DIVISION DE ESTUDIOS DE POSGRADO
FACULTAD DE INGENIERIA

VNIVERSIDAD NACIONAL
AVTONOMA DE
MEXICO

ORIENTOR FIELDS AND CONTROL SYSTEMS

S. RACZYNSKI

REPORTE TECNICO № 85001

SUBJEFATURA DE INGENIERIA ELECTROMECANICA
APARTADO POSTAL 70-256
04510 MEXICO D.F.
MEXICO

S. Raczynski

ORIENTOR FIELDS AND CONTROL SYSTEMS

Contents

Introduction

It is known that the time-optimal controls of linear finite-dimensional systems are of so-called "bang-bang" type, i.e. the optimal control signals switch between some finite set of values. Dealing with nonlinear systems, it is important to know whether optimal controls are of the "bang-bang" type or of some type of functions being a generalization of that type. The problem of applicability of some kind of generalized "bang-bang" controls for nonlinear systems in R^n had already been solved in the 1960s. The main contributions in this field (Ważewski [1], [2], [3]) were based on the results of Marchaud [8] and Zaremba [9], who (independently) pointed out the fundamental properties of solutions to so-called contingent equations. It is interesting that, in fact, Zaremba and Marchaud had already described the main properties of trajectories and reachable sets of control systems with convex sets of controls, in the years 1934-36. In 1961 Ważewski [1] pointed out that the contingent equation is equivalent to certain differential inclusion. In [2] one can find a generalization of those results to the case of nonconvex sets of admissible directions, called orientors. The applications to control problems follow directly from the further papers of Ważewski [3], [5]. One of the most important notions used by Ważewski is the notion of quasitrajectory, being a weak solution to certain differential inclusion. It can be seen that the limit functions of some sequences of trajectories with "bang-bang" control are quasitrajectories. Such functions are well known in automatic control as "sliding regimes". The theorems of closedness of reachable sets for quasitrajectories lead directly to existence theorems for optimal quasitrajectories. On the other hand, the existence of sliding regimes approximating optimal quasitrajectories follows from the theorem of Turowicz [10].

For much profound study of differential inclusions and more references the reader is refered to [38].

In the present paper some attempt was made to generalize the known results on orientor fields to the case of Banach space. It should be noted that the fundamental theory of contingent equations in Banach space has been published by Sui-nee Chow and Y.D.Shuur [12].

The right restriction on generalizations of the results mentioned above is the existence and measurability (in the sense of Lusin) of the so-called tendor field (def. I.1.5). As indicated in the sequel, the relevant assumption which must be imposed on systems under consideration is that the tendor field associated with the system exists and is measurable in the sense of Lusin. This is fulfielled automatically if the system state space X is separable and locally compact and 2^X is separable. If it is not the case, the appropiate assumptions must be imposed. These assumptions are included in the definitions of normal and regular control systems introduced in this paper.

Let us note that we do not consider any stability properties. The control systems and orientor fields appearing in this contribution are defined over some finite time-interval $J = [0, T]$. Let us note also that the considerations of the sections I.1. - II.1 concern systems in a real reflexive separable Banach space X and in II.2 we assume, in addition, that the dual X^* of X is uniformly convex.

This contribution was designed to be used as a graduate text in nonlinear optimal control theory and was used as an additional material for the courses given at the Academy of Mining and Metallurgy in Krakow, Poland, and at the National Autonomical University of Mexico.

The author wishes to express gratitude to professor Henryk Górecki of the Academy of Mining and Metallurgy for the inspiration and his valuable advice, to professors Andres Buzo de la Peña and Federiko Kuhlmann-Rodríguez of the National Autonomical University of Mexico for their help, and to professor Gustavo Medrano of the same University for his suggestions and comments. The author is also most grateful to Teresa Quintero for her excellent typing of the manuscript.

Notation

Unless otherwise specified, the symbols used throughout this text have the following meanings

X	- a real separable and reflexive Banach space
ϕ_X	- the origin of X
X^*	- the dual of X; in section II.2 X^* is supposed to be uniformly convex
$V(A,q)$	- the q-neighbourhood of the set A
$d^*(\cdot,\cdot)$	- the Hausdorff metric
ϕ	- the empty set
G	- a space of all closed equibounded subsets of X (def. I.1.5)
N	- orientor field
Q	- tendor field
E	- convex orientor field
$a \in B$	- means "a is a member of B"
ℓ.s.c.	- lower semi-continuous
u.s.c.	- upper semi-continuous
$A \setminus B$	- $\{z: z \in A, z \notin B\}$
$\cup$	- union
$\forall$	- "fol all"
$\exists$	- "exists"
$\cap$	- intersection
$\subset$	- inclusion
$\|\cdot\|$	- norm
$X \times Y$	- the product space

$\dot{x}$ - dx/dt

t - time

x - the system state

(f,C) - a control system

C - control domain

D - the tendor kernel of a control domain

$\Gamma(\cdot)$ - the set of all trajectories of an orientor field or control system with fixed initial condition

$f|_L$ - the restriction of a function f to a set L

a.e. - "almost everywhere"

$Z(N)$ - the reachable set of N

$\langle \cdot, \cdot \rangle$ - scalar product

$\langle \cdot, \cdot \rangle_{\pm} \langle \cdot, \cdot \rangle_s$ - semi-inner products

$\tilde{N}(B,b)$ - the set of all normals to a set B at the point $b \in \partial B$ (def. II.2.6)

∂A - the boundary of a set A

Gy/z - the parallel algorithm with y parallel, z of them of object type

J - a closed interval $[0, T]$, where $T > 0$

Ω - a subset of the space X (see I.2)

I. General concept of orientor field

I.1. Notation and definitions

Let X be a real separable and reflexive Banach space with a norm $\|\cdot\|$ and the origin ϕ_X . All multifunctions considered in the sequel are supposed to be bounded-valued.

I.1.1. Definition

The point-to-set distance between a point $x \in X$ and a set $B \subset X$ is defined as

$$d(x,B) = \inf_{b \in B} \| x - b \| .$$

I.1.2. Definition

By the closed q-neighborhood of a set $A \subset X$ we mean the set

$$V(A, q) = \overline{\{ X \ni x : \ d(x,A) \leq q \}} .$$

I.1.3 Definition

Let H be the space of all closed, bounded and nonempty subsets of X. Let us define a metric in H as follows

$$d^* (A,B) = \inf \{s : A \subset V(B, s), \ B \subset V(A,s), \ s \in R_+\} .$$

I.1.4 Remark

The function d^* is called the Hausdorff metric in H. It can be easily seen that if the sets of H were not closed, d^* would not be a metric, but merely a semi-metric in H.

I.1.5. Definition

Let Y be a topological space, let $M\varepsilon(0, \infty)$ be a constant and

$G = \{Z : Z\varepsilon H,\ d^*(\{\phi_x\}, Z) \leq M\}$ (here Z is an element of H and, at the same time, a subset of X). A mapping $N : Y \to G$ is called an orientor field iff:

(i) the set conv N(y) (i.e. the smallest convex hull of the set N(y)) is closed for all $y\varepsilon Y$.

(ii) the set tend N(y) (i.e. the smallest closed subset B of N(y), such that conv B= conv N(y)) exists, for all $y \varepsilon Y$.

The mappings $E: Y \ni y \longrightarrow$ conv N(y) and $Q: Y \ni y \longrightarrow$ tend N(y) are called convex field and tendor field generated by N, respectively. The set N(y) is called the orientor and Q(y) is called the tendor of N(y).

I.1.6 Remark

In the finite-dimensional case the conditions (i) and (ii) of I.1.5 are fulfilled automatically, and so is when N(y) is compact (by the Krein-Milman theorem). Unfortunately, in the general case of Banach space it is not the case. Let us consider, for example, the space ℓ_2 (of sequences), and the set $U\varepsilon\ \ell_2$ consisting of the sequences {1,0,0,...}, {0,1,0,...}, {0,0,1,0,...} e.t.c. Obviously U is closed in ℓ_2 and the sequence $U_n = \{1/n, 1/n, \ldots, 1/n, 0, 0, \ldots\}$ (containing n times the value 1/n) belongs to conv U. But if n tends to infinity, then $u^* = \lim u_n = \{0,0,0,\ldots\}$. The sequence u* is not a linear combination of any points (sequences) of U, so conv U is not closed.

Thus, the term "orientor field" denotes some class of multifunctions. Let us observe that we do not require compactness of $N(y)$. Of course, one can define an orientor field in Banach space is the same way as it was done in R^n (Ważewski [1]), i.e. as a multifunction $X \to G$. However, the lack of compactness of $N(y)$ makes any considerations of such fields difficult, and in order to get some applicable results we have to impose strong additional assumptions. It seems to be reasonable to restrict the class of orientor fields to those fulfilling (i) and (ii) of I.1.5 in order to deal with some natural generalization of orientor field in X. On the other hand, let us observe that in many practical cases orientor fields associated with control systems are compact. Such situation occurs when the control domain (the set of all permissible controls) is compact (e.g. a control system with finite number of scalar control signals), even if the system state space is infinite-dimensional.

In the fig. 1 some possible shapes of orientors N in R^2, its convex hull E and the tend N are indicated.

I.1.7. Definition

A multifunction $Y \longrightarrow G$ is said to be continuous if it is continuous as a mapping $Y \longrightarrow G$, in the topology of G induced by the Hansdorff metric d^*.

I.1.8. Definition

A multifunction $N:Y \longrightarrow G$ is lower semi-conductinuous (l.s.c.) et a point $y_0 \varepsilon Y$ if

$$y_i \varepsilon Y,\ y_i \longrightarrow y_0 \implies N(y_0) \subset \{b : b \varepsilon X,\ d(b, N(y_i)) \longrightarrow 0\},$$

where d is the point-to-set distance. This "sequential" definition is equivalent to the following.

I.1.9. Definition

A multifunction $N:Y \longrightarrow G$ is lower semi-continuous at a point y_0 if for any open set $M \subset X$ such that $N(y_0) \cap M \neq \phi$ a neighborhood ρ of y_0 exists such that $N(y) \cap M \neq \phi$ for all $y \varepsilon \rho$ (ϕ stands for an empty set).

I.1.10. Definition

A multifunction $N:Y \longrightarrow G$ is upper semi-continuous (u.s.c.) at y_0 if for any open set $M \subset X$ such that $N(y_0) \subset M$ a neighborhood ρ of y_0 exists such that $N(y) \subset M$ for all $y \varepsilon \rho$.

I.1.11. Remark

Let us consider, as an example, the following orientor field: $N:R_+ \longrightarrow G$, where G is the set of all closed and equi-bounded subsets of R^2. Let $N(y)=\{a,b,c,d,\}$, where $a=(-1,0)$, $b=(1,1)$, $c=(1,-1)$ and $d=(y,0)$. The tendor Q of N consists of the points a, b, c for all $y \leq 1$, and $Q= \{a, b, c, d\}$ for all $y > 1$. The field Q is obviously continuous for all $y \neq 1$. It is easily seen that, due to I.1,9, Q is l.s.c. at $y = 1$, but, due to I.1.10 it is not u.s.c. at $y = 1$. An orientor field (multifunction) is said to be continuous (l.s.c. or u.s.c.) over some set $W \subset Y$ if it is continuous (l.s.c. or u.s.c. respectively) for all $y \varepsilon W$.

In the case when X is metric compact space it was proved [7] that continuity of N implies lower semi-continuity of the corresponding tendor field Q, and, provided Y is measurable, the field Q is Lusin-measurable. Recall that a function $f:T \longrightarrow U$ (where T is a topological Hansdorff space with Radon measure μ

and U is topological) is said to be Lusin-measurable if for any compact $K \subset T$ and $\varepsilon > 0$ a compact set $L \subset K$ exists such that $\mu(K \setminus L) < \varepsilon$ and the restriction $f|_L$ is continuous.

I.1.12. Definition

Let π be a σ-additive field of subsets of Y. An orientor field $N: Y \longrightarrow G$ is said to be π-measurable if $N^{-1}(U) \varepsilon \pi \,\forall\, U \subset X$, U open. The $N^{-1}(U)$ is defined as $\{Y \varepsilon y: N(y) \cap U \neq \phi,\ U \subset X\}$.

I.1.13. Definition

A mapping $s: Y \longrightarrow X$ is called a selector of a field $N: Y \longrightarrow G$ if $s(y) \subset N(y)$ for each $y \varepsilon Y$.

I.1.14. Remark

The function s defined above is also called selection. Such terminology is used for example by Castaing and Valadier in the book [14] containing many valuable results on selections of certain multifunctions. We shall use the term "selector" as being used far earlier by Kuratowski and Ryll-Nardzewski [13].

It was pointed out [13] that if Y is a topological space and S is a countably additive family of subsets of Y, then from the statement $\{y: N(y) \cap A \neq \phi\} \varepsilon S$ whenever $A \subset X$ is open, follows that a selector s of the field N such that $s^{-1}(A) \varepsilon S$ exists. Moreover, if N is continuous modulo a first category set, then a selector of this field continuous modulo a first category set exists. The following theorem of Castaing and Valadier gives some good characterization of selectors.

I.1.15. Theorem

Let (Y,π) be a measurable space, G be as in I.1.5, and let $N:Y \longrightarrow G$ be an orientor field. Then the following are equivalent conditions.

(i) $N^{-1}(U)\varepsilon\pi$ for any open U, $U\subset X$

(ii) $d(x, N(\cdot))$ is a measurable function for any $x\varepsilon X$,

(iii) a countable set of measurable selectors $\{f_n\}$ of the field N exists, such that $N(y) = \overline{\{f_n(y)\}}$ for all $y\varepsilon Y$, i.e. $\{f_n(y)\}$ is dense in $N(y)$.

I.1.16. Remark

The proof of the above, as well as of the following two theorems, may be found in the book of Castaing and Valadier [14]. These theorems are quoted here for their particular importance in the control theory. Before formulating the next two theorems, let us recall that a topological Hausdorff space P and a mapping $P \longrightarrow S$ exist (we follow the definition of Suslin space as in [14], III.17). The symbol $\otimes$ will denote a quotient σ-additive field of subsets, see Warga [15], I.4.G. By $\hat{\pi}$ we mean $\bigcap\pi_\mu$ for all positive finite measures μ.

I.1.17. Theorem

Let (T,π) and (U,u) denote spaces with measures, let S be a Suslin space, let Σ be a multifunction defined on T with values in the space of non-empty subsets of S which graph belongs to $\pi\otimes\beta(S)$ (where $\beta(\cdot)$ is a borel field of subsets), and let Θ be a multifunction defined on T with values in the space of nonempty subsets of U, which graph belongs to $\pi\otimes u$. Suppose that $g: TxS \longrightarrow U$ is a $(\pi\otimes\beta(S),u)$-measurable function and that $g(t,\Sigma(t)\cap\Theta(t) \neq \phi$ for each $t\varepsilon T$. Then a $(\hat{\pi}, \beta(S))$ - measurable selector σ of Σ exists, such that $g(t, \sigma(t))\varepsilon\Theta(t)$.

This selector is at the same time a limit of a sequence of π-measurable functions with a finite number of values.

I.1.18. Theorem

Let (T,π) be a measurable space, let E denote a separable Banach space, let $f:T \longrightarrow E$ be a continuous function and suppose that N is a multifunction defined on T with values in a space of nonempty closed subsets of E. Then the mapping

$$R \ni t \longrightarrow \Sigma(t) \overset{df}{=} \{N(t) \ni x : \| f(t) - x \| = d(f(t), N(t))\}$$

is a multifuction measurable in the sense of definition I.1.12, if $\Sigma(t) \neq \phi$.

I.1.19. Remark

Let us abserve that if measurable fields usualy have measurable selectors, continuous fields do not need to have continuous selectors. Let us consider the following counterexample, constructed by Cellina . Let the multifunction $F:R \longrightarrow 2^{R^2}$ be described as $F(t)=\{(x,y):x=a(t)\sin(\omega/t),\ y=\cos(\omega/t),$ $\omega \varepsilon [0, 2\pi - h(t)]\}$, where a and h are continuous functions, $a(t)>0$ and $h(t)>0$ if $t \neq 0$, $a(t) \longrightarrow 0$ and $h(t) \longrightarrow 0$ if t tends to zero. The set F(t) for t=0 degenerates to the section $\{(x,y):x=0,\ y\varepsilon[-1,1]\}$ and it can be easily seen that F is continuous for all $t\varepsilon[-1,1]$. However, no selectors continuous at t=0 exist.

I.1.20. Definition

Let H be the space of all closed subsets of the space of linear continuous operators $L(Y^1, X)$ where Y^1 is some measurable Banach subspace of Y, and let G^L be defined as G is I.1.5 while replacing X by $L(Y^1, X)$. An absolutely continuous function $x\colon Y^1 \to X$ is called a trajectory of a multifunction $N\colon Y^1 \to G^L$ if $\dot{x}(y) = v(y)$ a.e. in Y^1, where $y \in Y^1$ and v is a measurable selector of the multifunction $N(\cdot, x(\cdot))\colon Y^1 \to G^L$.

I.1.21. Remark

Let us observe that $\dot{x}(y)$ is a continuous linear operator $Y^1 \longrightarrow X$ and $\dot{x}(y) \in L(Y^1, X)$. Therefore N should have values in the space of subsets of $L(Y^1, X)$. We define the trajectory in a subspace Y^1 of Y in order to replace (in the next section) Y^1 by an interval of R_+, and to interpret y as the time and X as a state space of some control system. Let us note that considering diverse subspaces Y^1 we get diverse corresponding sets of trajectories.

I.1.22 Definition

Let $Y = Y^1 x X$, Y^1 measurable. An absolutely continuous function $x\colon Y^1 \longrightarrow X$ is called a quasitrajectory of a multifunctions $N\colon Y \longrightarrow G^L$ if there exists a sequence of absolutely continuous and equibounded functions $\{x_i\colon Y^1 \longrightarrow X\}$, such that $x_i(y_0) = x(y_0)$ for some $y_0 \varepsilon Y^1$ and

(i) $x_i(y) \longrightarrow x(y) \ \forall\ y \varepsilon Y^1$,

(ii) $d\,(\dot{x}_i(y), N(y, x_i(y)) \longrightarrow 0$ a.e. in Y^1,

(iii) $\dot{x}_i(y)$ are equibounded a.e. in Y^1.

I.2 Orientor fields associated with control systems

From now on we shall restrict our attension to the case when $Y = J \times \Omega$(see I.1.5), where $J \subset R_+$ denotes an interval $[0,T]$, $T > 0$, and Ω is a set $\{ X \ni x : \| x \| \leq M_x, M_x > 0 \}$. The variable $t \in R_+$ will be interpreted as the time, and $x \in X$ as the system state variable. We have also Y' (of definition I.1.20) = J.

I.2.1. Definition

Let $f: J \times \Omega \times U \longrightarrow X$ be a bounded continuous function, where U is a separable real Banach space. Let $C: J \longrightarrow W$ be a multifunction, where W is the space of all nonempty closed and bounded subsets of U. The pair (f,C) is called control system, the set $C(t)$, $t \in J$ is called the control domain.

I.2.2. Definition

An absolutely continuous function $x: J \longrightarrow X$ is called a trajectory of a control system (f,C) on J with the initial condition $x_0 \in X$ if $x(0) = x_0$ and

$$\dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. on } J, \tag{I-1}$$

where $u: J \longrightarrow U$ is a measurable selector of the multifunction C.

I.2.3. Remark

Here $\dot{x}(t)$ denotes $dx(t)/dt$. Fortunately the Frechet's derivative in the case when $Y' = J \subset R_+$ is the limit of the usual difference quotient and belongs to X.

I.2.4. Definition

An absolutely continuous function $x: J \longrightarrow X$ is a quasitrajectory of a control system (f,C) on J with the initial condition x_0, if sequences of functions $\{x_i\}$ and $\{u_i\}$ exist, such that $x_i : J \longrightarrow X$ is absolutely continuous, $u_i : J \longrightarrow U$ is measurable, $x_i(0) = x_0$ for $i = 1,2,\ldots$, and

(i) $x_i(t) \longrightarrow x(t)$ on J,

(ii) $\| \dot{x}_i(t) - f(t, x_i(t), u_i(t)) \| \longrightarrow 0$ a.e. on J,

(iii) $u_i(t) \in C(t)$ on J.

I.2.5. Definition

Let (f,C) be a control system. The orientor field defined as follows

$$N(t, x) = \{X \ni v : v = f(t, x, u),\ u \in C(t)\}\ \forall\ t \in J,\ x \in \Omega,$$

is called the multifunction associated with the control system (f,C). The set $N(t, x)$ is called the control counterdomain.

I.2.6. Definition

A control system (f, C) is said to be normal if the multifunction $N: J \times \Omega \longrightarrow X$ associated with it is an orientor field.

I.2.7. Remark

Linear control system with continuous function f and compact and continuous control domain is obviously normal. There are, however, systems (even linear) which do not fulfil conditions of definitions I.2.1 and I.2.6. Consider, for example, a distributed control system with $f(t, x, u) = Ax + Bu$, where A is a differentialtion eperator (as in the heat or mass transfer equations). Then A is linear but not continuous operator and the system (f, C) is not a control system in the sense of def. I.2.1. The theory of distributed systems, described by partial differential equations, has been considered by the huge number of authors and will not be discussed here in it's classical form (see for example Lions [16], [17]). We are rather going to replace a non-continuous operator A by other, continuous one, which approximate A in an appropriate sense. It seems, that such approximation may be constructed in most practical cases, due to particular physical interpretation (see section III.3).

On the other hand, it is easy to find many examples of normal control systems. It is sufficient to consider a system with continuous function f and continuous, compact-valued control domain C(t). Then the associated multifunction N(t, x) is continuous and compact-valued, the field E(t, x) is closed and, by the Krein- Milman theorem, Q(t, x) = tend N(t, x) exists (as a closure of the set of the extremal points of N(t, x)). In this case (f, C) is a normal control system.

I.2.8. Assumption

Let us assume that $x_0 \in X_0 \overset{df}{=} \{X \ni x: \|x\| + T \cdot M < M_X\}$, where M appears in I.1.5 and M_X determines the set Ω (the beginning of I.2).

I.2.9. Remark

It follows from I.2.8 that if a local solution of (I.1) exists, then a solution of (I.1) exists on the whole interval J(Deimling [18], § 3). The inclusion $x_0 \in X_0$ implies that $gr(x) \subset int(J \times \Omega)$, where x is a trajectory of a system (f,C) and gr(x) denotes the graph of x.

