

## **SIMULATION OF BEHAVIOR AND INTELLIGENCE**

### **Contra-monotonic systems in the analysis of the structure of multivariate distributions**

J. E. Mulla

Department of Economics, Tallinn Technical University (1973 – 1980)  
Translated from *Avtomatica i Telemekhanika*, No. 7, pp. 167 – 175, July, 1981.  
Original article submitted June 9, 1980.

005 – 1179/81/4207 – 0986 \$7.50 1982 Plenum Publishing Corporation

1179

We alert the readers' obligation with respect to copyrighted material.

**Correspondence Address:** Byvej 269  
2650 Hvidovre  
Copenhagen Denmark  
Phone (45) 36 49 33 05  
E-mail: [joseph\\_mullat@email.dk](mailto:joseph_mullat@email.dk)

#### **Abstract**

The problem of distinguishing condensations in multivariate space of measurements based on a qualitative vector criterion is presented. We find solutions by a special parameterization of functions, the values of which decrease in all regions of the definition in inverse proportion to the values of the parameters.

Keywords: monotonic, distributions, equilibrium, Nash, cluster

## 1. Introduction

The analysis of the structure of the probability density function of measurements in an  $n$ -dimensional space is a traditional topic of investigation in such applied fields as experimental design [1], image analysis [2], the analysis of decision making [3], pattern recognition [4], etc.

At the conceptual level, the structure of a distribution is customarily represented by a set of data clusters, sometimes called modes [5]. The analysis of such a structure, indirectly if not explicitly, is usually reduced to an optimization variational problem. i.e., the maximization of some scalar performance indexes characterizing the identified clusters. Instead of scalar performance index, in this article we use a vector index, and base the concept of optimality on the so-called Equilibrium State in the sense of Nash [6].

Approaching the analysis of the structure of a measurement density function in  $n$ -dimensional space, our standpoint is the equilibrium state concept. It is justified by the fact that, essentially, what happens, is the replacement here of a single multidimensional problem by many “almost one-dimensional” problems in projections onto the coordinate axes. On each axis a cluster is delineated in such a way as to “bind” the axes together in a rigorously defined way. So, exposed to such a “bind” the cluster on a given axis cannot be “nudged” without in some measure deteriorating itself on the other axes in the sense of investigated performance index, subject to the condition that these others are fixed.

The superiority of the proposed approach is not restricted to the indicated “technical detail” of replacing one multidimensional space by one-dimensional projections. Indeed, an equilibrium state identified by means of the given vector index is parameterized by so-called thresholds, which satisfy the density levels of the clusters. In certain special cases, at any rate, an equilibrium state as the solution of a system of equations can be expressed analytically in the form of threshold functions, whereupon the identified clusters can be fully scanned in the spectrum of possible density levels.

The proposed theory for the identification of clusters of the probability density of measurements in  $n$ -dimensional space is set forth in two parts. In the first part (Sec. 2) the theory is not taken beyond the scope of customary multivariate functions and it concludes with a system equations, namely the system whose solution in the form of threshold functions makes it possible to scan the identified clusters. In the second part (Sec. 3) the theory now rests on a more abundant class of measurable functions specified by the class of sets represented on the coordinate axes by at most countable set of unions or intersections of segments. Overall the construction described in this part is so-called contra-monotonic system; actually, the first part on multiparameter contra-monotonic systems is also discussed in these terms (special case).

The fundamental result of the second part does not differ, in any way, from the form of the system of equations in the first part; the essential difference is in the space of admissible solutions. Whereas in the system of equations of the first part the solution is a numerical vector, in the second part it is a set of measurable sets containing the sought-after measurable density clusters. As the solution of the system of equations, the set of measurable sets serves as a fixed point of special kind mapping of subsets of multidimensional space. This particular feature is utilized in an iterative solving procedure.

