WHITTLE, P. (1953). Estimation and information in stationary time series. *Ark. Mat. Astron. Fys.* **2** 423–434.

WHITTLE, P. (1961). Gaussian estimation in stationary time series. *Bull. Inst. Internat. Statist.* **33** 105–130.

WIENER, N. (1958). *Nonlinear Problems in Random Theory*. M. I. T. Press, Cambridge.

DEPARTMENT OF STATISTICS,
UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA 94720

DISCUSSION ON PROFESSOR BRILLINGER'S PAPER

D. R. COX (*Imperial College, London*) My comments concern the statistical aspects of Dr. Brillinger's interesting paper. First, when it is required to study the dependence of a process $\{N\}$ on an explanatory process $\{M\}$, there are often strong arguments for arguing conditionally on the observed process $\{m\}$. In particular, assumptions about $\{M\}$ itself are avoided; even its stationarity is not required so long as the interrelations are time-invariant.

Secondly, some qualification seems desirable of. Dr. Brillinger's blanket recommendation that $\{M\}$ should, where possible, be chosen to be Poisson. Will not much depend on the constraints on observation and on the nature of the interrelations? For instance, one can envisage situations where it would be more informative to take $\{M\}$ as a regular sequence of widely spread points, supplemented, perhaps, by some pairs of points close together to examine linearity.

Thirdly, an alternative to the study of interrelations is via the modulation of simple models for $\{N\}$ (Cox, 1972). In this the intensity of the $\{N\}$ process is modified by a factor depending on relevant aspects of the $\{M\}$ process. Two advantages of this approach are that in certain cases likelihood functions can be obtained and that simple relations, nonlinear in Dr. Brillinger's special sense, can be accommodated; for example, the backward recurrence time in the $\{M\}$ process may be particularly relevant. An advantage of Dr. Brillinger's approach is that special assumptions about $\{N\}$ are avoided.

REFERENCES

COX, D. R. (1972). The statistical analysis of dependences in point processes. In *Stochastic Point Processes*, P. A. W. Lewis, ed. Wiley, New York.

P. Z. MARMARELIS (*California Institute of Technology*) Professor Brillinger's well-written paper on the identification of point process systems fulfills, among others, a long-standing need for such work in the field of neurophysiological system analysis. I expect that many applications of these techniques on point process systems (certainly on neural systems) will come to fruition following Brillinger's work.

Motivated by this pioneering work, several additional points will need discussion and analysis in making this methodology practical and useful to experimental research:

1. Delimitation of the class of point process systems which can be represented by a Volterra expansion (3.6).

On this point, paralleling the Wiener analysis for continuous input-output systems, we may note that the systems must have finite memory (effect of input point on output must diminish to zero in finite time). What about point process systems exhibiting a "hysteresis" effect? Can the formulation be easily modified to include them?

2. Formulation of similar identification schemata for systems having (i) input a continuous process, output a point process, and (ii) input a point process, output a continuous process.

For case i, use of a Gaussian white-noise signal as input presents the advantages discussed by Wiener (1958). For case ii, use of a point process with independent increments, such as the Poisson, has the advantages discussed by Brillinger, i.e., $q_{MM}(u)$, $q_{MMM}(u, v)$, $q_{MMMM}(u, v, w)$ are all identically zero. However, in both cases, stochastic process parameters such as the second-order cross product density must be redefined, e.g. in case i with input $x(t)$

$$E\{dN(t + u)x(t)\} = P_{Nx}(u)\,du\,dt \tag{1}$$

with loss of interpretation of this quantity as a probability. Similar models as (3.6) can be written in both cases i and ii and the kernels r_0, $r_1(u)$, $r_2(u, v)$ similarly evaluated. For example, in case i we would have

$$\begin{aligned} r_0 &= P_N \\ r_1(u) &= q_{Nx}(u)/A_x \\ r_2(u, v) &= q_{Nxx}(u, u - v)/(2A_x^2) \end{aligned} \tag{2}$$

where $q_{xx}(\tau) = A_x\delta(\tau)$. Similarly for case ii.

3. Nonlinear formulation for systems with multiple inputs (the linear case is covered in the paper).

For a multiple-input point process system a model can be written down which accounts for interaction between the different input point processes: for an m-input ($M_i(t)$, $i = 1, \cdots, m$), one output system it will be:

$$\mu_M(t) = s_0 + \sum_{k=1}^{k} \sum_{i_1,\cdots,i_k}^{m} \int\int \cdots \int_{u_1,\cdots,u_k;\text{distinct}} s_{i_1,\cdots,i_k}(t - u_1, \cdots, t - u_k) \prod_{l=1}^{k} dM'_{i_l}(u_l) \tag{3}$$

where $M_i'(u) = M_i(u) - uP_{Mi}$.

