Biol. Cybern. 56, 261–268 (1987)

Biological Cybernetics

Identification of MGB Cells by Volterra Kernels

III. A Glance Into the Black Box*

Y. Yeshurun[1,**], N. Dyn[1], and Z. Wollberg[2]

[1] School of Mathematical Sciences, Tel Aviv University, Israel
[2] Department of Zoology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Israel

Abstract. Neuronal systems can be described by their transfer functions, which can be represented by a Volterra series expansion. While the high level of abstraction which characterizes this representation enables a global description, it is problematic, to some extent, in the context of linking the formal representation of the system to its actual structure. The formal representation is unique, yet there are multiple physical realizations of this representation. Separating the system's output into its logical components (linear, cross-linear, and self nonlinear, in this study), and inspecting the relative contribution of these components, might provide a key towards a linkage between the formal and actual representations. Based on results drawn from identification of MGB cells of the squirrel monkey, it is shown that the relative contributions can be described in neurobiological terms such as excitation and inhibition and thus be attributed to actual subsystems.

Introduction

In an attempt to characterize the role of the Medial Geniculate Body (MGB) of the squirrel monkey in auditory signal processing, we have recently described a method for obtaining the first and second order kernels of the system which culminates in a single MGB cell (Yeshurun et al. 1985), and presented some findings based on it (Yeshurun et al. 1987). According to our approach, data concerning early auditory processing (i.e. spectral separation taking place between the ear and the auditory nerve) is incorporated into the model, so that the input to the system is not the raw signal impinging on the ear, but rather a multi input representation, consisting of 6 inputs or 18 inputs, for a spectral resolution of 1 or $^1/_3$ octave, respectively. Obviously, any process which can be directly described and be removed from the "black box", reduces the complexity of the system and facilitates the interpretation of the results, obtained through the process of identification.

In neurobiological applications of system identification methods, attempts were made to define the relationships between the kernels and the inner logical structure of the systems under study (e.g., Poggio and Reichardt 1976). These efforts concern mainly systems where the information about the inner structure is available. For example, if it can be assumed that a system consists of two building blocks, say a rectifier and an adder, then it is possible to decide according to the kernels, if the rectifier precedes the adder or vice versa (Hung et al. 1977). This can be done mainly when dealing with a small number of neurons, whose connections are well studied (Marmarelis and Naka 1973).

The process of exploring and understanding the nature of information processing in a "Black Box" by Volterra methods, usually faces a fundamental difficulty. These methods are aimed at a mathematical description of the input to output transformations in the explored system. In an ideal case, when the system satisfies the related mathematical requirements (Volterra 1930) and can formally be represented by its transfer function (kernels), this representation is valid and unique. On the other hand, the relation between these kernels and the inner structure of the system is not at all unique: while the kernels are unique in the functional aspect, there might be numerous realizations of the implied transformation, all of them valid

* Supported by a grant No. 84/B from "The Israel center for psychological biology", Charles E. Smith foundation, and by the R. and J. K. Field foundation for studies in Neurophysiology
** Present address: Brain Research, Computational Neuroscience Labs, Department of Psychiatry, NYU Medical Center, 550 First Ave., New York, NY 10016, USA

and having the potential of being the actual realization. However, when data concerning the inner mechanisms of the system is available, the gap between the functional and the structural descriptions can be narrowed, and speculations concerning this structure, based on the kernels, might be proposed, and be used as a basis for further working hypotheses.

It is possible to investigate the inner mechanism of a system not only by its kernels, but also by its kernels' products, by the predicted responses. Theoretically, a description of a system in the kernel space is equivalent to its description by the predicted responses space. The kernel description is more convenient, since it has a compact form, and we avoid the need to consider a large set of possible outputs of the system. On the other hand, the main advantage of using the predictions as a tool for peeking into the system, resides on the fact that we are more familiar with the system's types of responses and usually known how to "read" their meaning much better.

Due to these reasons, namely, lack of detailed information concerning the inner structure, and the experience gained in the analysis of auditory neurons' responses to vocalizations, we investigate the inner mechanisms of the system under study mainly by the "prediction space" approach, although some analysis of the kernels themselves is also represented.