## **2. Contra-monotonic systems over a family of parameters**

Here a monotonic system represents first a one-parameter and then a multiparameter family of functions defined on real axis. This type of representation is a special case of a more general monotonic system described in the next section.

We consider a one-parameter family of functions  $\mathbf{p}(x;h)$  defined on the real axis, where  $h$  is a parameter. For definiteness, we assume that an individual copy  $\mathbf{p}$  of the indicated family is a function integrable with respect to  $x$  and differentiable with respect to  $h$ . The family of functions  $\mathbf{p}$  is said to be contra-monotonic if it obeys the following condition: for any pair of quantities  $l$  and  $g$  such that  $l \leq g$  the inequality

$$\mathbf{p}(x;l) \geq \mathbf{p}(x;g)$$

holds for any  $x$ .

The specification of a multiparameter family of functions  $\mathbf{p}$  is reducible to the following scheme. We replace the one function  $\mathbf{p}$  by a vector function  $\mathbf{p} = \langle \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n \rangle$ , each  $j$ -th component of which is a copy of the function depending now on  $n$  parameters  $h_1, h_2, \dots, h_n$ , i.e.,  $\mathbf{p}_j = \mathbf{p}_j(x; h_1, h_2, \dots, h_n)$ . The contra-monotonicity condition for any pair of vectors  $l = \langle l_1, l_2, \dots, l_n \rangle$  and  $g = \langle g_1, g_2, \dots, g_n \rangle$  such that  $l_k \leq g_k$  ( $k = 1, 2, \dots, n$ ) is written in the form of  $n$  inequalities:

$$\mathbf{p}_j(x; l_1, l_2, \dots, l_n) \geq \mathbf{p}_j(x; g_1, g_2, \dots, g_n).$$

We note that this condition rigorously associates with the family of vector functions a componentwise partial ordering of the vector parameters.

We give special attention to the case of a so-called de-coupled multi-parameter family of functions  $\mathbf{p}$ . The family is said to be de-coupled if the  $j$ -th component of a copy of vector function  $\mathbf{p}$  does not depend on the  $j$ -th component of the vector of parameters  $h$ , i.e., on  $h_j$ . Therefore, a copy of function  $\mathbf{p}$  of a de-coupled multi-parameter family is written in the form  $\mathbf{p}_j(x, h_1, \dots, h_{j-1}, h_{j+1}, \dots, h_n)$  ( $j = 1, \dots, n$ ).

We now return to the original problem of analyzing a multi-modal empirical distribution in multidimensional space. We first investigate the case of one axis (univariate distribution).

Let  $p(x)$  be the probability density function of points in the  $x$  axis. For the contra-monotonic family  $\mathbf{p}$  we can choose, for example, the functions  $\mathbf{p}(x; h) = p(x)^h$ . It is easy verified that the contra-monotonicity condition is satisfied.

We consider the following variational problem. With respect to an externally specified threshold  $u^o$  ( $0 \leq u^o \leq 1$ ) let it be necessary to maximize the functional

$$\mathbf{P}(h) = \int_{-h}^{+h} [\mathbf{p}(x; h) - u^o] dx.$$

It is clear that for small  $h$  the quantity  $\mathbf{P}(h)$  will be small because of the narrow interval of integration, while for the large  $h$  it will be small by the contra-monotonicity condition. Consequently, the value of  $\max_h \mathbf{P}(h)$  will necessarily be attained for certain finite nonzero  $h^o$ .

It is readily noted that if  $p(x)$  is a unimodal density function with zero expectation, then the maximization of the functional  $\mathbf{P}(h)$  implies the identification of an interval on the axis corresponding to a concentration of the density  $p(x)$ . But if  $p(x)$  has a more complicated form, then the maximum of  $\mathbf{P}(h)$  specifies an interval in which is concentrated the “essential part”, in some definite sense, of the density function  $p(x)$ .