I.2.10 Theorem

Let (f,C) be a control system (not necessarily normal) and let N be a multifunction associated with (f,C). We put $Y' = R_+$ and $Y = R_+ \times X$ in definitions I.1.20 and I.1.22. We suppose also that N and C are measurable and that I.2.8 holds. Then

(i) the set of all trajectories of N with an initial condition x_0 defined on J and that of all trajectories of (f, C) defined on J with the same initial condition are identical to each other

(ii) the set of all quasitrajectories of N with an inital condition x_0 defined on J and that of all quasitrajectories of (f, C) defined on J with the same initial condition are identical to each other.

Proof

Let us denote by $\Gamma(N)$, $\Gamma(f,C)$, $\Gamma_q(N)$ and $\Gamma_q(f,C)$ the sets of trajectories of N, trajectories of (f,C), quasitrajectories of N and quasitrajectories of (f,C), respectively (with the same initial condition x_0).

(i) If $\Gamma(N)$ and $\Gamma(f,C)$ are empty, the theorem is trivially fulfilled. Let us assume that $\Gamma(f,C)$ is nonempty. Then from I.22 and I.2.5 follows that a measurable selector v of the field N exists, such that $v(t) = \dot{x}(t) = f(t, x(t), u(t)) \in N(t,x)$ a.e. on J, where x is an absolutely continuous function and u is a measurable selector of C. This means that x is a trajectory of N and thus $\Gamma(f,C) \subset \Gamma(N)$. To prove the opposite inclusion, let us assume that $\Gamma(N)$ is nonempty and that x is a trajectory of N. Then x is a measurable selector of N (we identify the functions which differ over a set of measure zero). Let us consider a nonempty set $P(t)=\{C(t) \ni v: f(t,x(t),v)=\dot{x}(t)\}$.

From I.1.17 follows that P have a measurable selector u. Indeed, let us substitute in I.1.17 $g(t,\sigma) = f(t,x(t), \sigma)$, $\Sigma(t) = C(t)$, $\Theta(t) = \{\dot{x}(t)\}$ (one-point set), and $u(t)=\sigma(t)$. Then, from I.2.2 we conclude that x is a trajectory of (f,C) and that $\Gamma(N) \subset \Gamma(f,C)$ which completes the proof of (i).

(ii) If $\Gamma_q(N)$ and $\Gamma_q(f,C)$ are empty, the theorem holds trivially. Let us suppose that $\Gamma_q(f,C)$ is nonempty. Observe that from I.2.4 and I.2.5 follows that $f(\cdot, x_i(\cdot), u_i(\cdot))$ is a measurable selector of $N(\cdot, x(\cdot))$. Consequently from the fact that $\| \dot{x}_i(t) - f(t, x_i(t), u_i(t)) \| \to 0$ follows that $d(\dot{x}_i(t), N(t, x_i(t))) \to 0$ and that $\Gamma_q(f,C) \subset \Gamma_q(N)$. Let us prove the apposite inclusion. Suppose that $\Gamma_q(N)$ is nonempty. Then, a sequence $\{x_i\}$ exists fulfilling I.1.22. Let us denote $B_i(t)=B(\dot{x}_i(t), c_i)$, $N_i(t) = N(t, x_i(t))$, $d_i(t) = d(\dot{x}_i(t), N(t, x_i(t)))$, where $c_i \varepsilon R_+$ for $i = 1, 2, \ldots,$ $B(a, b)$ is a ball with centre in a and radius b.

Let us consider a multifunction $P_i(t)= \{C(t) \ni v: f(t,x_i(t), v) \varepsilon B_i(t)\}$ an substitute $c_i(t)=2d_i(t)$. It is clear (see I.2.5) that $B_i(t) \cap N_i(t) \neq \phi \; \forall \; t \varepsilon J$ and, consequently, $P_i(t)$ is nonempty. We apply now once more theorem I.1.17, substituting $g(t,\sigma) = f(t,x_i(t), \sigma)$, $\Sigma(t) = C(t)$, $\mathcal{B}(t) = B_i(t)$, $u_i(t) = \sigma(t)$. By this way we couclude that a measurable selector u_i of the multifunction P_i exists. Taking into account that $d_i(t) \to 0$ we see that x, x_i and u_i fulfil the conditions of I.2.4 which means that $\Gamma_q(N) \subset \Gamma_q(f,C)$, which completes the proof.

I.2.11. Theorem

The set of quasitrajectories of an orientor field defined on J with initial condition x_0 is closed in the sense of the norm $\| \cdot \|_\infty$, where by $\|x\|_\infty$ or sup norm of x we mean sup $\| x(t) \|$, $t \varepsilon J$.

Proof

If $\Gamma_q(N)$ is empty, the theorem holds. Let us assume that $\Gamma_q(N)$ is nonempty. Let $x_n(0) = x_0$, $x_n \longrightarrow x$ in the norm $\|\cdot\|_\infty$, where $\{x_n\}$ is a sequence of quasitrajectories of N. Then for each n there exists a sequence $\{x_{i,n}\}$ of absolutely continuous functions, such that $x_{i,n}(t) \to x_n(t)$, $r_{i,n}(t) \overset{\text{df}}{=} d(\dot{x}_{i,n}(t), N(t, x_{i,n}(t))) \longrightarrow 0$ if $i \longrightarrow \infty$, for almost all $t \in J$. Then, this is true for all $t \in J \setminus J_n$, where J_n is of measure zero. Let us denote $J' = \cup J_n$. It is clear that $\mu(J') = 0$. Observe that for each $t \in J \setminus J'$ we have $\forall\, n, c\ \exists\ i(n,c): \forall\, k \geq i(n,c)$, $\| x_{k,n}(t) - x_n(t)\| + r_{k,n}(t) \leq c$. Substituting $c = 1/n$ and taking into account that $x_n \longrightarrow x$ we conclude that

$$x_{i(n,c),n}(t) \longrightarrow x(t)$$

$$r_{i(n,c),n}(t) \longrightarrow 0$$

if $i \longrightarrow \infty$. The convergence of $\{x_n\}$ is uniform (as a convergence in the sup norm), and therefore x is absolutely continuous. Thus, it is clear that x is a quasitrajectory of N, which completes the proof.

I.3. Existence of trajectories and quasitrajectories

It is known that continuity of the right-hand side is not sufficient for existence of solutions to an ordinary differential equation in a Banach space. To achieve existence, various assumptions are imposed on the right-hand side, e.g. the Lipsnitz condition (Maurin [20]) or conditions of dissipative or compactness type (Deimling [18]). Let us introduce the following.

1.3.1. Assumption

Let (f,C) be a control system and let θ be some neighbourhood of the set $\bigcup_{t\in J} C(t)$, where $C(t)$ is the control domain. We assume that C is measurable and that the function f fulfils the following condition.

$$\| f(t,x_1,u) - f(t,x_2,u)\| \le \omega\,(t,\|x_1 - x_2\|) \qquad (I\text{-}2)$$

for all $x_1, x_2 \in \Omega$, $t \in J$, $u \in \theta$, where ω: $R_+ \times R_+ \to R_+$ is such a function that $v(t) \equiv 0$ is the unique solution with the initial condition $v(0) = 0$ of the differential equation $\dot{v}(t) = \omega(t,v(t))$, over the interval J

I.3.2. Theorem

Let $N:J \times \Omega \to G$ be the multifunction associated to a control system (f,C) and let I.2.8 and I.3.1 hold. Then the sets $\Gamma(N)$, $\Gamma_q(N)$, $\Gamma(f,C)$ and $\Gamma_q(f,C)$ are nonempty.

Proof

Due to I.1.15 a measurable selector u of the multifunction C exists. Let us now fix the function u. Applying the theorem of Lusin we conclude that a sequence of measurable functions $\{\varepsilon_i\}$ exists, such that $w_i = u + \varepsilon_i$ is continuous and $\varepsilon_i \to 0$ in the sense of measure. Let us denote $g_i(t,x) = f(t,x,u(t) + \varepsilon_i(t))$. It is clear that g_i fulfills condition (I-2). Consequently, a solution x_i of the equation $\dot{x}(t) = g_i(t,x(t))$ with the initial condition x_0 exists on J. Let us consider two of such solutions, say x_i and x_j. Denoting $r_{ij} = x_i - x_j$, $e_{ij}(t,x) = g_i(t,x) - g_j(t,x)$ we have $e_{ij}(\cdot,x) \to 0$ in the sense of measure on J for each x and $\| g_i(t,x_i) - g_i(t,x_j)\| \leq \omega(t,\|r_{ij}\|) + \|e_{ij}\|$. Consequently, we have

$$\frac{d}{dt}\|r_{ij}\| \leq \|\dot{r}_{ij}\| \leq \omega\,(t,\ \|r_{ij}\|\,) + \|e_{ij}\| \tag{a}$$

Let ϕ_{ij} be the upper integral of the equation $\dot{v} = \omega(t,v) + \|e_{ij}\|$ with the initial condition zero. It is known that any solution $z_{ij} = \|r_{ij}\|$ to the differential inequality (a) is mayorized by ϕ_{ij}. It is clear that $\{\phi_{ij}\}$ is a compact set of functions and has a subsequence $\{\phi_{ij}\}^1$ convergent to some function ϕ. If ϕ were not equal zero a.e., the equation $\dot{v} = \omega(t,v)$ should have a non-unique solution passing through (0,0) on J, which contradicts our assumptions. Thus, $\lim \{\phi_{ij}\}^1 = \phi \equiv 0$ a.e. on J and we conclude that $\{r_{ij}\}$ has a subsequence tending to zero in the sense of measure, and that $\{x_i\}$ has a subsequence convergent in this sense. For $\dot{x}_i$ are equibounded and x_i absolutely continuous, it is clear that the limit $\bar{x}$ of this subsequence is absolutely continuous and that $\dot{\bar{x}} = f(t,\bar{x},u)$ a.e. on J. This means that to any measurable selector u of the control domain C, corresponds a trajectory of the control system (f,C). Thus we conclude that $\Gamma(f,C)$ is nonempty and, due to I.2.10 $\Gamma(N)$ is nonempty. For each trajectory is at the same time a quasitrajectory, it is clear that $\Gamma_q(N)$ and $\Gamma_q(f,C)$ are nonempty, which completes the proof.

I.3.3. Remark

Let us observe that existence of trajectories do not implies existence of time-optimal trajectories (to be defined in the sequel), because $\Gamma(f,C)$ need not be closed. Moreover, closedness of $\Gamma_q(f,C)$ in the sup norm also do not implies existence of optimal quasitrajectories, because the set of points belonging to graphs of quasitrajectories (the union of graphs) need not be closed as a subset of $R_+ \times X$ in the norm induced by $|\cdot|$ and $\|\cdot\|$ of X. To consider the problem of existence of optimal quasitrajectories, we must establish some properties of non-convex control systems.

I. 4. Systems with nonconvex counterdomain of control

I.4.1. Definition

Let C be the control domain of a normal control system (f,C) and let N be the orientor field associated with (f,C). The set

$$D(t,x) = \text{the smallest set } V \subset C \text{ such that } f(t,x,V) = \text{tend } N(t,x)$$

is called the tendor kernel of the control domain C.

I.4.2. Remark

In the theory of linear systems the term "nonconvex control" often refers to systems with nonconvex set of permissible controls, i.e. to systems with nonconvex control domain. The "relaxed" controls are often defined as controls belonging to the convex hull of control domain. It should be noted that in the case of nonlinear system the most important property appears to be convexity of the control counterdomain and not of the control domain. It will be easy to observe that the convex hull of control domain C have no meaning here and do not appear in the sequel.

In practical applications systems with nonconvex control counterdomain are of particular importance. Those are mainly systems with switching regulators generating so-called "bang-bang" controls. The optimal controls are often of "bang-bang" type with finite or infinite number of switchings, called sliding regimes. The class of controls belonging to the tendor kernel of control domain is a generalization of the "bang-bang" controls. The following questions arise:

(i) is controlling with the full control domain C equivalent (in some sense) to controlling with the tendor kernel of C?

(ii) what are relations between trajectories and quasitrajectories of the systems (f,C), (f,D) and of the associated fields N and Q?

In analogy with the Ważewski's approach, let us formulate some theorems connected with nonconvex control.

I.4.3. Theorem

Let $P:J \rightarrow G$ be a multifunction convex and continuous in the Hansdorff sense. By $x: J \rightarrow X$ we shall mean an absolutely continuous function. If a sequence of trajectories $\{v_i\}$ of P exists, such that $v_i(t) \rightarrow x(t)$ on J, then $\dot{x}(t) \varepsilon P(t)$ a.e. on J, i.e. x is a trajectory of the field P.

Proof

Let us assume, on the contrary, that a set $M \subset J$ and a constant $q > 0$ exist, such that $\mu(M) > 0$ and

$$d(\dot{x}(t), P(t)) \geq q \quad \forall\, t \,\varepsilon\, M. \tag{I-3}$$

For x is absolutely continuous and X is reflexive, $\dot{x}$ exists a.e. on J and is measurable. Hence, J is a sum of a null set and a countable family of mutualy disjoint compact sets J_i such that $\dot{x}|_{J_i}$ is continuous and $\mu(J_i) > o$ for each $i = 1,2,\ldots$ (see Bourbaki [19], IV, §5). Let $B \subset J$ be a compact set, $\mu(B) > 0$, and $s \in B$ be a fixed point. Assume now that a neighbourhood α of s exists such that $\mu(\alpha \cap B) = 0$. For s is an arbitrary point of B, the set B is locally null. It has a finite external measure and, consequently, we should have $\mu(B) = 0$. Hence, a neighbourhood with the above properties can not exist for arbitrary $s \in B$, unless B is of measure zero. Thus, for each nonzero compact $B \subset J$ we have

$$\exists\, B \ni s:\ \mu(\alpha \cap B) > 0 \quad \forall\, \alpha:\ \ s \in \alpha, \quad \alpha \text{ open.} \tag{I-4}$$

It is clear that a number k must exist, such that $\mu(J_k \cap M) > 0$. Let $J_k \cap M = Z$. From (I-4) it follows that a point $\eta \in Z$ exists such that $\mu(\alpha \cap Z) > 0$ for any neighbourhood α of η. The function $\dot{x}|_Z$ is continuous at η. Hence, from (I-3) it follows that a convex set A and a neighbourhood ζ of η exist such that $\dot{x}(t) \in A \ \forall\, t \in \zeta' = \zeta \cap Z$ and $\inf\{ r:\ r = d(\beta, P(\eta)),\ \beta \in A\} > q/2$. For P is continuous, a neighbourhood δ of η exists such that $P(t) \in V(P(\eta), q/4)\ \forall\, t \in \delta$, where V stands for neighbourhood of a set. Let $\zeta' \cap \delta = \gamma$. We have $d^*(A, P(t)) > 0 \ \forall\, t \in \gamma$. Applying the Banach theorem on separation of convex sets we see that the continuous linear functional $z: X \to R$ exists, such that $z(y) < 1 \ \forall\, y \in V(P(\eta), q/4)$ and $z(y) > 1 \ \forall\, y \in A$. The functions v_i are trajectories of P and, consequently, $z(\dot{v}_i(t)) < 1$ a.e. on γ and $z(\dot{x}(t)) > 1$ a.e. on γ. Taking into account that $\mu(\gamma) > 0$ and integrating the real-valued functions $z(\dot{x}(\cdot))$ and $z(\dot{v}_i(\cdot))$ over the set γ we obtain

$$\int_\gamma z(\dot{x}(t))\, dt > \mu(\gamma), \quad \int_\gamma z(\dot{v}_i(t))\, dt < \mu(\gamma)$$

Differently speaking, a constant $c > 0$ exists such that

$$\int_{\gamma} z(\dot{x}(t) - \dot{v}_i(t))\, dt > c \tag{I-5}$$

Let us observe that $v_i(t) \to x(t)$ on J and, consequently,

$$\lim_{i \to \infty} \int_a^{\sigma} z(\dot{x}(t) - \dot{v}_i(t))\, dt = 0$$

for arbitrary $\sigma \in [a, b] = J$. On the other hand, for any Lebesgue-measurable function $y : J \to R$ the equality holds

$$\lim_{i \to \infty} \int_J y(t)\, z(\dot{x}(t) - \dot{v}_i(t))\, dt = 0$$

(see Klambauer [34], chapter 3, theorem 3). Taking as y the characteristic function of γ we obtain

$$\lim_{i \to \infty} \int_{\gamma} z(\dot{x}(t) - \dot{v}_i(t))\, dt = 0 ,$$

which is the contradiction to (I-5). Thus, the constant q and the set M satisfying (I-3) do not exist. Taking into account that P(t) is closed we conclude that $\dot{x}(t) \in P(t)$ a.e. on J, which completes the proof.

I.4.4. Remark

The property proved above is analogycal to some property of solutions to contingent equations (Zaremba [9]). The similar theorem for contingent equations in Banach spaces was given by Chow and Shuur [12].

I.4.5. Theorem

Let $E: J \times X \to G$ be a convex-valued multifunction, continuous is the Hansdorff sense. Then each quasitrajectory of E on the interval J is at the same time a trajectory of this field.

Proof

From def. I.1.22 it follows that a sequence of functions $\{x_i\}$ exists, such that

$$x_i(t) \to x(t),$$

$$d(\dot{x}_i(t), E(t,x_i(t)) \to 0 \quad \text{a.e. on } J,$$

where x_i are absolutely continuous and $\dot{x}_i$ are equibounded on J.

Suppose that x is not a trajectory of E. Then, a nonzero set $L \subset J$ and a constant q must exist such that

$$d(\dot{x}(t), E(t,x(t)) \geq q > 0 \quad \text{a.e. on } L \tag{I-6}$$

Let us consider a multifunction $\tilde{E}(t) = V(E(t,x(t)), q/2)$, $\tilde{E}(t)$ being the $q/2$ neighbourhood of $E(t)$, where x is some fixed quasitrajectory of E. The multifunction $\tilde{E}$ is convex and continuous on J. Let $T_i = \{J \ni t: \dot{x}_i(t) \notin \tilde{E}(t)\}$. It is known that the convergence "almost everywhere" on J implies the convergence in measure on J. Consequently, $\mu(T_i) \to 0$. Let y be a measurable selector of E, and let $w_i(t) = \dot{x}_i(t)$ a.e. on $J \setminus T_i$ and $w_i(t) = y(t)$ on T_i. Next, let us define a function z_i such that $z_i(0) = x_i(0)$ and $\dot{z}_i(t) = w_i(t)$ a.e. on J. Observe that z_i is absolutely continuous, $z_i(t) \to x(t)$ on J and $\dot{z}_i(t) \in \tilde{E}(t)$ a.e. on J. Thus, the functions z_i, x and the multifunction $\tilde{E}$ satisfy the hypotheses of theorem I.4.3. Consequently, we have $\dot{x}(t) \in \tilde{E}(t)$. This is a contradiction to (I-6), which completes the proof.

I.4.6. Lemma

Let S be a multifunction defined on J and let $S(t) = \operatorname{conv} R(t)$, where $R(t)$ is a finite set of points of the space X. Let T_i be a countable set of mutualy disjoint subsets of J such that $\mu(J) = \mu(\cup T_i)$. Let $R(t)$ be constant on cach of the subsets T_i and let $u: J \to X$ be a function which is constant on each T_i and $u(t) \in S(t)$ on J. Then a measurable function $w : J \to X$ exists such that $w(t) \in$ tend $S(t)$ a.e. on J and $\int_J u(t)\, dt = \int_J w(t)\, dt$.

Proof

Let us denote $c_i = u(t)$, $S_i = S(t)$, $R_i = R(t)$ for $t \in T_i$. We have $c_i \in S_i = \operatorname{conv} R_i(t)$ on T_i. For S_i is the set of all convex combinations of points belonging to R_i, the sets of points $s_{i,1}, s_{i,2}, \ldots, s_{i,k(i)}$ and of numbers $b_{i,1}, b_{i,2}, \ldots, b_{i,k(i)}$ must exist such that

$$C_i = \sum_j b_{i,j} s_{i,j} \quad , \quad \sum_j b_{i,j} = 1 \; , \quad b_{i,j} \geq 0 \; ,$$

where $s_{i,j} \in R_i$. Let us divide T_i into $k(i)$ measurable mutualy disjoint subsets $T_{i,1}, T_{i,2}, \ldots, T_{i,k(i)}$ such that $\mu(T_{i,j}) = b_{i,j}\, \mu(T_i)$. Substituting $w(t) = s_{i,j}$ on $T_{i,j}$ we obtain the function with required properties, which completes the proof.

I.4.7. <u>Remark</u>

In the finite-dimensional case the tendor field Q corresponding to a continuous orientor field N is σ-measurable and, in the same time, Lusin-measurable (see Remark I.1.11). In the case of a Banach space X the field Q need not be measurable in the sense of Lusin. This follows from the fact that separability of X does not imply separability of G. Then, the theorem of Lusin on measurable separable-valued functions does not hold for the function $Q : J \times \Omega \to G$. This is the relevant difficulty while considering control systems in Banach spaces. Let us observe, that if we restrict our attension to compact-valued fields, this difficulty disapperar. Let us, however, do not impose any compactness assumptions, requiring, however, that the field Q is Lusin-measurable.

Before formulating the next assumptions and theorems, let us note that the integrals of vector-valued functions appearing in our considerations are of

Bochner type (see Maurin [20], Warga [15] I.4.D). Let us recall, that if X is reflexive, then an absolutely continuous function $f: R \to X$ is differentiable almost everywhere, and

$$f(b) - f(a) = \int_a^b \dot{f}(t)\, dt,$$

where $[a, b]$ is an interval of R.

I.4.8. Definition

Let (f,C) be a normal control system and let N be the orientor field associated with (f,C). The system (f,C) is said to be regular if Q = tend N is ℓ.s.c. and measurable in the sense of Lusin as a function $Q(\cdot, x(\cdot)): J \to G$ (G equipped with the Hansdorff metric), for any absolutely continuous function $x: J \to X$.

I.4.9. Theorem

Let (f,C) be a regular control system with associated fields N, E = conv N, Q = tend N. Let D be the tendor kernel of the control domain C and let $x : J \to X$ be an absolutely continuous function such that $x(0)= x_0$, where x_0 fulfils assumtion I.2.8.Then the following are quivalent conditions

(i) x is a quasitrajectory of (f,C),

(ii) x is a quasitrajectory of (f,D),

(iii) x is a quasitrajectory of E,

(iv) x is a quasitrajectory of N,

(v) x is a quasitrajectory of Q,

(vi) x is a trajectory of E.