For example, for a two-input one-output point process system we will have

$$\begin{aligned} \mu_M(t) = s_0 &+ \int s_1(t - u)\,dM_1'(u) + \int s_2(t - u)\,dM_2'(u) \\ &+ \int\int s_{11}(t - u_1, t - u_2)\,dM_1'(u_1)\,dM_1'(u_2) \\ &+ \int\int s_{22}(t - u_1, t - u_2)\,dM_2'(u_1)\,dM_2'(u_2) \\ &+ \int\int s_{12}^*(t - u_1, t - u_2)\,dM_1'(u_1)\,dM_2'(u_2) \end{aligned} \tag{4}$$

where $s_{12}^*(t_1, t_2) = s_{12}(t_1, t_2) + s_{21}(t_1, t_2)$ is the "cross-term" giving the second order nonlinear interaction between inputs $M_1(t)$ and $M_2(t)$. The system characterizing kernels $s_{i_1,\cdots,i_k}(\cdot)$ can be found (with some modification) by formulas analogous to the one-input case (3.9). For purposes of identification it is convenient to choose $M_1(t)$ and $M_2(t)$ to be independent processes with independent increments, such as Poisson; then e.g. $q_{M_1M_2}(u) = 0$ for all u. Then, for example (cf. also Marmarelis and Naka (1974)),

$$\begin{aligned} s_{11}(u, v) &= q_{NM_1M_1}(u, u - v)/(2P_{M_1}^2) \\ s_{12}^*(u, v) &= q_{NM_1M_2}(u, u - v)/(P_{M_1}P_{M_2}) . \end{aligned} \tag{5}$$

These relationships can also be conveniently formulated in terms of spectral functions, as indicated in the paper for the one-input case (3.9).

4. Study of the statistical variance of the estimates of r_0, $r_1(u)$, $r_2(u, v)$ due to record lengths, noise sources (at input and/or output), etc.

On this point the usual statistical methods of estimating the variance should prove directly applicable—even though the expressions for systems with a high degree of nonlinearity, i.e., $s_k(\cdot) \neq 0$ for large k's, will tend to become unwieldy. Nevertheless, such a study would be very useful in practical applications of these techniques.

5. Study of the mean-square-error of the predictor for a truncated representation (3.6).

This is a point of fundamental importance. Many point process systems will be described by models (3.6) for K a fairly large number. The question arises: If we truncate the series at some $k = m$ where $m < K$, and compute $s_k(\cdot)$, $k = 1, \cdots, m$, from cross-covariances such as (3.9), for Poisson input, do we obtain the best mean-square-error predictor of form (3.6) for m terms? The answer would be affirmative if the terms of series (3.6) are orthogonal for a Poisson input $M(t)$. To see this, we can write the truncated (3.6), with $M'(u) = M(u) - uP_M$, a.s.

$$\begin{aligned} \mu_M{}^{(m)}(t) &= s_0 + \textstyle\sum_{k=1}^m F[s_k(\cdot); M'(t)] \\ &= s_0 + \textstyle\sum_{k=1}^m \int \cdots \int_{u_1,\cdots,u_k;\text{distinct}} s_k(t - u_1, \cdots, \\ &\qquad t - u_k)\, dM'(u_1) \cdots dM'(u_k) . \end{aligned} \tag{6}$$

The error due to the truncation is

$$e_m(t) = \mu_M(t) - \mu_M{}^{(m)}(t) .$$

Then, paralleling Schetzen (1974) for continuous process systems with a Gaussian white input, it can easily be shown that, if $M(t)$ is a Poisson process, then

$$E\{e_m(t) \cdot F[f_r(\cdot); M(t)]\} = 0 \qquad \text{for} \quad r = 0, 1, 2, \cdots, m \tag{7}$$

for any causal function $f_r(\cdot)$.

Now, let us consider a different set $\{s_r{}^*(\cdot)\}$ and

$$\mu_M{}^{*(m)}(t) = s_0{}^* + \textstyle\sum_{k=1}^m F[s_k{}^*(\cdot); M(t)] . \tag{8}$$

Then the error for this characterization is

$$e_m^*(t) = \mu_M(t) - \mu_M^{*(m)}(t) = \mu_M(t) - \mu_M^{(m)}(t) + \mu_M^{(m)}(t) - \mu_M^{*(m)}(t) \tag{9}$$

and

$$E\{(e_m^*(t))^2\} = E\{e_m^2(t)\} + E\{[\mu_M^{(m)}(t) - \mu_M^{*(m)}(t)]^2\} + 2E\{e_m(t)[\mu_M^{(m)}(t) - \mu_M^{*(m)}(t)]\}\,. \tag{10}$$

The last term of this expression is zero in view of (7). Therefore, since the second term is nonnegative, the mean-square-error is minimum when

$$\mu_M^{(m)}(t) = \mu_M^{*(m)}(t) \tag{11}$$

that is, when $s_k^*(\cdot) = s_k(\cdot)$. This completes the proof that the functions $s_k(\cdot)$ computed from cross-covariances of the system output and input point processes, with a Poisson input, provide the best mean-square-error predictor model (3.6) of the output processes N based on input process M, for any truncation of the series (3.6).