Directly from the form of function  $\mathbf{P}(h)$  we deduce the following necessary condition for local maximum (the zero equation of the derivative with respect to  $h$ :  $\frac{\partial}{\partial h}\mathbf{P}(h)=0$  or, in the expanded form, the equation

$$\mathbf{p}(-h;h)+\mathbf{p}(h;h)+\int_{-h}^{+h}\frac{\partial}{\partial y}\mathbf{p}(x;y)|_{y=h}dx=2u^o. \quad (1)$$

The root of the given equation will necessarily contain one at which  $\mathbf{P}(h)$  attains a global maximum. We have thus done with the problem: we found the central cluster points of the density function on one axis in terms of a contra-monotonic family of functions.

To find the central clusters of a multivariate distribution in  $n$ -dimensional space we invoke the notion of a multiparameter contra-monotonic family of functions  $\mathbf{p}$ . Let the family of functions  $\mathbf{p}$  in vector form be written, say, in the form  $\mathbf{p}_j(x;h_1,\dots,h_n)=p_j(x)^{h_j}$ , where  $h=\sum_{k=1}^nh_k$ , and  $p_j(x)$  is a projection of the multivariate distribution on the axis  $j$ -th axis. In the stated sense the goodness of the delineated central cluster is evaluated by the multivariate (vector) performance index  $\mathbf{P}=\langle\mathbf{P}_1,\dots,\mathbf{P}_n\rangle$ , where

$$\mathbf{P}_j(h_1,h_2,\dots,h_n)=\int_{-h_j}^{h_j}[\mathbf{p}_j(x;h_1,\dots,h_n)-u_j]dx \quad (2)$$

and  $u_j$  is the component of the corresponding externally specified multidimensional threshold vector  $u$ :  $u=\langle u_1,u_2,\dots,u_n\rangle$ . As in the one-dimensional case, of course, it is meaningful to use the given functional only distributions  $p_j(x)$  with zero expectation.

Once the goodness of a delineated cluster has been evaluated by the vector index, it must be decided, based on standard [7] vector optimization principles, what is an acceptable cluster. In this connection it desirable to indicate simultaneously a procedure for finding an extremal point in the space of parameters. It turns out that for so-called Nash-optimal

Equilibrium State there is a simple technique for finding solutions at least in de-coupled family of contra-monotonic functions  $\mathbf{p}$ .

En equilibrium situation (Nash point) in the parameter space  $h = \langle h_1, \dots, h_n \rangle$  with indices  $\mathbf{P}_j$  is defined as a point  $h^* = \langle h_1^*, h_2^*, \dots, h_n^* \rangle$  such that for every  $j$  the inequality

$$\mathbf{P}_j(h_1^*, \dots, h_{j-1}^*, h_j, h_{j+1}^*, \dots, h_n^*) \leq \mathbf{P}_j(h_1^*, \dots, h_j^*, \dots, h_n^*)$$

holds for any value of  $h_j$ . In other words, if there are no sensible bases in the sense of index  $\mathbf{P}_j$  on the one ( $j$ -th) axis, then the equilibrium situation is shifted with respect to the parameter  $h_j$ , subject to the condition that the quantities  $h_k^*$ ,  $k \neq j$ , are fixed on all other axes.

Clearly, a necessary condition at a Nash point in the parameter space (as in the one-dimensional case) is that the partial derivatives tend to zero, i.e., the  $n$  equalities  $\partial / \partial h_j \mathbf{P}_j(h_1^*, \dots, h_n^*) = 0$  must hold. The sufficient condition comprises the  $n$  inequalities  $\partial^2 / \partial h_j^2 \mathbf{P}_j(h_1^*, \dots, h_n^*) \leq 0$ .

