, 1, \cdots, k - 1$ are accordingly of state 1 and the spike number k is of state 2. At the time point $t = x_1 + \cdots + x_k$, we consult the distribution $\{p_2(k)\}$ for information on the length of the first run of state 2 spikes (denoted by Y_{21}). Let the result be $Y_{21} = j$. Next we sample j independent observations from the interval distribution $f_2(x)$. At the time point $t = x_1 + \cdots + x_{k+j}$, we fix Y_{12} from $\{p_1(k)\}$ and so on. The runs thus form an alternating sequence Y_{11}, Y_{21}, Y_{12}, Y_{22}, Y_{13}, Y_{23}, $\cdots$. Note that all samplings from the distributions $f_1(x)$, $f_2(x)$, $\{p_1(k)\}$, and $\{p_2(k)\}$ are independent. The part of a realization we have just simulated is illustrated in Fig. 1, with $k = 3, j = 2$.

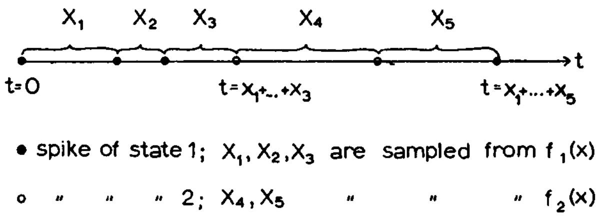

FIGURE 1 The realization described in the Description of the PM Process.

In contrasting the PM model with the semi-Markov and the renewal models, let it be pointed out that in simulating the two state semi-Markov model we would choose the type of each interval separately at the preceding spike. The probability would depend only upon the type of the preceding interval in accordance with the following matrix of transition probabilities

$$\begin{array}{lc|cc|}
 & & \multicolumn{2}{c}{\text{Type of the}} \\
 & & \multicolumn{2}{c}{(k+1)\text{th interval}} \\
 & & 1 & 2 \\ \hline
\text{Type of the} & 1 & \alpha_1 & 1-\alpha_1 \\
k\text{th interval} & 2 & 1-\alpha_2 & \alpha_2 \\ \hline
\end{array}\ .$$

From the mathematical point of view, this is equivalent to specifying that both distributions of runs $\{p_v(k)\}$, where $v = 1, 2$, are geometrical (cf. equation 2). This is to say that formally the two state semi-Markov model is specified by the geometrical form of the distribution of runs. But from the point of view of interpretation, the semi-Markov model is distinguished by the absence of a "memory"; the choice of the type of interval to follow is made anew for each separate interval.

In simulating a renewal process there would be no choice of type of intervals since all intervals are drawn from the same distribution independently of each other.

The Imbedded and the Full Process

The analytical treatment of the process will be built on the observation that mapping the time points of spikes in the full process on the nonnegative integer of their serial number gives us an alternating renewal process (cf. Cox, 1962, p. 80) on a discrete axis. The dimension of this axis is the ordinal number of the spikes, and we shall accordingly refer to it as the ordinal axis. The relationship between the full process and the discrete alternating renewal (DAR) process is shown in Fig. 2.

The DAR is composed of the alternating sequence Y_{11}, Y_{21}, Y_{12}, $\cdots$. For a run of spikes of state $v = 1, 2$, irrespective of the serial number we write Y_v, and we then have for $k = 1, 2, \cdots$,

$$\text{Prob}\,\{Y_v = k\} = p_v(k), \qquad v = 1, 2, \tag{3}$$

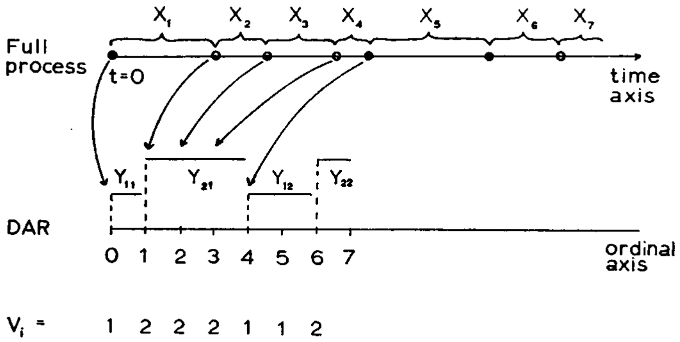

FIGURE 2 The mapping of the full process on the integers.

where

$$0 \leqq p_v(k) \leqq 1, \qquad \sum_{k=1}^{\infty} p_v(k) = 1. \tag{4}$$

We assume that the expectations of the runs are finite and we denote them

$$EY_v = \lambda_v, \qquad v = 1, 2. \tag{5}$$

For the probability-generating functions of the distributions $\{p_v(k)\}$, we write $P_v(s)$, $v = 1, 2$, where $-1 < s < 1$. The tails of the distributions are

$$q_v(k) = p_v(k + 1) + p_v(k + 2) + \cdots, \tag{6}$$

and we note (Feller, 1957, p. 249) that their generating functions are

$$Q_v(s) = (1 - s)^{-1}(1 - P_v(s)). \tag{7}$$

We shall say that the DAR is in state $v = 1, 2$ at point i if the ongoing run is of type v. We include by convention the left, but not the right, end point of each interval with the interval (cf. Fig. 2). We indicate the state of the DAR by defining a sequence of two-valued random variables

$$V_i = \begin{cases} 1 \text{ if the DAR is in state 1 at point } i \\ 2 \text{ if the DAR is in state 2 at point } i \end{cases} \tag{8}$$

where $i = 0, 1, \cdots$.

The sequence $\{X_i ; i = 1, 2, \cdots\}$ of times between events in the full process is constructed by letting the value of V_i determine from which distribution X_{i+1} is sampled. The full process is thus a set of random functions defined on a DAR. The theory of semi-Markov processes could be similarly based on a Markov chain (cf. Cox and Miller, 1965, p. 352).

We want the sequence $\{X_i; i = 1, 2, \cdots\}$ to be stationary (cf. Cox and Lewis, p. 60), and this is reduced to the requirement that the sequence $\{V_i; i = 0, 1, \cdots\}$ is stationary. In the simulation example the DAR was started deterministically in state 1, and all the runs of the same type had the same distributions. We shall refer to that model as the ordinary DAR. For the ordinary DAR the sequence $\{V_i; i = 0, 1, \cdots\}$ is not stationary and we shall be forced to modify the start of the DAR to construct an equilibrium DAR (cf. Cox, 1962, p. 85).

The Marginal and the Transition Probabilities

It is obvious from the specification of the sequence $\{X_i; i = 1, 2, \cdots\}$ that the marginal pdf of the X_i's is a mixture of the pdfs $f_v(x)$, $v = 1, 2$, with the weights dependent upon the marginal probabilities, Prob $\{V_i = v\}$, of the equilibrium DAR.

The serial correlations of the X_i's will depend upon the "transition probabilities," Prob $\{V_{i+k} = w \mid V_i = v\}$, of the equilibrium DAR in a rather straightforward way. This can be seen by the following argument. Write $X_iX_{i+k} \mid (v, w)$ as a shorthand for X_iX_{i+k} under the condition that $V_{i-1} = v$ and that $V_{i+k-1} = w$, $v, w = 1, 2$. It follows from the independence assumption in the construction of the sequence $\{X_i; i = 1, \cdots\}$, that for $k = 1, 2, \cdots$,

$$E(X_iX_{i+k} \mid (v, w)) = \mu_v\mu_w, \qquad v, w = 1, 2. \tag{9}$$

This reduces the problem of finding the serial correlations to that of finding the marginal and the transition probabilities of the equilibrium DAR.

