NONLINEAR STRUCTURES FOR SYSTEM IDENTIFICATION

By

R. HABER and L. KEVICZKY
Department of Automation, Technical University Budapest

(Received March, 1974)

Presented by Prof. Dr. F. CSÁKI

Introduction

The identification has three phases: approximation of the structure, parameter estimation and checking, but only the latter two have satisfactory — but not definite — theory. [1, 2]. Nowadays the main problem is to find the suitably fitting structure. The estimation of model order has already been investigated for linear systems [1, 3, 4, 5] but this is not the case for nonlinear dynamic systems. The reason may be the many variations of the model and noise structures. The task is of topical interest because the description of complex chemical and biological systems is based more and more on input-output modelling since the establishment and solution of reaction, material and energy equations are often difficult. In case of nonlinear structures the mathematical description and description for identification purpose are contradictory, namely each nonlinear dynamic process can be characterized by a VOLTERRA series of infinite number of parameters while only a finite number of parameters can be identified. Since different models can be established in case of a given number of parameters and among them the VOLTERRA series of finite parameter gives the poorest approximation, our aim was to elaborate different structures for identification purposes and at the same time to see how they approximate the general VOLTERRA series description.

Several algorithms have been elaborated for linear parameter estimation under noisy conditions. Since these assume only the parameters to be linear, they can also be used for nonlinear identification. Therefore our aim was to find (nonlinear) models which are linear in parameters. The algorithms have already been published in [6, 7].

The VOLTERRA-series expansion

A nonlinear static function can be approximated by its TAYLOR series in the vicinity of the working point, which is actually a polynomial of infinite order. Similarly an impulse response of "infinite extension" describes the linear dynamic systems. Generally the so-called VOLTERRA integral operator is used to describe nonlinear, dynamic, continuous systems. Since the discrete form is preferable for data processing by digital computers, further on we deal with such a description:

$$y(t) = r_0 + \sum_{i=0}^{\infty} w_i\, u(t-i) + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_{ij}\, u(t-i)\, u(t-j) + \dots$$
$$\dots + \sum_{i=0}^{\infty} \dots \sum_{k=0}^{\infty} w_{i\dots k}\, u(t-i) \dots u(t-k). \tag{1}$$

Here $u(t)$ and $y(t)$ are the input and output signals, respectively, at the t-th moment, assuming that the sampling time $T = 1$, r_0 is the constant term, w_i, w_{ij}, . . . are the generalization of the linear weighting function series, the kernels of the series.

The series is "twice infinite in dimension"; both in time, which derives from the series expanded form of the transfer function of dynamic part, and in the number of sums, which derives from the more and more accurate approximation of the nonlinear, static characteristics. At the same time, only a finite number of parameters can be estimated.

Further on, TAYLOR series of the static characteristics of the process is assumed to exist (that can be differentiated any times, has no break point or discontinuity). So the closeness of approximation depends only on the number of sums. A quadratic approximation better than the linear one can be obtained if only the kernels of second order are considered. Besides this the time-memory has also to be restricted.

In such a way "the finite order VOLTERRA weighting function model" is obtained (FVW):

$$y(t) = r_0 + \sum_{i=1}^{m} w_i\, u(t-i) + \sum_{i=1}^{m} \sum_{j=1}^{m} w_{ij} u(t-i)\, u(t-j). \tag{2}$$

Here $m < \infty$ is the degree of time-memory and the summing up is started from 1 which is usual for the difference equations describing the dynamic processes. Be

$$W(z^{-1}) = w_1 z^{-1} + \dots + w_m z^{-m} \tag{3}$$

where z^{-1} is the backward shift operator and $\mathbf{W}$ is an $m \times m$ symmetrical matrix. Rewriting Eq. (2)

$$y(t) = r_0 + W(z^{-1}) u(t) + \mathbf{f}^T(u(t)) \mathbf{W} \mathbf{f}(u(t)) \tag{4}$$