An essential issue here, however, is the fact that the necessary condition (equalities) acquires a simpler form for de-coupled family of contra-monotonic functions than in the general case. Thus, by the decoupling of the family  $\mathbf{p}$  the partial derivative  $\partial \mathbf{P}_j / \partial h_j$  is identically zero, and the system of equations, see (1) by analogy, with respect to the sought-after point  $h^*$  is reducible to the form

$$\mathbf{p}_j(-h_j; h_1, \dots, h_{j-1}, h_{j+1}, \dots, h_n) + \mathbf{p}_j(h_j; h_1, \dots, h_{j-1}, h_{j+1}, \dots, h_n) = 2u_j \quad (3)$$

Now the sufficient condition is satisfied automatically for any solution  $h^*$  of Eqs. (3).

In conclusion we write out the system of equations for two special cases of a de-coupled family of contra-monotonic functions  $\mathbf{p}$ .

1. Let  $\mathbf{p}_j(x; h_1, \dots, h_{j-1}, h_{j+1}, \dots, h_n) = p_j(x)^{s-h_j}$ , where  $s = h_1 + h_2 + \dots + h_n$ . Then the system of equations (3) is reducible to the form

$$p_j(-h_j)^{s-h_j} + p_j(h_j)^{s-h_j} = 2u_j \quad (j = 1, \dots, n).$$

2. Let the role of  $\mathbf{p}_j(x; h_1, \dots, h_{j-1}, h_{j+1}, \dots, h_n)$  be taken by the function

$$p_1(x)^{h_1} p_2(x)^{h_2} \dots p_{j-1}(x)^{h_{j-1}} p_{j+1}(x)^{h_{j+1}} \dots p_n(x)^{h_n}.$$

The system of equations (3) for finding a solution, i.e., an equilibrium situation (Nash point)  $h^*$ , is written

$$p(-h_j)/p_j(-h_j)^{h_j} + p(h_j)/p_j(h_j)^{h_j} = 2u_j \quad (j=1,\dots,n),$$

where  $p(x) = p_1(x)^{h_1} p_2(x)^{h_2} \dots p_n(x)^{h_n}$  is the product of univariate density functions.

We conclude this section with an important observation affecting the vector of thresholds  $u = \langle u_1, u_2, \dots, u_n \rangle$ . By straightforward reasoning we infer that each component  $h_j^*$  of the equilibrium situation  $h^*$  is a function of thresholds and  $h^*$  can be represented by a vector function of thresholds in the form  $h_j^* = h_j^*(u_1, u_2, \dots, u_n)$ . If the solution of the system of equations (3) can be expressed analytically, then prolific possibilities are afforded for scanning the equilibrium situations in the parameter space and, accordingly, selecting an “acceptable” cluster in the spectrum of existing densities of measurements in a multidimensional space of thresholds. A similar approach can be used when solutions of Eqs. (3) are sought by numerical methods.

### 3. Contra-monotonic systems over a family of segments

A multi-parameter family of contra-monotonic functions used for the analysis of multivariate distributions, unfortunately, has one substantial drawback. Generally speaking, there is no way to guarantee the identification of homogeneous distribution clusters in projection onto the  $j$ -th axis, because the segment  $[-h_j, h_j]$  can contain several distinct modes. On the other hand, it is sometimes desirable to identify modes by merely indicating a family of segments containing each mode separately. The construction proposed below enlarges the possibilities for the solution of such a problem by augmenting the contra-monotonic systems of the proceeding section in natural way.

Thus, on real axis we consider subsets represented by at most countable set of operations of union, intersection, and difference of segments. The class of all such subsets is denoted by  $B$ , and each representative subset by  $H \in B$  (which we call a  $B$  set) is distinguished from like sets by length  $m$  (by measure zero). A set  $L$  is congruent with  $G$  ( $G = L$ ) if the

measure of the symmetric difference  $GDL$  is equal to zero ( $mGDL=0$ ); a set  $L$  is contained in  $G$  ( $L \subseteq G$ ) with respect to measure  $m$  if  $mG \setminus L = 0$ . A measure on the real axis, being an additive function of sets (the length), is determined by taking to the limit the length of the sets in the set of unions, intersections, and differences of segments forming the  $B$  set. Then set-theoretic operations over  $B$  sets will be understood to mean up to measure zero. By convention, all  $B$  sets of measure zero are indistinguishable.