We first define for the ordinary DAR with $k = 0, 1, \cdots$,

$$\pi_{vw}(k) = \text{Prob}\,\{V_k = w \mid V_0 = v\}, \qquad v, w = 1, 2, \tag{10}$$

and denote the corresponding generating functions, $\Pi_{vw}(s)$. We shall find it convenient to first calculate $\pi_{11}(k)$. The DAR is in state 1 at k either because it did not leave state 1 until after k or because a full cycle of two runs is completed at $j \leqq k$, and the DAR is in state 1 at k, having started afresh at j. Formally this gives for $k = 0, 1, \cdots$,

$$\pi_{11}(k) = \text{Prob}\,\{Y_{11} > k\} + \sum_{j=0}^{k} \text{Prob}\,\{Y_{11} + Y_{21} = j\}\pi_{11}(k - j). \tag{11}$$

It follows immediately from Feller's (1957, p. 290) renewal theorem that

$$\Pi_{11}(s) = \frac{Q_1(s)}{1 - P_1(s)P_2(s)}, \tag{12}$$

and further that

$$\lim_{k\to\infty} \pi_{11}(k) = \frac{\lambda_1}{\lambda_1 + \lambda_2}. \tag{13}$$

From the fact that $\pi_{v2}(k) = 1 - \pi_{v1}(k)$, $v = 1, 2$ and the symmetry, it follows that the matrix of the limiting values is for $v, w = 1, 2$

$$[\lim_{k\to\infty} \pi_{vw}(k)] = (\lambda_1 + \lambda_2)^{-1} \begin{bmatrix} \lambda_1 & \lambda_2 \\ \lambda_1 & \lambda_2 \end{bmatrix}. \qquad (14)$$

Once the starting state does not influence the limiting values of the transition probabilities, we call them equilibrium probabilities and denote them simply π_v, $v = 1, 2$, so that

$$\pi_v = \lambda_v(\lambda_1 + \lambda_2)^{-1}. \qquad (15)$$

The equilibrium DAR is constructed by choosing the starting state in accordance with the probability vector (π_1, π_2). Given that the first run is of state $v = 1, 2$, we denote its length Z_v, that is, Z_v takes on the values $1, 2, \cdots$. The distribution of Z_v is denoted $g_v(k)$ and defined for $k = 1, 2, \cdots$ as

$$g_v(k) = \lambda_v^{-1} q_v(k - 1), \qquad v = 1, 2. \qquad (16)$$

That this construction really makes $\{V_i;\ i = 0, 1, \cdots\}$ a stationary sequence is proved at some length in Ekholm (1968) but is fairly evident when one notes that $\{g_v(k)\}$ can be derived from $\{p_v(k)\}$ as the distribution of a "forward recurrence run," (cf. Cox and Lewis, 1966, p. 62).

For the transition probabilities of the equilibrium DAR, we shall write $\pi'_{vw}(k)$ with generating functions $\Pi'_{vw}(s)$. The same kind of logic as was used for equation 11 shows that for $k = 0, 1, \cdots,$

$$\pi'_{11}(k) = \text{Prob}\,\{Z_1 > k\} + \sum_{j=0}^{k} \text{Prob}\,\{Z_1 + Y_{21} = j\}\pi_{11}(k - j). \qquad (17)$$

Note that at the right-hand side we have a transition probability of the ordinary DAR. Turning to generating functions gives, upon using equations 12 and 16,

$$\Pi'_{11}(s) = \frac{1}{1 - s} - \frac{1}{\lambda_1} \cdot \frac{sQ_1(s)Q_2(s)}{1 - P_1(s)P_2(s)}. \qquad (18)$$

For the second term on the right-hand side we shall need a special notation. Let

$$T(s) = \frac{sQ_1(s)Q_2(s)}{1 - P_1(s)P_2(s)}, \qquad (19)$$

and let the sequence with generating function equation 19 be $\{t(k)\}$. Since $(1 - s)^{-1}$ is the generating function of the sequence $1, 1, \cdots$, it is clear from

$$\pi'_{11}(k) + \pi'_{12}(k) = 1$$

that

$$\pi_{12}(s) = \lambda_1^{-1}T(s). \qquad (20)$$

From reasons of symmetry it follows that $\Pi'_{21}(s)$ and $\Pi'_{22}(s)$ are acquired from $\Pi'_{12}(s)$ and $\Pi'_{11}(s)$ on substituting λ_2 for λ_1. This argument gives us the matrix of the transition probabilities in terms of the $t(k)$'s as

$$\pi'_{vw}(k) = \begin{bmatrix} 1 - \lambda_1^{-1}t(k) & \lambda_1^{-1}t(k) \\ \lambda_2^{-1}t(k) & 1 - \lambda_2^{-1}t(k) \end{bmatrix}$$

$$k = 0, 1, \cdots, \qquad v, w = 1, 2. \qquad (21)$$

Note that from equation 19, on putting $s = 0$, we have $t(0) = 0$.

For the unconditional probabilities of the equilibrium DAR we can write

$$\text{Prob}\,\{V_i = v,\ V_{i+k} = w\} = \text{Prob}\,\{V_i = v\}\ \text{Prob}\,\{V_{i+k} = w \mid V_i = v\}$$

$$= \pi_v \pi'_{vw}(k), \qquad (22)$$

where the last equation is justified by the stationarity of the sequence $\{V_i\,;\, i = 0, 1, \cdots\}$. The matrix of the unconditional probabilities is then, by equation 15 and 21, for $k = 1, 2, \cdots,$

$$[\text{Prob}\,\{V_i = v,\ V_{i+k} = w\}] = (\lambda_1 + \lambda_2)^{-1}\begin{bmatrix} \lambda_1 - t(k) & t(k) \\ t(k) & \lambda_2 - t(k) \end{bmatrix} \qquad (23)$$

Marginal Distribution and Serial Correlations

The definition of the equilibrium DAR implies that the pdf of the marginal distribution of the X_i's is

$$f(x) = \pi_1 f_1(x) + \pi_2 f_2(x), \qquad (24)$$

and the mean and variance of the marginal distribution are then

$$\mu = \pi_1\mu_1 + \pi_2\mu_2 \qquad (25)$$

$$\sigma^2 = \pi_1\sigma_1^2 + \pi_2\sigma_2^2 + \pi_1\pi_2(\mu_1 - \mu_2)^2.$$

The product moments $E(X_iX_{i+k})$ are obtained by weighting the product moments of equation 9 by the probabilities of equation 23. When we further subtract the square of the marginal mean and divide by the marginal variance, the serial correlations are found to be, for $k = 1, 2, \cdots,$

$$\rho(k) = (\mu_1 - \mu_2)^2\pi_1\pi_2\sigma^{-2}(1 - [\lambda_1 + \lambda_2][\lambda_1\lambda_2]^{-1}t(k)). \qquad (26)$$

In principle, equation 19 gives a full account of the relationship between the sequence $\{t(k)\}$ and the run distributions $\{p_\nu(k)\}$, $\nu = 1, 2$. To expand $T(s)$ in a power series is, however, difficult in most cases of practical interest, and we shall therefore derive a more direct way of computing $\{t(k)\}$ from $\{p_\nu(k)\}$, $\nu = 1, 2$. Let, for $k = 0, 1, \cdots,$

$$\begin{aligned} p(k) &= p_1(k) * p_2(k) \\ q(k) &= q_1(k) * q_2(k) \end{aligned} \tag{27}$$

where the star denotes the convolution operation. Note that from $p_1(0) = p_2(0) = 0$ follows $p(0) = p(1) = 0$. The generating functions $P(s)$ and $Q(s)$ of $\{p(k)\}$ and $\{q(k)\}$, respectively, are

$$\begin{aligned} P(s) &= P_1(s)P_2(s) \\ Q(s) &= Q_1(s)Q_2(s). \end{aligned} \tag{28}$$

We then define a new sequence $\{r(k)\}$ for $k = 0, 1, \cdots$ as

$$r(k) = q(k) + p(2)r(k-2) + p(3)r(k-3) + \cdots + p(k)r(0). \tag{29}$$

Since $p(0) = p(1) = 0$, $r(k)$ is uniquely defined and equation 29 is a recursive relation. Equation 29 is, further, of the form implied by Feller's (1957, p. 290) renewal theorem, and writing $R(s)$ for the generating function of $\{r(k)\}$, we thus have

$$R(s) = \frac{Q_1(s)Q_2(s)}{1 - P_1(s)P_2(s)} \tag{30}$$

and

$$\lim_{k\to\infty} r(k) = \frac{\lambda_1\lambda_2}{\lambda_1 + \lambda_2}. \tag{31}$$

Comparing equation 30 with equation 19, which was obtained on probabilistic grounds for $T(s)$, and remembering that $t(0) = 0$, we find that

$$r(k) = t(k+1). \tag{32}$$

Equations 26, 27, 29, and 32 provide an easy and powerful way of computing the serial correlations from numerically specified distributions $\{p_\nu(k)\}$, $\nu = 1, 2$. From equations 31 and 32 it follows that $\lim_{k\to\infty} \rho(k) = 0$.

