

# Signal Transformation and Coding in Neural Systems

VASILIS Z. MARMARELIS MEMBER, IEEE

**Abstract**—The subject of signal transformation and coding in neural systems is fundamental in understanding information processing by the nervous system. This paper addresses this issue at the level of neural units (neurons) using nonparametric nonlinear dynamic models. These models are variants of the general Wiener–Bose model, adapted to this problem as to represent the nonlinear dynamics of neural signal transformation using a set of parallel filters (neuron modes) followed by a binary operator with multiple real-valued operands (equal in number to the number of modes). The postulated model constitutes a reasonable compromise between mathematical complexity and current neurophysiological evidence. It incorporates nonlinear dynamics and spike generation mechanisms in a fairly general, yet parsimonious manner. Although this study has objectives limited to a single unit and represents a small contribution in a vast and complex research area, it is hoped that it will facilitate progress in the systematic study of the functional organization of neural systems with multiple units.

## I. INTRODUCTION

THE nervous system is a complex structure composed of a multitude of functional components (sensory receptors, neurons, axons, muscle fibers) interconnected in cascade, parallel, lateral, and feedback configurations that process information (i.e., receive, transform, code, transmit, and decode) in order to perform specific vital tasks in its interactions with the environment and in preserving physiological homeostasis. Developments in systems science and cybernetics allow us to formulate the study of information flow to, from, and within the nervous system as a “systems problem” whereby signals, representing information, travel between neural system components and are dynamically transformed by them. The use of the systems approach is predicated on appropriate conceptual and mathematical formalism in describing the “transfer characteristics” of the neural system components, their interconnections, and the transformed neural signals. In the case of the nervous system, the level of system integration/decomposition (i.e., at what level of the natural hierarchy of functional organization do we focus) may range from the subcellular (molecular) to the behavioral; however, in this paper, we limit ourselves to the level of individual neurons (viewed as the basic operational unit) which receive certain input signals and generate, in a causal manner, certain output signals. The hope is that accurate understanding of the functional properties of the “neural unit” will facilitate the study of neural groups and networks and allow the construction of “integrated”

neural systems of greater organizational complexity with specified functional characteristics.

In this paper, we conceptualize each neuron as a “black box” that receives certain input signals and produces certain output signals on the basis of a nonlinear dynamic “rule” described by an explicit mathematical expression (model). The use of the black-box concept for the representation of a single neuron allows us to bypass the complex biophysical mechanisms that are active at the subcellular level, and simply concentrate on the input–output transformation. This, in turn, allows the development of reduced-complexity models for aggregates of neurons and the study of their functional properties.

In developing this mathematical model, we seek a compromise between mathematical complexity and biological relevance in order to obtain tractable mathematical formulations of the problem while preserving the essential functional features that have been observed experimentally. Our goal is to search for unifying operational principles that explain the largest possible amount of current experimental evidence and conceptually organize our understanding about the system function. One challenging aspect of neural system modeling concerns the role of nonlinearities. There is no doubt that nonlinearities are omnipresent in neural systems and their role is essential at least for some aspects of their function. Compressive, decompressive, and sigmoid nonlinearities observed in sensory receptors, certain types of facilitatory or occlusive interaction, synaptic gap transmission, and the generation mechanism of action potentials are some of the most widely recognized examples. The challenge is the actual identification of these dynamic nonlinearities and their modeling in a manner that is practical and meaningful from the point of view of advancing our understanding of neural function. Finally, one should note the presence of extraneous and intrinsic noise that places the modeling problem in a stochastic framework.

In the case of dynamic nonlinearities of black-box systems, the possible identification and modeling methods can be classified as either parametric or nonparametric. The parametric methods require *a priori* knowledge (or postulation) of a specific model in the form of differential or difference equations. Nonparametric methods employ general model forms of integral equations or functional expansions valid for a broad class of dynamic nonlinearities. The most widely used nonparametric methods are referred to as the Volterra–Wiener approach, and are based on functional expressions. One such method has been adapted to the case of spike data, suitable for iden-

Manuscript received February 2, 1988; revised July 1, 1988. This work was supported by Grant RR01861 from the National Institutes of Health.

The author is with the Department of Biomedical Engineering, University of Southern California, Los Angeles, CA 90089.

IEEE log Number 8824413.

tification and modeling of a wide class of neural systems by utilizing explicit threshold operators. It is this latter method, and its use for understanding signal transformation and coding in neural systems, that constitute the focus of this paper.

## II. STRUCTURE AND FUNCTION OF NEURAL UNITS

As mentioned above, we focus on the neuron as the basic operational neural unit from the viewpoint of signal transformation and coding. The term "transformation" is used to indicate the dynamic process of combining input information to produce a composite intracellular potential at the site of the axon hillock where the generation of action potentials (spikes) takes place; the latter representing the "coding" of neural information. Note that the representation of information flow in the nervous system (neural signals) is either in the form of "graded potentials" (GP) or "action potentials" (AP). The latter are often referred to as "spikes" (i.e., impulse-like waveforms) and it is accepted that their information content is in the interspike time intervals and not in their magnitude or shape. The graded potentials are continuous-time signals, whose values (like those of the AP's) are measured relative to the "resting potential" of the neuron. The generation and transformation of these potentials is accomplished through complex biophysical mechanisms of excitable tissues (ionic currents arising from changing membrane permeabilities), release (and diffusion) of neurotransmitters in synaptic gap junctions, electrical coupling in "tight" junctions, and passive "electrotonic" spread of potentials. These biophysical mechanisms have been studied extensively and are the subject of a vast literature (for partial review, see [2], [3], [14]). In this paper, we will limit ourselves to those notions, arising from this extensive biophysical knowledge, that are relevant in the "systems and signals" context of our formulation.