Linear and Nonlinear Kernels

The system being identified is represented in this study (see Yeshurun et al. 1985) by the two leading terms of the Volterra representation:

$$Y(t)=Y_l+Y_q, \tag{1}$$

where

$$Y_l=\sum_{r=1}^{n}\int_0^M H_r(\tau)x_r(t-\tau)\mathrm{d}\tau \tag{2}$$

$$Y_q=\sum_{r=1}^{n}\sum_{|s-r|\leq 1}\int_0^M\int_0^M H_{rs}(\tau_1,\tau_2)x_r(t-\tau_1)x_s(t-\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2\,. \tag{3}$$

Here M denotes the length of the memory, and n is the number of inputs. Identification of multi input systems by kernels of first and second order, consist of the computation of three main types of kernels (Marmarelis and Naka 1974; Windhorst et al. 1983). These kernels are: linear ($H_r(\tau)$), cross quadratic ($H_{rs}(\tau_1,\tau_2)$ $|s-r|=1$ in this study), and self quadratic ($H_{rr}(\tau_1,\tau_2)$). The cross quadratic kernels are convolved with two components (x_r and x_s) of the input, and the self kernels are convolved twice with the same single component (x_r). Notice that since the time resolution used for the inputs is 3 ms (Yeshurun et al. 1985), we consider only global effects of the kernels, and not time-dependent fine details.

The linear kernels (Fig. 1) have, in most cases, the form of damped oscillations, with the peak near zero (up to 9 ms) and vanishing time of about 5–10 ms. The linear kernels resemble each other, in general, and differ only in the amplitude and the sign of its extremum values. The comparison between the contribution of the linear and the second order ones is discussed elsewhere (Yeshurun et al. 1987).

Regarding the second order kernels, cross and self kernels differ from each other in several aspects. When compared by eye, it is very clear that the amplitudes of the self kernels are greater than those of the cross kernels. Quantifying this difference, and taking

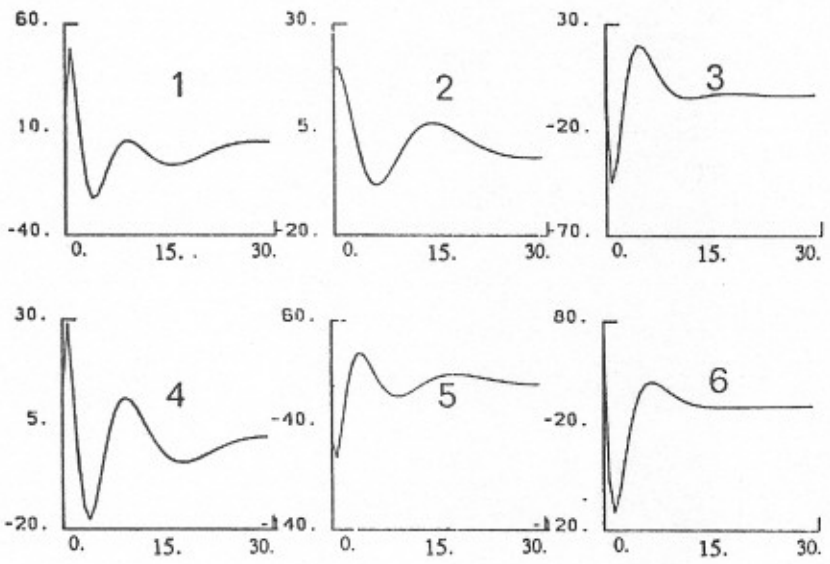

Fig. 1. Linear kernels of a cell. Denote the spectral band of the kernel (here $n=6$). The units on the *horizontal axis* are milliseconds, and on the *vertical axis* relative amplitude of kernels. Notice that the vertical axis is scaled for each kernel