To sum up these results, we note that the marginal pdf is a mixture of $f_1(x)$ and $f_2(x)$, with the weights proportional to the means of the respective runs. The serial correlations depend upon the distributions $f_1(x)$ and $f_2(x)$ only through their means and variances, and the expression dependent upon these serves, in fact, as a scaling factor. The nonconstant factor of the serial correlations depends in a rather complicated way upon the distributions $\{p_\nu(k)\}$, $\nu = 1, 2$. Only for the lag 1 is a simple result available. For all admissible choices of the distributions $\{p_v(k)\}$, we will have $t(1) = 1$. Thus, $\rho(1)$ depends only on the means λ_v, $v = 1, 2$ of the runs and $\rho(1) > 0$ for $\lambda_1\lambda_2 - \lambda_1 - \lambda_2 > 0$. Since, in any case, λ_1, $\lambda_2 \geqq 1$, the serial correlation of lag 1 is positive except for points in the (λ_1, λ_2) plane which fall on the hyperbola $\lambda_1\lambda_2 - \lambda_1 - \lambda_2 = 0$ in which case $\rho(1) = 0$, or between this hyperbola and its asymptotes $\lambda_1 = 1$, $\lambda_2 = 1$ in which case $\rho(1) < 0$.

For higher lags more information about the $\{p_v(k)\}$ will be involved but in no very obvious way. Examples can, however, easily be constructed and computed according to the method developed above. A few examples of correlograms for different distributions of runs are given in connection with the empirical results.

The second-order properties of the process of counts under the PM hypothesis are derived in Ekholm (1968). The Laplace transform of the intensity function is obtained but is difficult to invert. By numerical methods, however, the spectrum of counts can be reached. The most important general result is that the second-order properties of counts are more complex than those of the intervals. The former involve the full information about the distributions $f_1(x)$ and $f_2(x)$ and the distributions $\{p_v(k)\}$, $v = 1, 2$. We shall not penetrate into these questions since in the empirical work we will compute the theoretical intensity function by a simple shuffling method.

EMPIRICAL RESULTS

Experimental Method and Data

The most important single decision to be made when applying the PM model to empirical spike trains is whether to use an implicit or an explicit way of analysis. We call the analysis explicit when each impulse interval is classified as either belonging to the distribution with pdf $f_1(x)$ or to the distribution with pdf $f_2(x)$. An implicit analysis would not try to discover the type of the individual intervals, but would instead use mathematically specified hypotheses about the forms of the distributions $f_v(x)$ and $\{p_v(k)\}$, $v = 1, 2$.

We have chosen the explicit way of analysis since in the recordings at our disposal (Hyvärinen, 1966) a decent number of bimodal histograms could be found. Although it would not be illogical to use the idea of two different states when the states are not physically discernible, it seems more relevant to use this idea when the states are fairly differentiated. At least, it seems sensible to start the application of the PM model with an explicit analysis to find what kind of distributions there can occur for the runs.

We selected 17 spike trains with bimodal histograms which did not fit the renewal model in respect to either their serial correlograms or their intensity functions.

The spike trains were chosen from the recordings of spontaneous spike potential activity in developing brains (Hyvärinen, 1966). The recordings were made with glass capillary microelectrodes in different diencephalic nuclei under light urethane anaesthesia in rabbits ranging in age from newborn to adult. The interspike intervals were measured with an ordinary slow general purpose computer with a resolution of 2.16 msec (Halme and Hyvärinen, 1968).

Statistical Methods and Results

Stationarity. To ascertain the stationarity of the spike trains, we used first a photographic dot display of the intervals. This was supplemented by a one-way analysis of variance of the impulse intervals (Rao, 1965, p. 202). For this purpose we divided the intervals into groups of 50, or 20 for cases where the total number of observed spikes was less than 500. Here and in all other significance tests reported in this paper we use the 5% significance level.

The analysis of variance does not efficiently detect mild but consistent trends in the means of the consecutive groups. To guard against this possible source of error, we plotted the means against their serial numbers. A number of series were discarded because of unstationarity and some were cut shorter. The number (n) of observed impulse intervals of the 17 series left are reported in Table I.

Interval Histograms. In the construction of the marginal histograms, we varied the grouping interval, (τ), depending upon the shape of the histogram and the number of observed intervals. The histograms were then scrutinized carefully to determine the cutting point between the two hypothesized distributions. We did not in these selected cases pay attention to the possible overlap of the two pdfs $f_\nu(x)$, $\nu = 1,2$.

In Fig. 3 four histograms are seen with the cutting points (c) indicated. The histogram gives an empirical estimate of the pdf $f(x)$ in the sense that $f(k\tau + \frac{1}{2}\tau)\cdot\tau$ is the probability of an interval of length between $k\tau$ and $(k + 1)\tau$ and the histogram reports these probabilities. To characterize the appearance of the histograms not reproduced here we report in Table I some basic figures about the units studied. The mean, (μ), and the coefficient of variation, (σ/μ), of the intervals are commonly used, and we report these though they are somewhat out of place in the case of bimodal histograms. More to the point are the means, (μ_1 and μ_2), of the short and long intervals computed separately. We further report σ_1/μ_1 and $\sigma_2/(\mu_2 - c)$. The coefficient of variation σ_1/μ_1 gives a rough measure of the departure of the pdf $f_1(x)$ from the exponential one. For three units (94109, 53402, and 60105) $f_1(x)$ is close to the exponential according to this criterion. For the distribution of long intervals it is more relevant to compute

TABLE I

BASIC DATA ABOUT THE HISTOGRAMS

Unit	n	μ	σ/μ	μ_1	σ_1/μ_1	μ_2	$\sigma_2/(\mu_2-c)$	d
		msec		*msec*		*msec*		
49101	700	46	0.96	29	0.30	148	0.49	0.86
53402	710	229	1.61	57	1.09	968	0.41	0.94
58201	919	392	1.17	184	0.77	1230	0.60	0.84
62501	397	492	1.12	3	0.27	888	0.70	0.64
64201	485	189	1.32	37	1.37	524	0.68	0.81
77601	500	430	0.97	3	0.30	830	0.51	0.98
81102	1421	252	1.58	76	0.82	1070	0.43	0.93
91301	508	302	1.78	155	0.53	2172	0.43	0.92
46403	393	94	0.53	35	0.29	114	0.66	0.49
88301	341	838	0.69	93	0.62	1109	0.49	0.62
89102	421	766	1.18	60	1.19	1623	0.62	0.74
53301	601	131	0.90	36	0.52	234	0.61	0.71
60105	845	1027	0.90	18	1.09	1498	0.57	0.55
94109	259	851	0.98	41	0.92	1163	0.78	0.36
77105	225	375	1.40	141	0.79	1426	0.56	0.89
81105	348	194	2.23	82	0.72	1229	0.86	0.63
88302	309	589	1.74	121	0.84	2530	0.48	0.88