Let us first review briefly the fundamentals of neuronal structure and function that are common (with variations in size, geometry, and specific functional attributes) in most neurons. The "prototypical" neuron is comprised of a cell body (soma), a tree-like structure of fibers (dendrites), and a long fiber (axon) with occasional branches (collaterals). The axon is attached to the soma at the "axon hillock" and, along with its collaterals, ends at synaptic terminals (boutons) that are used to pass information onto other neurons through "synaptic junctions." The soma contains the nucleus and is attached to the trunk of the dendritic tree from which it receives incoming information. The dendrites are conductors of input information to the soma (input ports) and exhibit usually a high degree of tree-like arborization (up to several thousand branches). Input information arrives at various sites in the dendritic tree and the soma through synaptic terminals of other neurons or electrical (tight) junctions. The synaptic terminals release chemical transmitters upon stimulation by arriving AP's or GP's, which diffuse across the gap (typically 200 Å) of the synaptic junction and induce the generation of a "postsynaptic potential" (PSP) at the

postsynaptic site of the dendrite or soma. The transmitter can be excitatory or inhibitory, which determines the relative polarity between pre- and postsynaptic potentials. The PSP propagates away from its junction in both directions in an electrotonic manner (i.e., a way analogous to current flow in passive electric cables) leading to attenuation of amplitude and spreading of the waveform [13]. Although passive dendritic fibers are more common, active (or semiactive) ones have also been found [7]. These active fibers have voltage dependent permeabilities and allow at least partial regeneration of the propagating signal, leading to less attenuation of amplitude. This partial regeneration property of dendritic membrane should be contrasted with the full regeneration property of axonal membrane that fully maintains the amplitude of the propagating AP. The resulting dendritic potentials (DP) merge at the various nodes of the dendritic tree and eventually arrive at the soma where, combined with potentials generated directly at the soma by possible somatic junctions, they produce the composite intracellular potential (CIP) at the site of the axon hillock. Note that the merging of DP's is, most likely, a nonlinear operation because the generation of these potentials is due primarily to migration of sodium ions, which may lead to sigmoid saturation (clipping).

There are two types of neurons: those which generate AP's at the axon hillock through a threshold mechanism, and those which do not. The latter transmit through their axon the CIP as a graded potential and the former generate a sequence of AP's (spike train) if the values of the CIP exceed a certain threshold. The transmission of GP's through the axon, as carriers of information, is not effective except over very short distances ( $< 0.3$  mm) owing to rapid attenuation of GP's with distance. Therefore, the use of this communication mode is limited to neurons with very short axons (e.g., photoreceptors, horizontal, and bipolar cells in the retina). On the other hand, AP's can propagate practically unaltered over long distances because they are regenerated by the excitable membrane tissue, thus offering an effective mode of communication in the nervous system. It is this latter type of neuron that our paper will concentrate on.

This brief phenomenological review of the basic structural and functional characteristics of the "prototypical" neuron serves as the neurophysiological foundation of our analytical study in that it provides the guide for meaningful modeling assumptions and formulation of the associated system problem.

## III. SIGNAL TRANSFORMATION IN NEURAL UNITS

The problem of signal transformation in a neural unit (neuron) is defined as the modeling study of a "black-box" system that transforms received input information into the composite intracellular potential (CIP) at the axon hillock, as discussed in the previous section.

In general, a neuron receives input information through synaptic or electrical junctions (excitatory or inhibitory) at various sites of its dendrites and soma. These input sig-

nals are provided by other neurons (or external stimuli in the case of sensory receptors) and can be either GP's or AP's. The input signals are first transformed at the junctions, an operation which in the synaptic junctions involves mild low-pass filtering (time-constants of about 2–4 ms) and a sigmoid nonlinearity reflecting molecular transmitter-receptor interactions [5]. The resulting dendritic potentials (DP) propagate along the dendritic branches and merge together in a nonlinear fashion due to saturation of sodium ion currents and the complex geometry of the dendritic tree. The propagation of DP's along a passive fiber follows Rall's linear cable theory [13] and results in amplitude attenuation and spreading of the waveform with time (distance). Although the latter transformation is linear, this is not true if the ionic conductances of some dendritic fibers are (subcritically) voltage dependent leading to nonelectrotonic conduction in dendrites, as suggested by Leibovic and Sabah [6] and observed experimentally by Llinas *et al.* [7]. Furthermore, the dynamic transformation of propagating DP's and their convergence to the soma is complicated by the complex geometry of the dendritic tree and the variability in ionic species involved, membrane properties and junction sites and types. The following general statement can be made with some degree of confidence: the overall dynamic transformation of input signals into the CIP is nonlinear and of a fairly broad bandwidth (up to a maximum of about 100 Hz).

The generality of the above statement is useful in that it establishes the nonlinear and broad-band dynamic nature of the transformation, and points to the suitability of the nonparametric modeling approach to this problem. Consequently, the Volterra–Wiener modeling approach is selected to represent this transformation since it satisfies the stated requirements. This approach has been extensively used in studies of nonlinear neural systems in recent years (for partial review, see [8] and [10]). According to this approach, the input–output relation of a broad class of nonlinear dynamic systems can be expressed as a functional series (Volterra series):

$$u(t) = \sum_{n=0}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} k_n(\tau_1, \cdots, \tau_n) \cdot x(t - \tau_1) \cdots x(t - \tau_n) d\tau_1 \cdots d\tau_n \quad (1)$$

where  $x(t)$  and  $u(t)$  are the input and output signals respectively, and  $\{k_n(\tau_1, \cdots, \tau_n)\}$  are the system kernels that characterize its nonlinear dynamics, i.e., they can be viewed as generalized nonlinear filters. An associated functional series (Wiener series) that is orthogonal for Gaussian white-noise inputs has been used extensively in identification studies of physiological systems [8], [10]. In this paper, we are not concerned with the identification problem (although that would be important if experimental data were analyzed), but we wish to explore modeling issues of the aforementioned transformation. To this purpose, we consider an alternative model formulation of the Volterra–Wiener class of systems known as the Wiener–

Bose model [1]. The Wiener–Bose model utilizes a complete set of orthogonal linear filters  $\{L_n\}$  and a static (zero-memory) nonlinearity ZMN to propose the general block-structured model shown in Fig. 1 that is equivalent to the functional series of (1). The orthogonality of the filters  $\{L_n\}$  is essential for system identification purposes; however, a model having the structure shown in Fig. 1 remains valid, in general, for any complete (but not necessarily orthogonal) set of filters spanning the entire system dynamics. In practical terms, a set of linearly independent filters that captures the essential characteristics of the system dynamics is adequate. Based on this observation, we propose that the dynamic neural transformation be modeled by a finite set of filters that capture the important dynamic system characteristics followed by an appropriate static nonlinearity ZMN. These key filters will be termed “modes” of the neuron and their outputs  $\{v_i\}$  will be termed “internal variables” (IV). To retain mathematical simplicity and physiological interpretability, the number of modes should be kept to the necessary minimum in each case. The form of the static nonlinearity must also be determined to complete the model. This formalism can be easily extended to the multi-input case [9].