REFERENCES

[1] MARMARELIS, P. and NAKA, K. (1974). Identification of multi-input biological systems. *IEEE Trans. Biomedical Engineering* **21** 88–101.

[2] SCHETZEN, M. (1974). A theory of nonlinear system identification. *Internat. J. Control* **20** 577–592.

[3] WIENER, N. (1958). *Nonlinear Problems in Random Theory.* Wiley, New York.

J. P. SEGUNDO (*University of California, Los Angeles*) Individual nerve cells manifest their activity in several ways (Segundo, 1968), one of which is the generation of events referred to as action potentials, spikes or impulses. The latter, when recorded through a small electrode inserted into the cell, appear commonly as recurrent events which are identical to each other and are separated by intervals quite longer than their individual duration. These experimental facts justify the assimilation of a sequence (i.e., a train) of spikes with a point process. Furthermore, when nerve cells are connected and interact through junctions referred to as synapses, the spike activity in one, called presynaptic, has repercussions upon the spike activity in the other, called postsynaptic. The study of this transformation from pre- to postsynaptic spikes constitutes an important chapter of neurophysiology. There are other manifestations of, and interactions between, nerve cells (Segundo, 1968); the fact that a bulk of knowledge is on spikes may reflect not just their importance but also the ease with which they can be recorded, and even a trend of scientific fashion.

Most work on the synapse has concentrated on its mechanisms (e.g., chemical, biophysical, etc.). Also meaningful, however, is the dynamic study of how the one intensity of generation (i.e., that of presynaptic spikes) influences the other (i.e., that of postsynaptic spikes). This is in essence an "identification" problem, where the "black box" is all that which is interposed between one spike train and the other. Identification procedures apply, for the most part, to systems that are linear and whose inputs and outputs are continuous or at best defined at equispaced instants. This work, as well as other by the same author, allows use of precise and rigorous procedures to the understanding of synaptic transfer. It remains now to apply them to data from living cells, and to coin the quantitatively inferred conclusions in terms that have physiological meaning.

Special care is required when mathematical techniques are applied to biology (Segundo, 1971). Firstly, one must not be carried away by a natural desire for explicit solutions into applying procedures (e.g., linear systems analysis) outside of their justified domain, into unrealistic idealization of the observed phenomena, into ignoring nonstationarities (e.g., cycles, learning, aging), or into applying to living matter concepts (e.g., desirable, parasitic) developed for man-made machines.

REFERENCES

SEGUNDO, J. P. (1968). Functional possibilities for communication and coding of neuronal properties and interactions. *Acta Neurol. Latinoamer.* **14** 340-344.

SEGUNDO, J. P. (1971). Considerations regarding the study of central integration. *Acta Cientifica Venezolana* **22** 134-136.

* * * * * * * * * *

The remarks of the discussants are gratefully appreciated. I agree with Professor Cox that it is often sensible to argue conditionally, especially on an observed input process. In the present situation it might further be sensible to argue conditionally on $N(0, T]$, the observed number of output points. I also agree that if it is possible to "design" the input process, consideration shouldn't be restricted to the Poisson. In the paragraph below expression (3.9) I have suggested the use of an input process with widely spaced points, in order to estimate the average impulse response of a linear system. Professor Cox's remark that such input may also be used to examine for nonlinearity is well taken. One could use the model of (3.6) to discuss this suggestion in a formal manner. His final suggestion of fitting simple models, of the type indicated, would appear to be especially useful in the case of nonstationary processes.

Professor Marmarelis raises a number of interesting questions. I hope that some colleagues will become interested in examining his first one of the degree of accuracy of Volterra expansions. Hida (1970) is a relevant reference. I understand Professor Marmarelis to mean by "hysteresis effect" the occurrence of multi-valued output. It occurs to me that this problem hasn't really been seriously addressed even in simple regression situations, despite the existence of relevant data sets such as those of earthquake travel times. Professor Marmarelis remarks that interpretations as probabilities are typically lacking for the parameters of hybrid point-continuous processes. This is a disadvantage. However one probability parameter that is proving useful in empirical analysis is that defined by

$$\text{Prob}\,\{dN(t+u) = 1 \ \text{ and } \ x < X(t) < x + dx\} = p_{NX}(u, x)\, dt\, dx\,.$$

Interesting graphs result from plotting estimates of $p_{NX}(u, x)$ as a function of x for fixed u.

Professor Segundo's remarks put into proper perspective what is involved in the leap from a probabilistic idealization to a real world phenomenon. I hope that the readers of these *Annals* are aware that neurophysiologists are currently collecting extensive point process data sets under well-controlled experimental situations and are raising a host of interesting analytic questions. It would clearly be of mutual benefit to probabilists and neurophysiologists to collaborate in the analysis of this data.

DAVID R. BRILLINGER