We associate with every  $B$  set  $H$  a nonnegative function  $\mathbf{p}(x;H)$ , which is Borel measurable (or simply measurable) and whose domain of definition is on the real axis.<sup>1</sup> In other words, in contrast with the one-parameter family of contra-monotonic functions of the preceding section, the parameter  $h$  is now generalized, namely, it is extended to the  $B$  set  $H$ . As before, we say that a family of measurable functions  $\mathbf{p}$  is contra-monotonic if it obeys the following condition: for any pair of sets  $L$  and  $G$  such that  $L \subseteq G$  the inequality

$$\mathbf{p}(x;L) \geq \mathbf{p}(x;G)$$

holds for any  $x$ .

The scheme of specification of a multi-parameter family of functions is analogous to the previous situation. In place of a scalar function  $\mathbf{p}$  we now specify a vector function  $\mathbf{p} = \langle \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n \rangle$ , each  $j$ -th component of which is a copy of a function depending at the outset on  $n$  parameters  $\langle H_1, H_2, \dots, H_n \rangle$  ( $B$  sets), i.e.,  $\mathbf{p}_j = \mathbf{p}_j(x;H_1, H_2, \dots, H_n)$ . Again, the contra-monotonicity condition is reducible to the statement that for any pair of vectors (ordered sets of  $B$  sets) of the form  $L = \langle L_1, \dots, L_n \rangle$  and  $G = \langle G_1, \dots, G_n \rangle$  such that  $L_k \subseteq G_k$  ( $k = 1, 2, \dots, n$ ), the following  $n$  inequalities are satisfied:<sup>2</sup>

$$\mathbf{p}_j(x;L_1, \dots, L_n) \geq \mathbf{p}_j(x;G_1, \dots, G_n).$$

These inequalities associate a partial ordering of sets of  $B$  sets with a family of vector functions  $\mathbf{p}$  in a rigorously defined way.

---

<sup>1</sup> A function  $\mathbf{p}(x;H)$  is Borel measurable if for any numerical threshold  $u^o$  the set of all  $x$  of the real scale for which  $\mathbf{p}(x;H) > u^o$  is measurable:  $\{x : \mathbf{p}(x;H) > u^o\}$  is  $B$  set.

<sup>2</sup> Here  $x$  is a point on the  $j$ -th axis. This is tacitly understood everywhere.

In the case of a de-coupled family of contra-monotonic functions, where the  $j$ -th component of a copy of the vector function  $\mathbf{p}$  does not depend on the parameter  $H_j$ , or  $B$  set on the  $j$ -th axis of definition of the function  $\mathbf{p}_j$ , this component  $\mathbf{p}_j$  of the vector function  $\mathbf{p}$  is written  $\mathbf{p}_j = \mathbf{p}_j(x; H_1, H_2, \dots, H_n)$ .

Following again the order of discussion of Sec. 2, we now consider the original problem of analyzing the structure of a multi-modal empirical distribution in a multidimensional space. We first investigate the case of a one-dimensional (univariate) distribution.

Let  $p(x)$  be the density function of points on the  $x$  axis. In the role of the contra-monotonic family of functions  $\mathbf{p}$ , we adopt functions of the form  $\mathbf{p}(x; H) = p(x)^{F(H)}$ , where  $F(H) = \int_H p(x) dx$  is the probability of a random variable occurring in a  $B$  set under the probability density function  $p(x)$ . It is clear that the contra-monotonicity condition is satisfied.

We consider the following variational problem. Given the externally specified threshold  $u^o$  ( $0 \leq u^o \leq 1$ ), maximize the functional

$$\mathbf{P}(H) = \int_H [\mathbf{p}(x; H) - u^o] d\mathbf{m}.$$

The integral here is understood in the Lebesgue sense with respect to measure  $\mathbf{m}$ , where  $\mathbf{m}$ , as mentioned before, is the length of the  $B$  set on the  $x$  axis.