To illustrate the proposed model, let us consider some examples. A neural system that exhibits two distinct modes of dynamics: a (leaky) integrating mode and a (slow) differentiating mode, shown in Fig. 2 as impulse response functions ( $m_1, m_2$ ) of the two filters ( $L_1, L_2$ ) corresponding to these modes where

$$m_1(\tau) = \frac{\tau^2}{12} [\exp(-\tau/6) - \exp(-\tau/2)] \quad (2)$$

$$m_2(\tau) = -(\tau - 10) \exp[-(\tau - 10)^2/16] \quad (3)$$

where the time unit is numerically set to one (dimensionless). Consider the static nonlinearity ZMN that has a sigmoid characteristic in terms of a linear combination of the two IV's ( $v_1, v_2$ ):

$$u = \frac{1}{1 + \alpha \exp(-\beta_1 v_1 - \beta_2 v_2)} \quad (4)$$

where  $u$  is the output CIP. For  $\alpha = 0.4, \beta_1 = 0.2, \beta_2 = 0.4$  the resulting (hyper) sigmoid surface  $u = u(v_1, v_2)$  is shown in Fig. 3. The first- and second-order kernels of this nonlinear system are shown in Fig. 4, and they reflect the combined dynamics of the two modes  $m_1$  and  $m_2$  under the (hyper) sigmoid nonlinear transformation of (4). Kernels of this approximate form have been measured in actual neural systems through white-noise stimulation and analysis. Therefore, our suggestion is that actual kernel measurements be used in each particular case to select the minimum number of necessary dynamic modes. If the modes are picked correctly, then the form of the static nonlinearity ZMN can be easily determined from white-noise experimental data owing to the ergodic properties of the stimulus.

The problem of identifying the minimum number of required neuron modes from actual input–output experi-

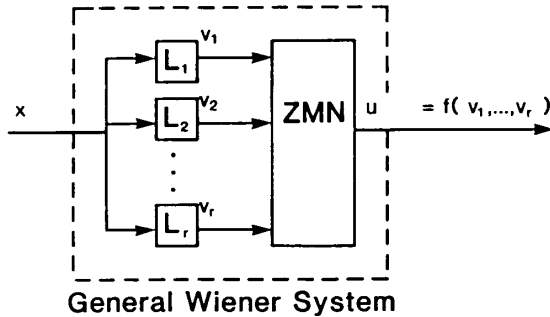


Fig. 1. The Wiener-Bose model of nonlinear dynamic systems.

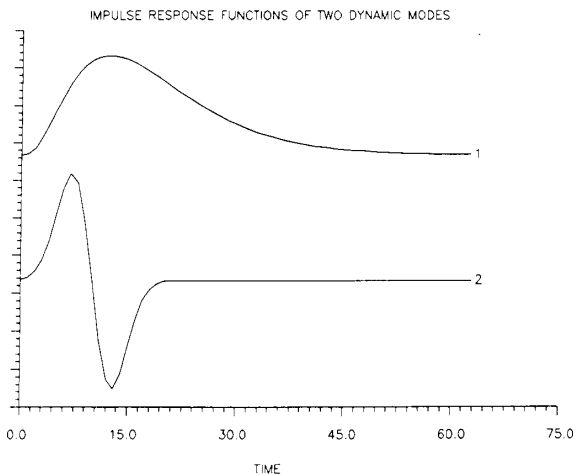


Fig. 2. The impulse response functions of modes  $L_1$  (trace 1) and  $L_2$  (trace 2).

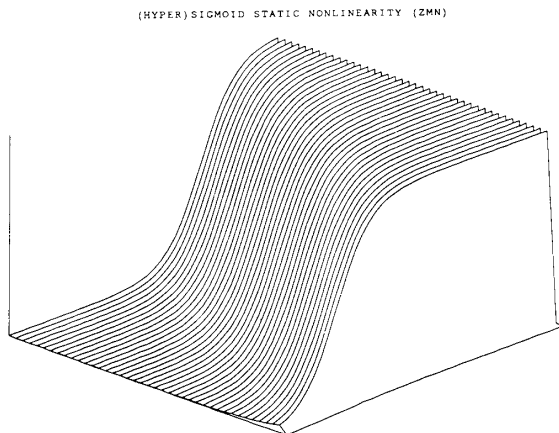


Fig. 3. The form of the static nonlinearity defined by (4).

mental data can be rather formidable in practice. In the general case, a set of orthonormal modes (corresponding to the filter bank of Fig. 1) and the corresponding coefficients can be obtained following the Wiener-Bose approach when the input is Gaussian white noise [1], [8].

However, reducing this set of general modes to the minimum necessary for a specific neuron is a task for which no general approach is currently available. Nonetheless, if we limit ourselves to measurements of the first two kernels (which is the usual practice in applications to date), then one such method can be devised by diagonalizing the (symmetric) matrix of the obtained expansion coefficients for the second-order kernel. Then inspection of the absolute values of the diagonal entries (i.e., eigenvalues of the original matrix) indicates the minimum number of required modes (i.e., the most significant ones) and allows determination of these modes as the inner products of the corresponding eigenvectors with the original general modes. This approach is analogous to singular value decomposition in determining minimum order realizations of linear systems from noise-contaminated experimental data. The significant modes thus selected from the second-order kernel, along with their counterparts in the transformed first-order kernel representation, enter in the model as quadratic terms. An additional linear term may be selected that corresponds to the remaining significant terms of the first-order kernel representation. The complexity of this task is compounded by experimental limitations in applying and measuring the appropriate input-output signals (especially in the multi-input/multi-output case), as well as by the presence of data-contaminating noise. All these practical limitations point to the feasibility of obtaining only rough (but hopefully adequate) approximations of the proposed model in practice.