where

$$\mathbf{f}^T(u(t)) = [u(t-1), \dots, u(t-m)]\,. \tag{5}$$

The form (4) is linear in the parameters, i.e. the output can be given as a scalar product of the parameter vector **p** and the situation vector **g**:

$$y(t) = \mathbf{g}^T \mathbf{p} \tag{6}$$

where

$$\mathbf{p} = [r_0, w_1, \ldots, w_m, w_{11}, w_{12}, \ldots, w_{mm}]^T \tag{7}$$

and

$$g^T = [1, u(t-1), \ldots, u(t-m), u^2(t-1), u(t-1)u(t-2), \ldots, u^2(t-m)] \tag{8}$$

This was the first general form applied to estimate the nonlinear systems [9, 10, 11].

Further on, only quadratic models will be dealt with.

Simple nonlinear dynamic structures

A static second order polynomial function can be given in the form:

$$y(t) = c_0 + c_1 u(t) + c_2 u^2(t). \tag{9}$$

The weighting function series of a linear, dynamic, discrete-time model can be approximated by fractional polynomials

$$y(t) = \sum_{i=1}^{\infty} w_i u(t-i) \approx \frac{B(z^{-1})}{A(z^{-1})} u(t) \tag{10}$$

where

$$B(z^{-1}) = b_1 z^{-1} + \ldots + b_m z^{-m} \tag{11}$$

$$A(z^{-1}) = 1 + \tilde{A}(z^{-1}) = 1 + a_1 z^{-1} + \ldots + a_n z^{-n}. \tag{12}$$

These parts can be used to construct the simple cascade models i.e. the "simple WIENER model" (SW) (Fig. 1) and the "simple HAMMERSTEIN model" (SH) (Fig. 2). These models used to be described by introducing the auxiliary variable $v(t)$, and accordingly their identification is done by iteration [12, 8]. The HAMMERSTEIN model is seen to be linear in the parameters [13, 8]. It can be provided that this model is a special case of the so-called "generalized HAMMERSTEIN model" (GH) (Fig. 3) which may be considered a multiple-input single-output linear, dynamic system in the new vector space obtained by transforming the input signals [6, 14]. The output is described by

$$y(t) = c_0 + \frac{B_1(z^{-1})}{A(z^{-1})} u(t) + \frac{B_2(z^{-1})}{A(z^{-1})} u^2(t) \tag{13}$$

or

$$y(t) = -\tilde{A}(z^{-1}) y(t) + c_0^* + B_1(z^{-1}) u(t) + B_2(z^{-1}) u^2(t) \tag{14}$$

where

$$c_0^* = c_0 \left(1 + \sum_{i=1}^{n} a_i\right). \tag{15}$$

Including the functions of the output in the right side of (14), there is a nonlinear feedback and we get the "extended HAMMERSTEIN model" (EH) (Fig. 4):

$$y(t) = -\tilde{A}(z^{-1})y(t) + c_0^* + B_1(z^{-1})u(t) + B_2(z^{-1})u^2(t) + D(z^{-1})y^2(t). \tag{16}$$

Considering the original VOLTERRA series (1), also the cross-products of the input signals are seen there to occur, but not to occur in (13).

This fact inspired us to involve these terms, by analogy to the "finite order VOLTERRA weighting function model" (2), possible by including a shift-storage of finite elements. The so-called "finite order VOLTERRA model" (FV) is shown in Fig. 5, described as:

$$y(t) = -\tilde{A}(z^{-1})y(t) + c_0^* + B_1(z^{-1})u(t) + \mathbf{f}^T(u(t))\mathbf{B}_2\mathbf{f}(u(t)) \tag{17}$$

where $\mathbf{B}_2$ is an $m \times m$ symmetrical matrix. Similarly, the "extended finite order VOLTERRA model" (EFV) is (Fig. 6):

$$y(t) = -\tilde{A}(z^{-1})\, y(t) + c_0^* + B_1(z^{-1})\, u(t) + \mathbf{f}^T(u(t))\, \mathbf{B}_2\, \mathbf{f}(u(t)) + \\ + \mathbf{f}^T(y(t))\, \mathbf{D}\, \mathbf{f}(y(t)). \tag{18}$$

Here

$$\mathbf{f}^T(y(t)) = [y(t-1), \ldots, y(t-l)] \tag{19}$$

and $\mathbf{D}$ is an $l \times l$ symmetrical matrix.