Proof

Let us start with the implication (vi) $\Rightarrow$ (v) $\Rightarrow$ (iv) $\Rightarrow$ (iii).

For the orientor $E(t, x(t))$ is bounded on J, a constant K exists such that $\| v \| < K \; \forall v \in E(t, x) \; \varepsilon \; J \times X$. The field Q is Lusin-measurable and, consequently, for any constant $q > 0$ a family of compact mutually disjoint subsets $\{F_i\}$ of J exists such that $\mu(J) = \mu(\cup F_i)$, $\| \dot{x}(t) - \dot{x}(s) \| \leq q$, $d^*(Q(t, x(t)), Q(s, x(s)) \leq q$ for all $t \in F_i$, $s \in F_i$.

Let us choose a point $t_i \in F_i$ and denote $u(t) = \dot{x}(t_i)$, $\tilde{Q}(t) = Q(t_i, x(t_i))$ for all $t \in F_i$, $i = 1,2,\ldots$. We have

$$\| u(t) - \dot{x}(t) \| < q, \quad d^*(\tilde{Q}(t), Q(t, x(t))) \leq q \quad \text{a.e. on } J. \qquad \text{(I-7)}$$

Observe that $E(t_i, x(t_i))$ is the set of all convex combinations of points belonging to $Q(t_i, x(t_i))$. Thus, a finite set R_i of points exists such that $R_i \subset Q(t_i, x(t_i))$, $\dot{x}(t_i) \in S_i = \operatorname{conv} R_i$. Denoting $S(t) = S_i$ for all $t \in F_i$ we have $u(t) \in S(t)$, tend $S(t) \subset \tilde{Q}(t)$ and $\| u(t) \| \leq K$ a.e. on J. Let us divide J into m subintervals $J_p = [a_{p-1}, a_p]$ in such a way that $a_p = pT/m$, $\mu(J_p) \leq q$ (recall that $T = \mu(J)$). Observe that the functions u and S are constant on each of the sets $J_p \cap F_i$. Now we apply lemma I.4.6. Denoting $I(u,A) = \int_A u(t)\, dt$ we conclude that a measurable function $w\colon J \to X$ exists such that $\| w(t) \| \le K$, $w(t) \in \text{tend } S(t)$ a.e. on J and

$$I(u, J_p) = I(w, J_p) . \tag{I-8}$$

Denoting $k(t) = I(w, [0, t])$ we have $\| k(t) \| < KT$, $\dot{k}(t) = w(t)$, $\| \dot{k}(t) \| \le K$ and

$$\dot{k}(t) \in \tilde{Q}(t) \quad \text{a.e. on } J \tag{I-9}$$

Observe that $\| k(t) - x(t) \| \le I(w-u, [0, t]) + I(u - \dot{x}, [0, t])$, and $I(w - u, [0, t]) = I(w - u, [a_{p-1}, t])$ where p is such that $t \in J_p$. Taking into account that $|t - a_{p-1}| \le \mu(J_p) \le q$ we can point out that

$$\| k(t) - x(t) \| \le (2K + T)q \ \forall\ t \in J. \tag{I-10}$$

From (I-9) and (I-10) we have

$$d(\dot{k}(t), Q(t, x(t))) \le q \quad \text{a.e. on } J. \tag{I-11}$$

Let us introduce a sequence of numbers $\{q_i\}$ such that $q_i \to 0$ and denote by k_i the function k defined above with $q = q_i$. Substituting k_i and q_i for k and q in (I-10) and (I-11) we obtain

$$k_i(t) \to x(t) \quad \text{on } J \tag{I-12}$$

$$d(\dot{k}_i(t), Q(t, x(t))) \to 0, \quad \| \dot{k}_i(t) \| \le K \quad \text{a.e. on } J \tag{I-13}$$

The lower semi-continuity of the field Q results in the inclusion $Q(t, x(t)) \subset V(Q(t, k_i(t)), r)$ which holds for an arbitrary $r > 0$ and i big enough. By virtue of (I-13) it is clear that

$$d(\dot{k}_i(t), Q(t, k_i(T))) \to 0 \quad \text{a.e. on } J. \tag{I-14}$$

Taking into account (I-12), (I-14) and the fact that $\dot{k}_i(t)$ are equibounded on J we see that x is a quasitrajectory of the field Q. For $Q(t, x) \subset N(t,x) \subset E(t,x)$ it is clear that x is a quasitrajectory of N and E. Thus we have (vi) $\Rightarrow$ (v) $\Rightarrow$ (iv) $\Rightarrow$ (iii).

From the definition of the tendor kernel D(I.4.1) it follows that (iv)$\Leftrightarrow$(ii). The equivalence (iii) $\Leftrightarrow$(i) follows from theorem I.2.10, and from I.4.5 we have (iii) $\Leftrightarrow$ (vi), which completes the proof.

I.4.10. Remark

The above proof is a slightly modified version of the proof given by Ważewski [2] for the case $X = R^n$. See also [3, 5, 7].

From the practical point of view it is interesting to establish under what conditions quasitrajectories might be approximated by trajectories. This follows from the following.

I.4.11. Theorem

Suppose that (f,C) is a control system and that I.3.1 holds. Then, each quasitrajectory of (f,C) is the pointwise limit of a sequence of trajectories of (f,C).

I.4.12. Remark

As was proved earlier (I.3.2), I.3.1 implies that trajectories and quasitra jectories of (f,C) exist. We omit the proof of the above theorem, because it is similar to that of I.3.1 and is nearly identical as in the finite-dimensional case which can be found in the paper of Turowicz [10].

Now, let us make some remarks on tendor control in practical applications. As mentioned earlier, by tendor control we mean some generalization of control signals generated by switching (so-called "bang-bang") controllers used in many systems of automatic control. Fig. 2a illustrates a possible shape of orientor N in R^2. In this case the tendor of N consists of the four points A, B, C and D.

This is an example of typical "bang-bang" control, i.e. control which is composed of a finite number of signals and finite or infinite number of switchings between those signals. The case indicated in fig. 2b is a little bit more complicated. Now, the tendor of N consists of infinite number of points, namely A, D and the whole arc B C. The control signal which could be generated in this case might also be obtained by swithing between the points of the tendor set, but it can not be realized by a simple "four-point" controller, as in the case 2a. This is what we call tendor control.

As for the approximation theorem I.4.11 it should be noted that the asumption I.3.1 holds for all functions fulfilling the Lipshitz condition. On the other hand, even in R^n counterexamples might be constructed for functions which do not fulfill I.3.1. Let us quote the example given by Pliś [33].

Let a system (f,C) has a following associated orientor: $N(t,x) = N(x) \subset R^2$ which consists of the points

$$a(x) = (a_1(x),\ a_2(x)),\ b(x) = (b_1(x),\ b_2(x)),$$

where

$$a_1(x) = 1,\quad b_1(x) = -1,\quad a_2(x) = b_2(x) = x_1^2 + |x_2|^{1/2} \tag{I-15}$$

where $x = (x_1,\ x_2) \varepsilon R^2$.

The convex orientor $E(x) = \text{conv } N(x)$ consists of the points v belonging to the section $|v_1| \leq 1,\ \ v_2 = x_1^2 + |x_2|^{1/2}$.

A function $y(t) = (y_1(t),\ y_2(t))$ such that

$$\dot{y}_1(t)\ \varepsilon [-1,\ 1] \tag{I-16}$$

$$\dot{y}_2(t) = y_1(t)^2 + |y_2(t)|^{1/2} \tag{I-17}$$

almost everywhere on J is a trajectory of E and, consequently, it is a quasi-trejectory of N. Thus, the function $\tilde{y}(t) \equiv 0$ is a quasitrajectory of N, whith the initial condition zero.

Let us observe that $\tilde{y}$ is not a trajectory of N, because on each trajectory we should have $\dot{y}_1 \neq 0$ and $y_1 \neq 0$ on any subinterval of J. Moreover, the quasitrajectory $\tilde{y}$ can not be approximated by any sequence of trajectories of N. This follows from (I-17), because we have $\dot{y}_2(t) > |y_2(t)|^{1/2}$ and, consequently, $y_2(t) > 1/4\ t^2$ and no trajectory passes through a sufficiently small neighbourhood of $\tilde{y}$. Of course, the above system does not fulfil the assumption I.3.1.

It is easy to verfy that if we put in (I-15), for example, $a_2=b_2= x_1^2 + x_2$, then $\tilde{y} \equiv 0$ is a quasitrajectory of N, $\tilde{y}$ is not a trajectory, but it can be approximated by a sequence of trajectories of N, being a simple "sliding regime". It should be noted, that if many of linear systems fulfill our assumptions, there are also linear systems which does not. Firstly, a linear operator (function f) in a Banach space need not be continuous. Consequently, we can not apply directly our results to evolution equations with differentional operators. This problem will be discussed in the sequel and some applications to nonlinear diffusion processes will be considered. Secondly, we assume that f is bounded. It might appear, from the theoretical point of view, that this is a strong assumption. Let us observe, however, that we are dealing here with nonlinear systems, and a simplest nonlinearity, which always exists in real control systems is saturation. Thus, the boundedness assumption seems to be reasonable from a practical point of view.

I.5. Existence of time-optimal trajectories

It was pointed out (I.2.11) that the set of quasitrajectories of orientor field is closed in the sup norm. Hence, we might expect that optimal quasi-trajectories (i.e. quasitrajectories with some "extremal" properties, defined in the sequel) exist. Before formulating appropiate existence theorem, let us prove the following.

I.5.1. Theorem

The union of graphs of all trajectories $x: J \to X$ of a convex multifunction E with the initial condition x_0 is a closed subset of $J \times X$.

Proof

Let us denote the above union of graphs by Z, and let $Z_p = Z \cap A_p$, where A_p is the hyperplane $t = p$. Let $\{z_n\}$ be a sequence of points belonging to Z_p, converging to a point z. To each point z_n corresponds a trajectory x_n of E, such that $x_n(p) = z_n$. Each of x_n is differentiable on $J \setminus I_n$, where I_n is of measure zero. Let us denote $J' = J \setminus I$, where $I = \cup I_n$ (of course $\mu(I) = 0$). Then, a sequence of measurable functions $v_n : J \to X$ exists such that $v_n = \dot{x}_n(t)$ on J'. Let $w_n = 1/n \;\; \Sigma_{i=1}^{n} v_i$. The orientor $E(t, x)$ is bounded. Hence, w_n is a Cauchy sequence in the sup norm. Functions

$$y_n^{(s)} = \int_0^s w_n(t)\,dt$$

are trajectories of E, which follows from convexity of this field. From linearity and continuity of the Bochner integral follows that $y_n(p) \to z$ and $y_n(s)$ is a Cauchy sequence for all $s \in [0, p]$. Thus, $\{y_n(s)\}$ is precompact for all $s \in [0, p]$. For $\dot{y}_n$ are equibounded, $\{y_n\}$ is a set of equicontinuous functions. From the theorem of Ascoli follows, that $\{y_n\}$ is compact in the sup norm. Hence, a subsequence $\{y_n\}'$ of $\{y_n\}$ exists, which converge to some function h in this norm. For $y_n(p) \to z$, we have $h(p) = z$. Due to I.4.9 and I.2.11 the function h is a trajectory of E, which completes the proof.

I.5.2. Remark

From the above theorem and from I.4.9 follows that the union of graphs of quasitrajectories of corresponding fields N, Q and systems (f,C) and (f,D) are also closed. This fact is closely connected with the existence of time-optimal quasitrajectory.

I.5.3. Definition

Let $\bar{x} : J \to X$ be a trajectory (a quasitrajectory) of a control system (f,C) with initial condition x_0. Let $g : J \to X$ be a continuous function, called the goal trajectory and let $t' \in J$ exists, such that $\bar{x}(t') = g(t')$ and $t' = \inf \{J \ni t : x(t) = g(t)\}$, where $x \in \Gamma(f,C)$ ($x \in \Gamma_q(f,C)$ respectively). Then $\bar{x}|_{[0, t']}$ is called a time-optimal trajectory (quasitrajectory) of (f,C) whith respect to the goal g and the initial condition x_0.

I.5.4. Theorem

Let (f,C) be a regular control system, let $g : J \to X$ be a goal trajectory and $(\bigcup \{ gr(x) : x \in \Gamma_q(f,C)\}) \bigcap gr(g) \neq \phi$. Then, a time-optimal quasitrajectory of (f,C) exists.

I.5.5. Remark

The proof of this theorem is omitted, for it is a simple consequence of the theorem I.5.1. It should be observed that the theorem I.5.1 is relevant. Let us note that optimal quasitrajectory is defined on $[0, t']$, i.e. on the part of J on which the goal is reached. Consider, for example, the family of functions $f_a(t) = |\sin t|^{1/a}$ $t \in [0, 2\pi]$, where $a \in [1, \infty]$. This family is closed in the sup norm, but $Z = \bigcup gr(f)$ is not, because the points $(\pi, 1)$ $(2\pi, 1)$, $(0,1)$ do not belong to Z. In this example the derivative of f_a is not bounded. Let us construct a similar counterexample using functions with bounded derivatives. Let $X = \ell_2$ and consider the following set of functions $\{f_k : [0, \pi] \to \ell_2\}$

$$f_1(t) = (\sin t + t,\ 0,\ 0,\ \ldots)$$
$$f_2(t) = (0\ ,\ \sin t + t/2,\ 0,\ 0,\ \ldots)$$
$$\vdots$$
$$f_k(t) = (0, \ldots, 0,\ \sin t + t/k\ ,\ 0,\ 0,\ \ldots),$$
$$\vdots$$

f_k having the non-zero component at the k-th position. The set $\{f_k\}$ is closed in the sup norm, as a set of isolated points in $C([0,\pi], \ell_2)$. However, we have $f_k(\pi) \to (0, 0, 0, \ldots)$, but the point $(0, 0, 0, \ldots)$ does not belong to graph of any function of $\{f_k\}$, i.e. $\cup\, gr(f_k)$ is not closed.

I.5.6. Definition

Let $I : \Gamma_q(f,C) \to R$ be a functional defined as

$$I(x) = \int_J e(t, x(t), u(t))\, dt$$

A function $\bar{x}: J \to X$ such that $I(\bar{x}) = \inf\{I(x) : x \in \Gamma_q(f,C)\}$ is said to be a quasitrajectory optimal in the sense of criterion I.

I.5.7. Remark

Of course, we can define an optimal trajectory in the same way. Extending the state variable x to $(x, y) \in X \times R$, where $\dot{y} = e(t, x(t), u(t))$ with the initial condition $y(0) = 0$, we obtain new control system (g, C) where $g = (f, e)$. If the (scalar) function e is sufficiently regular, the theorem I.5.1 can be applied to this system to prove that a quasitrajectory optimal in the sense of criterion I exists.

II. Reachable sets and optimal control

II.1. Some properties of reachable sets

In this section we formulate some, fairly general, version of the Pontriagin's maximum principle. This is preceded by certain considerations concerning properties of reachable sets. It appears, that, in these conside_rations, the relevant role play vectors normal to the boundary of a reachable set. These vectors will be defined without introducing any inner product in the state space. It should be noted, that vector normal to the reachable set boundary will be an element of the state space X and not of the dual space X* as often stated in the classical maximum principle. In the classical approach in R^n or Hilbert space it is convenient to identify the normal vector with a functional belonging to X^*. In our case such interpretation is not correct, unless X is a Hilbert space.

II.1.1. Definition

By the reachable set of a control system (f, C) (of an orientor field N) with the initial condition (t_0, s) we mean the set $\{(t,v):(t,v)\varepsilon[t_0,T] \times X,$ $v = x(t)$, $x(t_0) = s$, x is a trajectory of (f,C)(of N respectively)$\}$. The reachable set will be denoted as $Z((f,C), (t_0, s))$ or $Z(N, (t_0, s))$. If $t_0=0$ the t_0 will be omitted an we will write $Z((f,C), s)$ or $Z(N, s)$. If $s = x_0$ is fixed, this notation will be abbreviated to $Z(f, C)$ or $Z(N)$. Replacing the word "trajectory" by "quasitrajectory" we define the reachable sets for quasitrajectories, being denoted as $Z_q((t, C), (t_0, s))$ or $Z_q(N,(t_0, s))$.

II.1.2. Definition

Let K_σ be the hyperplane $t = \sigma$. The set $Z^\sigma(f,C) = Z(f,C) \cap K_\sigma$ is called a time-section of a reachable set $Z(f,C)$. We have, respectively,

$Z_q^\sigma(f,C) = Z_q(f,C) \cap K_\sigma$, $Z_q^\sigma(N) = Z_q(N) \cap K_\sigma$.

II.1.3. Remark

In the theory of contingent equations the reachable set is called the emission zone.

II.1.4. Theorem

Let (f,C) be a normal control system and let I.2.8 and I.3.1 hold. Then the set $\overline{Z^\sigma(f,C)}$ is connected for each $\sigma \in J$.

Proof

Let us suppose, on the contrary, that $\overline{Z^\sigma(f,C)}$ is not connected for certain $\sigma \in J$, $\sigma > 0$. From theorem I.3.2 follows that $Z^\sigma(f,C)$ is nonempty. Then, sets A and B and a constant $\varepsilon > 0$ should exist, such that $Z^\sigma(f,C) = A \cup B$ and

$$\inf \{d(a,b) : a \in A,\ b \in B\} > \varepsilon \qquad \text{(II-1)}$$

Let us define a function $\phi : X \to R$ such that $\phi(x) = d(x,A) - d(x,B)$. This function is continuous , $\phi(x) < -\varepsilon$ on A and $\phi(x) > \varepsilon$ on B. Hence, we have

$$|\phi(x(t))| > \varepsilon \; \forall \; t \in J \qquad \text{(II-2)}$$

where x is a trajectory of (f,C)

The sets A and B are included in the reachable set, and, consequently trajectories x_1 and x_2 exist, such that $x_1(\sigma)\varepsilon$ A, $x_2(\sigma)\varepsilon$ B. Let u_1 and u_2 denote such controls that $\dot{x}_i = f(t, x_i(t), u_i(t))$ a.e. on J, i= 1, 2. Let us define the multifunction

$$M(t,x) = \{f(t, x, u_1(t)), f(t, x, u_2(t))\} .$$

We have $M(t,x) \subset N(t,x)$, where $N(t, x)$ is the orientor field associated to (f,C). The convex hull of $M(t,x)$ is the set

$$P(t,x) = \text{conv } M(t,x) = \{X \ni v : v = f(t,x,u_1(t)) \cdot \lambda +$$
$$+ f(t,x, u_2(t)) \cdot (1-\lambda), \lambda \varepsilon [0, 1] \} .$$

The field P is associated with the system (g, [0,1]), which trajectories fulfill the equation

$$\dot{x}(t) = g(t,x(t),\lambda) = f(t,x(t),u_1(t)) \cdot \lambda + f(t,x(t),u_2(t)) \cdot (1-\lambda) \qquad \text{(II-3)}$$

with the initial condition $x(0) = x_0$. The solution $x(t,\lambda)$ of (II-3) exists and depends continuously on the parameter λ (Maurin [20], part I). Let us observe that

$$\phi(x(\sigma,\lambda)) = \phi(x_1(\sigma)) < \varepsilon \quad \text{for} \quad \lambda = 1 \quad \text{and}$$
$$\phi(x(\sigma,\lambda)) = \phi(x_2(\sigma)) > \varepsilon \quad \text{for} \quad \lambda = 0 .$$

From continuity of $x(t,\lambda)$ follows that a constant $\tilde{\lambda} \varepsilon [0,1]$ exists, such that $\phi(x(\sigma,\tilde{\lambda})) = 0$. The function $x(t,\tilde{\lambda})$ is a trajectory of the field P. Observe, that $P = \text{conv } M$ and $M(t, x)$ consists of two points, moving continuously. Thus, the system (g, [0, 1] is regular and theorem I.4.9 may be applied, which means that $x(t,\tilde{\lambda})$ is a quasitrajectory of N and of (f,C). Moreover, from I.4.11 follows that $x(t,\tilde{\lambda})$ is a limit of sequence $\{x_n\}$ of trajectories of N and of (f, C). Thus, for any $\eta > 0$ a trajectory $x_k \varepsilon \{x_n\}$ exists, such that $\phi\ (x_k(\sigma)) < \eta$, which contradicts (II-2) and completes the proof.

II.1.5. Remark

A similar property of reachable set was pointed out by Zaremba [9] in the case $X = R^n$. In that paper also was pointed out that the reachable set is upper semi-continuous as a function of the initial condition. The similar theorems are given by Chow and Shunr [12] and by Castaing and Valadier [14]. With somewhat different assumptions we are going to prove continuity of reachable sets, following the analogycal theorem of Turowicz [11] (for the case $X = R^n$).

II.1.6. Theorem

Let (f, C) be a regular control system and let I.3.1 hold. Then the set $Z_q^t((f, C), x)$ is continuous in the Hausdorff sense as a function of the initial condition x.

Proof

From I.3.2 follows that $Z_q((f, C), x)$ is nonempty. Let us choose a trajectory $\bar{x}$ of (f,C) with an initial condition x_0, and corresponding control $\bar{u}$. Let $z = \bar{x} - x$, $r(t) = \| z(t) \|$, where x is the trajectory with the control $\bar{u}$ and with an initial condition x_1. From I.3.1 follows that

$$\dot{r}(t) \leq \| \dot{z}(t) \| \leq \| f(t, \bar{x}(t), \bar{u}(t)) - f(t, x(t), \bar{u}(t) \| \leq$$
$$\leq \omega(t, \| \bar{x}(t) - x(t) \| = \omega(t, r(t)) \quad \text{a.e. on } J.$$

From the above differential inequality follows that

$$r(t) \leq \phi(t, \| x_0 - x_1 \| , \qquad \text{(II-4)}$$

where $\phi(t, v_0)$ is the upper integral of the equation $\dot{v}(t) = \omega(t, v(t))$ with the initial condition $v_0 = v(0) = \| x_0 - x_1 \|$. From the properties of ω follows that $\phi(t, v_0) \to 0$ if $v_0 \to 0$. Observe that (II-4) is valid for all pairs of trajectories with identical controls and diverse initial conditions. Thus, to each point of $Z^t((f, C), x_0)$ corresponds a point of $Z^t((f, C), x_1)$ and vice versa, the two points belonging to trajectories with identical control. Let us observe, that (II-4) holds uniformly, i.e. do not depend on u and do not depend on the corresponding pair of points of $Z^t((f,C), x_0)$ and $Z^t((f,C), x_1)$ Thus, $r(t) \to 0$ uniformly if $x_1 \to x_0$. Consequently we have

$$d^* \, (\overline{Z^t((f,C),x_0)}, \overline{Z^t((f,C),x_1)}) \to 0 \quad \text{if } x_1 \to x_0 \, .$$

From I.4.11 follows that $Z_q^t(f,C) = \overline{Z^t(f, C)}$. This means, that $Z_q^t((f,C),x)$ is continuous in the Hansdorff sense with respect to x, which completes the proof.