The form of the static nonlinearity need not be limited to the one given by (4) and we expect a great variety depending on neuron characteristics. Even if we stay with the (hyper) sigmoid expression of (4), the exponent term may be a different function of  $v_1, v_2$ . For instance, a neurophysiological phenomenon known as "shunt inhibition" may result in amplitude modulation of one mode by another, leading to bilinear terms in the exponent. In the previously discussed example of a (hyper) sigmoid ZMN, we have

$$u = \frac{1}{1 + \alpha \exp(-\beta_1 v_1 - \beta_2 v_2 + \gamma v_1 v_2)}. \quad (5)$$

The shape of this static nonlinearity is shown in Fig. 5 for  $\alpha = 0.5$ ,  $\beta_1 = 0.5$ ,  $\beta_2 = 0.5$ ,  $\gamma = 0.1$ , exhibiting regions of mutual facilitation and occlusion. Another example could be a case of full-wave rectification of one mode or both modes giving rise to quadratic terms in the exponent of the (hyper) sigmoid (shown in Fig. 6). Full-wave rectification in the presented phenomenological context may arise from combined synaptic inputs by two other neurons that have previously applied half-wave rectification to their (same) input. We must emphasize that the form of ZMN is a matter of pure conjecture in these examples.

#### IV. SIGNAL CODING BY NEURAL UNITS

The signal coding problem by a neural unit (neuron) is defined as the process of generating a sequence of AP's,

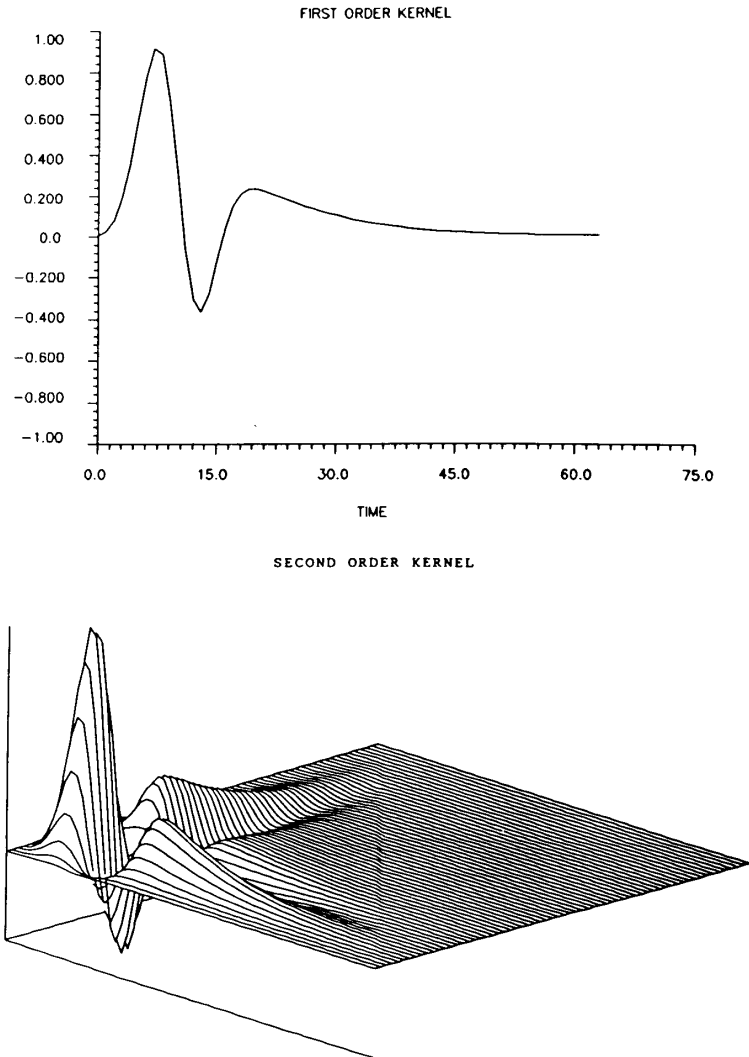


Fig. 4. The first and second order kernels of the nonlinear system defined in the text.

on the basis of the CIP and a threshold mechanism, that carries the desired information content. The biophysical mechanisms of the generation of AP's have been extensively studied starting with the pivotal work of Hodgkin and Huxley [4]. Since the detailed biophysical issues involved are not critical for this paper, we borrow only those developed notions that are relevant in our systems formulation of the problem. In its simplest form, the AP generation mechanism can be viewed as a threshold-trigger (TT) device that produces a spike (AP) whenever its input (the CIP) exceeds a certain threshold value. There is a "reset" mechanism in TT (analogous to a gas-tube circuit) that returns the output to the resting level after each spike firing and prevents the generation of another spike for a short period of time following each firing (refractory period). There are dynamics involved in this process, but

they are so fast relative to the previously discussed dynamics that they can be ignored in first approximation.

There are several important issues regarding modeling of the AP generation mechanism that will be left untreated in this paper. To name a few: the distinction between absolute and relative refractoriness (refractory dynamics), the experimentally observed characteristics of fatigue, adaptation and habituation, the spontaneous activity, and the stochastic threshold variations. These issues must and will be incorporated in this modeling framework in the future. At present, we limit ourselves to the simplified postulate of the TT and (by employing discrete-time representation of our signals to bypass the issue of exact AP shape and refractoriness) we focus on what can be learned from this simplified model in terms of coding of incoming neural information into a sequence of AP's (spike train).

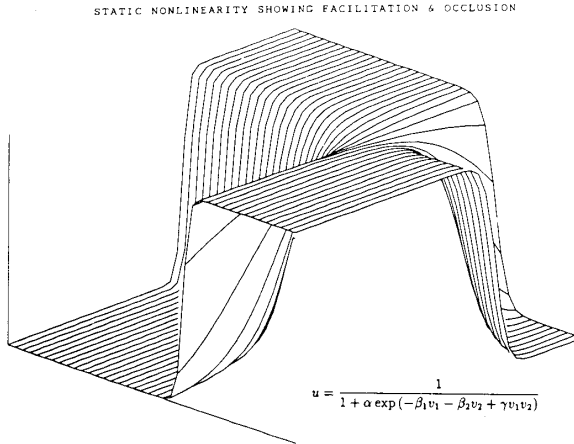


Fig. 5. The form of the static nonlinearity defined by (5), exhibiting regions of facilitation and occlusion.