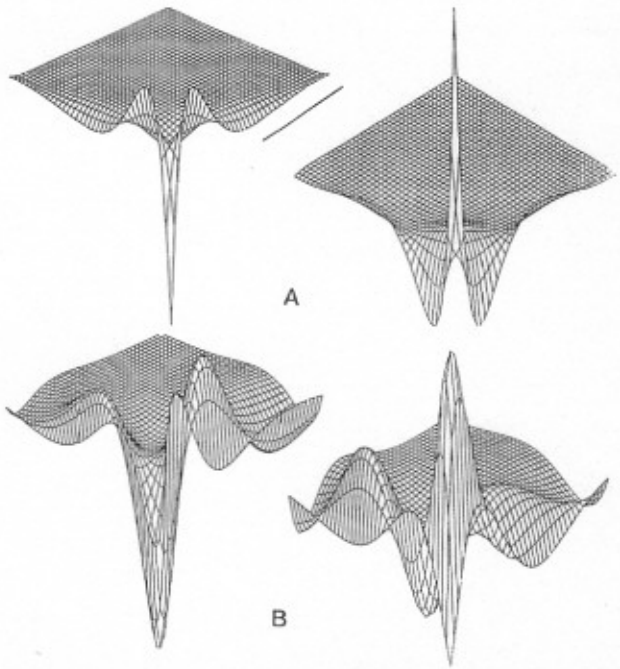

Fig. 2. A Typical self kernels. Two (different) kernels are described by a 3*D* perspective view. The view is from behind the origin of the $\tau_1-\tau_2$ plane, τ_1 is to the right. Time scale -9 ms. **B** Cross kernels, details as in **A**

$\max|H(\tau_1,\tau_2)|$ as a measure of the amplitude, it turned out that the self kernels are larger by 1–2 orders of magnitude then the cross kernels, and this holds for all the 41 cells in our sample. The self kernels are "bigger" than the cross kernels even if some other norms (as the L_2) are employed.

Another aspect, by which the cross and self kernels differ from each other, is their structure. The self (Fig. 2A) and the cross (Fig. 2B) have a typically different form. In most of the cases, the self kernels have their extremal amplitude (positive or negative) immediately after the rise time (up to 9 ms), with oscillations which are rapidly damped in all directions (since the self kernels are symmetrical). Such kernels can be considered as a crude approximation of a two dimensional δ function, which represents, as a kernel, a mechanism whose output resembles a mere time shift of the input.

On the other hand, the cross kernels are more complex, and have no common outstanding feature is the self kernels. The extremal amplitude does not always occur immediately after the rise time, and even if it is, attenuation is much slower, and in some cases it lasts more than 10 ms. The kernel demonstrated in Fig. 2B is characterized by the fact that its amplitudes in $\tau_1<\tau_2$ are greater than in $\tau_1>\tau_2$. Such a structure, in its extreme form, namely, when $H_{rs}(\tau_1,\tau_2)=0$ for $\tau_2\neq 0$ represents a system where the contribution of the interaction between input's components x_r and x_s at time t depends on the past values of x_s, $x_s(\tau)$, $\tau\leqq t$, and on the instanteneous value of x_r, $x_r(t)$.

The major differences in amplitude and structure between cross and self kernels, are apparently pointing towards the existace of two types of transformation taking place in the processing of single component channels (self kernels), and multi component channels (represented by cross kernels). In single component channels (i.e. limited band), the transformation is relatively simple, and has a short memory. A possible neuronal realization would consist of a relatively small number of cells through which the information is processed, until it is reflected and recorded in the MGB. The realization of the cross component channels, involves more complex transformations and has a longer memory, and therefore consists of more "processors" (cells?) along its pathway. The possible neuronal correlate of these "processors" is, obviously, speculative in nature, and may consist of cells spatially located at various sites of the auditory pathway, including the MGB.

Relative Contributions of Cross and Self Kernels

As mentioned earlier, the maximal amplitudes of the cross kernels are about 1–2 orders of magnitude smaller than these of the self kernels. A question thus arises, what does this finding mean, concerning the inner mechanisms of the system? At a first glance, it would suggest that the contribution of the cross kernels to the system, compared to that of the self kernels, is negligable. In order to test this hypothesis, we separated the total predicted response of the system into its components, and evaluated the relative contributions of the cross and self components:

$$Y_c=Y_c(t)=\sum_{r=2}^{n}\int_0^M\int_0^M H_{rr-1}(\tau_1,\tau_2)x_r(t-\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2\,,$$

$$Y_s=Y_s(t)=\sum_{r=1}^{n}\int_0^M\int_0^M H_{rr}(\tau_1,\tau_2)x_r(t-\tau_1)x_r(t-\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2\,.$$