II.1.7. Remark

Let us observe that $\overline{Z^t((f, C), x)} = Z_q^t((f, C), x)$ and, consequently, $Z^t((f, C), x)$ is upper semi-continuous with respect to x.

II.1.8. Theorem

Let (f, C) be a regular control system and let I.2.8 and I.3.1 hold. Let (t_1, x_1) be a point belonging to the boundary of $Z_q(f, C)$, where $t_1 \varepsilon (0,T)$. Then

(i) at least one quasitrajectory of (f, C) passes through the point (t_1, x_1),

(ii) the graph of each quasitrajectory of (f, C) which passes through (t_1, x_1) belongs to $\partial(Z_q(f, C))$ for all $t \varepsilon [0, t_1]$ ($\partial(\cdot)$ denotes the boundary of a set).

Proof

(i) This assertion follows from theorem I.5.1.

(ii) Let us suppose, on the contrary, that a quasitrajectory $\bar{x}$ of (f, C) exists, such that

$$\bar{x}(\tau) \not\in \partial(Z_q(f, C)) \text{ for some } \tau \, \varepsilon \, (0, t_1) \text{ and } \bar{x}(t_1) = x_1 \quad \text{(II-5)}$$

From I.4.9 follows, that the quasitrajectories of (f, C) are at the same time, trajectories of the field $E = \operatorname{conv} N$, when N is the field associated to (f, C). Let us observe, that from I.2.8 follows that $Z(E) \subset \operatorname{int} R \times \Omega$. This means that at any point $(t, c) \varepsilon Z(E)$ a neighbourhood $\gamma(t, x)$ exists, such that $f(t, x, u)$ fulfills the assumptions of this theorem in $\gamma(t,x)$ for all $u \varepsilon \Theta$. Let us consider a field

$$B(s,x) = \{X \ni v: - v \varepsilon E(t_1 - s, x)\} .$$

Observe, that to each trajectory $x : [0, t_1] \to X$ of E corresponds a trajectory $y : [0, t_1] \to X$ of B such, that $s = t_1 - t$ and $y(s) = x(t)$. From (II-5) follows that a trajectory $\bar{y}$ of B exists, such that $\bar{y}(0) = x_1 \varepsilon \partial Z(E)$ and $\bar{y}(t_1 - \tau) = \bar{x}(\tau) \varepsilon \operatorname{int} Z(E)$. Trajectories are continuous functions. Consequently, we can choose τ in such a way that $x(t) \varepsilon \gamma(t_1, x_1)$ on $t \varepsilon [t_1 - \tau, t_1]$ for each trajectory x passing through (t_1, x_1). From the definition of B follows that

$B(s, x) = \operatorname{conv} \{X \ni v: v = -f(t_1 - s, x, u), u \varepsilon C(t_1 - s\}$. Thus, $B(s,x)$ is the convex field associated with the control system (f', C'), where $f'(s,x,u) = -f(t_1 - s, x, u)$, $C'(s) = C(t_1 - s)$. The system (f',C') is regular and II.1.7 hold. This means that the reachable set $Z^{t_1 - \tau}(B, y)$ depends continuously on its initial condition $y = y(0)$. Consequently, a neighbourhood η of x_1 exists, such that $Z^{t_1 - \tau}(B,y) \cap \operatorname{int} Z(E, x_0) \neq \phi$ for each $y \varepsilon \eta$, where we identify the points (t,x) and (s,y) of the time-state spaces of the systems (f,C) and (f',C') respectively, such that $x = y$ and $s = t_1 - t$. Choosing an initial condition $y(0) \varepsilon \eta$ such that $y(0) \notin Z(E, x_0)$ we conclude that a trajectory y' of B exists which passes through the point $s=0$, $y(0) \varepsilon (\eta \setminus Z^{t_1}(E, x_0))$ and through the point $s = t_1 - \tau$, $y(t_1 - \tau) \varepsilon \operatorname{int} (E)$. These two points belong to the graph of some trajectory x' of E, i.e. to the graph of some quasitrajectory of (f, C). Thus, x' is a quasitrajectory with the initial condition x_0, which reaches the point $x'(t_1) = y(0) \notin Z_q(f,C)$. This contradicts the definition of reachable set and completes the proof.

II.1.9. Remark

Fig 3 indicates an interpretation of the above proof in the case $X = R^2$.

II. 2. Optimal control

In this section we assume that X is a real reflexive and separable Banach space and that X^* (the dual of X) is uniformly convex.

II.2.1. Definition

Let $J' = [t', t''] \subset J$. A trajectory (quasitrajectory) $x : J' \to X$ of a control system (f, C) passing through a point (t', s) is said to be optimal over J' if the graph of x belongs to the boundary of the reachable set $Z((f,C), (t', s))$ (of $Z_q((f,C), (t',s))$ respectively) over the interval J'. Replacing (f,C) by an orientor field N we define an optimal trajectory of an orientor field.

II.2.2. Remark

In the above definition we do not declare whether the trajectory or quasitrajectory is time-optimal one, or optimal in the sense of some particular criterion of optimality. It is clear that II.2.1 concerns time-optimal (quasi-) trajectories if some goal trajectory is defined. On the other hand, many problems of optimal control can be converted into time-optimal problems, by appropriate extension of the system state space (see I.5.3, I.5.6 and I.5.7).

II.2.3. Definition

Let Y be a normed vector space, let $A \subset Y$ be a nonempty set and let $K(y, \rho)$ be an open ball with radius ρ and center y. A vector $v \in Y$ is said to be normal to A in the point $z \in \partial A$ if $v \neq \phi_Y$ and $K(z + v, \| v \|) \cap A = \phi$. The set of all normals to A at a point $z \in \partial A$ will be denoted as $\tilde{N}(A, z)$.

II.2.4. Definition

Let X^* be the dual of X and let $F: X \to 2^{X^*}$ be the duality mapping, defined as follows: $F(x) = \{x^* \in X^*: x^*(x) = \| x \|^2 = \| x^* \|^2\}$. The functionals $<\cdot,\cdot>_+$, $<\cdot,\cdot>_-$: $X \times X \to R$ defined as follows

$$< x, y >_+ = \sup \{z(x): z \in F(y)\} ,$$

$$< x, y >_- = \inf \{z(x): z \in F(y)\} ,$$

are called semi-inner products in X.

II.2.5. Remark

The semi-inner products have the following properties

$$< x + y, z >_\pm \leq < x, z >_\pm + < y, z >_\pm ,$$

$$| < x, y >_\pm | \leq \| x \| \cdot \| y \| ,$$

$$< x + ay, y >_\pm = < x, y >_\pm + a \| y \|^2 \quad \forall a \in R,$$

$$< ax, by >_\pm = a \cdot b < x, y >_\pm \ \forall a, b \in R, a \cdot b \geq 0.$$

If X is a Hilbert space, then $<\cdot,\cdot>_+ = <\cdot,\cdot>_- = <\cdot,\cdot>$, where $<\cdot,\cdot>$ is the usual inner product.

If X^* (the dual space of X) is uniformly convex then $<\cdot,\cdot>_+ = <\cdot,\cdot> \stackrel{df}{=} <\cdot,\cdot>_s$. The semi-inner product $<\cdot,\cdot>_s$ exists and has the following properties [31]

$$<\alpha x + \beta y, z>_s = \alpha <x, z>_s + \beta <y, z>_s \ \forall \alpha, \beta \varepsilon R; \ x, y, z \varepsilon X,$$

$$|<x, y>_s| \leq \| x \| \cdot \| y \| \ \forall \ x, y \varepsilon X, \ <x, x>_s = \| x \|^2 .$$

Let us consider, for example, the space $\ell_p(R)$ of all sequences of real numbers with the usual norm $\| \cdot \|_p$. Let q be the conjugate exponent of p, i.e. $p^{-1} + q^{-1} = 1$. Then the semi-inner product of $x, y \varepsilon \ell_p(R)$ is defined by

$$<x, y>_s = c(y) \ \Sigma \ x_i \ |y_i|^{p-2} y_i ,$$

where $c(y) = \| y \|_p^{q^{-1}(q-p)}$. It is clear that the above semi-inner product becomes the usual inner-product of ℓ_2 if $p = 2$ and it is easy to verify that it has the properties indicated above.

II.2.6. Theorem

Let E be a convex and continuous orientor field and let Y be the Banach space $Y = R \times X$ equipped with the norm $\| t, x \|_Y = \left(| t |^2 + \| x \|^2\right)^{1/2}$. Suppose that a trajectory x of E with the initial condition $(0, x_0)$ passes through a point $(t', x') \varepsilon \partial Z(E)$, where $t' \varepsilon$ int J, Z(E) is the reachable set of E (see II.1.1). Let $\tilde{N} = \tilde{N}(Z(E), (t', x'))$ be nonempty, and let $v \varepsilon \tilde{N}$. We denote by $\dot{x}^-(t)$ and $\dot{x}^+(t)$ the left- and right-hand derivatives of x at a point t. Then

(i) $<(1, \dot{x}^+(t')), v>_s \leq 0$,

(ii) $\langle(1, \dot{x}^{-}(t')), v\rangle_s \geq 0$,

provided $\dot{x}^{+}$ and $\dot{x}^{-}$ exist at t' .

Proof

The norm $\|\cdot\|_Y$ will be denoted shortly as $\|\cdot\|$. This will not cause ambiguity if we will bear in mind that v and (t, x) belong to $Y = R \times X$. (i) Let us consider the function

$$\phi(s) = \| p - (t' + s, x(t' + s)) \| ,$$

where $p \varepsilon Y$ is such that $p-(t', x(t'))= v$. We have

$$\phi(s) = \| p - (t', x(t')) - s(1, \dot{x}^{+}(t')) \| + o(s)=$$

$$= \| v + s(-1, -\dot{x}^{+}(t')) \| + o(s), \tag{II-1}$$

if $s > 0$, where $o(s)/s \to 0$ if $s \to 0$.

Let us observe that $\phi(0) = \| v \|$ and

$$D^{+}\phi(0) = \lim_{s \to 0_+} \frac{\| v + s(-1, -\dot{x}^{+}(t')) \| - \| v \|}{s} , \tag{II-2}$$

where D^{+} denotes the right-hand derivative. We apply now the equality

$$\langle a, b\rangle_s = \| b \| \lim_{s \to 0_+} \frac{\| b + s \cdot a \| - \| b \|}{s} \tag{II-3}$$

(Deimling [18], proposition 5.2).

From (II-1), (II-2) and (II-3) we obtain

$$\phi(0)\, D^{+}\phi(0) = \langle(-1, -\dot{x}^{+}(t')), v\rangle_s \tag{II-4}$$

Let us note that $D^+\phi(0) \geq 0$. Indeed, if $D^+\phi(0)$ were negative, a neighbounrhood ρ of $s = 0$ would exist, such that $\|p - (t, x(t))\| < \|v\|$ in ρ, i.e. the point $(t, x(t)) \varepsilon Z(E)$ would belong to $K(p, \|v\|)$ which contradicts the assumption that v is normal to $Z(E)$. Thus we have

$$<(-1, -\dot{x}^+(t')), v> \geq 0$$

and, due to the definition of the semi-inner product

$$<(1, \dot{x}^+(t')), v> \leq 0, \qquad \text{(II-5)}$$

which completes the proof of the part (i).

(ii) To prove this part, we have to consider the function $\Psi(s) = \| p - (t' - s, x(t' - s))\|$ instead of $\phi(s)$. Applying an argument similar as in (i) we can point out that $<(1, \dot{x}^-(t')), v>_s \geq 0$, which completes the proof.

II.2.7. Theorem

Let (f, C) be a regular control system (see I.4.8) with associated orientor field N, fulfilling assumptions I.2.8 and I.3.1. Let E and Q denote a convex field and tendor field generated by N, respectively. Let $x: J \to X$ be an optimal quasitrajectory of (f, C) with an initial condition x_0. We denote

$$A = \{t : t \varepsilon \operatorname{int} J, \tilde{N}(Z(E), (t, x(t))) \neq \phi\}.$$

Suppose that $A \neq \phi$ and $\mu(A) > 0$. Then

$$\dot{x}(t) = \sigma(t, x(t)) \text{ a.e. on } J, \qquad \text{(II-6)}$$

where σ is a selector of E, such that $\sigma(\cdot, x(\cdot))$ is measurable and

(i) $<(1, \sigma(t, x(t)), v>_s = 0$ a.e. on A, and, moreover

(ii) $\max \{c : c = <(1, z), v>_s, z \in E(t, x(t))\} = 0$ (II-7)

for all $v \in \tilde{N}(Z(E), (t, x(t)))$, $t \in A \setminus I$, where I is the null set over which $\dot{x}(t)$ does not exist.

Proof. It is known that, under the above assumptions, $Z_q(f, C) = Z(E)$ and, consequently, x is a trajectory of E. Since x is differentiable almost everywhere on J, we have $\dot{x}^+(t) = \dot{x}^-(t)$ a.e. on J. Thus, the equality (i) follows from (i) and (ii) of the Theorem II.2.6.

The prove (ii) let us suppose that a point $z \in E(t', x(t'))$ exists such that $t' \in A \setminus I$ and $<(I, z), v>_s > 0$. Consequently, a convex neighborhood η of z exists, such that $<(I, w), v>_s > 0$ for all $w \in E'(t, x) = E(t, x) \cap \eta$. The multifunction E' is continuous and convex-valed. Hence, a continuous selector s(t, x) of E' exists. Let y be the trajectory of E' which corresponds to s, with the initial condition (t', x(t')). Thus, $\dot{y}^+(t')$ exists and $<(I, \dot{y}^+(t')), v>_s > 0$. Thus, $\dot{y}^+(t')$ exists and $<(1, \dot{y}^+(t')), v>_s > 0$. But once (t', x(t')) belongs to the reachable set of E with the initial condition $(t_0, x(t_0))$, the trajectory y can not leave the reachable set. Using the same argument as in the proof of the point (i) of Theorem II.2.6. we conclude that $<1, \dot{y}^+(t'), v>_s \leq 0$ which is a contradiction to the previous inequality and completes the proof.

II.2.8 Remark

Let us observe that Theorem II is formulated in terms of quasitrajectories of (f, C), without any convexity and compactness assumptions imposed on control domain and/or counterdomain.

The similar geometrical interpretations are given by Markus and Lee [23], Leitmann [21], Boltiansky [22] and others. It should be noted that the optimal quasitrajectory of (f, C) need not be regular in the sense defined in [21], i.e. the plane tangent to the reachable set along with the quasi-trajectory needs not to exist. We require only that the set of normals $\tilde{N}$ be nonempty. However, it seems to be difficult to prove that it is the case, i.e. that $\mu(A) \neq 0$. To overcome this difficulty, let us consider the following property which might help us in this connection.

II.2.9 Property (P)

A constant η exists such that for each $(t, x) \in J \times X$ the reachable set $Z_q^s((f, C), (t, x))$ is convex for all $s \in (t, t+\eta)$.

Of course, not every control system has this property . One might expect that if the orientor set N(t, x) is convex then the property holds. In this case the following corollary may be useful.

II.2.10 Corollary

Let (f, C) be a regular control system having property (P) and let the assuptions of Theorem II.2.7 hold. Let us denote

$$H(t) = \bigcup [\tilde{N}(Z(E, (s, x(s))), (t, x(t))): s \in (o, t)]$$

$$B = \{t\colon t \in \text{int } J,\ H(t) \neq \emptyset\}\ .$$

Then $\mu(B) = \mu(J)$ and Theorem II.2.7 holds while replacing in it the set A by B and $\tilde{N}(Z(E), (t, x(t)))$ by $H(t)$.

Proof. Let $x\colon J \longrightarrow X$ be an optimal quasitrajectory of (f, C) with an initial condition $(0, x_0)$. Obviously $x'\colon [s, T] \longrightarrow X$ being the restriction of x to $[s, T]$, is an optimal quasitrajectory of (f, C) with the initial condition $(s, x(s))$, i.e. it lies on the boundary of $Z(E, (s, x(s)))$, where $s \in \text{int } J$. By the property (P) the set $\tilde{N}(Z(E, (s, x(s))), (t, x(t)))$ is nonempty for each $s \in (t - \eta, t)$, which means that $\mu(B) = \mu(J)$. It is clear that the Theorem II.2.7 holds for any initial condition $s \in \text{int } J$ and any interval $[s, T']$, $s < T' < T$. This observation completes the proof.

II.2.11 Adjoint inclusion and possible applications

In order to construct any useful algorithm it is necessary to calculate the "adjoint trajectory" $v\colon J \longrightarrow X$. One might expect that the variable v (the vector nomral to the boundary of the reachable set) satisfies certain differential equation or differential inclusion. We will not consider here this problem in details. Let us, however, observe that the part (ii) of Theorem II.2.7 provides the "control rule" which can be used to choose the value of $\sigma(t, x)$. On the other hand, the part (i) of that theorem leads to the equality

$$\langle \bar{\sigma}(t, x(t)), v(t) \rangle_s = 0 \qquad \text{(II-8)}$$

where $\bar{\sigma} = (1, \sigma)$ and σ can be determined due to the maximum principle (II-7). The equality (II-8) may provide additional conditions for v. Let us consider, for example, the case when X is a Hilbert space and σ does not depend explicitely on t. Thus, form (II-8) we get

$$\frac{d}{dt} < \bar{\sigma}(x(t)), v(t) = < \bar{\sigma}_X (\dot{x}(t)), v(t)> + <\bar{\sigma}(x(t)), \dot{v}(t)> = 0,$$

where $<\cdot , \cdot>$ stands for the inner product of $Y = R \times X$.

Let us observe that $\bar{\sigma}_X$ is the Frechet's derivative of $\bar{\sigma} : Y \to Y$, i.e. $\bar{\sigma}_X$ is a continuous linear operator. Taking into account that $\dot{x}(t)=\sigma(x(t))$ we get

$$< \bar{\sigma}(x(t)), \dot{v}(t)> \;\; = - < \bar{\sigma}(x(t)), \bar{\sigma}_X^*(v(t)) > ,$$

where $\bar{\sigma}_X^*$ is the adjoint operator of $\bar{\sigma}_X$. Consequently, the desired differential inclusion for v becomes

$$\dot{v}(t) \varepsilon V(t, v(t), x(t)) \overset{df}{=\!=} \{X \ni z : <\bar{\sigma}(x(t)), z> = -<\sigma(x(t)), \bar{\sigma}_X^*(v(t))>\} \qquad \text{(II-9)}$$

Obviously, an adjoint vector which satisfies the equation

$$\dot{v}(t) = - \bar{\sigma}_X^*(v(t))$$

is a solution to (9), as in the classical maximum principle.

In practical applications the most common case is that of regular control systems. It results from the physical nature of control variable. Even if the system is described by partial differential equations with distributed (function-valued) control, its technical realization can hardly be done with infinite number of control variables. Once the control is of finite dimensionality, the corresponding control counterdomain N(the orientor set) is compact by the theorem of Mazur. In that case it is easy to show that the control system under consideration is regular one. Another question is the computational aspect of Theorem II.2.1. It might appear that this version of the maximum principle does not provide any useful algorithms. Indeed, in the general case it is difficult to determine the normal vector η and the selector σ of the "relaxed" control counterdomain E. But, let us observe that the general case refers to the situation when the optimal control need not exist. Consequently, the control variable does not appear in the conditions (i) and (ii) of the Theorem II.2.1. This property seems to be interesting one. Namely, the philosophy of optimization algorithm is not to look for the optimal control, but rather for the optimal quasitrajectory. No uniquess and/or existence of optimal control is required. The main task is to find an optimal quasitrajectory and then to determine corresponding control, corresponding sliding regime, or to approximate it by certain sequence of trajectories. Of course, if the control counterdomain is convex, then the set E is equal to N and can be parametrized by the control variable. In that case the search is to be made in the control space directly.

'tq.

Some important applications of the maximum principle concern systems governed by partial differential equations.

Let us give now some general remarks. We consider control system whose state variable is a function of time t and of some spacial variable x. In technical applications we have $x \varepsilon R^3$, but here we let $x \varepsilon R^1$, to simplify the notation. The system is governed by a nonlinear partial differential equation, and the problem is posed as some nonlinear version of the Neuman problem

$$\frac{\partial z(t, x)}{\partial t} = A(t, z)(x) \tag{II-10}$$

$$z(0, x) = \phi(x) \tag{II-11}$$

$$z_x \Big|_{\ell-} = u(t) \ , \ z_x \Big|_{-\ell+} = -u(t), \tag{II-12}$$

where $z_x \Big|_{\ell\pm} = \lim \partial z(t, x) / \partial x$ with $x \to \ell$, $x \gtrless \ell$. The function z(t, x) is the state variable defined over the set $S = R_+ \times L$, $L = [-\ell, \ell]$, u is the control $u(t) \varepsilon C \subset R$, C compact, and A(t, z) is some differential operator. Moreover, we require that

$$z_{v+} - z_{v-} = B(z(t, v)), \tag{II-13}$$

where $z_{v\pm}$ are the righ- and left-hand derivatives at x=v and x=v is a solution to some algebraic equation $x=E(z(t,x))$. Conditions like (II-13) often appear in models of processes of mass and/or heat transfer with some chemical reactions, solidication e.t.c. Let us confine our attension on the case

$A(t, z)(x) = a(t, z(t, x))\dfrac{\partial^2 z(t, x)}{\partial x^2}$ for $x \in (\text{int } L)\setminus\{v\}$, $a(t, z)$ real-valued.