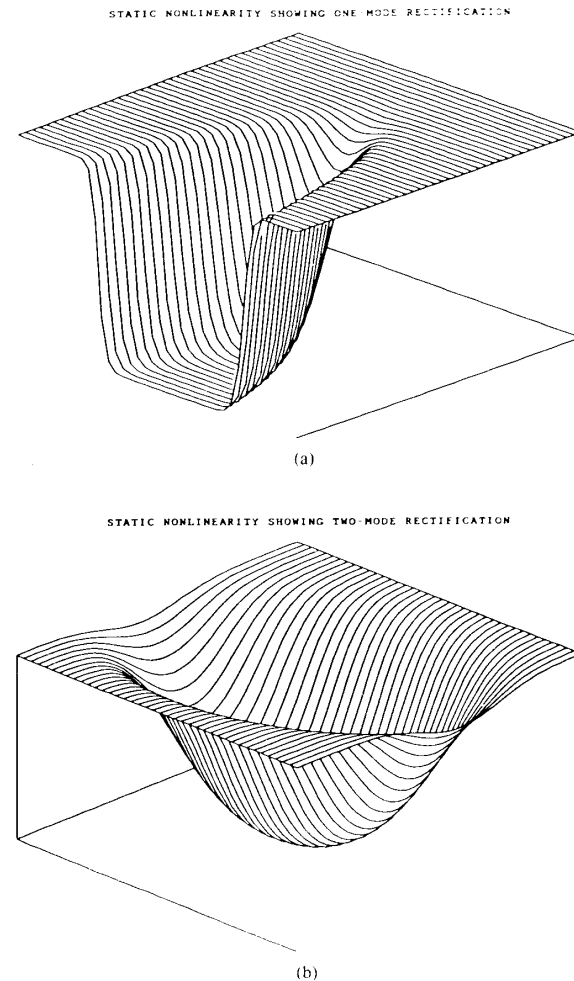


Fig. 6. The form of the static nonlinearity containing quadratic terms of (a)  $v_2$ , (b)  $v_1$  and  $v_2^2$  in the exponent of the basic (hyper) sigmoid expression (full-wave rectification characteristics).

First, we discuss the effect of appending TT at the end of the model discussed in Section III. In a recent paper regarding theoretical aspects of Wiener modeling of spike-output systems [11] we demonstrated that the addition of a TT operator at the end of the general Wiener model reduces considerably the number of required Wiener kernels for the full representation of the system. This reduction is of critical importance for the system identification and modeling tasks, and it is especially dramatic when the nonlinearity ZMN is monotonic. To demonstrate this, consider a system with one mode  $L$  followed by a static nonlinearity ZMN and a threshold-trigger TT, as shown in Fig. 7(a). The cascade of ZMN and TT can be viewed as a single static nonlinearity NTT that produces a spike whenever its input  $v(t)$  attains values within specific ranges determined by the combination of ZMN and TT. For instance, if the characteristic of ZMN and the threshold value  $\theta$  of TT are as shown in Fig. 7(b) with a solid line, then the resulting characteristic of the composite static nonlinearity NTT is as shown in Fig. 7(c). This indicates that a third-order Wiener model [i.e., model order equal to the number of intersections between threshold line and  $f(v)$ ] would be sufficient in this case regardless of the degree of nonlinearity of the function  $f(v)$ . A more dramatic example would be a threshold  $\theta'$  (shown also in Fig. 7(b) with a dashed line) resulting in a single intersection point at  $v = v'$  and reducing the order of the required Wiener model to one (NTT characteristic is shown in Fig. 7(c) with a dashed line). Computer simulations illustrating this point can be found in Marmarelis *et al.* [11].

According to Wiener's theory, the general nonlinear system of the Wiener class (i.e., a system having finite output variance for Gaussian white-noise input) can be represented as in Fig. 1. If this general Wiener-Bose model is followed by a threshold-trigger TT (shown in Fig. 8) then we have the general model for Wiener systems with spike outputs. The nonlinear characteristic of NTT will be, in general, an  $r$ -dimensional binary function:

$$y = \frac{1}{2} + \frac{1}{2} \operatorname{sgn} \{ f(v_1, v_2, \dots, v_r) - \theta \} \quad (6)$$

where  $u = f(v_1, v_2, \dots, v_r)$  is the ZMN characteristic and "sgn" is the signum function. In other words, NTT will produce a spike whenever the combination of values  $(v_1, v_2, \dots, v_r)$  is such that  $f(v_1, v_2, \dots, v_r) \geq \theta$ . These "trigger" values of  $(v_1, v_2, \dots, v_r)$  define "trigger regions" that are demarcated by the solutions of the equation

$$f(v_1, \dots, v_r) - \theta = 0. \quad (7)$$

The solutions of (7) are the "trigger lines" corresponding to the trigger points of the one-dimensional case shown in Fig. 7. In direct analogy with the one-dimensional case discussed before, the form of these trigger lines determines the required (minimum) order of the Wiener model.

In actual applications, these trigger regions can be determined experimentally by computing the values of  $(v_1,$

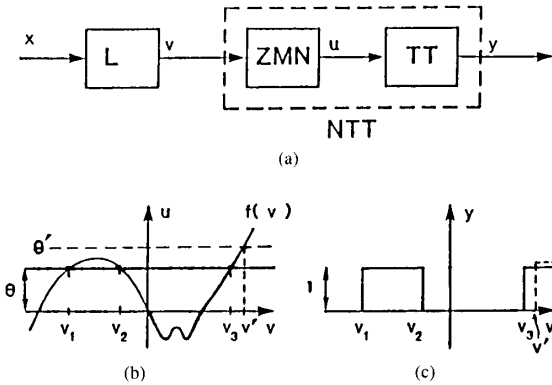


Fig. 7. (a) The L-ZMN-TT cascade model. (b) ZMN characteristic and threshold values  $\theta$  and  $\theta'$  of TT, discussed in the text. (c) The resulting NTT characteristics for the two cases in (b).