These contributions were computed for all the cells. It was found, that in spite of the differences in the kernels' amplitudes, the contribution of the cross kernels to the output was not negligable, as compared to the contribution of the self kernels. Actually, in some cases it was even greater (Fig. 3). The fact that the contributions Y_c and Y_s were not proportional to the amplitudes of the kernels can be explained by the oscillatory nature of the kernels which might have a cancelation effect in the convolution with the vocalizations. It should be stressed that there might be types of inputs for which this cancelation does not happen (these can be arbitrarily tailored input functions), and then, an amplitude of a given kernel actually reflects its corresponding

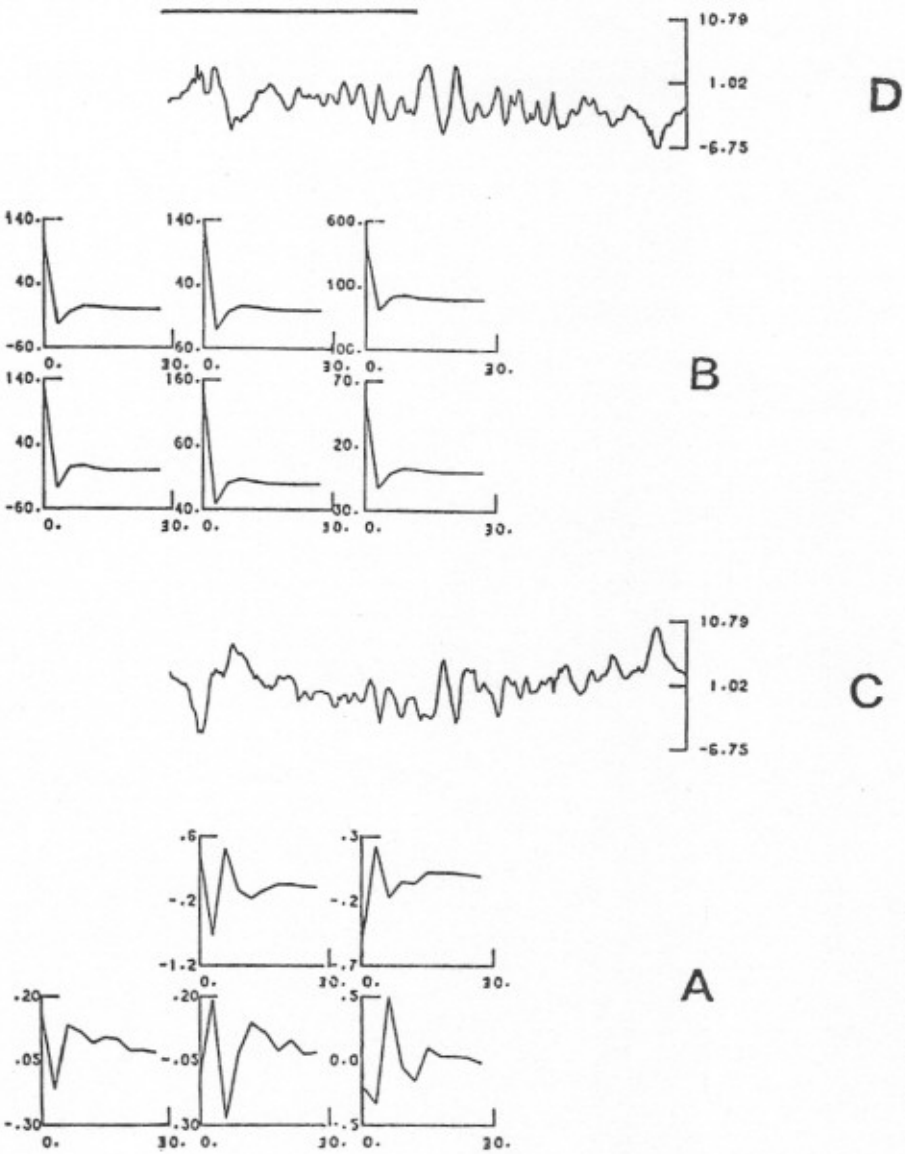

Fig. 3A–D. Relative contributions of self and cross kernels to the output of the model. **A** and **B** amplitude of cross and self kernels respectively (for each τ_1, the extremal value of $H(\tau_1, \tau_2)$ over τ_2 is drawn. Time is ms. **C** and **D** summed contribution of the cross and self kernels, respectively. Time scale (denoted in **D**) −400 ms

contribution. However, natural vocalizations comprises a highly significant subset of the possible inputs to the studied system, hence, it can be concluded that there is no trivial connection between the amplitude of a kernel and its relative contribution.