Now let us return to the other cascade model, i.e. to the WIENER model. Its generalization, the "generalized WIENER model" (GW) is shown in Fig. 7. The "extended WIENER model" (EW) results from assuming different dynamic transfer lags before the multiplier. (Fig. 8).

After having reviewed the above eight simple nonlinear structures let us consider some typical criteria.

1. The extension can easily be given for higher order polynomial terms.

2. If $\mathbf{B}_2$ or $\mathbf{D}$ are diagonal matrices then the HAMMERSTEIN models are special cases of the corresponding VOLTERRA models.

3. The HAMMERSTEIN and the finite order VOLTERRA models are linear in the parameters, and can be reduced to form (6) contrary to WIENER models which are nonlinear in the parameters.

4. The simple and generalized WIENER models include the corresponding HAMMERSTEIN model (to be proved subsequently).

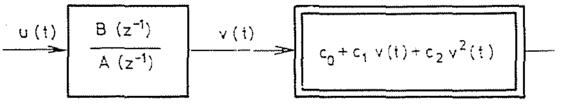

Fig. 1

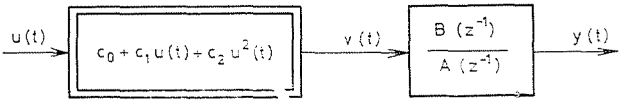

Fig. 2

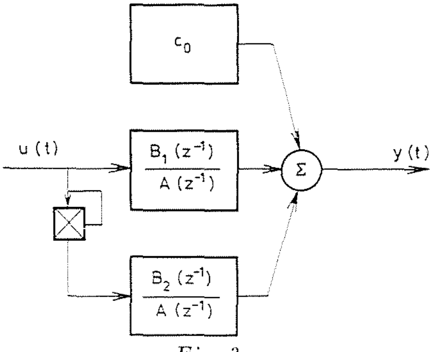

Fig. 3

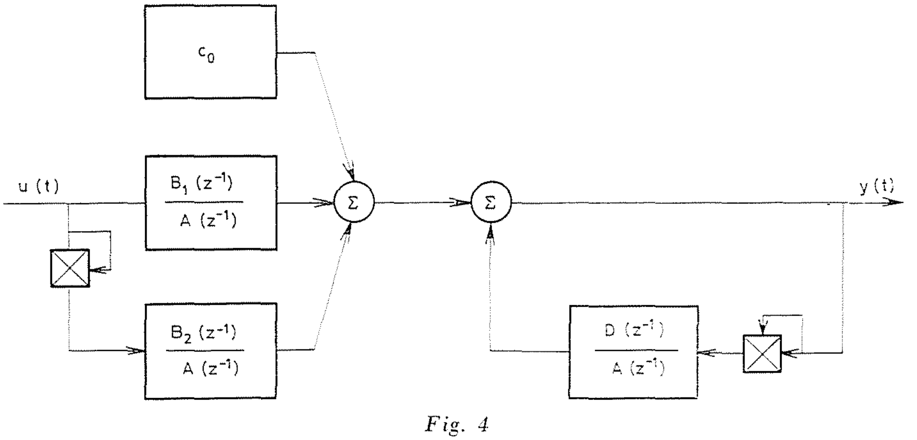

Fig. 4

5. Every mentioned structure — excepting the finite order VOLTERRA models — is a so-called separable model, i.e. nonlinear static and linear dynamic elements occur separately. Remark that only some of the physical systems can be written in this way, e.g. the inductance of the exciting coil of a generator which determines the time constant depends on the flux. If the structure of dependence of the model is known it may be transformed into a separable one that is, however more difficult to estimate.