In order to describe this system by one equation of evolution let us extend $z(t, x)$ to the entire half-plane $R_+ \times R$, put $z(0, x) = 0$ for all $x \notin L$ and let z be treated as a generalized function, z_x being the distributional derivative. Observe that

$$z_{xx} = z'_{xx} + \delta(x+\ell)z_x|_{-\ell+} - \delta(x-\ell)z_x|_{\ell-} + \delta(x-v)\ (z_{v+} - z_{v-})$$

where $\delta(\cdot)$ is the distribution of Dirac, z'_{xx} is the usual derivative $\partial^2 z/\partial x^2$ in $(R_+ \times R)\setminus\{-\ell, v, \ell\}$ and $z'_{xx}(t, x) = \lim z_{xx}(t, y)$ with $y \to x$, $y < x$, for $x \in \{-\ell, v, \ell\}$. We suppose that z is continuous in $R_+ \times R$, $\phi(x)$ is of class C^2 and $\phi(-\ell) = \phi(\ell) = 0$.

Let us define

$$A'(t,z,u) = (z_{xx} + \delta(x+\ell)u(t) + \delta(x-\ell)u(t) - \delta(x-v)B(t,v))\, a(t,z) \qquad \text{(II-14)}$$

Observe that $A = A'$ in $(\text{int } L)\setminus\{v\}$ and A' is an extension of A to the set $R_+ \times R$. Thus (II-10, 11, 12, 13) can be expressed by one equation

$$\frac{\partial z(t, x)}{\partial t} = A'(t, z, u)(x), \quad z(0, x) = \phi'(x) \qquad \text{(II-15)}$$

which includes all boundary conditions. Here $\phi'(x) = \phi(x)$ in L, $\phi(-\ell) = \phi(\ell) = 0$, $\phi(x) \equiv 0$ in $R\setminus L$. Let us introduce the following notation. We denote by ψ: $R_+ \to L^2(R, R)$ the function of time, whose value (for some fixed time) is the function $z(t, \cdot) \varepsilon L^2(R, R)$. We have

$$z: R_+ \times L \to R\ , \quad \psi: R_+ \to L^2(R, R),$$

$$\psi(t) = z(t,\cdot) \varepsilon L^2(R, R), \quad \psi(t)(x) = z(t, x) \varepsilon R$$

Thus (II-15) becomes an ordinary differential equation in the space $L^2(R, R)$

$$\frac{d\psi(t)}{dt} = A'(t, \psi(t), u(t)) \tag{II-16}$$

where the operator A' is uniquely defined by (II-14). Unfortunately, A' is not continuous in L^2, and we have to know if (and with what assumptions) a solution to (II-16) exists. We shall not consider here this problem. The right way to solve optimization problems for systems described by linear evolution equations was given by Lions [17] and it seems possible to solve nonlinear problems similarly. Let us give some another approach, using some "smoothed" operator, in certain analogy to weak solutions of (II-16). The concept of smoothing is based on some physical observation. Namely, the main idea is that such physical parameters as temperature and concentration can not be measured in a point and they always are some global properties of a finite set of particles. Consequently, we replace the right-hand side of (II-15) by the expected value of some function of a random variable y with density $g(x - y)$, i.e.

$$\frac{\partial z(t, x)}{\partial t} = \int_{-\infty}^{\infty} A'(t,z,u)(y)\, g(x-y)\, dy \tag{II-17}$$

This, in $L^2(R, R)$ becomes

$$\frac{d\psi(t)}{dt} = F(t, u, \psi(t)), \tag{II-18}$$

where the operator F is defined by the right-hand side of (II-17) and (II-14). As $g(\cdot)$ we can take a density of normal distribution with zero mean and sufficiently small variance. Provided $E(z)$ and $B(z)$ are regular enough, the operator F is continuous and Lipshitz. Since the control domain $C \subset R$ is compact, the control counterdomain of the system (F, C) is also compact. It implies that (F,C) is a regular control system and the results of previous sections can be applied. Observe that if the variance of g tends to zero, the right-hand side of (II-17) tends to that of (II-15) but, it should be emphasized that in our approach $g(\cdot)$ <u>is fixed</u> and we do not need that (II-17) be satisfied with diverse densities $g(\cdot)$.

III. Some remarks on computational aspects

III.1. Local algoritm for infinite-dimensional system

III.1.1. General observations

To pose the problem we have to determine what means "infinite dimensionality" in a computer program. It might appear that, for technical reasons, computer programs realized in practical applications must be finite-dimensional. On the other hand, there are algorithms solving partial differential equations and performing quite complicated operations in spaces of functions (e.g. speech signal processing, prediction e.t.c.) with considerable efficiency. Taking into account recent rapid development of supercomputers with paralled processing it seems to be possible to develop standart software tools for infinite dimensional optimization problems. As for the "infinite dimensionality" in computer programs, let us observe that such properties of state and control space as noncompactness and lack of measure affect program performance causing heavy difficulties which grow rapidly with the space dimensionality. For example, the known difficulties in generating realizations of random variables having multidimensional distributions are obviously caused by the lack of probability measure in the infinite-dimensional case. The noncompactness of some sets of fuctions causes difficulties of another kind. It is impossible, for example, to construct a finite ε - grid in a set of functions which are not equicontinuous. Consequently, it is rather difficult to create catalogs of pattern functions used in data-base of signal recognition systems. On the other hand algorithms utilizing such calculations as operations over vectors or evaluation of inner-product, i.e. the permissible operations of some infinite-dimensional vector spaces, are quite effective even if the dimensionality of a problem is high. It is why the linear programming algorithms do work satisfactority even for problems of many thousands of unknowns.

Any computer realization of an infinite-dimensional problem must include some kind of truncation or projection into a finite-dimensional space or spaces. The main requirement which differs the standart finite and infinite-dimensional program is that the truncated dimensionality or the space of projection be not visible by the user, i.e. that the dimensionality does not appear in formal parameters of procedures called directly by the user. Let us give some remarks on an approximation scheme which may be used for this purpose.

III.1.2. Galerkin approximations

Let X_n denotes a finite-dimensional subspace of a real reflexive Banach space X, and let P_n be a projection from X to X_n. Let $\{P_n\}$ be a sequence of projections, such that $\| P_n \| = 1$, $P_n = X \to X_n$ for $n = 1, 2, \ldots,$ and $P_n x \to x \ \forall\, x \varepsilon X$. The sequence $\{X_n, P_n\}$ is called the projection scheme. It allows us to construct a sequence of solutions to ordinary differential equations in spaces X_n, converging to a solution of the corresponding equation in X. The following is the basic theorem on the convergence of such solutions.

III. 1.3. Theorem

Let X be a Banach space with a projection scheme $\{X_n, P_n\}$, such that $\| P_n \| = 1$ for $n = i, 2, \ldots,$ and $P_n x \to x \;\forall\, x \in X$. Let $f: J \times \Omega \to X$ be a continuous and bounded function such that $\| f(t, x) \| < M$ in $J \times \Omega$, M being a positive constant, and

$$< f(t, x) - f(t, y) >_- \leq \omega\,(t, \|x - y\|) \cdot \|x - y\| \qquad \text{(III-1)}$$

for all $x, y \in \Omega$, $t \in J$, where $< \cdot, \cdot >_-$ is a semi-inner product in X and $\omega : J \times R_+ \to R$ is a function which satisfies the following condition: for any $\gamma > 0$ there exists $\delta > 0$, $\{t_i\}$ and $\{\rho_i : J \to R_+\}$ such that $t_i > 0$, $t_i \to 0$, ρ_i are continuous and

$$\rho_i(t_i) \geq \delta t_i, \; D^-{}_i(t) \geq \omega(t, \rho_i(t)), \;\; 0 < \rho_i(t) \leq \gamma \quad \text{(III-2)}$$

Then the equation

$$\dot{x}(t) = f(t, x(t)), \;\; x(0) = x_0 \,, \;\; x \in X \qquad \text{(III-3)}$$

has the unique solution x over the interval J, the equation

$$x(t) = P_n f(t, x(t)), \;\; x(0) = P_n x_0, \; x \in X \qquad \text{(III-4)}$$

has the unique solution x_n over $[0, b]$, where $b = \min\{\mu(J), r/M\}$ (r denotes the diameter of Ω) for n big enough, and $x_n(t) \to x(t)$ uniformly on $[0, b]$.

III.1.4. Remark

The proof of the above theorem can be fund in [18]. Some interesting application concerns a distributed system described, for example, by equation

(II-17). The sketch of a possible approach is as follows.

Let the state space be $X = L_2$ and let $g(\cdot)$ be some truncated density function, being, however, sufficiently regular. Thus, we can restrict the bounds of integration to some finite interval K, and we have $X = L_2(K, R)$. Let S be a countable dense subset of K, and let $S_n = \{s_1, s_2, \ldots, s_n\}$ be a finite subset of points of S. Let us define a sequence $\{S_n\}$ of subsets of S in such a way that

$$\max_k \min_j |s_k - s_j| \to 0 \text{ with } n \to \infty \text{ and } k, j \leq n .$$

For example, if $K = [0, 1]$ we can choose $S_1 = \{1/2\}$, $S_2 = \{1/4, 2/4, 3/4\}$, $S_3 = \{1/8, 2/8, \ldots, 6/8, 7/8\}$ e.t.c. with the following enumeration of points: $s_1 = 1/2$, $s_2 = 1/4$, $s_3 = 3/4$, $s_4 = 1/8$, $s_5 = 3/8$, e.t.c. By X_n we denote the space of all step functions being constant in each of the intervals constituting the set $K \setminus S_n$. We have $X_1 \subset X_2, \ldots, \subset X$. The projection P_n: $X_n \to X$ is defined as $P_n z = z_n$, where z_n is a step function such that $z_n \in X_n$, $z_n|_{S_n} = z|_{S_n}$, and $z(x) = z(s_i)$ for all $x \in [s_i, s_k]$, s_i and s_k being points of S_n with consecutive positions (not numbers) and $s_i < s_k$. Let now x be a point of K in which $\partial^2 z/\partial x^2$ exists. We can express the derivative $\partial z/\partial x$ in term of points belong ing to S_n, as follows

$$\frac{\partial z(x)}{\partial x} = \frac{z(s_k) - z(x)}{s_k - x} + \eta_n(x)$$

where s_k is a point of S_n such that $s_k > x$ and $|s_k - x| = \min_i\{|s_i - x|, s_i \varepsilon S_n\}$ and $\eta_n(x)/|s_k - x| \to 0$ with $n \to \infty$. Similarly

$$\frac{\partial^2 z(x)}{\partial x} = q(s_k, s_p, x) + \eta_n(x), \quad s_k, s_p \varepsilon S_n ,$$

where $q(s_k, s_p, x)$ is an appropiate difference quotient, $s_k > s_p > x$, s_p is the nearer right-hand neighbour of x and s_k is that of s_p. Thus, with the above notation we can express the projection of the second derivative as

$$P_n \frac{\partial^2 z(x)}{\partial x^2} = q(s_k, s_p, s_r) + \eta_n(s_r), \tag{III-5}$$

where $s_k > s_p > x > s_r$, and s_r, s_p and s_k are points of S_n with consecutive values.

We are now in a position to apply the Galerkin's projection sheme to the system described by the equation (II-17). To do this, let us observe that

$$\int_{-\infty}^{\infty} \delta(y + \ell)\, u(t)\, g(x - y)\, dy = u(t)\, g(x + \ell)$$

and the distributional terms in (II-17) disappear, being replaced by quite regular functions of the form u(t) g(x). Substituting the corresponding projections we obtain a sequence of finite-dimensional equations in X_1, X_2, ... which is, in fact, a sequence of Galerkin approximations.

It is perhaps appropiate to mention that, from the computational point of view, the above scheme isnothing more thensome finite-difference algorithm with decreasing step of the space-grid. Let us observe, however, that the enumeration of points in the set S provides that the particular, say the k-th, equation corresponds to the physical parameter in the same point in all iterations (in all the spaces X_1, X_2, ...). Thus, the solutions reached in the iterations can be easily compared to each other in order to check the rate of convergence and some stop-criterion.

III.1.5. The algorithm

Let us recall that the main results of the maximum principle II.2.7 are

$$< (1, \sigma(t, x(t)), v >_S = 0 \text{ a.e. on } A \qquad \text{(III-1)}$$

and

$$\max \{ c : c = \langle (1, z), v \rangle_S : z \varepsilon E(t, x(t)) \} = 0 \qquad \text{(III-2)}$$

where σ is a measurable selector of E , E is the convex orientor field generated by the control system under consideration, v is a vector normal to the boundary of the reachable set at the point (t, x(t)), x: J → X is an optimal trajectory and A denotes the set over which v exists. With this notation and under the hypotheses of II.2.7, III-1 and III-2 are necessary conditions for optimality. Considering projections into a plane t = const we get, due to II.2.11 and II.7

$$H = < \sigma(t, x(t)), \eta >_S = 0 \quad \text{a.e. on } A, \qquad \text{(III-3)}$$

where η is the projection of v. Consequently, any algorithm based on these results consists in some searching process which allows us to determine $\sigma(t, x)$ which maximizes H.

What we obtain as a result, is a trajectory which graph belongs to the boundary of the reachable set. Let S denote some given set of the space Y, such that at least one trajectory of E reaches the set S, and let a point $y' \varepsilon S \cap Z(E)$ exists such that $P_J y' = \min \{P_J y : y \varepsilon S \cap Z(E)\}$, where $P_J y$ is the projection of y into J (recall that $Y = J \times X$, thus $P_J y$ is real-valued). In this case an optimal trajectory exists which reaches the set S in optimal time. Suppose that the set μ of normals to the set $\overline{X \setminus S}$ at y' and the set of normals to Z(E) at y' are nonempty. It is easy to point out that in this case a normal v to Z(E) at y' must exist such that $v \in \mu$. This is the condition of transversality, well known in the finite-dimensional case. Choosing diverse sets S and, eventually, extending the state space we can define diverse optimization problems which provide, or not, corresponding transversality conditions.

The computer algorithm in the finite-dimensional case with convex compact orientor N(t, x) is well known. The standard approach is to apply some of conjugated gradients or variable metric algorithm, where the gradient $\partial H/\partial u$ is evaluated as follows. We integrate, with given initial point x_0 and given control function $\bar{u}$, the state equation $\dot{x} = f(x, \bar{u}, t)$, storing the trajectory $\bar{x}$ in memory. Then, we integrate the equations of the conjugated vector η (see II-9) with the final condition p(T), backward in time. Once the values of $\bar{x}$ and $\bar{u}$ are known, it is possible. Moreover, if we evaluate p(t) for some $t \in J$, we can evaluate H in this point, and, simultaneously, $\partial H/\partial u$ at time t. This latter value is to be stored as a value of gradient (search direction) in the space of control functions (note that p(t) need not be stored). Thus, one integration of the state equations and one of the conjugate equations give us the search direction and the search process can be continued [28].

A similar procedure may be applied in the infinite-dimensional case. The main difference is that the integration processes for the state and conjugated variables have no dimensionality parameter specified. The integration process for each trajectory consists of a number of integrations of finite-dimensional

Galerkin approximations which result in a (truncated) solution to the infinite system. Thus, the user do not control the dimensionality specifying, perhaps, only some upper bound for this parameter. It should be noted that, at any rate, the memory needed and time consumption might be huge and the known difficulties, arising while integrating infinite systems, can not be overcome. Application of computers which high level of parallelism might cause some progress in this field.

III.2. Parallel algorthm

Most of the published applications of paralled computers concern the problems of linear algebra, linear programming and parallel evaluation of arithmetic expressions [24], [25]. If the number of papers devoted to applications of that kind is fairly great, the published results on parallelism in non-linear optimization problems and optimal control are rather small/the scarce examples are i.g. [26], [27]. It seems that the technological development of parallel computers has been so rapid that the theoretical research in the field of parallelism in optimal control was left behind. It should be noted that both hard ware and software tools for such research are available. The main software tool is the Ada language which undoubtedly will stimulate the research.

Our task is to formulate some parallel optimization process and to assess its speed-up in relation to certain classical sequentional algorithm. It shouldbe noted that, if the convergence of an optimization algorithm may, in most cases, be proved, its speed can hardly be established without strong assumptions imposed on the object function. As we do not intend to do so, this section is, from the theoretical point of view, merely a problem statement. Namely, we try to investigate the speed of convergence of the proposed procedure, simulating the process or parallel computations.

III.2.1. The statement of the problem

Let us recall the basic optimization algorithm of gradient type. Let $F: R^m \to R$ be a continuously differentiable function called the object function and grad

$F(z)$ be the gradient of F evaluated at a point $z \varepsilon R^m$. The problem is to find $z_o \varepsilon R^m$ such that $F(z_o) \leq F(z)$ in some neighbourhood of z_o. The gradient algorithm is as follows

Algorithm G.

Step 0. Choose $z_o \varepsilon R^m$ such that the set $\{z: F(z) \leq F(z_o)\}$ is bounded.

Step 1. Set $i = 0$.

Step 2. Evaluate $h_i = D(z_i)$ grad $F(z_i)$.

Step 3. If $h_i=0$ then stop run. Else go to step 4.

Step 4. Evaluate the smallest non-negative number λ_i such that $F(z_i + \lambda_i h_i) = \min_{\lambda} \{F(z_i + \lambda h_i\}$:

Step 5. Set $z_{i+1} = z_i + \lambda_i h_i$, set $i = i+1$ and go to step 2.

In the above algorithm $D(z)$ is a positively defined matrix which elements are continuous functions of z. If $B = I$/the unit matrix/ then we have the steepest descent method. If $D \neq I$ and it depends on gradients evaluated in previous steps, then G is a conjugated gradient algorithm. A comprehensive discussion of such optimization algorithms can be found in [28].

The Ada language parallel facilities make it possible to use parallel computations at the algorithmic level. That is what we call parallelism in this paper. From this point of view algorithm G is sequential one, despite of the fact that the vector and matrix operations of steps 2, 4 and 5 might be executed with concurrent computations.

The most constly operations in the above algorithm are evaluation of grad F and evaluation of the object function. Each "big" iteration of the algorithm requires one evaluation of grad F and several evaluations of F in the step 4. Both gradient and object function evaluation will be treated as tasks in the sequel, i.e. as program units which can be exacuted concurrently. The main idea is to launch a number of such tasks and not to separate (in real time) object function and gradient evaluation. The whole process of optimization will be said to be of type Gy/z where y is the number of tasks (both of gradient and object function type) which might be executed in parallel and z is the number of tasks of object function type, called OFT in the sequel. Thus, the maximal number of the gradient-type tasks /GT/ being executed concurrently with OFTs is equal to y-z. The sequential algorithm G is of type G1/1 where no parallelism exists.

Before formulating the search rules for a, possible realization of a Gy/z process, let us specify a set of data contained in the common pool visible from within all the tasks.

The common data are as follows.

G, U - arrays of the same type as the argument of the object function,

X - the lower value of F achieved in the search,

NF - the number of active OFTs,

NG - the number of active GTs,

Ln - logical variables, n = 1, 2, 3 .

To initiate the process of optimization we set L1 = true, NF = 0, NG = 0, $X = \infty$, L2 = false, L3 = false. Now let us specify the search process actions being some realization of a process Gy/z. We denote by ϕ the array with all elements are zeros. By "if Ln" we mean "if Ln is true". The process actions are as follows.

-i- If $L1$ and $NG < y - NF$ then new GT starts, $L1 := \text{false}$ and $NG := NG + 1$.

-ii- If a GT ends then G is given new value, evaluated by this GT, $NG := NG - 1$ and $L2 := \text{true}$.

If $G = \phi$ then stop run.

-iii- If $L2$ and $NF < z$ then new OFT starts. It evaluates the value of $F(W)$, where $W(I) := U(I) + Y * G(I)$ for all components of W. Y is a value obtained by random drawing with normal distribution, mean H and standard deviation V. The distribution is truncated so that $Y > 0$.

-iv- If an OFT ends and $F(W) \geq X$ then $NF := NF - 1$.

If an OFT ends and $F(W) < X$ then $NF := NF - 1$, $U := W$, $L1 := \text{true}$.

Moreover, in both cases parameters H and V are changed due to some updating rule.

-v- If L3 then stop run. L3 is a stop condition, not specified here in details.

In the above process GT has higher priority then OFT, i.e. the event /i/ has a higher priority then that of event/iii/. For GTs and OFTs are asynchronous, the shared arrays G and U might be updated at some irregular time instants depending on whether a GT or an OFT ends. Therefore G and U are transfered to GTs and OFTs by value, i.e. GT takes a copy of U and OFT takes a copy of G and U at the time instant it is activated.

Let us observe that in the case of process G 1/1 with $V = 0$, $D = I$ and $H = \lambda_i$ we get the algorithm G. Thus, the process Gy/z is some extension of G to the parallel searching process. Moreover, under the hypotheses introduced in this section, G 1/1 converges to the solution of the optimization problem, due to

[28] /section 2.1/ . Process G1/1 with the above parameters has, however, little practical meaning, for the parameter λ_i is not known. The evaluation of λ_i requires several evaluations of the object function. To do this in the sequential algorithm we have to apply some sequential algorithm of one-dimensional minimization. This could be replaced by some parallel search, for example such as formulated in the point /iii/. Then we get a process of type Gy/z. It should be noted, that the rules /i/+/v/ concern some particular realization of a serach process of type Gy/z.

One might expect that the speed-up of Gy/z is as high as the degree of parallelism y. It, however, is not the case. If the "quality criterion" of the seraching process is equal to $t \cdot m$ /where m is the amount of central memory needed and t denotes the CPU time/, then the sequential algorithm is better than the process Gy/z. It results from the fact that the information which have been gathered in the consecutive steps of the algorithm G is properly utilized at the next step of G. In the parallel search a new OFT is launched before completing all already active OFTs. Consequently, while generating its parameters, the process could not make use of the information obtained from the OFTs active at this instant of time. As a result, the speed-up is not proportional to y.

To improve the speed of the search, we could apply the approach similar to the "branch and bound" method known in the operational research. Namely, let us assume that the result of an OFT can be assessed before completing this OFT. If this assessment shows that no improvement of the object function can be attained in this OFT, then it may be terminated before its completion and a new OFT may be launched immediately.