While analysing these relative contributions, we noticed that the cross and self components of the predicted response are related in a rather typical manner (e.g., Fig. 4). In many regards, though not totally, they look like a mirror image of each other, and this is even clearer if one considers only the baselines of the responses and neglectw the superimposed high frequencies.

Obviously, Y_c and Y_s are not exactly symmetrical (for they would cancel each other completely), but the trend is clear. In Fig. 4 one can also see that the amplitudes of $|Y_c(t)|$ and $|Y_s(t)|$ are high, relative to the total predicted responses $Y(t)$, while $Y_c(t) + Y_s(t)$ is of the order of magnitude of the total response, as is the linear contribution Y_l. It should be stressed, that since all the numerical calculations are made with 14 (and in critical loops with 28) significant digits, and the ratio of the partial contribution to the net outcome is less than 10:1, it is not probable that the cancellation and "mirror" effect cause severe numerical errors.

It is convenient to classify a component (e.g., Y_c, Y_s, Y_l) of the predicted response as "excitatory" if it resembles the positive part of the envelope of the input which elicits it, and as "inhibitory" if it resembles the negative part of this envelope (Fig. 5). The meaning we

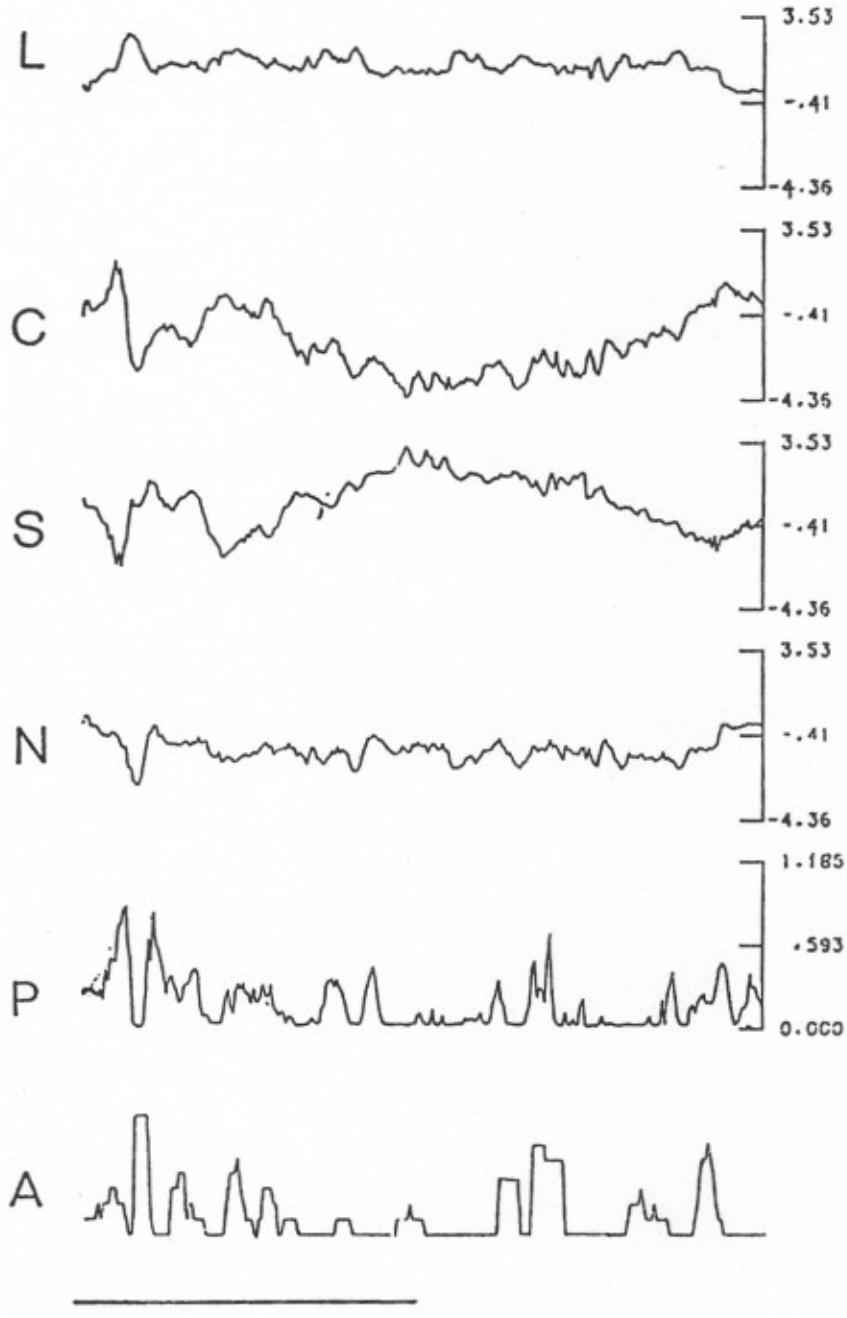

Fig. 4. Components of a predicted response. *A* actual response, *P* predicted response, *N* nonlinear contribution to *P*, *S* self contribution, *C* cross contribution, *L* linear contribution. Time scale −400 ms