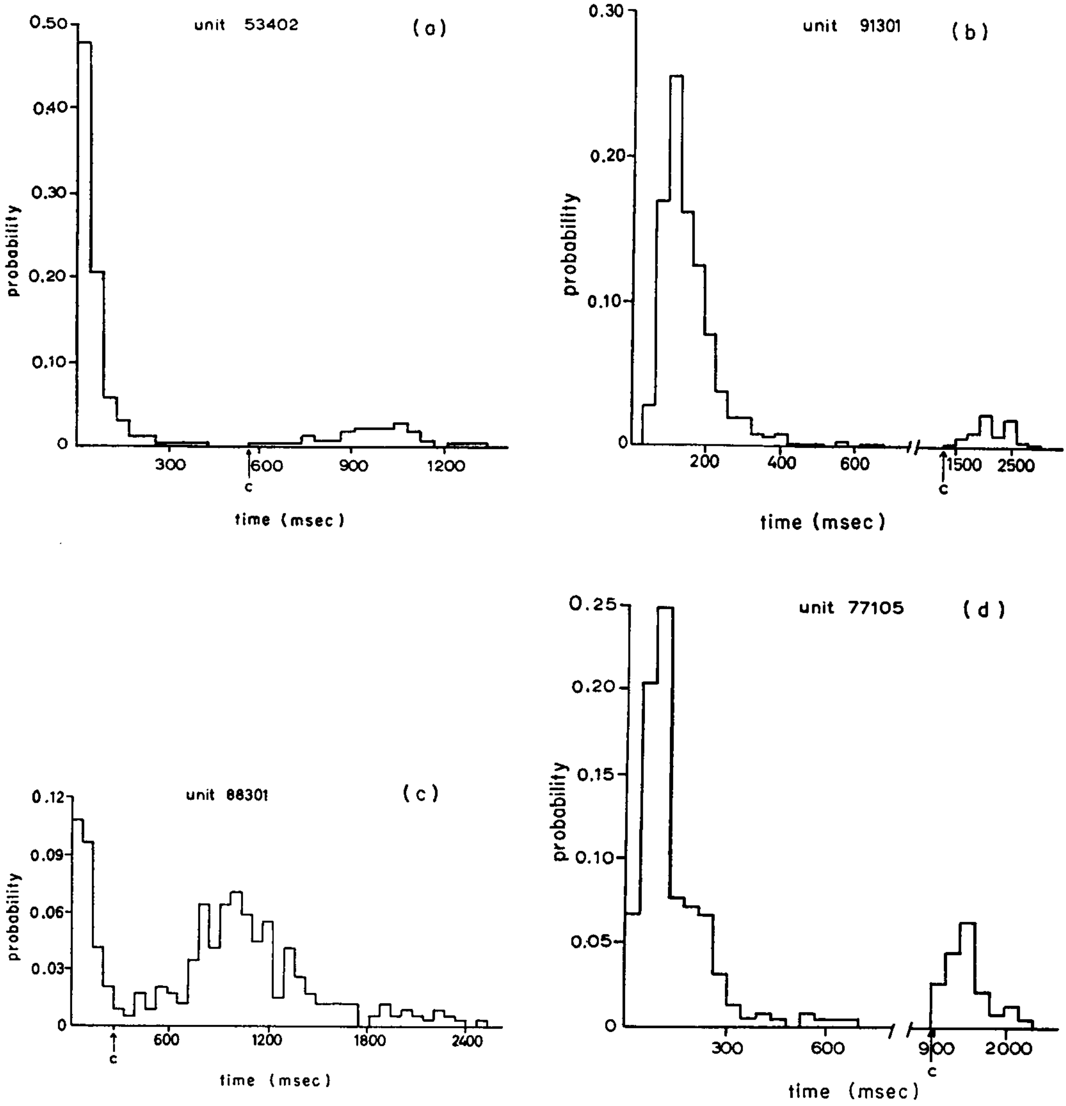

FIGURE 3 The marginal histogram of four units.

$\sigma_2/(\mu_2 - c)$ than σ_2/μ_2. The former expression would acquire the value 1 for the delayed exponential pdf $\theta \exp(-\theta(x - c))$. Table I shows, that for all 17 units the distribution of long intervals is more concentrated than the delayed exponential pdf would be.

A measure of the separation between the two hypothesized distributions is provided by $d = (\mu_1 - \mu_2)^2\pi_1\pi_2\sigma^{-2}$. This is the constant factor of equation 26 for the serial correlations. Note that $0 < d < 1$, and that d approaches its lower bound when $\mu_1 - \mu_2$ approaches zero, and its upper bound when either $\mu_1 - \mu_2$ approaches infinity and/or σ_1 and σ_2 approach zero. The numerical values of d are reported in Table I.

The Distribution of Runs. Having determined the cutting point, we counted the number

TABLE II

BASIC DATA ABOUT THE DISTRIBUTIONS OF RUNS

Unit	λ_1	π_1	G = geomet.	mode = k	$p(k)$	λ_2	π_2
49101	5.9	0.86		5	0.29	1.0	0.14
53402	4.4	0.81		4	0.34	1.0	0.19
58201	4.5	0.80		3	0.17	1.1	0.20
62501	1.3	0.45		1	0.69	1.6	0.55
64201	2.4	0.69		2	0.40	1.1	0.31
77601	1.2	0.48		1	0.82	1.3	0.52
81102	4.9	0.82		3	0.19	1.1	0.18
91301	12.4	0.93		12	0.24	1.0	0.07
46403	1.2	0.26	G	1	0.86	3.4	0.74
88301	1.1	0.28	G	1	0.94	2.9	0.72
89102	1.8	0.55	G	1	0.46	1.5	0.45
53301	1.7	0.52	G	1	0.51	1.6	0.48
60105	1.3	0.31	G	1	0.72	2.9	0.69
94109	1.3	0.28	G	1	0.77	3.3	0.72
77105	4.6	0.82		3	0.25	1.0	0.18
81105	14.2	0.90		3	0.18	1.5	0.10
88302	4.5	0.80		4	0.28	1.1	0.20

of spikes in each run and tabulated the distributions of runs. We did not include the first and the last run since they are not observed in full. The means of the runs λ_1 and λ_2 and the equilibrium probabilities π_v, $v = 1, 2$ calculated according to equation 15 are given in Table II.

We tested the observed distributions of runs for deviation from the best fitting geometrical distribution by the ordinary χ^2 test (Rao, 1965, p. 325) using the maximum-likelihood estimator for the parameter. For the distributions of runs of long intervals the result is uniform. They all fit well with the hypothesis of a geometrical distribution. In some cases the number of degrees of freedom did not suffice for a formal test but there, too, the fit by eye was good.

Among the distributions of runs of short intervals, six units fit the geometrical distribution. These units are marked by a G in Table II. Two units (62501, 77601) have a distribution of runs close to the geometrical distribution but differ significantly from it by the absence of the tail. Unit 81105 has too long a tail but does have a geometric appearance.

The remaining distributions of runs of short intervals differ markedly from the geometrical. They are all more or less symmetrical though not always with a single pronounced peak. Three examples are seen in Fig. 4. We have not tried to fit any other family of distributions to the 11 distributions that fail to be geometrical. To depict the peakedness of the distributions, we indicate in Table II which integer (k) has the greatest probability and the numerical value of this maximum probability ($p(k)$).

The Serial Correlations. The serial correlation coefficients of the first 50 lags were estimated using the estimator given as (5.2.17) in Cox and Lewis, (1966, p. 91). The first 700 observations were used for this purpose. From the estimates of $\{p_v(k)\}$, μ_v, and σ^2 the corresponding serial correlations of the PM model were computed according to equation 26. In Fig. 5 the empirical and model correlograms are given for the four example units. It should be remembered here that for the renewal model the serial correlations of all lags are zero.