The above mentioned models can be described by parameters of a finite number. Since these are to approximate unknown processes, it has to be known what general models can be handled by them. Even more than the analytical relation between the models and VOLTERRA series, it is of importance to know what types of models lend themselves to estimate certain terms of the VOLTERRA series.

Relationship between the model parameters and the kernels of the VOLTERRA-series

In case of linear systems, there are two possibilities to compare the transfer function consisting of fractional polynomials of finite elements to the weighting function series. The first possibility is the well-known division of polynomials

$$y(t) = \frac{B(z^{-1})}{A(z^{-1})} u(t) = b_1 u(t-1) + (b_2 - b_1 a_1) u(t-2) + \dots \tag{20}$$

We get the same result rewriting (10):

$$y(t) = -\sum_{j=1}^{n} a_j y(t-j) + \sum_{i=1}^{m} b_i u(t-i). \tag{21}$$

Replace the former values of the output into the right-hand-side of (21)

$$y(t) = -\sum_{j=1}^{n} a_j \left[-\sum_{j_1=1}^{n} a_{j_1} y(t-j-j_1) + \sum_{i_1=1}^{m} b_{i_1} u(t-j-i_1) \right] + \sum_{i=1}^{m} b_i u(t-i). \tag{22}$$

Carrying on the recursive procedure, it is seen that the argument of y tends to $-\infty$, and only the input signals remain in the equation in agreement with the fact, that the system is only excited by its input signal. Contracting the terms with identical arguments again leads to Eq. (20). The k-th component of the weighting function is

$$w_k = \sum_{s=0}^{k-1} b_{k-s} \sum \left[(-1)^{\mu} \prod_{(\Sigma i = s)} a_i \right] \tag{23}$$

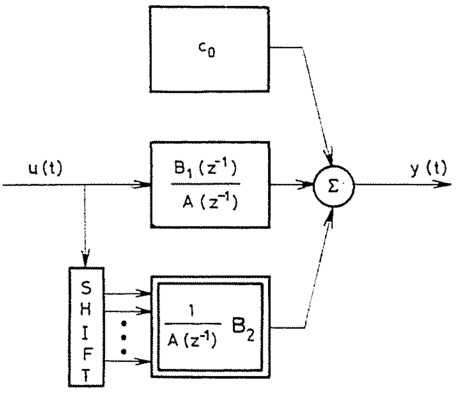

Fig. 5

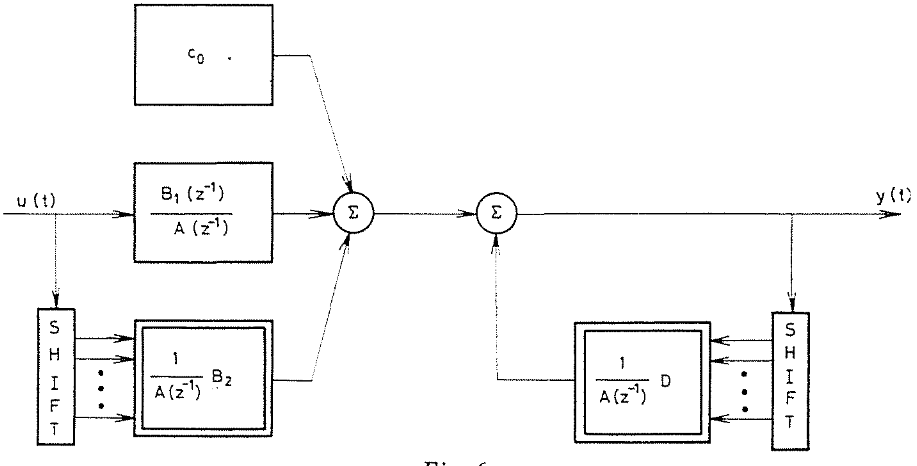

Fig. 6

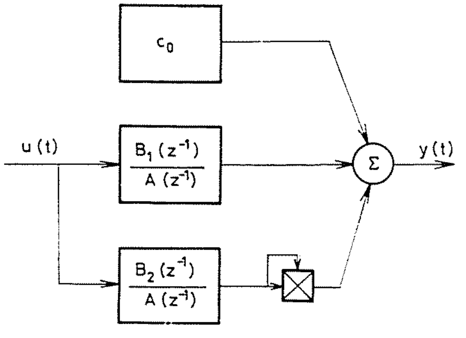

Fig. 7

where all possible products appear in the second sum, therefore some of them is repeated. μ is equal to 0 if the number of the a_i is even but it is equal to 1 if there are odd pieces of a_i in the product.