assign to the terms "excitatory" and "inhibitory" is not identical to the conventional neurobiological meaning, but is a generalization of these concepts, since actual inhibition can be manifested by "zero" activity and not by a "negative" activity. In most cases, Y_c and Y_s can be classified as "inhibitory" or "excitatory" according to this definition, yet there are cases where inhibitory and excitatory components are intermingled.

We found that in many cells the cross contribution, Y_c, is opposite, according to this classification, to the self contribution Y_s. This phenomenon is most prominent in those cases where the baseline of the input, namely, the envelope of the call, is clearly convex or concave (like the Shriek). However, when it exists and is manifested clearly in the prediction to this one call, it exists, though less prominently, in all other predicted response of the cell. This is also true in responses to vocalization which were used for the calculation of the kernels. The "symmetry" of the cross and self contribution is, therefore, a feature of the system, namely, it can be attributed to the neuronal system and the type of vocalizations which give rise to these responses.

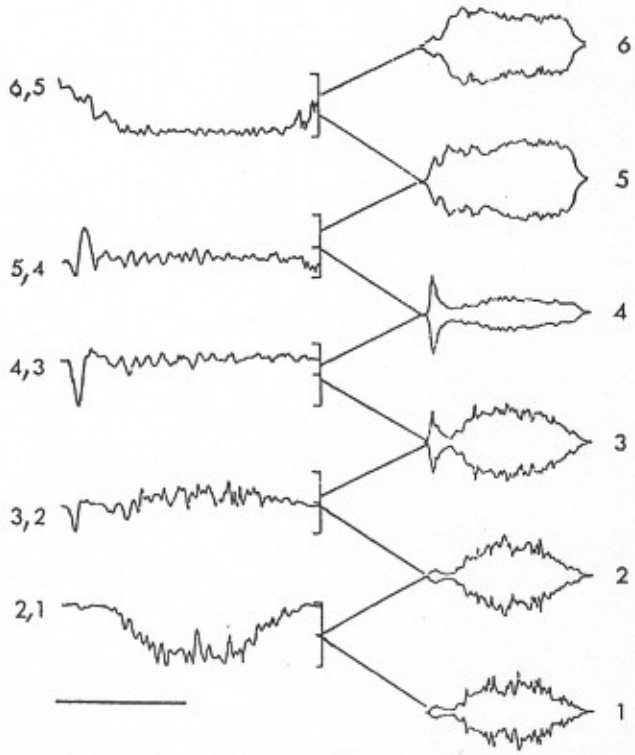

Fig. 5. Demonstration of the relations between envelopes of the input's components and response's components. *Left:* contributions of cross kernels, numbered on their left. (i,j) means the contribution of cross kernel of band i and band j. *Right:* filtered envelopes of spectral bands, bands' numbers on their right. Notice, for example, the influence of band 3 on the first part of (3,2) and band 4 on (4,3) and (5,4). Notice also that the similarity can be in reversed polarity [e.g. (4,3) and (4,5)]. Time scale −400 ms