Unfortunately, there is no satisfactory statistical test for the deviation of empirical correlo-

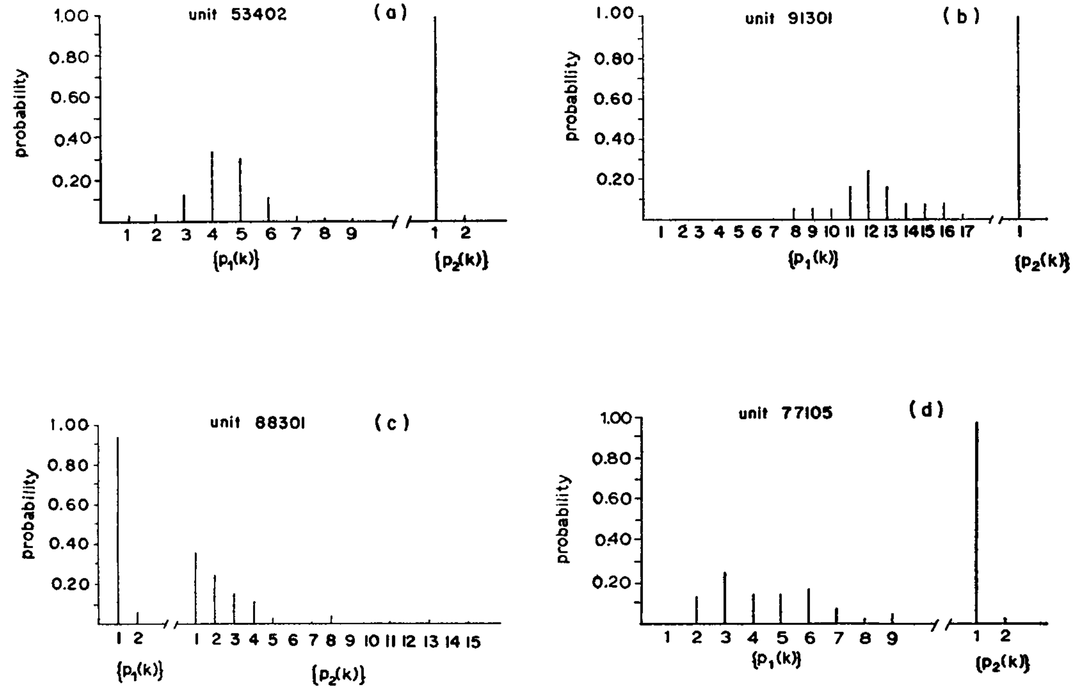

FIGURE 4 The distributions of runs of short $\{p_1(k)\}$ and long $\{p_2(k)\}$ intervals of four units.

grams from the correlograms of some postulated model (cf. Perkel, Gerstein, and Moore, 1967 *a*, p. 401). From the mathematical point of view, the information inherent in the correlogram can, however, be expressed also by the spectrum of the sequence $\{X_i\ ;\ i = 1, 2, \cdots\}$ of intervals (Cox and Lewis, 1966, p. 71). The spectrum of intervals of an observed series can be estimated, and there are tests for the deviation of an estimated spectrum from the theoretical spectrum of a postulated model. On the other hand, the spectrum of intervals is relatively inaccessible to interpretation. In short, the correlogram has a clear-cut interpretation but lacks a useful statistical theory and the opposite, more or less, is true of the spectrum of intervals. For this reason we resorted to the spectrum of intervals for a test of the deviation of the observed series from the PM model with regard to the second-order properties of intervals. The power spectrum of the empirical series was estimated by the periodogram estimator (cf. Cox and Lewis, 1966, Sect. 5.3). For this purpose only the 301 first observations were used. To use the full set of observations at this point would have led to rather formidable computations.

The spectral density of the PM model was evaluated by the basic formula (Cox and Lewis, 1966, formula 4.4.8, p. 71) using the 50 first model correlation coefficients. The convergence of the model correlations toward zero was slowest for unit 91301 (see Fig. 5). For most units the model correlations were already zero up to the fourth or fifth decimal at lags between 40 and 50. Thus it is obvious that the use of only 50 coefficients in calculating the spectral density did not introduce any error.

The test of the model followed the recommendations given by Cox and Lewis, (1966, Sect. 6.5 [*i*] and 6.4. [*ii*]), and was accordingly based on the ratio of the periodogram of the data to the spectral density of the model. We tried the appropriate χ^2 test with ten groups of size 15

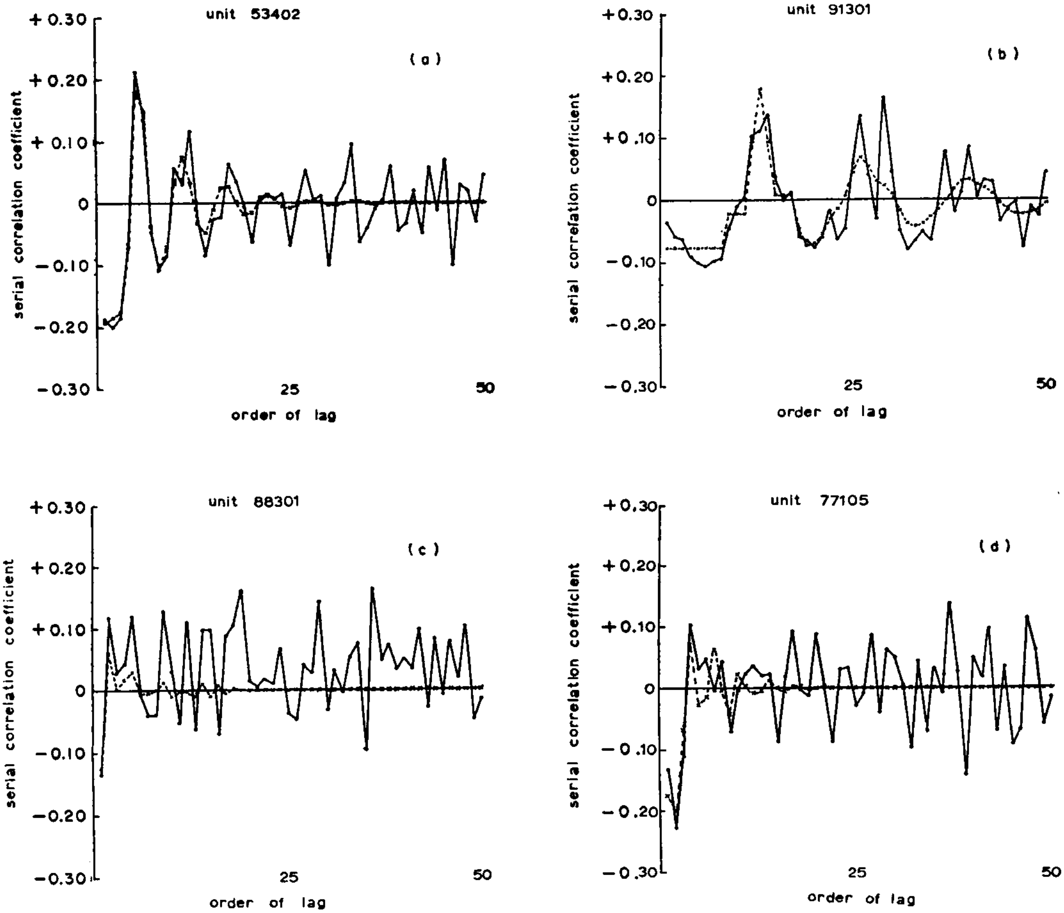

FIGURE 5 The serial correlation coefficients of the first 50 lags for four units. The dots connected by solid lines indicate the empirical estimates and the crosses connected by broken lines indicate the coefficients of the fitted PM model.

and six groups of size 25 and finally we tried the Kolmogorov-Smirnov test of the quantities (6.4.5.). All three procedures led to exactly the same conclusions at the 5% significance level.