w_k is seen to exist for all positive k, in spite of

$$b_{k-s} = 0; \text{ for } k - s > m\,. \tag{24}$$

The recursive replacing procedure may be applied for any difference equation, also for a nonlinear one [15, 16].

First let us apply the procedure for the "finite order VOLTERRA model" (17):

$$\begin{aligned} y(t) &= -\sum_{j=1}^{n} a_j y(t-j) + c_0^* + \sum_{i=1}^{m} b_i u(t-i) + \sum_{k=1}^{m}\sum_{l=1}^{m} b_{kl} u(t-k)\, u(t-l) = \\ &= -\sum_{j=1}^{n} a_j \Bigg[-\sum_{j_1=1}^{n} a_{j_1} y(t-j-j_1) + c_0^* + \sum_{i_1=1}^{m} b_{i_1} u(t-j-i_1) + \\ &\quad + \sum_{k_1=1}^{m}\sum_{l_1=1}^{m} b_{k_1 l_1} u(t-j-k_1)\, u(t-j-l_1) + c_0^* + \sum_{i=1}^{m} b_i u(t-i) + \\ &\quad + \sum_{k=1}^{m}\sum_{l=1}^{m} b_{kl} u(t-k)\, u(t-l)\,. \end{aligned} \tag{25}$$

Carrying on the substitutions, let us compare the result with the form (1) of VOLTERRA series. For the constant element we get:

$$r_0 = c_0^* \left[1 - \sum_{j=1}^{n} a_j + \left(\sum_{j=1}^{n} a_j \right)^2 \mp \ldots \right] = c_0^* \frac{1}{1 + \sum_{j=1}^{n} a_j} = c_0\,. \tag{26}$$

This was expected from (15) but it has to be completed the assumption that

$$\sum_{j=1}^{n} a_j \neq 1 \tag{27}$$

holds what is generally true for stable systems. For the linear term we get (23). Finally, the expression obtained for quadratic terms:

$$w_{kl} = \sum_{s=0}^{\min(k,\, l)-1} b_{k-s,\, l-s} \sum \Big[(-1)^{\mu} \prod_{(\Sigma i = s)} a_i \Big]\,. \tag{28}$$

It is remarkable that

$$w_{kl} = 0; \text{ for } |k - l| > m \tag{29}$$

since $\mathbf{B}_2$ is an $m \times m$ matrix.