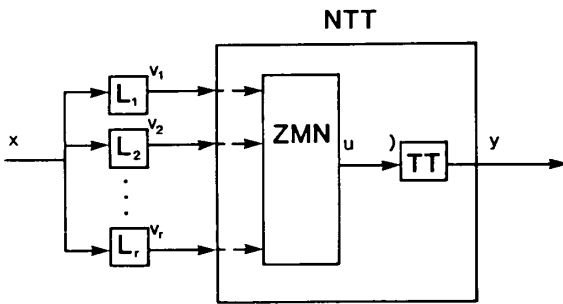


Fig. 8. The general Wiener-Bose model followed by a threshold-trigger TT, and the equivalent NTT static nonlinearity.

$\dots, v_r$ ) for which a spike is observed in the output. The locus of these values will form an estimate of the trigger regions of the system. This is, of course, practically possible only if a relatively small number of filters  $\{L_r\}$  can be found that span effectively the dynamics of the system under study.

This general formulation of the modeling problem for spike-output systems has important implications in the study of neural systems. A spike-generating unit (neuron) is seen as a dynamic element that codes input information into a sequence of spikes where the exact timing of the spikes contains the transmitted information. This coding operation is defined by a finite number of dynamical modes [corresponding to the internal variables  $v_i(t)$ ] and by the boundary of the trigger regions of the unit (trigger lines). Therefore, each neural unit can be represented by the (necessary minimum) dynamical modes and the corresponding trigger lines. This representation leads to a general and, at the same time, parsimonious description of the nonlinear dynamics of a neural unit. These units can be interconnected to form neural aggregates with specifiable functional characteristics. These aggregates may be composed of classes of units, with each class characterized by a specific dynamical representation. The issue of whether the use of units with fairly complex properties

offer an advantage over multiple units with simpler properties in the context of a neural aggregate (or network) remains outstanding at this juncture.

Let us now see how these ideas apply to the example discussed in the previous section. If the TT operator (with threshold  $\theta = 0.8$ ) is appended to the (hyper) sigmoid ZMN shown in Fig. 3, then the NTT characteristic shown in Fig. 9 results (i.e., the trigger line is a straight line). If the same is done for the ZMN characteristics shown in Fig. 6, then the NTT characteristics shown in Fig. 10 result. Note that the same NTT characteristics will result for all ZMN surfaces that have the same intersection line(s) with the threshold plane, i.e., they yield the same solution for (7). This clearly demonstrates that the detailed morphology of the ZMN surface in the subthreshold or suprathreshold region has no bearing on the pattern of generated spikes.

If the two dynamical modes of the system are the ones shown in Fig. 2, then the values of the internal variables  $v_1(t)$  and  $v_2(t)$  provide information through time about the integrated (magnitude) and differential (rate) characteristics of the input signal. The NTT characteristics then indicate which combinations of magnitude and rate values of the input signal will lead to the generation of a spike by the neuron (trigger regions). Furthermore, subregions of these trigger regions can be monitored by "downstream" postsynaptic neurons using the temporal pattern of the generated spike train and their own dynamic characteristics (modes and NTT). This "higher level" coding can provide the means for refined clustering of spike events that reflect specific "features" of the input signal, leading to specialized detection and classification of input information features and, eventually, to cognitive actions using these specialized features as "primitives."

In full awareness that this hypothesis of neural information processing is not yet tested and is still in a seminal stage, we nevertheless propose it as a plausible theory that exhibits some attractive characteristics, e.g., it incorporates nonlinear dynamics and the signal modalities found in neural systems in a manner that is not inconsistent with current experimental evidence.

We conclude with a simple example of signal coding based on the presented ideas. Consider a pulse input, shown in Fig. 11 along with the resulting internal variables  $v_1(t)$  and  $v_2(t)$  corresponding to the modes of Fig. 2. Application of a TT operator on  $v_1(t)$ ,  $v_2(t)$ ,  $-v_2(t)$  and  $v_2^2(t)$  separately yields the spike trains shown in Fig. 12 (traces, 1, 2, 3, and 4, respectively). These cases emulate the "on-sustained," "on-transient," "off-transient," and "on/off-transient" responses of neurons, respectively, that are often observed experimentally. They code an event of significant magnitude and its duration, the onset of an event, the offset of an event, and the onset and offset of an event, respectively. Application of TT (with appropriate threshold) on a linear combination of the internal variables  $[v_1(t) + 2v_2(t)]$  yields an "on-mixed" response shown in Fig. 13 along with the input pulse and the combined waveform of the internal varia-

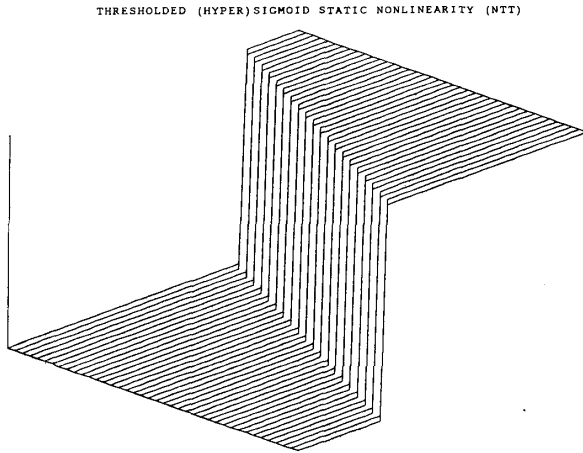


Fig. 9. The NTT characteristic for the ZMN of Fig. 3.

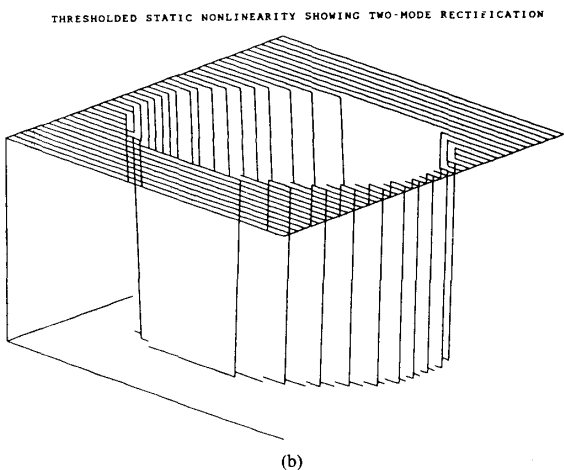
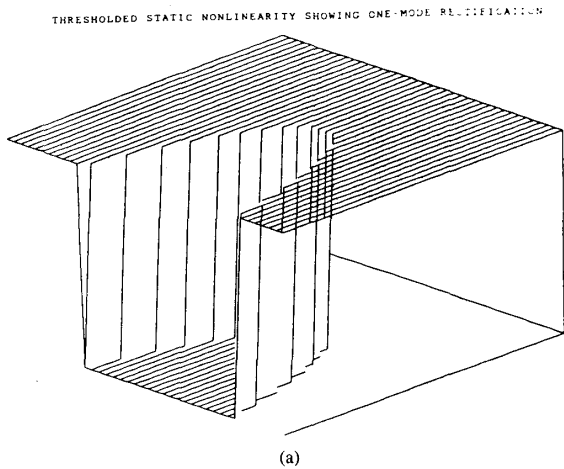


Fig. 10. The NTT characteristics for the ZMN's shown in Fig. 6 (displays (a) and (b), respectively).