To gauge not only the fit of our model but also its fit compared to the basic renewal model, we performed the Kolmogorov-Smirnov test of the empirical periodogram using the renewal model as a hypothesis. The results of these tests are reported in Table III in the following manner. If the PM model does not fit, this is marked by a minus sign; if the PM model fits and the renewal model does so, too, this is indicated by a zero; and if the PM model fits but the renewal model does not, a plus sign is used.

It was noted that the test for the deviation from the renewal model does not have good power in cases where the absolute values of the serial correlations are small (around 0.10), though the consistent pattern of the coefficients shows that the renewal hypothesis cannot be correct. Two such cases are units 81105 and 91301; the correlogram of the latter is given in Fig. 5. It is obvious from Fig. 5 *b* that the PM model is a much better explanation of the

TABLE III

THE FIT OF THE PM MODEL COMPARED TO THAT OF THE SIMPLE RENEWAL MODEL

Unit	Correlogram fit	Intensity fct fit	Combined fit
49101	+	+	+
53402	+	0	+
58201	+	+	+
62501	+	0	+
64201	+	+	+
77601	+	+	+
81102	+	+	+
91301	0	+	+
46403	+	0	+
88301	+	0	+
89102	+	0	+
53301	−	−	−
60105	−	0	−
94109	−	0	−
77105	+	−	−
81105	0	−	−
88302	+	−	−

empirical correlogram than the flat correlogram of the renewal model. The test based on the spectral theory does not work here, evidently because the amplitude of the correlogram is not very large. These two cases (81105 and 91301) are, however, the only ones where the test gives counter-intuitive results.

The Intensity Function. The intensity function of the data was estimated by the procedure developed in Cox, (1965), see also Cox and Lewis, (1966, Sect. 5.4. [v]). Here, again, we used only the 300 first observations to avoid excessive computation. A corresponding intensity function of the PM model was calculated by the following simulation method. Using the cutting point we classified each of the 300 first impulse intervals as short or long. The serial numbers of the short and long intervals were noted, and the short and long interval observations were then shuffled separately at random. A shuffled series was constructed by replacing each short interval with a short one and each long interval with a long one. The intensity function of the shuffled series was estimated in the same way as that of the empirical series.

The intensity function of the corresponding renewal model was calculated by using the integral equation of renewal theory (Cox, 1962, p. 54) as a starting point. Assuming that both the pdf of the marginal distribution $f(x)$ and the renewal density $h_R(x)$ are constant for the grouping intervals $(0, \tau)$, $(\tau, 2\tau)$, a recursive formula for $h_R(k\tau + \frac{1}{2}\tau)$, $k = 0, 1, \cdots$, can be derived. The input of this computation is then $f(k\tau + \frac{1}{2}\tau)$, $k = 0, 1, \cdots$, and for these we used the relative frequencies, divided by τ, of the marginal histogram.

In Fig. 6 the empirical intensity functions and the corresponding intensity functions under the PM and renewal hypotheses of the four example units are shown. The values of the intensity functions relate to the corresponding conditional probabilities (equation 1) in the

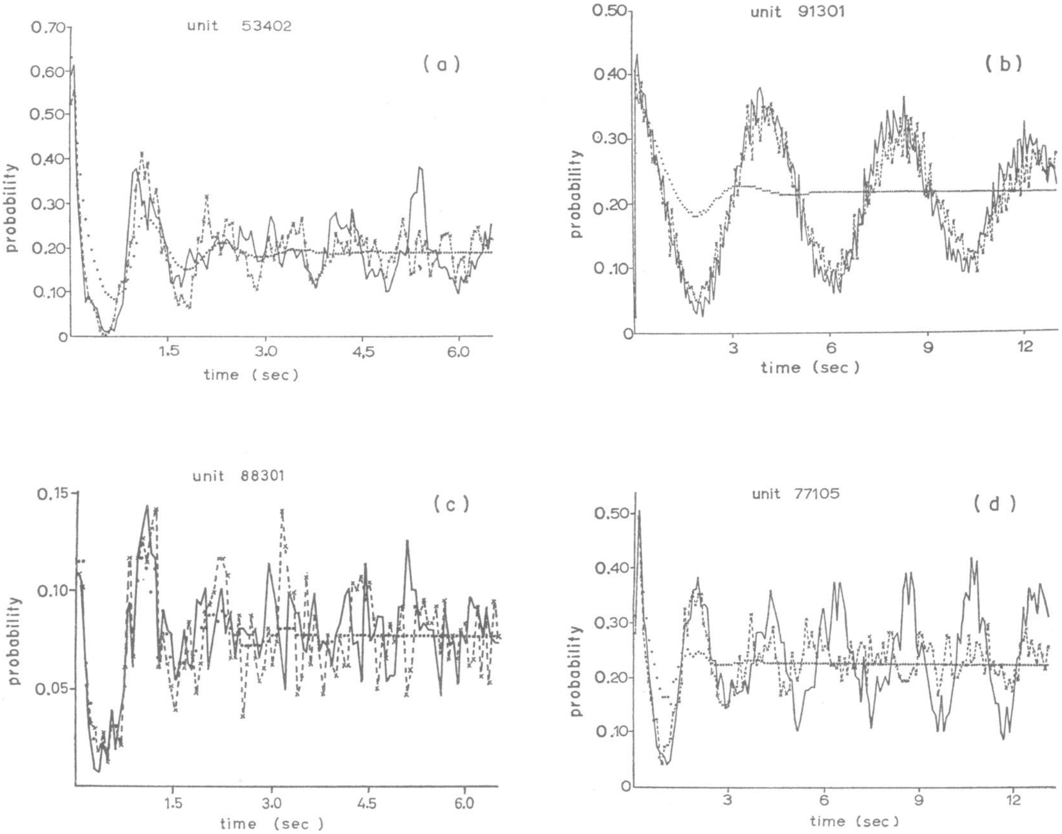

FIGURE 6 Intensity functions and renewal densities for four units. The solid line indicates the empirical estimate of the intensity function, the broken line indicates the intensity function of the fitted PM model, and the dotted line indicates the renewal density of a renewal process with the same marginal histogram.

same way as the histogram relates to the probabilities of interval length, cf. p. 784. A formal test of the goodness of fit would again call for turning to spectral theory, and the numerical work is heavy. Instead we have appraised the fit by eye, using the same three categories as for the serial correlations: the PM model fits better than the renewal model (+), both models fit (0), and the PM model does not fit (−). The classification was rather obvious; this is illustrated by the examples in Fig. 6. The full result is given in column 2 of Table III.

It is obvious that in cases where $f_1(x)$ is close to the exponential distribution, the intensity function converges rapidly to a constant value, so that not much difference can be found between the renewal and the PM hypotheses. A good example is unit 53402, for which it is clear from the correlogram test that the PM model is a much better description of the structure of the series than the renewal model. In this case the intensity function, however, gives no resolution. Note that $\sigma_1/\mu_1 = 1.09$ (see also Figs. 5 and 6). This remark, combined with the remark about the power of the periodogram test (p. 788), illustrates the point that the statistical analysis of spike trains should never be restricted to only the process of counts or to only the sequence of intervals.

The final column in Table III gives the combined result of the fit of the PM model. Units where either the serial correlations or the intensity function or both do not fit, are marked by a minus. Units where the fit in either aspect is better than for the renewal model and which are not unfit in either aspect are marked by a plus.

Before we turn to the conclusions, the following question should be posed and answered. What features of the model do we actually test? In the derivation of equation 26 for the serial correlations, the two necessary assumptions are that

$$E(X_i X_{i+k} \mid (v, w)) = \mu_v \mu_w, \qquad v, w = 1, 2 \qquad (1^\circ)$$

the discrete process is an alternating renewal process. (2°)

The discrete process means, of course, the series of 1's and 2's indicating from which distribution the intervals are drawn (cf. equation 8 and Fig. 2). That this process is an alternating renewal model implies, more particularly, that the consecutive runs of 1's and 2's are stochastically independent. A simple counter example would be that a short run of 1's tends to be followed by a long run of 2's.