III.2.2. Application to optimal control in R^n

Let us consider a time-discrete control system which trajectory is described by equation

$$x_k = f(x_{k-1}, u_k) \tag{III-4}$$

where $x_k = (x_{0,k}, x_{1,k}, \ldots, x_{n,k})$ is the system state vector, $u_k = (u_{1,k}, \ldots, u_{p,k})$ is the control vector, $f : R^n \times R^p \to R^n$ is a continuous vector-valued function, k stands for the instant of time, $k = 1, 2, \ldots, T$ and $u_k \in C_k$, where $C_k \in R^p$ is a set called control domain. The initial condition x_0 is given. The optimal control problem is to determine a control u_k, $k = 1, 2, \ldots, T$ such that the component $x_{0,T}$ reseaches its minimum. The function $f_{0,k}$ /being the first component of f_k/ is of the form $f_{0,k} = f_{0,k}(x_{1,k}, \ldots, x_{nk}, u_{1,k}, \ldots, u_{p,k})$. The other components of f_k also do not depend of $x_{0,k}$. The $x_{0,T}$ is the object function. To evaluate it, we must integrate the whole system trajectory, i.e. to solve equation (III-4) for $k = 1, 2, \ldots, T$. Thus, in this case the task OFT contains a step-by step procedure which integrates /summarizes/ equation (III-4). Let us observe that the array U of the process Gy/z has now $p \times T$ elements. The gradient of the object function may be evaluated due to the time-discrete version of the Pontriagin's maximum principle [28]. Namely, let $s_k = (s_{k,0}, s_{k,1}, \ldots, s_{k,n})$ be a vector /conjugated variable/ fulfilling the equation

$$s_{k-1} = \sum_{j=1}^{n} s_{j,k} \frac{\partial f_{j,k}(x_{k-1}, u_k)}{x_{j,k-1}} \tag{III-5}$$

for $k = 1, 2, \ldots, T$. The gradient of the object function $x_{0,T}$ is given by the formula

$$\frac{\partial x_{0,T}}{\partial u_{j,k}} = \sum_{i=1}^{n} s_{i,k} \frac{\partial f_{i,k}(x_{k-1}, u_k)}{u_{j,k}} \qquad \text{(III-6)}$$

To make use of this formula, we integrate (III-4) with a given initial condition x_0, storing the vectors x_k in memory. Next, we set $s_{0,T} = -1$, $s_{1,T} = s_{2,T} = \ldots = s_{n,T} = 0$ and integrate (III-5) in the reverse order of time. At each stage of integration one column of the gradient matrix $\partial x_{0,T}/u_{j,k}$ is evaluated, due to (III-6) The elements of this matrix could be enumerated to form one-dimensional array grad F in order to follow the notation of the previous section. Thus, we can realize the tasks OFT/integrating (III-4)/ and GT/integrating (III-4), (III-5) and evaluating (III-6). The other rules of search are as indicated in Gy/z.

Let us consider now the case when the object function in non-decreasing. This is often the case because the function $f_{o,k}$ is usually non-negative, e.g. quadratic one. In such case the assesment mentioned in the previous section is very simple. If the value of $x_{o,L}$ is greater than X (the current "best" value of the object function achieved in the search process) for some L < T, then $x_{o,T}$ can not be less then X and the integration of the trajectory in this OFT may be terminated inmediately.

The interesting problems which arise are

1. is the process Gy/z convergent to a/locally/ optimal solution?
2. what speed-up might be expected in comparision with G1/1?

As stated in the introduction, we shall not deal with very theoretical considerations here. Let us only observe, that the convergnece of G1/1 follows from the convergence of G. This enables us to suppose that Gy/z with $y > 1$ will also be convergent. As for the speed-up, some simulation experiments was carried out to investigate it.

III.2.3. Simulation experiments

The real-time parallel search process was simulated instead of realizing it in the real time. The main reason to simulate it was that the simulation model makes it possible to impose the /model/ time consumption charges on OFTs and GTs and to separate these charges form the time consumption of the other operations of the process. Moreover, this enables us to evaluate easily all statistics and tracing needed as a result.

The processes G and Gy/z were coded in the Simula 67 language. To evaluate λ_i in the algorithm G the typical golden-section algorithm was used, as described in [28]. It should be noted that the experiments concerned the simplest possible versions of both processes. Namely, the search directions were opposite gradient and not conjugated gradient ones. Of course, conjugated directions could lead to better performances of both processes. However, it seems that parallelism in Gy/z could be utilized in evaluating search directions in much more sophisticated way than in known sequential cojugated gradients algorithms. Consequently, parallel methods with conjugated search directions need some wide and profound treatment and are not considered in this paper.

The simulation experiments was carried out for some diverse control systems.

In all cases the process Gy/z was convergent and considerable speed-up was observed in comparison with the algorithm G. Let us show some typical example. The system equations are

$$\Delta x_{1,k} = x_{2,k}$$

$$\Delta x_{2,k} = u_k x_{1,k} - 0.5 x_{1,k} x_{2,k}^2$$

$$\Delta x_{0,k} = (x_{1,k} - c_1)^2 + (x_{2,k} - c_2)^2 ,$$

where

$$\Delta x_{i,k} = x_{i,k+1} - x_{i,k} , \quad u_k \varepsilon [-1, 1], \; x_{1,0}=1, \; x_{2,0}=-1,$$

$x_{0,0}= 0$, $c_1= 0, 5$, $c_2= 0, 5$. The problem is to determine u_k for k=0, 1, 2,..,T such that $x_{0,T+1}$ reaches its minimum at the end of the corresponding system trajectory. We set T = 19.

The processes were simulated with the /model/ CPU time consumption imposed as follows. One trajectory integration process was assumed to consume 20 time units and teh whole trajectory of conjugated variable plus the evaluation of the gradient was assumed to consume 40 CPU time units. The results are shown in fig. 5. The shared line indicates the value of $x_{0,20}$ reached by the algorithm G as a function of CPU /model/ time. The flat parts of the line correspond to the golden-section improvements of λ_i and to the conjugated variable trajectory and gradient evaluations. The solid lines indicate the values of $x_{0,20}$ reached by the process Gm+1/m for m = 2, 4 and 6. This case is a typical one and illustrates the possible speed up which may be achieved by the parallel search. The parallel computation of tasks is indicated in fig. 6 for the process G5/4.

The level G corresponds to the GT task, and the levels T1, T2, T3 and T4 to the four OFT tasks. The boldface sections of the lines correspond to active tasks and the numbers denote consecutive OFTs. Let us observe that the OFT tasks start simultaneously, but their further execution becomes asynchronous, because some OFTs /with "bad" controls/ are terminated before reaching the end of corresponding trajectories.

III.3. Determination of reachable sets and global optimization

III.3.1. Some remarks on global optimization

There are few cases when one can analytically determine the shape of the system reachable set. Usually some searching procedure must be applied. In general, the problem of qualification of a point as a member or non-member of a reachable set is equivalent, from the computational point of view, to some optimization problem . Thus, the problem of determination the shape of the reachable set is at least as difficult as the problem of global optimization. There are few results in this field which might be applied in practice. If the local properties of system trajectory are usually assumed to be known and quite regular, the global properties, correponding to great changes of controls, can hardly be predicted before the whole reachable set is known. Consequently, the local search techniques can not be adapted to the global optimization and to reachable set determination.

Let us consider some possible approach based on random search inside the reachable set. It is known that algorithms of stochastic type provide good results while used as a first step of optimization, both in local and global problems [30]. We shall not discuss here the question of superiority of stochastic or deterministc algorithms. Let us merely observe that any algorithm of computation is only one of the steps of the process of solving an optimization problem. The complete solving process might be characterized by the following chain of operations.

1. Given an optimization problem, the decision must be taken about the method

of solving; a particular algorithm is to be chosen.

2. The parameters of algorithm must be establish (e.g. the initial step of the search, the initial approximation, the stop criterion e.t.c.).
3. The corresponding computer program is to be run.
4. The results of computation are to be interpreted and accepted. Eventually the whole process, or its part, is to be repeated.

Observe that if the step 3 consists in execution of some, perhaps deterministic, computer program, the steps 1, 2 and 4 are realized by a man. These steps are not charged with any incertainty only if the shape of object function is known in advance. But in this case the optimization problem is obviously out of sense. If the shape of object function and the position of its extremum in the space of parameters are unknown, the decisions taken in steps 1, 2 and 4 are always stochastical, depending on subjective point of view and experience of particular person. Due to the above remarks, the division of optimization methods into deterministic or stochastic seems to be irrelevant. The practical experience shows that the stochastical algorithms do work satisfactory in global problems. Unfortunately, the lack of probabilistic measure make them usefulless in the general infinite-dimensional case. Let us describe, in the next section, some stochastic method which can provide a valuable information about the shape of a finite-dimensional projections of the system reachable set.

III.3.2. Density forming algorithm

This algorithm consists in some random search inside the system reachable set in the state space. It can be used to determine the reachable set as well as to assess the position of global minimum of the object function. Let us abstract, for a moment, from the system dynamics and recall the description of the density forming (DF) algorithm as given in [29]. Let (W, w) and (S, s) denote metric spaces with probabilistic measures w and s, and let $\phi : W \to S$ be a continuous mapping satisfying the Lipshitz condition. Let us denote $D = \phi(W) = \{S \ni v: \phi(z) = v, z \in W\}$. We assume also that the space W is bounded, compact and connected, and that the shape of the reachable set D can not be derived analytically, using some (known) formula for evaluation the value of ϕ. The problem is to determine, with some accuracy, the shape of D and the position of the minimal point u*, such that $\phi(u^*) = \min_u \{\phi(u) : u \in W\}$.

The simplest method of solving the above problem is the pure random search in the space W. In that search the probability density p in W is assummed to be constant in time and usually uniform in W. In the pure randon search a sequence of points $\{u_i \in W\}$, $i = 1, 2, \ldots$ is generated due to the density p and for each u_i its image $x_i = \phi(u_i)$ is evaluated. Having generated some sufficiently great number of points x_i we can assess the shape of the reachable set D and estimate the position of u*. Unfortunately, if the dimensionality of the space S is great, such search is very inefficient. Moreover, even if we can restrict our attension to some subspace of S with low dimensionality, the nonlinearity of ϕ may make the distribution of $\{x_i\}$ very nonuniform, so that some regions of D would be filled with some sufficiently dense subset of $\{x_i\}$ while some others would remain undiscovered at all. The aim of the DF algorithm is to achieve the uniform density of points x_i in D (or in some projection of D), by appropiate forming of the density p in the control space W.

Let us note that, under the above assumptions, the set D is compact. Hence, we can introduce a finite ε-grid in D. A spacial element of the grid will be treated as belonging to the reachable set D if the intersection of this element with D is nonempty. The union of all such elements will be denoted by D_1. It is clear that in order to determine the set D_1 by a random drawing we ought to look for some algorithm which provides the uniform density of points in D_1. In this case the probability that a point x_i does not fall in a particular spacial element of D_1 is equal to $(1 - 1/J)^L$, where J denotes the number of the spacial elements of D_1 and L is the number of points x_i generated by the algorithm. We neglect the behavior of the distribution of points x_i in some ε-neighbourhood of the boundary of D, letting the set D_1 be determined with some ε-accuracy. Obviously if the density q of points x_i in D_1 is not uniform, it exists at least one spacial element of D_1, for which the probability of non-hitting is greater that $(1 - 1/J)^L$, provided ε is small enough.

Let us divide the process of drawing into K stages, m of the points x be generated at each stage with the density $f(\cdot, b_i)$, where $i = 1, 2, \ldots, K$ and $b_i = (b_{1i}, \ldots, b_{pi})$ is some vector of parameters. Hence, the density of the set of all points generated by the algorithm in K stages is

$$p(\eta) = \frac{1}{N} \sum_{i=1}^{K} f(\eta, b_i)\, m, \quad \eta \varepsilon W$$

where $N = K \cdot m$. Let us suppose that a density p^* in W exists, for which the density q^* of the random variable $x = \phi(u)$ (where u has the density p^*) is such that the probability

$$P_j = \int_{E_j} q^*(x)\, ds,$$

where E_j is the j-th spatial element of D_1, is constant (does not depend on j). We suppose also, that the density p* may be expressed as

$$p^*(\eta) = \lim_{N\to\infty} \sum_{i=1}^{K} f(\eta, b_i)m.$$

In practice, the density f ought to be simple enough to be used while generating the points u in the searching process. On the other hand, it should ensure satisfactory approximation of p^*. In the proposed algorithm it was fixed as

$$f(\eta, b) = \begin{cases} c \cdot \prod_{j=1}^{\eta} g(b_j, d, \eta_j) & \text{for } \eta \varepsilon W \\ 0 \text{ for } & \eta \notin W \end{cases} \qquad \text{(III-7)}$$

where $\eta = (\eta_1, \ldots, \eta_\eta)$, $b = (b_1, \ldots, b_p)$, $g(b_j, d, \eta_j)$ is the density of one-dimensional normal distribution with main b_j and variance d, and the constant c serves for normalization. Here we suppose that W is a subset of a finite dimensional vector space. The DF algorithm consists of the following operations.

1. Substitute for p the density uniform in W.
2. Perform the operation 3 m times, go to 4.
3. Draw a point u as a realization of the random variable with the density p, calculate $x = \phi(u)$.
4. Determine a point $a \varepsilon S$, being some assessment of the point a^*, such that the density q (of points x in D) reaches its minimum in a^*, i.e. $q(a^*) = \min \{q(v): v\varepsilon D\}$, determine a point $b\varepsilon W$ such that $\phi(b) = a^*$.

5. If the density q is uniform in D, then stop.
6. Substitute for p the density f, due to(III-7), go to 2.

The operation 4 of the above algorithm is of particular importance. In this operation we have to assess the position of the point of minimal density q in D. Clearly we can not determine the exact position of a^*, because the only information we have is the finite set of points x calculated in the operation 3. Moreover, the set D is not known and we do not know where to look for the minimum.

III.3.3. Convergence of the DF algorithm

Let (u, x) denotes a pair of points determined in the operation 3 of the DF algorithm. We denote by $P_{j,k}$ the probability of hitting by x the j-th element of the grid belonging to D_1, in the iterations 1, 2, ..., k. By the iteration we mean one performance of the operations 2 ÷ 6. Let us denote by $R_{j,k}$ the probability of hitting by the point a the j-th grid element in the step k, and let us denote

$$S_{j,k} = \frac{1}{k} \sum_{i=1}^{k} P_{j,i} \tag{III-8}$$

$$T_{j,k} = \frac{1}{k} \sum_{i=1}^{k} R_{j,i} \tag{III-9}$$

By $E_{j,k}$ and $M_{j,k}$ we denote the expected values of $S_{j,k}$ and $T_{j,k}$ respectively, taken over the set of all possible realizations (runs) of the first k iterations. Let us consider the difference $\sigma_{j,k} = P_{j,k} - E_{j,k}$. For $P_{j,k}$ is bounded, the variance $V(\sigma_{j,k})$ is also bounded. It can be seen that $\sigma_{j,k} \xrightarrow[p]{} 0$, where $\xrightarrow[p]{}$ denotes the convergence in probability. To prove this, let us consider two cases: (i) the variance $V(\sigma_{j,k})$ does not tends to zero and (ii) $V(\sigma_{j,k}) \to 0$ with $k \to \infty$. (i) In this case we have

$$\sum_{i=1}^{k} V(\sigma_{j,i}) \to \infty \qquad \text{with } k \to \infty .$$

applying the theorem of Lindberg (Feller, [32]) we conclude that the distribution of the random variable

$$S = \frac{\sum_{i=1}^{k} \sigma_{j,i}}{\sum_{i=1}^{k} V(\sigma_{j,i})}$$

tends to the normal distribution with zero mean and unit variance, if $k \to \infty$. Consequently,

$$\frac{1}{k} \sum_{i=1}^{k} \sigma_{j,i} \xrightarrow[p]{} 0.$$

This imply that

$$\frac{1}{k} \sum_{i=1}^{k} E_{j,i} \xrightarrow[p]{} S_{j,k} \qquad \text{(III-10)}$$

and $\sigma_{j,i} \xrightarrow[p]{} 0$ if $k \to \infty$.

(ii) Observe that $\sigma_{j,k}$ has expected value zero. Thus, since $V(\sigma_{j,k}) \to 0$ we have $\sigma_{j,k} \xrightarrow[p]{} 0$.

By the similar way we can point out that

$$\frac{1}{k} \sum_{i=1}^{k} M_{j,i} \xrightarrow[p]{} T_{j,k} \tag{III-11}$$

Let us introduce now the following assumption about the estimate of the minimal density point.

Assumption (*)

The point a, being some estimate of the position of the minimal density point a* (see the operation 4 of the algorithm), belongs to the interior of spatial element of the ε-grid in S. Moreover, we suppose that a positive constant A exists, such that

$$M_{j,k} \geq A\left(\frac{1}{J} - E_{j,k-1}\right) + \frac{1}{J} \tag{III-12}$$

for all $j = 1, 2, \ldots, J$, $k = 1, 2, \ldots$, where J is the number of spatial elements of the ε-grid in D_1 .

The above assumption is introduced in order to make the estimate of a* "good enough". Due to (III-12) the expected value of the probability of hitting by a the j-th element in the first k iterations is greater then 1/J for the elements with low $E_{j,k-1}$ (i.e. with low density of points x). Let L denotes the number of the grid element containing the point a in the k-1-th iteration. Taking into account that a belongs to the interior of the grid element it is easy to point out that for any $0 < \gamma < 1$ a $d > 0$ exists such that the probability of hitting by x the L-th element, determined by (III-7), in the operation 3 of the k-th iteration is greater than γ. Consequently, we have

$$P_{L,k} \geq P_{L,k-1} \frac{k-1}{k} + \frac{\gamma}{k} . \tag{III-13}$$

For each $s \neq L$ we have

$$P_{s,k} \geq P_{s,k-1} \frac{k-1}{k} \tag{III-14}$$

Differently speaking, $P_{j,k}$ satisfies (III-13) with the probability $R_{j,k-1}$ and (III-14) with the probability $1 - R_{j,k-1}$. Let us introduce the random variable $\alpha_{j,k}$ as follows.

$$\alpha_{j,k} = \begin{cases} 1 & \text{with probability } R_{j,k} \\ 0 & \text{with probability } 1 - R_{j,k} \end{cases}$$

From (III-13) and (III-14) follows that

$$P_{j,k} \geq P_{j,k-1} \frac{k-1}{k} + \alpha_{j,k-1} \frac{\gamma}{k} \tag{III-15}$$

Let us define a sequence $d_{j,1}$, $d_{j,2}$, ... as follows

$$d_{j,1} = P_{j,1}$$

$$d_{j,k} = d_{j,k-1} + \alpha_{j,k-1} \frac{\gamma}{k} \quad \text{for } k = 2, 3, \ldots \tag{III-16}$$

For any realization of the above sequence we have $P_{j,k} \geq d_{j,k}$. Due to (III-16) we have

$$d_{j,k} = d_{j,1} \frac{1}{k} + \frac{\gamma}{k} \sum_{i=1}^{k-1} \alpha_{j,k} \tag{III-17}$$

Applying once more the Lindberg theorem we conclude that

$$\frac{1}{k} \sum_{i=1}^{k} \alpha_{j,i} \xrightarrow[p]{} \frac{1}{k} \sum_{i=1}^{k} R_{j,i} \tag{III-18}$$

with $k \to \infty$. Thus, by virtue of (III-17)

$$d_{j,k} \xrightarrow[p]{} \gamma T_{j,k} \tag{III-19}$$

for $j = 1, 2, \ldots, J$ and $k \to \infty$.

Let us suppose that

(**) a number k_0 exists such that $S_{j,k} \geq \frac{\gamma}{J}$ for all $k > k_0$.

Then, from (III-8) we obtain

$$\frac{1}{k} \sum_{i=1}^{k} P_{j,i} \geq \frac{\gamma}{J} \tag{III-20}$$

Let us consider now the case opposite to (**), namely suppose that

(***) for each $k_0 > 0$ a number $k_1 > k_0$ exists such that $S_{j,k_1} < \frac{\gamma}{J}$.

Consequently we have

$$\frac{1}{k_1} \sum_{i=1}^{k_1} E_{j,i} < \frac{\gamma}{J} .$$

Applying (III-11), (III-12) and the above inequality we get

$$P \left[T_{j,k_1} > \frac{A(1-\gamma) + 1}{J} \right] \xrightarrow[p]{} 1$$

with $k_0 \to \infty$. Thus, form (III-19) follows that

$$P \left[d_{j,k_1} > \gamma \frac{A(1-\gamma) + 1}{J} \right] \xrightarrow[p]{} 1.$$

Taking into account that $d_{j,k}$ is not greater then $P_{j,k}$ we obtain

$$P \left[P_{j,k_1} > \gamma \frac{A(1-\gamma) + 1}{J} \right] \xrightarrow[p]{} 1 \qquad \text{(III-21)}$$

with $k_0 \to \infty$

From (III-20) and (III-21) we conclude that for each realization of the sequence $P_{j,1}$, $P_{j,2}$, $P_{j,3}$ .... a subsequence $\{P_{j,k_n}\}$ exists such that

$$P \left[P_{j,k_n} \geq \frac{\gamma}{J} \right] \xrightarrow[p]{} 1 \qquad \text{(III-22)}$$

with $n \to \infty$, for all $j = 1, 2, \ldots, J$.

The convergence (III-22) means that P_{j,k_n} might be sufficiently close to

1/J (it depends on γ) in some probabilistic sense. Recall that $P_{j,k}$ is the probability of hitting by x the j-th element in all steps $i = 1, 2, .., k$. If $P_{j,k}$ is close to 1/J, the density of points x in D_1 is close to uniform one (with "ε-accuracy"). This property, however, holds with some probability which tends to 1 if n tends to infinity, where the probability $P[\cdot]$ in (III-22) means some probabilistic measure in the space of all possible realizations of the iterations 1, 2, ..., k_n. This means that a realization of the DF algorithm which does not converge (to the uniform density q in D) is not impossible. We merely conclude that the probability of such realization is small.

III. 3.4. Application to control systems

There are some essentially different methods of determination the reachable sets of control systems. It is known that the problem of classifying a point as a member or no-member of the system reachable set is equivalent to some optimization problem . The other possible approach is to apply the maximum principle in order to calculate trajectories belonging to the boundary of the reachable set and, having integrated sufficiently large number of such trajectories, estimate its shape. Such approach, hawever, occurs to be almost impracticable. The main difficulty which appears in its practical application is that we have no information about the possible range of the initial conditions for adjoint variables, which are needed to integrate trajectories belonging to the boundary. As we have to integrate many trajectories, we can not attempt to solve any two-point boundary value problem for each trajectory. Consequently, we need to have complete initial conditions for each one. The dependence of control on the adjoint vector is sharply nonlinear even in very simple cases, and for the wide changes of those initial conditions we may observe no dependence at all. This causes considerable difficulties in any procedure of searching the reachable set boundary.