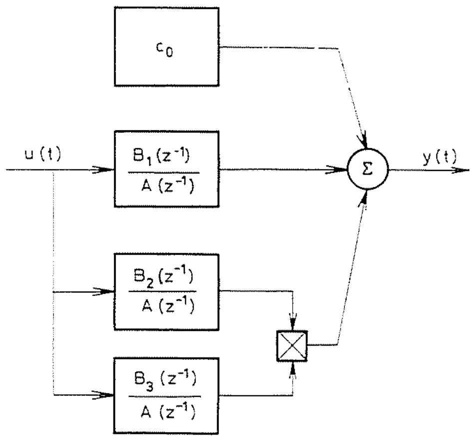

Fig. 8

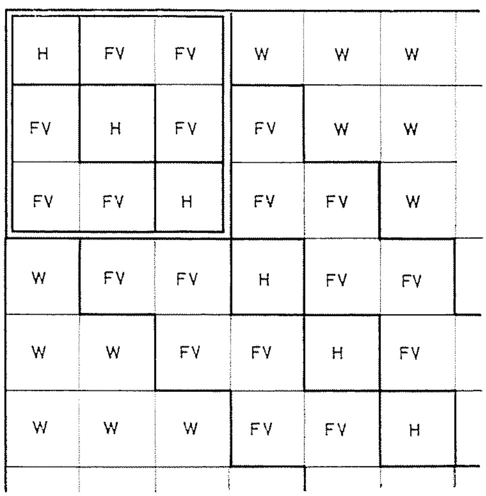

Fig. 9

Let us represent the elements w_{kl} in a two-dimensional screen (i.e. in an infinite matrix). The "finite order VOLTERRA model" takes account only of those quadratic terms of the VOLTERRA series for which Eq. (29) holds i.e. it estimates the band of width m along main diagonal line (see Fig. 9 – FV).

$\mathbf{B}_2$ being diagonal in the "generalized HAMMERSTEIN model":

$$w_{kl} = 0; \text{ for } |k - l| > 0. \tag{30}$$

Eq. (14) gives no estimation of terms of cross-product type of the input signal, only of the diagonal elements (in Fig. 9).

The "finite order VOLTERRA weighting function model", Eq. (2) is obtained from Eq. (17) under condition $n = 0$, so the model has no infinite time-memory any more, i.e. it estimates only m linear elements and the quadratic terms up to m order (see the framed square in Fig. 9). For $n > 0$, this model estimates with infinite time-memory independently of n but more precisely.

Let us consider "the extended WIENER model" (see Fig. 8). The constant and linear terms are obtained from (26) and (23) in this case, too. The quadratic terms can be established as a product of two linear series

$$w_{kl} = \sum_{s_1=0}^{k-1} \sum_{s_2=0}^{l-1} b_{2,k-s_1} b_{3,l-s_2} \tag{31}$$

$$\Big(\sum [(-1)^{\mu} \prod_{(\Sigma i = s_1)} a_i]\Big)\Big(\sum [(-1)^{\mu} \prod_{(\Sigma i = s_2)} a_i]\Big) .$$

This means that each quadratic term occurs, but the estimation has a low degree of freedom (see Fig. 9—W). In the case of the "generalized WIENER model" the number of parameter variations is decreasing:

$$b_{2i} = b_{3i} . \tag{32}$$

The extended models what are linear in parameters permit to approximate higher order kernels, by means of the quadratic terms of this estimation form.

Consider the "extended finite order VOLTERRA model" (18) by definition be:

$$x(t) = c_0^* + B_1(z^{-1})u(t) + \mathbf{f}^T(u(t))\mathbf{B}_2\mathbf{f}(u(t)) . \tag{33}$$

Introduce linear and nonlinear operators:

$$L(x) = -\tilde{A}(z^{-1})x(t) \tag{34}$$

and

$$N(x) = \mathbf{f}^T(x(t))\mathbf{D}\,\mathbf{f}(x(t)) \tag{35}$$

where

$$\mathbf{f}^T(x(t)) = [x(t-1), \ldots, x(t-l)] \tag{36}$$

and let:

$$F(x) = L(x) + N(x) . \tag{37}$$

Let us start to recursively replace the former values of $y(t)$ into the right-hand-side of (18). The k-th order approximation of the output (after k replacing steps) is obtained by omitting the former values of $y(t)$:

$$y_1(t) = x(t); \tag{38}$$

$$y_2(t) = x(t) + F(x) \tag{39}$$

$$y_k(t) = x(t) + F(x + F(x + \ldots + F(x))) . \tag{40}$$

Considering Eq. (37) and the fact that the superposition principle is valid for operator $L(x)$, we get:

$$y(t) = \sum_{i=0}^{\infty} L^i(x) + \Phi(x) \tag{41}$$

where the first term produces just the same kernels as the "finite order VOLTERRA model" (23), (26), (28), while $\Phi(x)$ contains the higher order kernels which can be computed recursively.

Without going into the details the products of the input signals are seen to occur at every degree but starting from ever earlier time-memories.