Calculating the intensity function of the model by the shuffling technique means that we take the observed pattern of 1's and 2's as given and assume that the lengths of the intervals depend only upon the state of the process, that is

$$\text{Prob}\,\{X_i \leqq x,\ X_{i+k} \leqq y \mid (v, w)\} = F_v(x) F_w(y). \qquad (3^\circ)$$

From 3° follows 1°, but the contrary is not true; accordingly 3° is a sharper assumption. It is an advantage of the explicit way of analysis (cf. in Experimental Method and Data) that the testing of assumption 3° can be made without interference from the possible fitness or unfitness of assumption 2°. It is thus clear that if the test based on the serial correlations shows significant deviation from the PM model, then either 1° or 2° or both are not fulfilled. If the empirical intensity function deviates from the intensity function of the model, then 3° is violated. The common feature of 1°–3° is that various quantities are assumed to be stochastically independent.

Conclusions

The third column of Table III shows that the PM model fits for 11 out of the 17 nonrenewal spike trains studied. Because of the arbitrary selection of the studied spike trains, the over-all result is of little interest aside from the fact that empirical series can be found for which the PM model is an appropriate statistical description of their structure. It is more revealing to find homogeneous subgroups both in the group of fits and in the group of misfits.

Let us first note that for all units the observed distribution of runs of long intervals is geometrical. The means (λ_2) of these distributions are with few exceptions close to 1, that is, the stay in the resting state seldom lasts longer than for one interval. After a rest of one long interval, the probability of starting a new burst is high. The interpretation of the geometrical distribution of runs of long intervals is, further, that this probability is not increased by resting for another interval or for any number of intervals. After a rest, whether the first or the second or the *n*th, the probability of changing to bursting is constant.

The most conclusive empirical finding is contained in units 49101, 53402, 58201, 62501, 64201, 77601, 81102, and 91301. These eight series are characterized by the facts that the PM model fits and that the distribution of runs of short intervals differs significantly from the geometrical distribution. These neurons function in the bursting state in a way that is clearly opposed to their function in the resting state. It seems as if the neuron possesses a counter which allows it to fire with short intervals for a predetermined number of times and then to turn to a rest. The counter works with a certain tolerance; the number of allowed short intervals varies to some extent. The preferred number of spikes in the bursting state varies between units in our material from 1 (unit 77601) to 12 (unit 91301). The degree of peakedness of the distributions of runs of short intervals also varies considerably. Unit 77601 has the highest concentration and the probability of only one short interval is 0.80; unit 58201 has the lowest peak and the probability of three short intervals is 0.17, but there is almost the same probability for two, four, and five intervals.

For the above mentioned eight pulse trains, the superior fit of the PM model over the renewal model is well documented. They have pronounced intensity functions and/or serial correlograms which the PM model is able to reproduce, (cf. Figs. 5 and 6). For these series, then, the assumptions 1°–3° seem to be well fulfilled.

The second group is comprised of units 46403, 88301, and 89102. These units are characterized by the fact that the PM model fits and the distributions of runs of both long and short intervals are geometrical. These series are the closest we have come to finding empirical examples of the semi-Markov model. The interpretation of the semi-Markov model is that both for the bursting and the resting state the mechanism of switching to the opposite state lacks a memory.

These three series are, however, not entirely unambiguous. Their intensity functions fit the renewal hypothesis and they differ, in fact, from the renewal model only in that their serial correlations of the first few lags are not zero. Unit 88301 is given as an example in Figs. 5 and 6. We are not convinced that these series could not be described equally well by some other model. The gain of the PM analysis performed lies, in these cases, mainly in giving a hint of the source of the correlations and less in giving a very definite description of the function. Under the renewal hypothesis the probability of a long interval succeeding a short one would be the same as the probability of a long interval succeeding a long one. For these three cases the serial correlations arise from the fact that the former probability is somewhat larger.

We do not want to stress the fact that both distributions of runs are found to be geometrical. That is, if one chooses a cutting point that is not related to the functioning of the cell, or that is related to the functioning but causes a large proportion of misclassifications of the intervals, then one could expect the observed distributions of runs to be approximately geometrical. For instance, if for some of the series in the first group we had mistakenly placed the cutting point in the middle of the distribution of short intervals, the distribution of both runs would have been approximately geometrical. This does, however, not invalidate the result that, given this cutting point, the series fulfill assumptions 1° and 2° and this combined with the aforementioned asymmetry of long and short intervals explains the serial correlations observed.

The first group of misfits is composed of units 53301, 60105, and 94109, characterized by the fact that the serial correlograms of the PM model do not fit the empirical correlograms. The empirical correlograms of these series have much the same appearance as those of the second group of fits. The observed distributions of runs of short intervals are also geometrical. Here, however, assumptions 1° or 2° or both are not fulfilled. For unit 53301 it is clear that 3° is not fulfilled, either, but for the other two units the intensity function is inconclusive at this point. The correct conclusion seems to be that the cutting point is totally unrelated to the functional patterns of these units. The analysis performed does not even detect the source of the serial correlations.

Units 77105, 81105, and 88302, which make up the second group of misfits, are more interesting. These series are characterized by the fact that their observed serial correlations differ markedly from zero for a number of lags. These correlograms are, further, well reproduced by the PM model. The observed distributions of runs of short intervals are not geometrical; those of units 77105 (see Fig. 4) and 88302, especially, are markedly peaked. The empirical intensity functions of these units are clearly periodic, but here the PM model fails definitely. Our conclusion is that for these units the cutting point is relevantly related to the functional pattern and assumptions 1° and 2° are fulfilled but 3° is not, that is, the lengths of the intervals are not dependent only on the state of the neuron. Superimposed on the dependence upon the state of the neuron, there seems to be a finer structure of dependence which we have not been able to identify.

DISCUSSION

We have found that for some spike trains the PM model is an appropriate statistical description. It was noted already in the section on Basic Mathematical Concepts that, if the PM model fits a spike train then the data are summarized by the estimate of the two pdfs $f_\nu(x)$, $\nu = 1, 2$ and the two discrete distributions $\{p_\nu(k)\}$. the description of the spike trains that fit the PM model is thus given by the histograms (Fig. 3), the distributions of runs (Fig. 4), and the assumptions of the model (Description of the PM Process). Though this is a more concise description than the full set of data it is still not a very sharp characterization and it is, further, clear that the assumptions we test by the present approach (cf. the last part of Statistical Methods and Results) are not very informative from the physiological point of view. This makes us question how the present approach could be sharpened.

It seems relevant here to start from the distinction between an explicit and an implicit statistical analysis (cf. Experimental Method and Data). We chose the explicit way for the study presented here in hope of finding some salient universal features in the distributions of runs. The only such result was the geometric form of all the distributions of runs of long intervals. This is not a very well-established result because of the general rarity of long runs of long intervals. Once the explicit analysis has not succeeded very well, we should contemplate the implicit possibility.

The task is then to set forth hypotheses about the mathematical form of both the distributions of intervals $f_\nu(x)$, and the distributions of runs $\{p_\nu(k)\}$, $\nu = 1, 2$. These hypotheses should, of course, be based on physiological considerations of the mechanism of spike production. An implicit statistical analysis based on such hypotheses would not need the questionable device of a cutting point. It would also be considerably stronger than the explicit analysis in the sense that the possibilities of the data to fail the PM model are greater, but the gain in information, in cases where the PM model still fits, is sharper. It is, however, no easy task to derive hypotheses about the distributions. We have no definite results to report here. The following two approaches based on some earlier work could be considered.