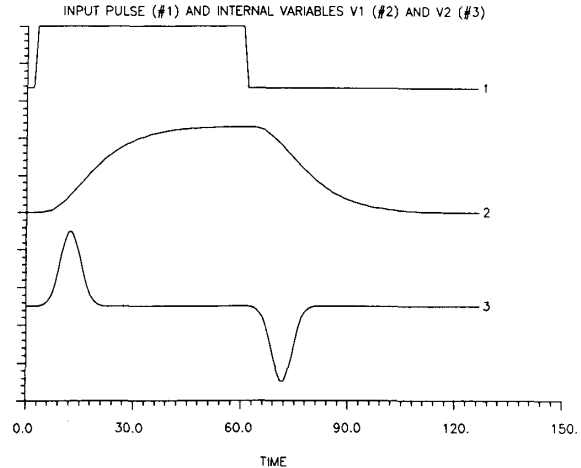


Fig. 11. A pulse input (trace 1) and the resulting internal variables  $v_1(t)$  (trace 2) and  $v_2(t)$  (trace 3), corresponding to the modes shown in Fig. 2.

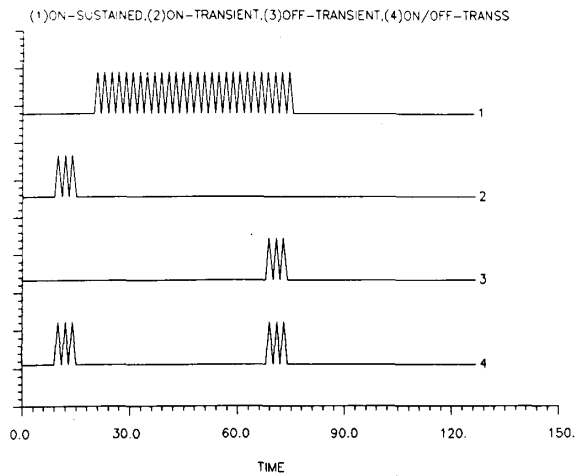


Fig. 12. Spike-train responses to the pulse input for four common types of neuronal responses: 1) on-sustained, 2) on-transient, 3) off-transient, 4) on/off-transient.

bles (this case would correspond to using the NTT characteristic of Fig. 9). The spike output codes the onset of the stimulus (event) and its duration in the same time-record. The higher level neurons (postsynaptic) must "know" that this is an "on-mixed" cell, otherwise they will mistakenly interpret the cell output as coding two distinct events. This can also be accomplished by cross examination of the outputs of several different type neurons receiving and coding the same input. To illustrate the idea of higher level decoding by monitoring different subregions of the NTT trigger regions for spike events (i.e., clustering of  $v_1, v_2$  values resulting in a spike), we consider the presented "on-mixed" spike train of Fig. 13 and plot the values  $(v_1, v_2)$  corresponding to an output spike on the  $(v_1, v_2)$  plane. The result is shown in Fig. 14 where



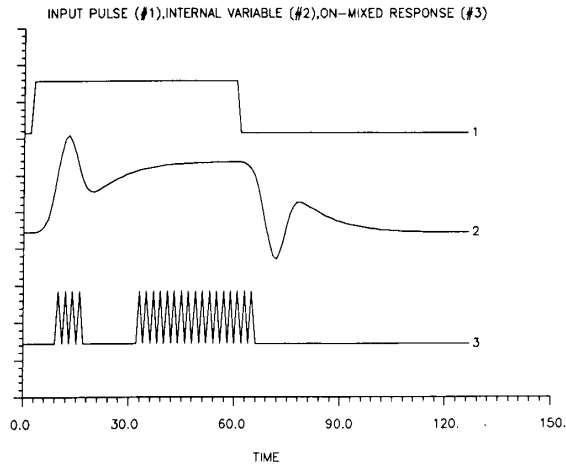


Fig. 13. The pulse input (trace 1), the combined internal variable (trace 2) and the "on-mixed" spike output (trace 3) (details in the text).

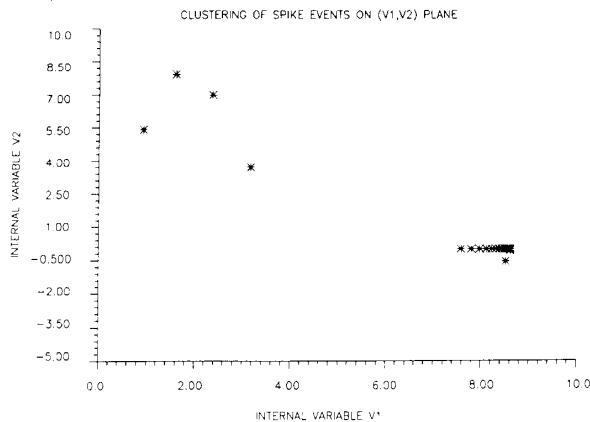


Fig. 14. Combinations of  $(v_1, v_2)$  values leading to spike generation for the example of Fig. 13. Two clusters of  $(v_1, v_2)$  trigger values appear.

the abscissa is  $v_1$  values and the ordinate  $v_2$  values. The combinations of  $(v_1, v_2)$  values that lead to spike generation do cluster in two groups. The upper left cluster corresponds to high  $v_2$  values (coding significant positive rate of change in the input signal) and the lower right cluster corresponds to high  $v_1$  values (coding significant magnitude of the input signal). These two clusters could be delineated by higher level neurons with appropriate dynamic characteristics, as suggested before, leading to extraction of input features. Of course, this simple case would be hardly worth the effort (i.e., this particular task could be accomplished with much simpler procedures); however, this example illustrates the possibilities for higher level decoding of more complex input information.