Clearly, the quantity  $\mathbf{P}(H)$  as a function of the length  $\mathbf{m}$  (measure of set  $H$ ) increases first and then, as  $\mathbf{m}H \rightarrow \infty$ , reverts to zero by the contra-monotonicity condition on the family of functions  $\mathbf{p}$ . Therefore, the value of  $\max_H \mathbf{P}(H)$  will necessary be attained on a certain  $B$  set of finite measure  $\mathbf{m}$  (see the analogous assertion in Sec. 2).

It is impossible in the same simple way to deduce directly from the form of the functional  $\mathbf{P}(H)$  any maximum condition comparable with the like condition of the preceding section (Eq. 1). To do so would require elaborating the notation of a "virtual translation" from a  $B$  set  $H$  to a set  $\tilde{H}$  similar to it in some sense, in such a way as to establish the necessary maximum condition. These circumstances exclude the case of a univariate distribution from further consideration. Nonetheless, as will be shown presently, for multivariate distribution

there are means for finding a  $B$  set that will maximize the function  $\mathbf{P}(H)$  at least in the case of a de-coupled family of contra-monotonic functions.

As in the preceding section, we evaluate the goodness of an identified central cluster by the multivariate (vector) performance index  $\mathbf{P} = \langle \mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n \rangle$ :

$$\mathbf{P}_j(H_1, H_2, \dots, H_n) = \int_{H_j} [\mathbf{p}(x; H_1, \dots, H_n) - u_j] d\mathbf{m},$$

where  $u_j$  is the coordinate of the corresponding multidimensional vector of thresholds  $u$ , specified externally:  $u = \langle u_1, u_2, \dots, u_n \rangle$ .

At this point we call attention to the fact that, in contrast with the analogous multivariate index of Sec. 2, the given functional now has significance for an arbitrary distribution, rather than only for the centered condition of zero-valuedness of the expectation. We again look for the required cluster in multidimensional space as an equilibrium situation according to the vector index  $\mathbf{P} = \langle \mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n \rangle$ . We regard a cluster as a set of  $B$  sets  $H^* = \langle H_1^*, H_2^*, \dots, H_n^* \rangle$  such that the following inequalities holds for every  $j$ :

$$\mathbf{P}_j(H_1^*, \dots, H_{j-1}^*, H_j, H_{j+1}^*, \dots, H_n^*) \leq \mathbf{P}_j(H_1^*, \dots, H_j^*, \dots, H_n^*) \quad (j = 1, \dots, n).$$

In a de-coupled family of contra-monotonic functions it is feasible (as in the multi-parameter case; see Eq. (3) ) to find an equilibrium situation. Equilibrium situations are sought to be a special technique of mappings of  $B$  sets onto real axes.

We define the following type of mappings of  $B$  sets onto real axes:

$$V_j(H_j) = \{ x : \mathbf{p}_j(x; H_j) > u_j \},$$

where  $u_j$  is the threshold involved in the expression for the functional  $\mathbf{P}_j$  ( $j = 1, 2, \dots, n$ ). Thus defined,  $n$  such mappings are uniquely expressible in the vector form

$$V(H) = \{ x : \mathbf{p}(x; H) > u \}.$$

Here  $H = H_1 \times H_2 \times \dots \times H_n$  denotes the direct product of sets  $H_j$ . We define a fixed point of the mapping  $V(H)$  as a set  $H^*$  for which the equality  $H^* = V(H^*)$  holds.

**Theorem 1.** For a de-coupled family of contra-monotonic functions  $\mathbf{p}$ , a fixed point of the mapping  $V(H)$  generates an equilibrium situation according to the vector index  $\mathbf{P} = \langle \mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n \rangle$ .