Let us consider some other method based on the random search of the inside of the reachable set, using the DF algorithm

Let U_p denotes the set of all measurable selectors of a multifunction C (control deomain) defined over the interval $J = [0, T]$. Apparently, using the terminology of control systems, the space W of the DF algorithm corresponds to the control set U_p (being the set of all permissible control functions) and the reachable set D to the set Z(f, C). We only say "corresponds" because in the general case, no probabilistic measures exist in U_p and/or Z(f, C). In order to apply the DF algorithm we must restrict its action to some measurable subsets or projections of U_p and Z(f, C). Let us consider a machine-man interface at the output and suppose that the user is interested to learn what is the shape of the system reachable set. Thus, the computer output must consist of some readable print-out in numerical or graphical form. It is clear that it can not be neither any infinite nor very huge amount of output data. This suggests that the interface implies considerable reduction of the dimensionality of the visible reprezentation of the reachable set. Namely, what is obtained as the result, is some finite-dimensional projection of the real reachable set. The finite-dimensional projection space may be equipped with appropriate measure and treated as the output space of the algorithm.

As for the input space W = U_p, it is always infinite-diemnsional, being some space of functions. To introduct a measure we must sharply restrict the space U_p. This might be done, for example, by considering only some family of input functions, parametrized by a finite set of parameters. To find such family we must be shure that this restriction does not influence significantly the shape of the reachable set. In the general case it is rather difficult problem. In the case of finite-dimensional regular system we can make use of the fact that the closures of the reachable sets of trajectories of the systems (f, C) and (f, D) are identical to each other, where D is the tendor kernel of C. Hence, the set U_p can be restricted to its subset U_q consisting of all permissible tendor controls. In many cases the functions of U_q are of "bang-

bang" type, which make it possible to parametrized U_q, provided some estimate of (finite) number of switchings is known. At any rate, the DF algorithm may be applied in this situations with diverse number of switchings and the sults analyzed in order to check the dependence of the shape of the resulting reachable set projections on the number of switchings.

Clearly the DF algorithm is rather coarse method and needs many trajectories to be calculated in each run. Note, however, its two considerable adventages. First, this is essentially paralled algorithm. The set of m trajectories calculated between consecutive changes of the density in control space (in subset U_q) can be completed in any order, or concurrently if the parallel computer is used. Thus, the parallel realization is at least m times quicker then the sequential one. Second, it does not require any adjoint trajectory equations. The only system specification needed is the standard description of the right-hand sides of system equations.

III. 3.5. Dynamical sensitivity analysis

Some useful application of the DF algorithm is determination of system sensitivity with respect to parameters which might change along with the trajectory. Our approach is quite different from the classical sensitivity problem. Namely, we are going to determine the influence of time-varying parameters, and the analysis is global, i.e. we permit great changes of parameters. Let the equation of system trajectory be as follows

$$\dot{x}(t) = f(x(t), u(t), p(t), t)$$

where p(t) is some vector of parameters. Treating p(t) as a stochastic process we obtain a stochastic control system. There exists huge literature in the field of such systems. The known results concern the transition of the distribution of the state vector along with a trajectory. In some cases, however, such analysis is hardly possible. Fistly, the numerical representation and processing of multidimensional distribution is difficult, particularly if the system is nonlinear. Secondly, to calculate the distribution of x for various time instants the exact description of stochastic properties of p is needed (provided u is given). Such data is not available in most of practical cases, and usualy some arbitrary assumptions are made.

Let us state some quite different problem, which might lead to certain assessment of the influence of changes of p. We suppose that u is some given control function (perhaps optimal one), and that we know the restrictions imposed on the changes of p. Let $p(t) \varepsilon W(t)$, where W is some given multifunction . Thus, to learn the influence of p we have to determine the reachable set of the system (f, W), the function p treated as control, u fixed.

III.3.6. Example

Let us consider a two-sector economical model indicated in fig 7 . The block P_1 represents production of investments goods V_1, P_2 is the sector producing consumer goods V_2 , C represents consumption, z stands for working force.

The variables u_1 and u_2 are controls determining the flows m_1, m_2, z_1 and z_2 (of investments and labor) which sypply the two sectors. D_1 and D_2 are some dynamical elements representing the accumulation and expiration of the invest_ ments. The total equipment (or capital) M_i installed in the sector i is assumed to be

$$M_i(t) = \int_0^\infty m_i(t - s)\, w_i(s)\, ds,$$

where i= 1, 2 and $w_i(s)$ is some weighting function representing the dynamical effectiveness of investiments. Let us suppose that the production v_i is described by the Kobb-Douglass function

$$v_i = A_i M_i^{S_i} z_i^{(1 - S_i)}$$

where S_i are some known constants, $S_i \varepsilon (0, 1)$, i = 1, 2. Let us consider the discrete-time version of the model. We have

$$M_{i,k+1} = h \sum_{j=1}^{L} m_{i,k-j+1}\, w_{i,j} \qquad \text{(III-23)}$$

where h denotes the time-step, the second index stands for the discrete time and i = 1, 2. We assume that the number L exists such that $w_{i,k}=0$ for all $k > L$. Let us suppose that the labor force is proportional to the consumption V_2, i.e. $z = cV_2$, $c > 0$. Taking into account the relations indicated above and those of fig 7 , we easily conclude that

$$V_1 = A_1\, M_1^{S_1}(u_2\, c V_2)^{(1 - S_1)} \qquad \text{(III-24)}$$

$$V_2 = \{A_2 \, M_2^{S_2} \, [(1 - u_2)c\,]^{(1 - S_2)}\}^{1/S_2} \, . \qquad \text{(III-25)}$$

In order to calculate $M_{i,k+1}$ due to III-23 we need the values of $m_{i,n}$ for $n = k,\ k-1, \ldots, k-L+1$. Let us denote

$$m_{1,k-1} = p_{1,k}\,,\ m_{1,k-2} = p_{2,k}, \ldots,\ m_{1,k-L+1} = p_{L-1,k}$$

$$m_{2,k-1} = r_{1,k},\ m_{2,k-2} = r_{2,k}, \ldots,\ m_{2,k-L+1} = r_{L-1,k}$$

Denoting $\Delta x_k = (x_{k+1} - x_k)/h$ we have

$$\Delta M_{1,k} = \sum_{j=1}^{L} w_{1,j} m_{1,k-j+1} - M_{1,k}/h = \sum_{j=2}^{L} p_{j-1,k} w_{1,j} + v_{1,k} u_{1,k} w_{1,1} - M_{1,k}/h.$$

Taking into account that $h\Delta p_{i,k} = p_{i,k-1} - p_{i,k}$ for $i = 2, \ldots, L-1$ and $h\,\Delta p_{1,k} = p_{1,k+1} - p_{1,k} = u_{1,k}\, v_{1,k} - p_{1,k}$ and omitting the index k we get

$$\Delta M_1 = \sum_{j=2}^{L} p_{j-1}\, w_{1,j} + v_1 u_1 w_{1,1} - M_1/h$$

$$\Delta M_2 = \sum_{j=2}^{L} r_{j-1}\, w_{2,j} + v_1(1 - u_1)\, w_{2,1} - M_2/h$$

$$\Delta p_1 = (u_1\, v_1 - p_1)/h$$

$$\Delta p_2 = (p_1 - p_2)/h \qquad \text{(III-26)}$$

$$\vdots$$

$$\Delta p_{L-1} = (p_{L-2} - p_{L-1})/h$$

$$\Delta r_1 = ((1 - u_1)\, v_1 - r_1)/h$$

$$\Delta r_2 = (r_1 - r_2)/h$$
$$\vdots$$
$$\Delta r_{L-1} = (r_{L-2} - r_{L-1})/h \quad .$$

The above system of 2L difference equations, together with III-24 and III-25 describes the system dynamics. We treat this system as some approxiamtion of a continuous system. Let us suppose that in the continuous case the weight function $w_i(s)$ is defined over some interval $s \in [0, t_1]$. Hence, its discretization $w_{i,k}$ is defined for $k = 1,\ldots,L$, where $(L+1)h > t_1 \geq Lh$. Since t_1 is fixed we see that if $h \to 0$ (if we take some sequence of discrete approximations) then the system dimensionality L tends to infinity. It is not any surprise. It is known that a (finite-or infinite-dimensional) system with time-delay can be treated as some system without delay in a Banach space. Thus, considering systems in Banach spaces we observe that no qualitative difference exists between systems with or without time-delay.

As one of possible applications of the DF algorithm, let us show some results of the dynamical sensitivity analysis applied to this model. It is interesting for example, what is the system sensitivity with respect to the values of S_1 and S_2. Fixing controls u_1, u_2 and other parameters and treating S_1 and S_2 as controls which can vary in some given intervals, we can determine the system reachable set. Fig 8 shows the projection of this set into the plane M_1,t. The results was obtained with $u_1 = u_2 = 1/2$, $S_1, S_2 \in [1/2(1-\varepsilon), 1/2(1+\varepsilon)]$, $L = 20$, the other parameters choosen as typical values appearing in ecomical literature. In fig 9 the time-section of the reachable set is indicated for the time-step $k = 20$.

The asterisks denote the end points of system trajectories belonging to the reachable set, the total of 500 trajectories integrated. It is interesting to compare this image with that of fig. 10 and 11, obtained by application of pure random search (without the DF mechanism) with the same number of trajectories integrated.

The images of fig. 9, 10 and 11 are the projections of the reachable set into the plane $V_1(T)$, $\int_0^T V_2(t)\,dt$ (investments in the sector P_1 and total consumption in the period T). The reachable set of fig. 10 was obtained with the "bang-bang" controls with 3 switching points, randomly generated. The image of fig. 11 corresponds to "bang-bang" controls, generated randomly at each time-step, with no restrictions imposed on the number of switchings.

III. 3.7. Example

In the previous example the system nonlinearity was rather weak and the adventage achieved by using the DF algorithm was not very significant. Let us consider another example, with stronger nonlinearities. Let the system equations are

$$\dot{x}_1 = x_2 + u_1$$

$$\dot{x}_2 = x_1 + x_2^2 + u_2 \quad ,$$

where u_1, u_2 $\varepsilon[-0.5, +0.5]$, $x_1(0) = x_2(0) = 0$, $t\varepsilon[0, 2.24]$. The projections of the reachable set at the plains (t, x_1) and (t, x_2) are shown in fig. 12. The time-section of the reachable set (the end-points of all integrated trajectories) for t = 2.24 is shown in fig. 13. A similar experiment has been done without the DF mechanism, with pure random switching times. The corresponding time-section is shown in fig. 14. It is easy to observe the sharply non-uniform density of the probability of hitting different regions within the reachable set in this case. A little bit better illustration of this fact can be obtained while calculating the probability of hitting as some smoothed continuous function. This can be done, for example, supposing that for each point of fig. 13 and 14 a neighbourhood exists in which the probability of being reachable is greater then zero. It was assumed that this probability is normal one centered in the corresponding point and having sufficiently small variance. The probability that some point of the (x_1, x_2) - plain is reachable was assumed to be the superposition of the above probabilities. The shape of the density function obtained in this way (with DF mechanism) is shown in fig. 15. An estimate of the reachable set can be obtained by truncation of this probability density to some small level, as shown in fig. 16. The corresponding density and its truncation to the same level, obtained without the DF mechanism are shown in fig. 17 and 18 respectively. It is clear that the estimated shape of the reachable set in the later case is incorrect.

References

1 Ważewski T., Sur une condition équivalente à l'equation au contingent. Bulletin de l'Academie Polonaise des Sciences, Ser.Sci.Math., Astr., Phys., vol. IX (1961), no. 2.

2 Ważewski T., Sur une généralisation de la notion des solutions d'une équation au contingent, Ibid., vol. X, no. 1, 1962.

3 Ważewski T., Sur les sýstèmes de commande nonlinéares dont le contredomaine de commande n'est pas forcément convexe, Ibid., vol. X, no.1, 1962.

4 Ważewski T., Sur une condition d'existence des fonctions implicites measurables, Ibid., vol. IX, no. 12, 1961.

5 Ważewski T., Sur un système de commande dont les trajectoires coincident avec les quasitrajectoires système de commande donné, Ibid., vol. XI, no. 3, 1963.

6 Ważewski T., On an optimal control problem / in conntion with the theory of orientor fields of A. Marchaud and S. Zaremba/, Proc. of the Conference "Differential Equations and Their Applicacions", Praga, 1962.

7 Ważewski T., Sur la sémicontinuité inférieure du "tendeur" d'un ensamble compact variant d'une façon continue, Bulletin de l'Academie Polonaise des Sciences, Ser. Sci. Math., Astr., Phys., vol. IX, no. 12, 1961.

8 Marchaud A., Sur les champs de demi-cones et les équations differentielles du primier ordre, Bulletin de la Soc. Math. de France, vol. 62, (1934).

9 Zaremba S. K., O równaniach paratingensowych, Dodatek do Rocznika Polskiego Towarzystwa Matematycznego, tom. IX, Kraków, 1935.

10 Turowicz A., Sur les trajectoires et quasitrajectoires des systèmes de commande nonlinéaires. Bull, de l'Academie Polonaise des Sciences, Ser. Sci. Math., Astr., Phys., vol. X, no. 10, 1962

11 Turowicz A., Sur les zones d'émision des trajectoires et des quasitrajectoires des systèmes de commande nonlinéaires, Ibid., vol. XI, no. 2, 1963.

12 Shui-Nee Chow, Shuur J. D., Fundamental theory of contingent equations in Banach space. Transactions of the American Mathematical Society, vol. 179, (1973).

13 Kuratowski K., Ryll-Nardzewski C., A general theorem on selectors, Bulletin de l'Academie Polonaise des Sciences, Série des Sci. Math., Astr. et Phys., vol. XIII, no. 6, 1965.

14 Castaing C., Valadier M., Convex analysis and measurable multifunctions, Lecture Notes in Mathematics, vol. 580, Springer-Verlag, Berlin-Heidelberg, New York, 1977.

15 Warga J., Optimal control of differential and functional equations, Academic Press, New York and London.

16 Lions J. L., Equations differentialles operationnelles et problémes aux limites. Springer-Verlag, Berlin-Göttingen-Heidelberg, 1961.

17 Lions J. L., Optimal control of systems governed by partial differential equations. Springer-Verlag, 1971.

18 Deimling K., Ordinary differential equations in Banach spaces. Lecture Notes in Mathematics, vol. 596, Springer-Verlag, Berlin-Heidelberg-New York, 1977.

19 Bourbaki N., Eléments de mathématique. Première partie, livre VI, Intégration.

20 Maurin K., Analiza. PWN, Warszawa, 1971.

21 Leitmann G., An introduction to optimal control. McGraw-Hill, 1966.

22 Pontriagin L. S., Boltijanskii V. G., Gamkrelidze R.V., Mishchenko E.F., Mathematical theory of optimal processes. Wiley, New York, 1962.

23 Lee E. B., Markus L., Foundations of optimal control theory. John Wiley and Sons, Inc., New York, London, Sydney, 1967.

24 Kuck D. J., Lawrie D. H., Sameh A. H., High speed computer and algorithm organization. Academic Press, 1977.

25 Kascic M. J., Vector processing on the cyber 200. Angew. Inf.: vol. 22, 1980.

26 Jones A. K., Schwarz P., Experience using multiprocessor systems - a status report. Comp. Surv., vol.12, 1980.

27 Gibson J. A., Gibbard R.W., Hybrid serial-parallel digital computation. Compt. Electr. Eng., vol.7, 1980.

28 Polak E., Computational methods in optimization. Academic Press, 1971.

29 Raczynski S., On the determination of reachable sets and optimal control by the random method. IFAC Symposium "Optimization methods", Varna 1974.

30 Rastrygin L. A., Al'goritmy y programmy slutshaynovo poiska. ZINATNE, Ryga 1969.

31 Martin R.H. Jr., Nonlinear operators and differential equations in Banach spaces. John Wiley & Sons, Inc., 1976.

32 Feller W., An introduction to probability theory and its applications. John Wiley and Sons, Inc., New York, 1957.

33 Plís A., Trajectories and quasitrajectories of an orientor field, Ibid., vol. XI, no.6, 1963.

34 Klambauer G., Real analysis, American Elsevier Publishing Company, 1973.

35 Clarke F. H., The maximum principle under minimal hypothesis, SIAM Journal of Control and Optimization, Vol.14, 1976.

36 Conti R., A convex programming problem in banach spaces and applications to optimum control theory, Journal of Computer and System Sciences, vol.4, No.1, 1970.

37 Rockafellar R. T., Duality in optimal control, in the book: Mathematical Control Theory, Springer-Verlag, Lectures Notes in Mathematics, No.680, 1978.

38 Aubin J.P., Cellina A., "Differential Inclusions", Springer-Verlag, 1984.

'tq.