Thomas (1966, pp. 154–162), starting out from some cell properties discussed by Burns (1955), proposes the following model of the clustered firing of neurons. There is a Poisson process of main spikes. These spikes are evoked by conditions external to the cell. Each main spike may be followed by a subsidiary spike which is then evoked by conditions internal to the cell. The probability for a subsidiary spike to follow a main spike is constant for all main spikes. Each subsidiary spike may, further, be followed by another subsidiary spike, which again is evoked by conditions internal to the cell. The probability for a subsidiary spike to follow a subsidiary spike is also constant through all subsidiary spikes, but different from the probability of a subsidiary spike to follow a main spike. If a subsidiary process is operative at the time of a main spike, it is then terminated and a new one may start, as after every main spike. The intervals between subsidiary spikes are drawn from a distribution which can have any pdf. For mathematical convenience Thomas assumes that this distribution is exponential in most connections.

From the statistical point of view Thomas's model can be described as a two state semi-Markov model with two types of intervals. (This is obvious from his Sect. 3.4, though not explicitly stated.) His model thus predicts that the serial correlations decrease exponentially in absolute value with the increasing order of the lag. In Hyvärinen's (1966) large material (cf. p. 774) very few such correlograms were found. Accordingly it seems to us that Thomas's model is not very realistic. The basic thinking of Thomas could, however, be combined with the more flexible PM model by abandoning the assumption that the process of main spikes is Poissonian. This assumption is, in fact, rather awkward. It is difficult to find a good reason why the process which reacts to outside stimulation (Thomas, 1966, p. 160) should be Poissonian. The Poisson assumption was obviously introduced by Thomas on mathematical grounds only. In the frame of the PM model it could be substituted by assuming that the pdf of the intervals in the "resting" state, that is $f_2(x)$, has any form whatsoever.

Thomas assumes that the probability that the cell turns to bursting is contant after every main spike. In the PM terminology this is equivalent to the runs of long intervals having a geometric distribution, that is, $\{p_2(k)\}$ is geometric. This is in good agreement with our empirical results and we think this assumption should be retained. To avoid ending up with the semi-Markov model, the distribution of runs of short intervals, that is $\{p_1(k)\}$, should not be assumed geometric but must have a memory. One should, accordingly, not assume that if a subsidiary process comes into operation that the probability of its ending is constant after every subsidiary spike. No other clues to its form are available. Neither do we have any clues to the form of the pdf $f_1(x)$ of intervals in the bursting state. There are thus many gaps to be filled in but it is an advantage of the PM scheme that they could be filled in on physiological grounds.

Many models for the mechanism of spike production share the common feature that, immediately after a spike, the membrane potential is supposed to return to a resting level. Through the effect of incoming excitatory and inhibitory postsynaptic potentials, the electrical state of the membrane is then changed. If the depolarization of the membrane reaches a threshold, the neuron fires. Many variants of this basic scheme are possible (cf. Stein [1967] and for a review up to 1965, see Moore, Perkel, and Segundo [1966]).

The assumption that the membrane potential returns to the resting level after each spike, is crucial from the statistical point of view. If every feature of the mechanism repeats itself after each spike in accordance with the same probabilistic rules, the spike train has the statistical properties of a renewal process. It is thus clear that to build a model of spike production which could produce spike trains having the statistical characteristics of a PM model, at least some of the assumptions inherent in the above scheme have to be abandoned. A natural way to do this is to retain the assumption that the membrane potential returns to the resting level after each spike, but to introduce the assumption that the ratio between the arrival rates of the excitatory postsynaptic potentials (EPSP) and the inhibitory postsynaptic potentials (IPSP) may vary for different intervals between spikes. The task is then to derive the probability distributions of the number of intervals in a run of intervals with a high ratio of EPSP to IPSP arrivals and in a run of intervals with a low ratio of EPSP to IPSP arrivals, that is to say, the distributions $\{p_1(k)\}$ and $\{p_2(k)\}$, respectively.

To make the PM model available for an implicit statistical analysis one should, of course, also specify the process by which the EPSP's and the IPSP's arrive and the mechanism by which they influence the membrane potential and the functioning of the threshold. These hypotheses should be at least detailed enough to make it possible to derive the form of the distributions of interval length $f_1(x)$ and $f_2(x)$.

The assumption that the membrane potential returns to the resting level after each spike implies that the length of the intervals do not directly influence each other. The only source of their serial correlations is through the fact that during some intervals the ratio of EPSP to IPSP arrivals is high and during others low; that is to say some intervals are more alike than others. This is exactly the content of the basic assumption of the PM model (cf. p. 776): the intervals are of two types, and conditional on the type, they are independent of each other.

Dr. Hyvärinen is a Research Fellow of the Finnish National Research Council for Medical Sciences.

Received for publication 23 January 1969 and in revised form 9 July 1969.

REFERENCES

Biscoe, T. J., and A. Taylor. 1963. *J. Physiol. (London)*. **168**:332.

Buller, A. J., J. G. Nicholls, and G. Ström. 1953. *J. Physiol. (London)*. **122**:409.

Burns, B. D. 1955. *J. Physiol. (London)*. **127**:168.

Cox, D. R. 1962. Renewal Theory. Methuen & Co. Ltd., London.

Cox, D. R. 1965. *J. Roy. Statist. Soc. Ser. B*. **27**:332.

Cox, D. R., and P. A. W. Lewis. 1966. The Statistical Analysis of Series of Events. Methuen & Co. Ltd., London.

Cox, D. R., and H. D. Miller. 1965. The Theory of Stochastic Processes. Methuen & Co. Ltd., London.

Ekholm, A. 1967. A generalization of the two-state semi-Markov model. Social Research Institute of Alcohol Studies. Helsinki. Report No. 27.

Ekholm, A. 1968. A pseudo-Markov model for stationary series of events. Social Research Institute of Alcohol Studies. Helsinki. Report No. 31.

Feller, W. 1957. An Introduction to Probability Theory and Its Applications. John Wiley and Sons, Inc., New York. 2nd edition. I.

Gerstein, G. L., and B. Mandelbrot. 1964. *Biophys. J.* **4**:41.

Hagiwara, S. 1954. *Jap. J. Physiol.* **4**:234.

Halme, A. J., and J. Hyvärinen. 1968. *Med. Biol. Eng.* **6**:563.

Hyvärinen, J. 1966. *Acta Physiol. Scand.* **68**: Suppl. 278.

Kiang, N. Y.-S. 1965. Discharge Patterns of Single Fibers in the Cat's Auditory Nerve. The M.I.T. Press, Cambridge.

Kuffler, S. W., R. Fitzhugh, and H. B. Barlow. 1957. *J. Gen. Physiol.* **40**:683.

Moore, G. P., D. H. Perkel, and J. P. Segundo. 1966. *Annu. Rev. Physiol.* **28**:493.

Perkel, D. H., G. L. Gerstein, and G. P. Moore. 1967 *a*. *Biophys. J.* **7**:391.

Perkel, D. H., G. L. Gerstein, and G. P. Moore. 1967 *b*. *Biophys. J.* **7**:419.

Pfeiffer, R. R., and N. Y.-S. Kiang. 1965. *Biophys. J.* **5**:301.

Poggio, G. F., and L. J. Viernstein. 1964. *J. Neurophysiol.* **27**:517.

Rao, C. R. 1965. Linear Statistical Inference and its Applications. John Wiley and Sons, Inc., New York.

Rodieck, R. W., N. Y.-S. Kiang, and G. L. Gerstein. 1962. *Biophys. J.* **2**:351.

Stein, R. B. 1967. *Biophys. J.* **7**:37.

Thomas, E. A. C. 1966. *Brit. Math. Stat. Psychol.* **19**:151.

Werner, G., and V. B. Mountcastle. 1963. *J. Neurophysiol.* **26**:958.