The proof of the theorem is simple. Thus, because  $\mathbf{p}_j$  is independent of the parameter  $H_j$ , the form of the function  $\mathbf{p}_j(x; H_1^*, \dots, H_{j-1}^*, H_{j+1}^*, \dots, H_n^*)$  does not depend on  $H_j$ . Also, the set  $H^* = H_1^* \times H_2^* \times \dots \times H_n^*$  in projection onto the  $j$ -th axis intersects the set  $H_j^*$  consisting exclusively of all points  $x$  for which  $\mathbf{p}_j(x; H_j^*) > u_j$ :  $H_j^* = \{x : \mathbf{p}_j(x; H_j^*) > u_j\}$ . It is immediately apparent that any  $H_j$  distinct from  $H_j^*$  the value of the functional  $\mathbf{P}_j(H_1^*, \dots, H_{j-1}^*, H_j, H_{j+1}^*, \dots, H_n^*)$  for immovable sets  $H_k^*$  ( $k \neq j$ ) cannot be anything but smaller than the quantity  $\mathbf{P}_j(H_1^*, \dots, H_{j-1}^*, H_j^*, H_{j+1}^*, \dots, H_n^*)$ .

It is important, therefore, to find the fixed points of the constructed mapping of  $B$  sets.

#### 4. Methods of finding equilibrium state for de-coupled families of contra-monotonic functions

The ensuing discussion rests heavily on the contra-monotonicity property of a function  $\mathbf{p}$ . To facilitate comprehension of the formulations and propositions we use the language of diagrams reflecting the structure of the relations involved in the constructed mappings of  $B$  sets, in particular the symbol  $\rightarrow$  denoting the relation “set  $X_1$  is nested in set  $X_2$  ( $X_1 \subseteq X_2$ ) “:  $X_1 \rightarrow X_2$ .

All diagrams of the relations between  $B$  sets are based on the following proposition: the relation  $X_1 \rightarrow X_2$  (as a consequence of the contra-monotonicity condition on  $\mathbf{p}$ ) implies that  $V(X_1) \leftarrow V(X_2)$ .

Now let the mapping  $V$  be applied to the original space  $W$  of axes on which the functions  $\mathbf{p}_j$  ( $j = 1, 2, \dots, n$ ) are defined. After the image  $V(W)$  has been obtained, we again apply the mapping  $V$  with the  $B$  set  $V(W)$  as its inverse image, i.e., we consider the image  $V^2(W)$ , and so on. In this way we construct a chain of  $B$  sets  $W, V(W), V^2(W), \dots$ , which we call the central series of the contra-monotonic system.



**Procedure for Solving the Equation  $X = V(X)$ .** A chain of  $B$  sets  $H_0, H_1, \dots$ , is generated recursively according to the following rule. Let the set  $H_k$  (where  $H_0$  is any  $B$  set of finite measure) be already generated in the chain. We use the mapping  $V(X)$  to transform the following  $B$  sets:

$$\begin{aligned} V\{V^2(H_k) \cup V(H_k)\}, & \quad V\{V(H_k) \cap H_k\}, \\ V\{V(H_k) \cup H_k\}, & \quad V\{V^2(H_k) \cap V(H_k)\}, \end{aligned}$$

which we denote, in order, by  $L_k^2, G_k, L_k, G_k^2$ . By the contra-monotonicity of the family of functions  $\mathbf{p}$  it turns out that  $L_k^2$  is a subset of  $G_k$  and that  $L_k$  is a subset of  $G_k^2$ . Picking any  $A_k$  based on the condition  $L_k^2 \subset A_k \subset G_k$ , and then  $B_k$  from the analogous condition  $L_k \subset B_k \subset G_k^2$ , we put the set  $H_{k+1}$  following  $H_k$  in the constructed series of  $B$  sets equal to  $A_k \cup B_k$ :  $H_{k+1} = A_k \cup B_k$ . The sets  $A_k$  and  $B_k$  can be chosen, for example, according to mapping rules in the class of  $B$  sets, namely,

$$\begin{aligned} A_k &= \{x : \frac{1}{2}[\mathbf{p}(x; L_k^2) + \mathbf{p}(x; G_k)] > u\}, \\ B_k &= \{x : \frac{1}{2}[\mathbf{p}(x; L_k) + \mathbf{p}(x; G_k^2)] > u\}. \end{aligned}$$

The conditions imposed on  $A_k$  and  $B_k$  are satisfied in this case.