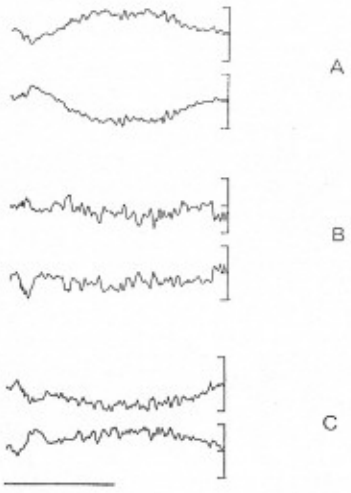

Fig. 6A–C. Categorization of 3 responses' components. Each of the 3 frames consists of cross (*upper*) and self (*lower*) contributions. **A–C** "excitatory", unidentified and "inhibitory" cross contribution, respectively. Time scale −400 ms

Table 1. Distribution of quality of predictions between 3 types of cells – "inhibitory cross", "excitatory cross", and unidentified. Mean quality (mean Q) is the sum of distances between the actual and predicted responses for each type, which is evaluated by a modified MSE (Yeshurun et al. 1985). *sd*: standard deviation. Significance level of the one way variance of analysis = 0.1

Cross type	*n*	Mean Q	*sd*
Inhibitory	25	4.92	0.77
Excitatory	10	5.51	0.82
Ambiguous	6	5.78	1.57

More formal approaches to the classification of Y_s and Y_c were also tested. We approximated the responses to the vocalization "Shriek" (which has a concave baseline) by a second order polynomial, fitted by least squares, and used the coefficient of the quadratic term of this polynomial as a measure to the convexity ("inhibitory") or concavity ("excitatory") nature of the response. This measure was found to yield identical results to these achieved by the described "manual" classification.

In the analysis of the cross and self contributions for all 41 sampled cells, inhibitory cross contributions were frequently encountered, along with excitatory self contributions. The cells were classified into three major classes: inhibitory cross, excitatory cross, and unidentified (Fig. 6). Table 1 summarizes the distribution of these three types. It can be seen that the impression of a relatively higher incidence of inhibitory cross contributions is indeed confirmed also if the averages are considered. When the quality of the prediction is tested vis a vis the distribution of inhibitory cross contribution, it can be seen (Table 1) that, on the average, inhibitory cross cells have better predictions (low distance values) than excitatory cross cells. Considering the quality of a predicted response as a measure of the degree to which a cell can be characterized by a second order Volterra representation, our findings suggest that whenever a cell is well described by the model – the cross contribution to the total response is inhibitory.

Subsystems of the Black Box

The Volterra kernels supply a functional description of systems only, and there are no clear cut rules which relate the features of the kernels to unique structures of the system. We face an even more complex situation when the borders of the system themselves are very broad, as is the case in our model. However, inspite of this uncertainty, it is possible to point at some structural conclusions.

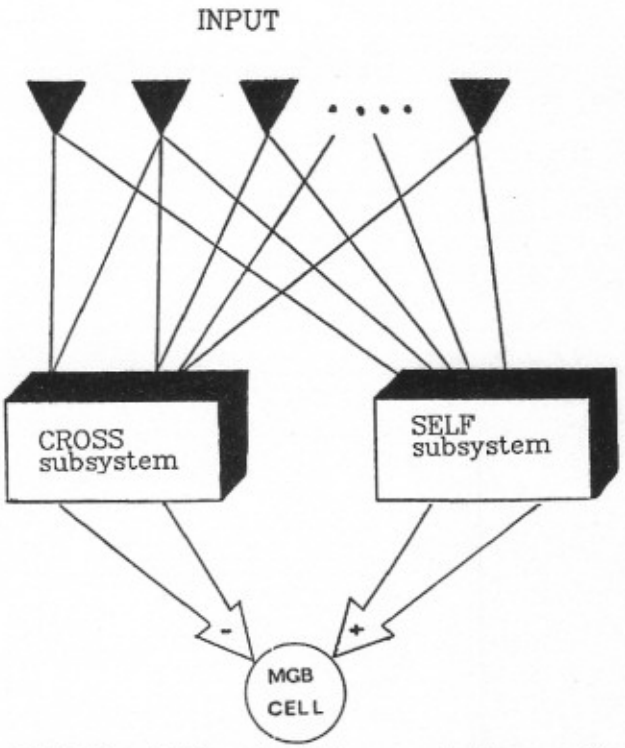

Fig. 7. Schematic illustration of the proposed subsystems. The input is passed through band pass channels (designated by the *triangles*). The "cross" subsystem involves multiplication of signals from multi channels (only nearest neighbor in our model) and summation. The "self" subsystem involves the multiplication of signals from a given channel by themselves and summation. From the viewpoint of the cell which is being recorded (designated by a *circle*), the output of the "cross" subsystem is inhibitory, while the output of the "self" subsystem is excitatory

In order to do it, we must assume the following:

1. The multi input system has a physical realization, namely, there are specific channels for band pass spectral components of the input signal. This assumption follows the current theory of the auditory system (e.g., Evans 1977), although reality is much more complicated then the octave bands, or even the non overlapping $^1/_3$ octave bands, of our model.