Conclusions

In our paper quadratic dynamic models are considered for identification purpose. Most of these models are to be linear in parameters and therefore the programs already available for multiple input — single output systems may be used to estimate them.

Relationships are given between the model parameters and the VOLTERRA series.

The extension for higher-order general polynomial forms follows logically from the involved statements.

In this paper no noise models have been dealt with but it has to be noted that the output noise models cannot be estimated by the extended models linear in parameters [17].

Summary

In this paper the simple nonlinear, dynamic process models are reviewed and the parameters of the equivalent VOLTERRA series are described.

The higher-order VOLTERRA series could be described with infinite time-memory by quadratic structures of finite elements linear in parameters.

References

1. ASTROM, K. J.—EYKHOFF, P.: System Identification.–A survey. IFAC, Prague. 1970.
2. BÁNYÁSZ, Cs.—KEVICZKY, L.: Identification of Linear Dynamical Processes Based upon Sampled Data I—II. Elektrotechnika 1975. (In press)
3. WOODSIDE, C. M.: Estimation of the Order of Linear Systems. Automatica, 1971. pp. 727—733.
4. UNBEHAUEN, H.—GÖHRING, B.: Application of Different Statistical Tests for the Determination of the most Accurate Order of the Model in Parameter Estimation, IFAC, Hague/Delft, 1973.
5. VAN DEN BOOM, A. J. W.—NAV DEN ENDEN, A. W. I.: The Determination of the Order of Process and Noise Dynamics. IFAC, Hague/Delft, 1973.

6. BÁNYÁSZ, Cs.—HABER, R.—KEVICZKY L.: Some Estimation Methods for Nonlinear Discrete Time Identification. IFAC, Hague/Delft, 1973.
7. BÁNYÁSZ, Cs.—KEVICZKY, L.—HABER, R.: Identification of Discrete Dynamic Systems with Separable Nonlinearity, 3rd All Union Conference on Statistical Methods in Control Processes. Vilnius, 1973.
8. HABER, R.—KEVICZKY L.: The Identification of the Discrete Time HAMMERSTEIN Model. Periodica Polytechnica — Electrical Engineering 1974. No. 1. pp. 71—84.
9. ALPER, P.: A Consideration of the Discrete VOLTERRA series. IEEE. Trans. on Automatic Control AC—10, 1965, pp. 322—327.
10. ZYPKIN, JA. S.: Adaptation und Lernen in kybernetischen Systemen. VEB Verlag Technik, Berlin 1970.
11. WESTENBERG, J. Z.: Some Estimation schemes for Nonlinear Noisy Processes. Th-Report 69-E-09, Eindhoven, 1969.
12. NARENDRA, K. S—GALLMAN, P. G.: An Iterative Method for the Identification of Nonlinear Systems Using a HAMMERSTEIN Model. IEEE Trans. on Automatic Control, AC—11, 1966, pp. 546—550.
13. CHANG, F. M. K.—LUUS, R.: A Noniterative Method for Identification Using HAMMERSTEIN Model. IEEE Trans. on Automatic Control, AC—16, 1971, pp. 464—466.
14. HABER, R.—KEVICZKY L.: The Extension of the Linear Discrete Time Identification Algorithms to Simple Nonlinear Plants. Measurement and Automation 1974 (In press).
15. CHRISTENSEN, G. S.—RAO, R. S.: On the Convergence of a Discrete Volterra Series. IEEE Trans. on Automatic Control 1970, pp. 40—41.
16. FU, F. C.—FARISON, J. B.: On the VOLTERRA Series. Functional Evaluation of the Response of Nonlinear Discrete-time Systems. International Journal of Control 1973, pp. 553—558.
17. HABER, R.—KEVICZKY, L.—FEDINA, L.—CSERNE, I.: Modelling of Ganglion by HAMMERSTEIN Model. 4-th IEE, Shiraz, 1974.

Róbert HABER
Dr László KEVICZKY } H-1521 Budapest