**Theorem 3.** For the series of sets  $V(H_k)$  to contain the limiting set  $V(H^*)$  as  $k \rightarrow \infty$ , which would be a solution of the equation  $X = V(X)$ , the following two conditions are sufficient:

$$\begin{aligned} \text{a)} \quad & \lim_{k \rightarrow \infty} \mathbf{m}G_k \setminus L_k^2 = 0, \\ \text{b)} \quad & \lim_{k \rightarrow \infty} \mathbf{m}G_k^2 \setminus L_k = 0. \end{aligned}$$

The plan of the proof is quickly grasped in the following nesting diagrams, which are consequences of the contra-monotonicity property of the functions  $\mathbf{p}$ , i.e.,

$$\begin{aligned} \text{I.} \quad & V^2(H_k) \leftarrow L_k^2 \rightarrow G_k \leftarrow V(H_k), \\ \text{II.} \quad & V(H_k) \leftarrow L_k \rightarrow G_k^2 \leftarrow V^2(H_k). \end{aligned}$$

Diagrams **I** and **II** imply the validity of the two chains:

$$\begin{aligned} 1) \quad & V^2(H_k) \setminus V(H_k) \subseteq V^2(H_k) \setminus G_k \subseteq L_k^2 \setminus G_k, \\ 2) \quad & V(H_k) \setminus V^2(H_k) \subseteq V(H_k) \setminus G_k^2 \subseteq L_k \setminus G_k^2. \end{aligned}$$

The first chain implies that for the limiting set  $H^*$  of the series  $H_0, H_1, \dots$ , the equality  $\mathbf{m}V^2(H_k) \setminus V(H^*) = 0$  holds, i.e.,  $V(H^*) \subset V^2(H^*)$ ; the second chain implies the opposite relation:  $V^2(H^*) \subseteq V(H^*)$ . Consequently,  $V(H^*)$  is the solution of the equation  $X = V(X)$ :  $V(H^*) = V(V(H^*))$ . Of course, the conditions of the theorem are sufficient for the existence of a solution of the equation  $X = V(X)$ , and their absence does not in any way negate some other solving technique, provided that solutions exist in general. The possibility that solution  $H^*$  of the equation  $X = V(X)$  do not exist should certainly not be dismissed.

### LITERATURE CITED

1. Finney, D.J., 1964, An Introduction to the theory of Experimental Design, Univ. Chicago Press.
2. Rosenfeld, A., 1969, Picture Processing by Computer, Academic Press, New York.
3. Fishburn, P.C., 1970, Utility Theory for Decision-Making, Wiley, New York.
4. Aizerman, M.A., Braverman, E.M., Rozonoer, L.I., 1970, The Method of Potential Functions in Machine Training Theory [in Russian], Nauka, Moscow.
5. Zagoruiko, N.G., and Zaslavskaya, T.I., 1968, Pattern Recognition in Social Research [in Russian], Sib. Otd. Akad. Nauk SSSR, Novosibirsk.
6. Owen, G. Game theory, 1968, Saunders, Philadelphia.
7. Becker, G.M., and McClintock, C.G., 1967, "Behavioral decision theory," in: Annual Review of Psychology, Vol. 18, Stanford, Calif.
8. Shilov, G.E., and Gurevich, B.L., 1967, Integral, Measure, and Derivative [in Russian], Nauka, Moscow.