Signal reconstruction of the input pulse on the basis of the aforementioned spike outputs can be easily accomplished through appropriate integration and linear combination of the resulting continuous-time signals (analog decoding). However, the importance of analog decoding

at higher levels of the nervous system is debatable, with the odds favoring a form of discrete (or symbolic) decoding such as the one outlined heuristically before, utilizing feature extraction and construction of (pre-)cognitive primitives.

## V. CONCLUSIONS

The issue of signal transformation and coding by neural units (neurons) was studied in the context of nonparametric nonlinear models and continuous-time (graded potentials) and spike-train (action potentials) signals. The postulated model is based on notions developed from existing neurophysiological knowledge, and constitutes a reasonable compromise between mathematical complexity and neurophysiological relevance. The model utilizes a small number of "dynamical modes" (representing the important overall dynamic characteristics of the neuron) and a multidimensional static nonlinearity (representing the overall nonlinear characteristics of the neuron, including the threshold mechanism by which action potentials are generated). The immense variety of individual neuron characteristics (in terms of geometric, synaptic, and biophysical characteristics) leads to a similar variety of dynamical modes and static nonlinearities that may be appropriate in different cases. Nonetheless, the postulated model offers a common analytical (and simulation) framework that allows a unified approach to the problem of signal transformation and coding by different types of neurons, and incorporates nonlinear dynamics and spike generation in a fairly general, yet parsimonious manner. A plausible theory of higher level coding was heuristically outlined based on "modal mappings" associated with the proposed model and leading to clustering of spike events.

The actual identification of the minimum number of required modes for the representation of a neural unit remains a formidable task in practice. The complexity of this task is compounded by the presence of data-contaminating noise and experimental limitations in applying and measuring the appropriate input-output signals. One possible approach to this problem, based on measurements of the first two kernels, is briefly outlined in Section III. However, the thrust of this paper is the proposed modeling approach and the specifics of the associated identification problem are not adequately addressed herein.

This paper was limited to neural units (neurons) with the rationale that accurate understanding of the functional properties of individual neurons will facilitate the study of neural aggregates and allow, ultimately, the understanding of the functional organization of "integrated" neural systems of greater complexity. This transition is, of course, a rather formidable challenge and this paper represents only a very small step in a very long journey.

## REFERENCES

- [1] A. G. Bose, "A theory of nonlinear systems," Tech. Rep. 309, Research Laboratory of Electronics, M.I.T., Cambridge, MA, 1956.

- [2] K. S. Cole, *Membranes, Ions, and Impulses*. Berkeley, CA: Univ. California, 1968.
- [3] S. Deutsch and E. Micheli-Tzanakou, *Neuroelectric Systems*. New York: New York Univ., 1987.
- [4] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *J. Physiol.*, vol. 117, pp. 500-544, 1952.
- [5] B. Katz and R. Miledi, "A study of synaptic transmission in the absence of nerve impulses," *J. Physiol.*, vol. 192, pp. 407-436, 1967.
- [6] K. N. Leibovic and N. H. Sabah, "On synaptic transmission, neural signals and psychophysiological phenomena," in *Information Processing in the Nervous System*, K. N. Leibovic, Ed. New York: Springer, 1969, pp. 273-292.
- [7] R. Llinas, C. Nicholson, J. A. Freeman, and D. E. Hillman, "Dendritic spikes and their inhibition in alligator purkinje cells," *Science*, vol. 160, pp. 1132-1135, 1968.
- [8] P. Z. Marmarelis and V. Z. Marmarelis, *Analysis of Physiological Systems: The White-Noise Approach*. New York: Plenum, 1978.
- [9] P. Z. Marmarelis and K.-I. Naka, "Identification of multi-input biological systems," *IEEE Trans. Biomed. Eng.*, vol. BME-21, pp. 88-101, 1974.
- [10] V. Z. Marmarelis, Ed., *Advanced Methods of Physiological System Modeling, Vol. 1*. Los Angeles, CA: Biomed. Simulations Res., 1987.
- [11] V. Z. Marmarelis, M. C. Citron, and C. P. Vivo, "Minimum-order Wiener modeling of spike-output systems," *Biol. Cybern.*, vol. 54, pp. 115-123, 1986.
- [12] N. Wiener, *Nonlinear Problems in Random Theory*. Cambridge MA: M.I.T. Press, 1958.
- [13] W. Rall, "Theory of physiological properties of dendrites," *Ann. NY Acad. Sci.*, vol. 92, pp. 1071-1092, 1962.
- [14] W. Rall, "Core conductor theory and cable properties of neurons," in *Handbook of Physiology, Section 1: The Nervous System, Vol. 1: Cellular Biology of Neurons*, E. R. Kandel, Ed., Bethesda, MD: Amer. Physiol. Soc., 1977, pp. 39-98.



**Vasilis Z. Marmarelis** (M'79) was born in Mytileni, Greece, on November 16, 1949. He received the Diploma in electrical and mechanical engineering from the National Technical University, Athens, in 1972 and the M.S. and Ph.D. degrees in engineering science (bioinformation systems) from the California Institute of Technology, Pasadena, in 1973 and 1976, respectively.

After two years of postdoctoral work at the Department of Bioinformation Systems, California Institute of Technology, he joined the faculty of Biomedical and Electrical Engineering, University of Southern California, Los Angeles, where he is currently an Associate Professor and Director of the Biomedical Simulations Resource. His main interests are in the areas of nonlinear and nonstationary system identification and modeling, with applications to biomedical systems. Other interests include spatio-temporal nonlinear and nonstationary signal processing with applications to physiological and artificial vision, and analysis of spike-train data encountered in neural systems. He is coauthor of the book *Analysis of Physiological Systems: The White-Noise Approach* (New York: Plenum, 1978; Russian translation: Moscow: Mir Press, 1981) and editor of the book *Advanced Methods of Physiological System Modeling* (Los Angeles: Biomedical Simulations Resource, 1987).