2. The self kernels and cross kernels represent some actual subsystems. This involves the existance of subsystems (not necessarily spatially adjoined, and not necessarily consisting of adjacent cells), where only inputs from a narrow spectral band is processed, and subsystems where interactions of inputs from neighboring spectral bands are processed. This assumption is not based on physiological evidence, but does not violate any known knowledge, and can be well integrated with the structure implied by the first assumption.

We assume that the MGB cell under investigation receives its inputs from several neural subsystems, and in particular, from a "cross" subsystem and a "self" subsystem (Fig. 7). These subsystems perform nonlinear transformations on interacting neighboring spectral bands (cross) and on single bands (self). The typical structure of the kernels imply that the single input transformation (self) is relatively simple (can be approximated by a two-dimensional δ function), and has a short memory. The processing of neighboring channels of input (cross), on the other hand, is more complex and has a longer memory.

The inhibition-excitation phenomenon of the cross and self kernels is very similar to its physiological correlates, even if it is not identical, and calls for a physiological explanation. The fact that in many cases, the output of the cross subsystem is inhibitory, can be related to a very known processing mode where cross interaction of neurons is inhibitory, namely, lateral inhibition. Our model does not handle the processing level of single axons, where the classical lateral inhibition occurs. However, as illustrated in Fig. 7, the interaction between several components of the input yields primarily inhibitory effects, whereas the processing of a single input component is mainly excitatory. This phenomenon can be operationally realized by a "lateral inhibition like" mechanisms. One should bear in mind, however, that linking the difference between cross and self contribution to excitatory and inhibitory responses of the suggested subsystems is speculative in nature, and while being a plausible explanation of the results, is by no means the only one.

As we deal with the inner structure of the "black box", an obvious direction is to search for the "real" subsystems where the actual processes takes place. The natural candidates for such a research are the Inferior Coliculus and the Auditory Cortex. Undoubtedly, at present, it is still premature to point at the exact neurobiological context of these subsystems. Yet, it is a plausible idea that such subsystems do exist, and that they can be described and detected by Volterra methods.

References

Evans EF (1977) Peripheral processing of complex sounds. In: Bullock TH (ed) Dahlem workshop on recognition of complex acoustic signals, pp 367–386

Hung G, Stark L, Eykhoff P (1977) On the interpretation of kernels. Ann Biomed Eng 5:130–143

Marmarelis PZ, Naka KL (1973) Non linear analysis and synthesis of receptive fields responses in the catfish retina, horizontal cell-ganglion cell chain. J Neurophysiol 36:605–618

Marmarelis PZ, Naka KI (1974) Identification of multiinput systems. IEEE Trans BE-21:88–101

Poggio T, Reichardt W (1976) Visual control of orientation behaviour in the fly. II: Towards the underlying neural interaction. Q Rev Biophys 9:377–448

Volterra V (1930) Theory of functionals. Blackie and Sons, Glasgow

Windhorst U, Niemann U, Koehler W (1983) Analysis of nonlinear physiological systems with single or multiple spike inputs and analogue or spike outputs. Biol Cybern 48:159–163

Yeshurun Y, Wollberg Z, Dyn N, Allon N (1985) Identification of MGB cells by volterra kernels. I. Prediction of responses to species specific vocalizations. Biol Cybern 51:383–390

Yeshurun Y, Wollberg Z, Dyn N (1987) Identification of MGB cells by volterra kernels. II. Towards a functional classification of cells. Biol Cybern 56:203–208

Received: March 12, 1986/First revision: August 10, 1986/ Second revision: November 22, 1986

Dr. Y. Yeshurun
Brain Research
Computational Neurosciences Labs
Department of Psychiatry
NYU Medical Center
550 First Avenue
New York, NY 10016
USA