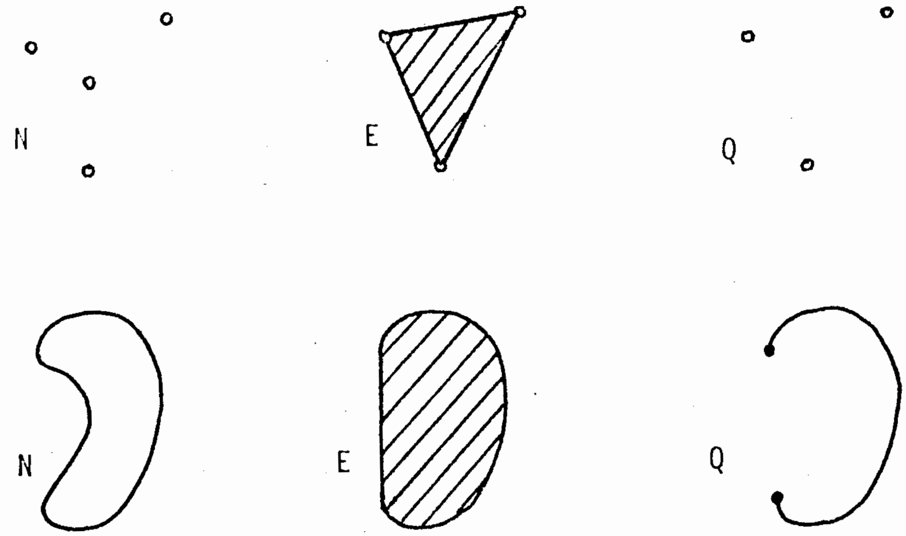

Fig 1

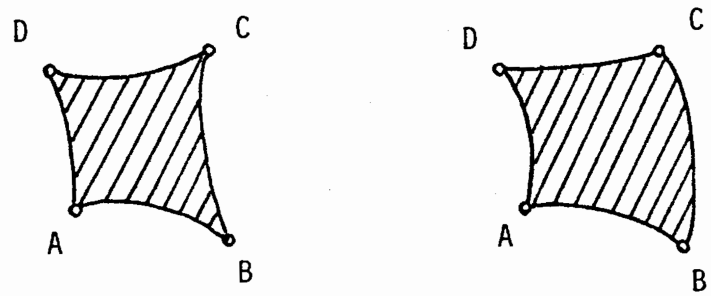

Fig 2

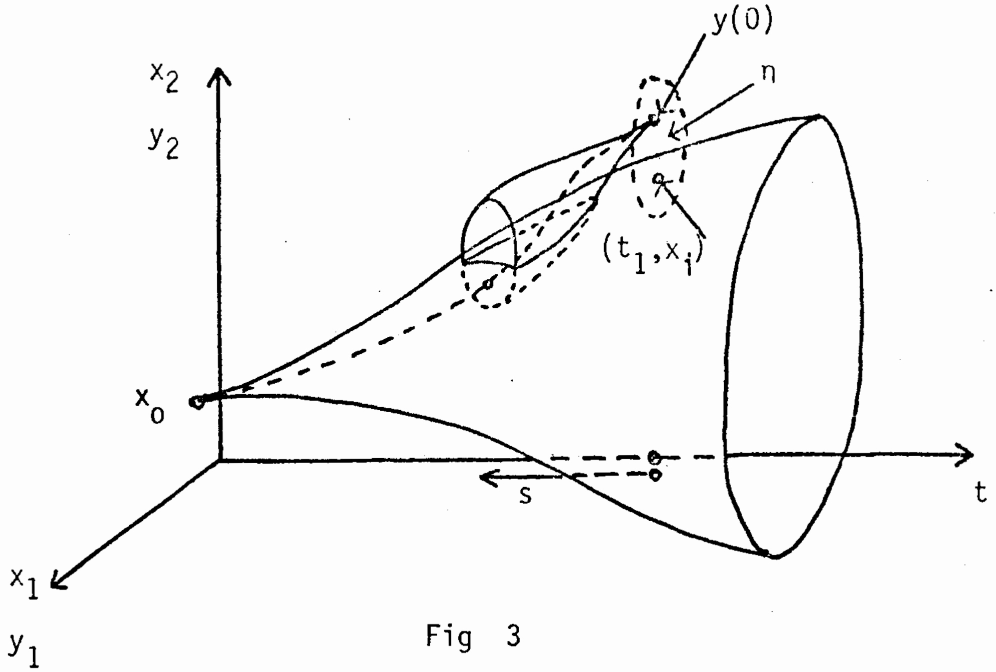

Fig 3

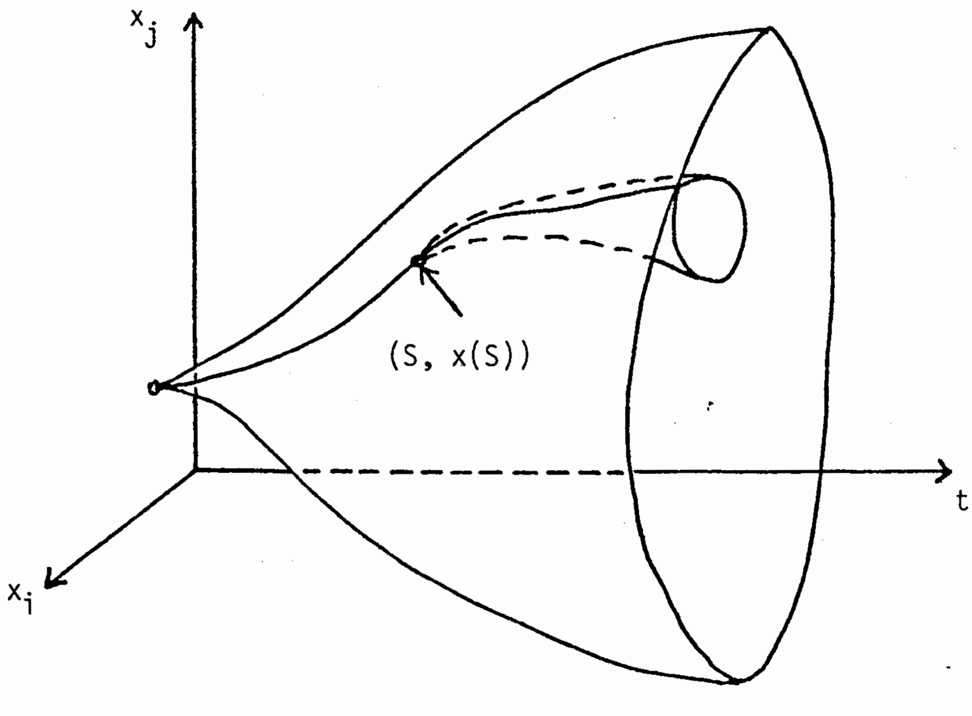

Fig 4

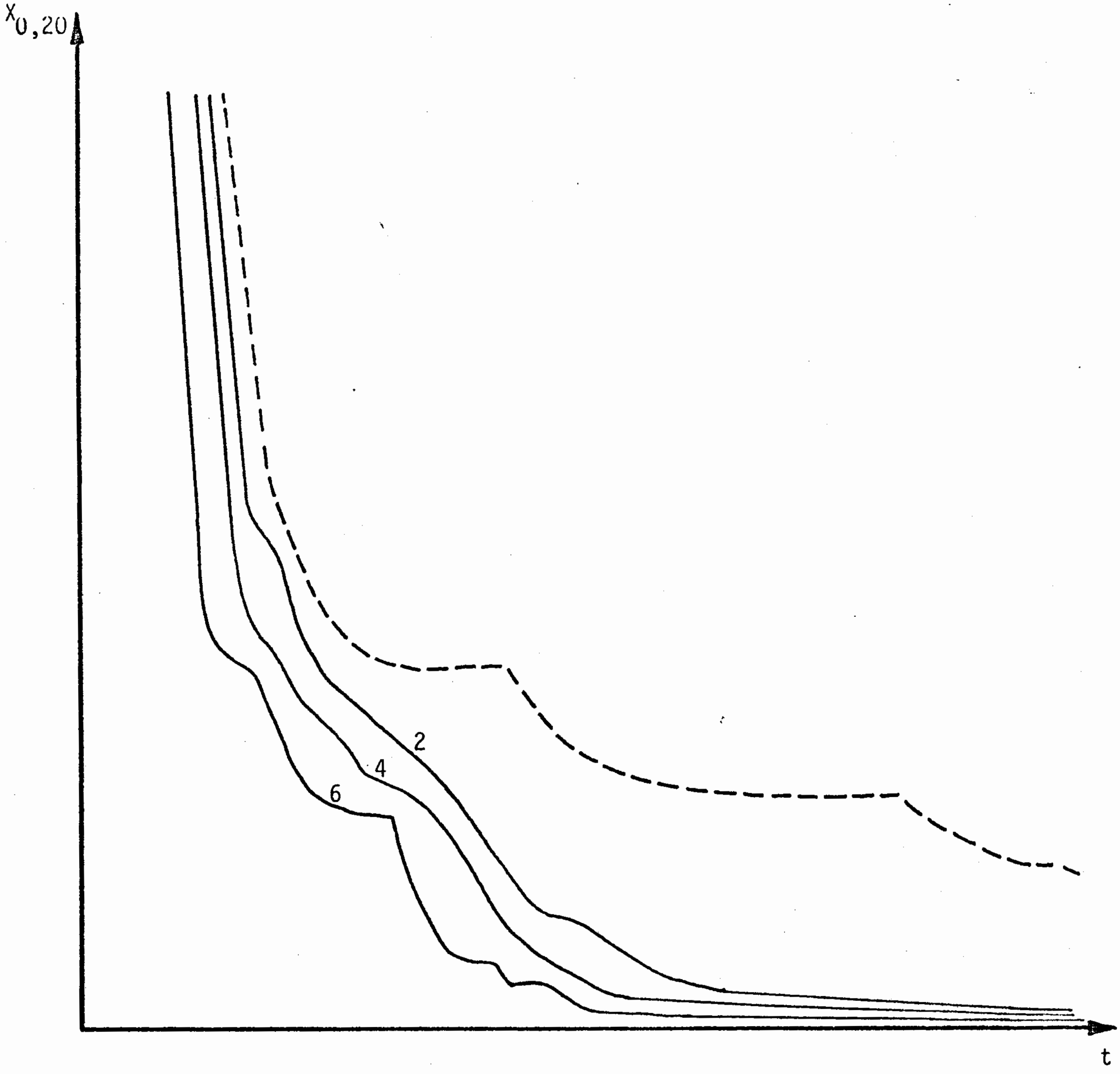

Fig 5

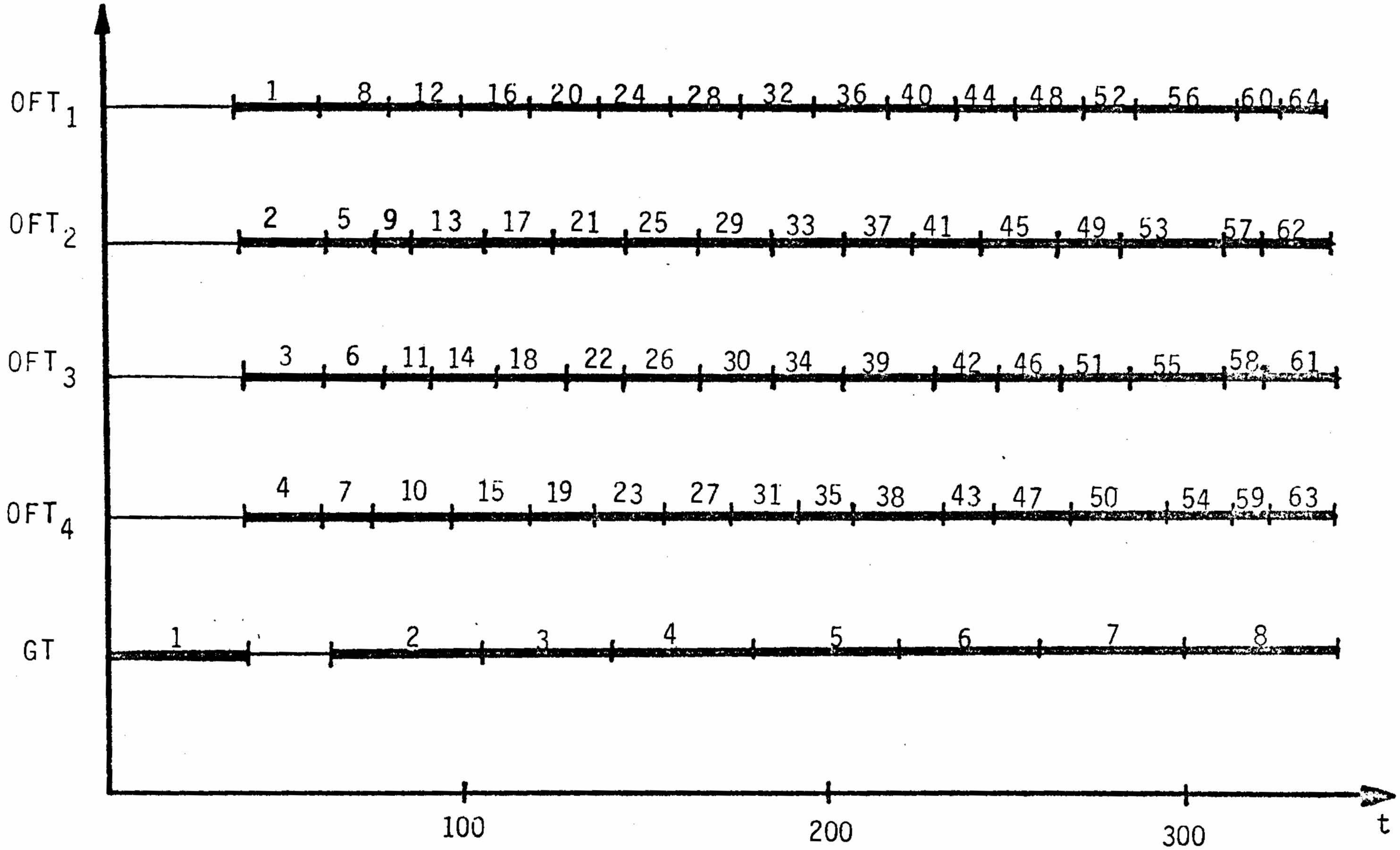
OFT$_1$
OFT$_2$
OFT$_3$
OFT$_4$
GT
1 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64
2 5 9 13 17 21 25 29 33 37 41 45 49 53 57 62
3 6 11 14 18 22 26 30 34 39 42 46 51 55 58 61
4 7 10 15 19 23 27 31 35 38 43 47 50 54 59 63
1 2 3 4 5 6 7 8
100
200
300
t

Fig 6

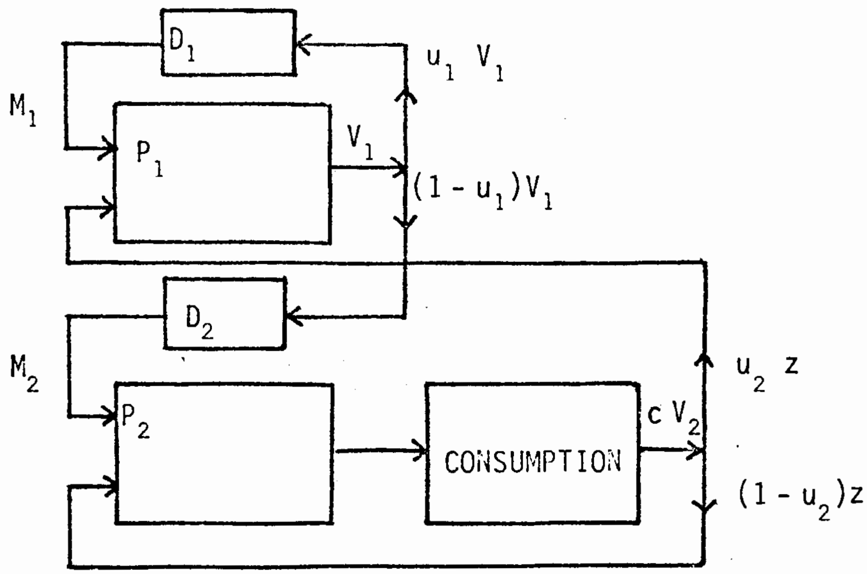

Fig 7

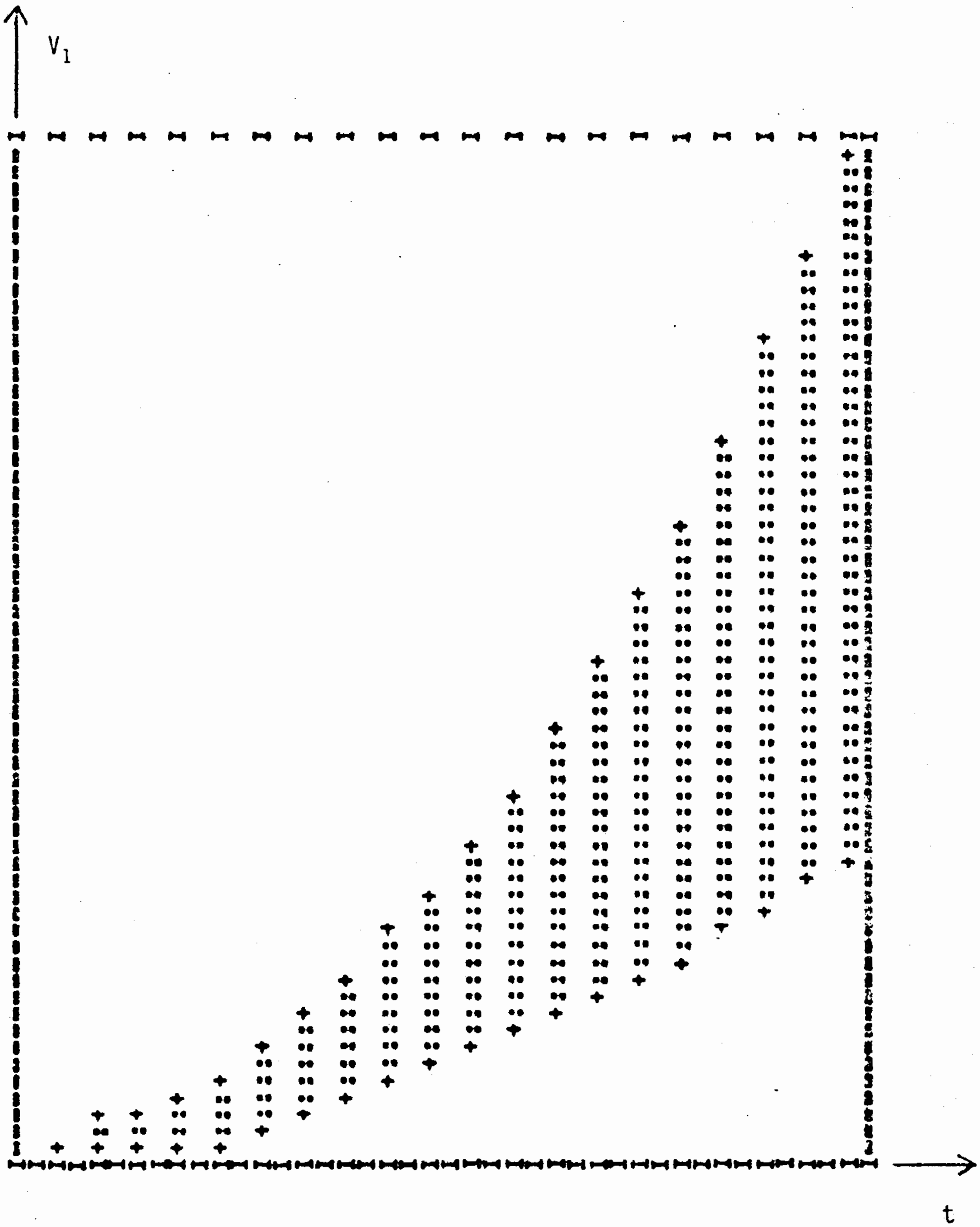

Fig 8

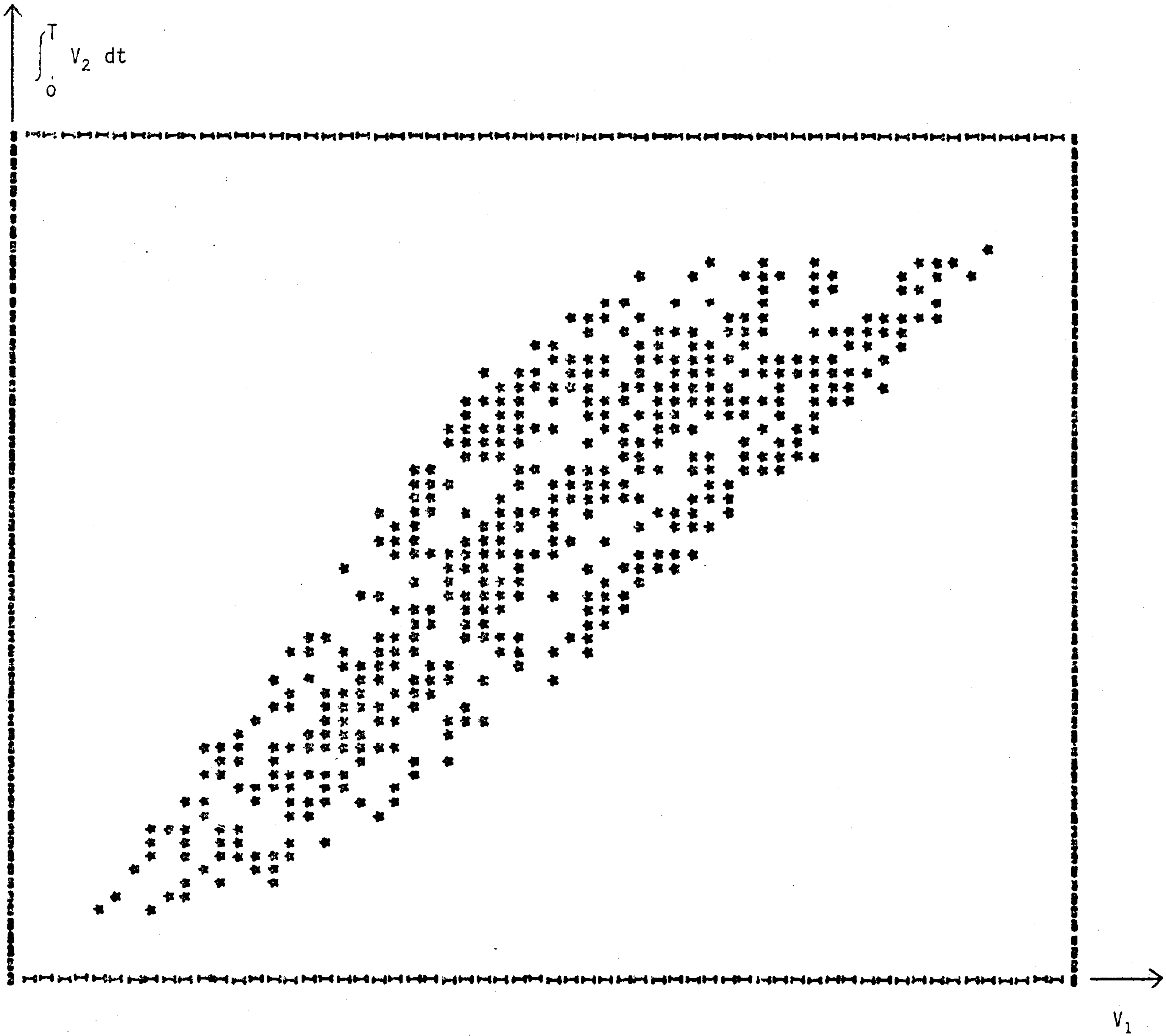

Fig 9 -

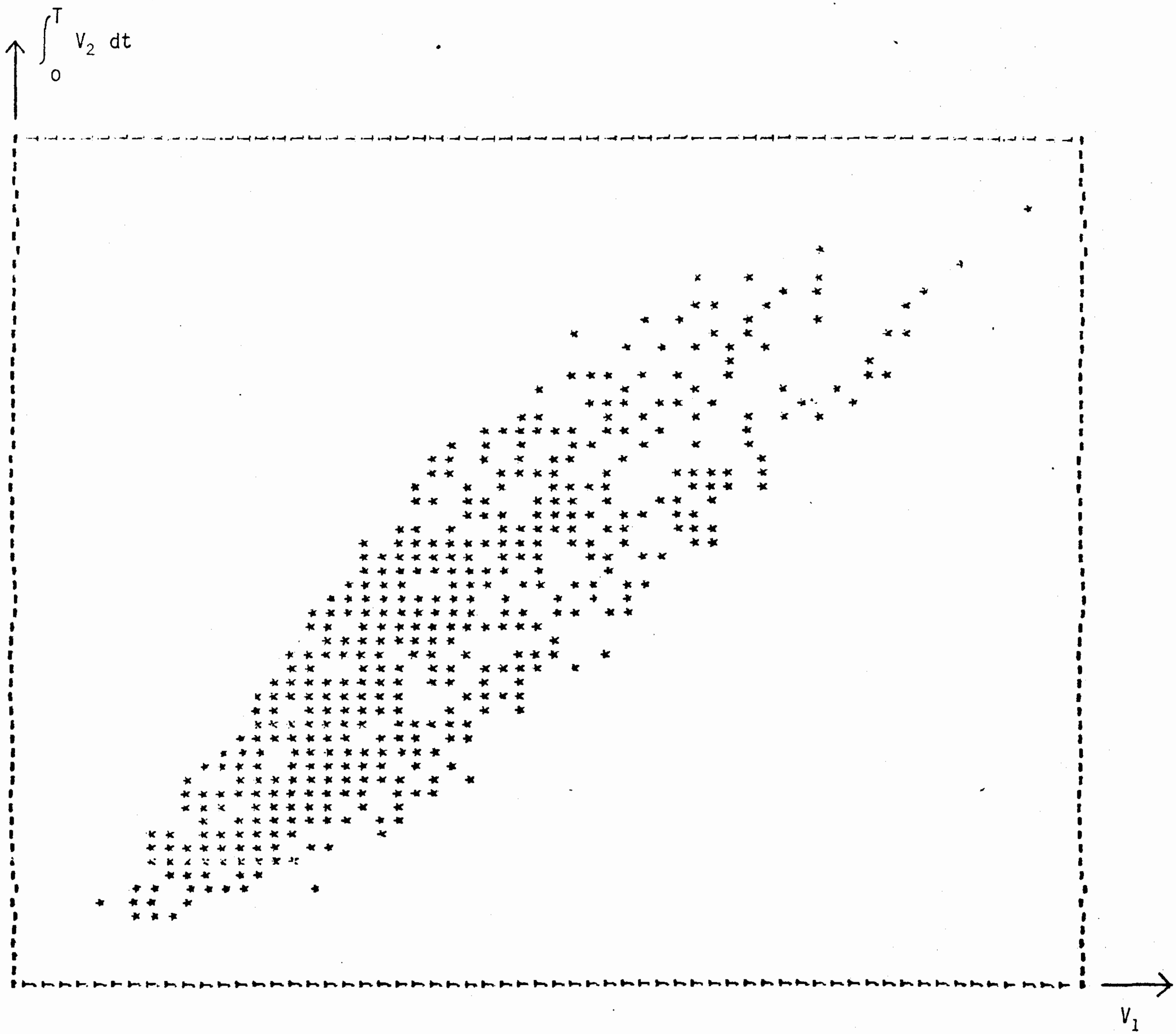

Fig 10

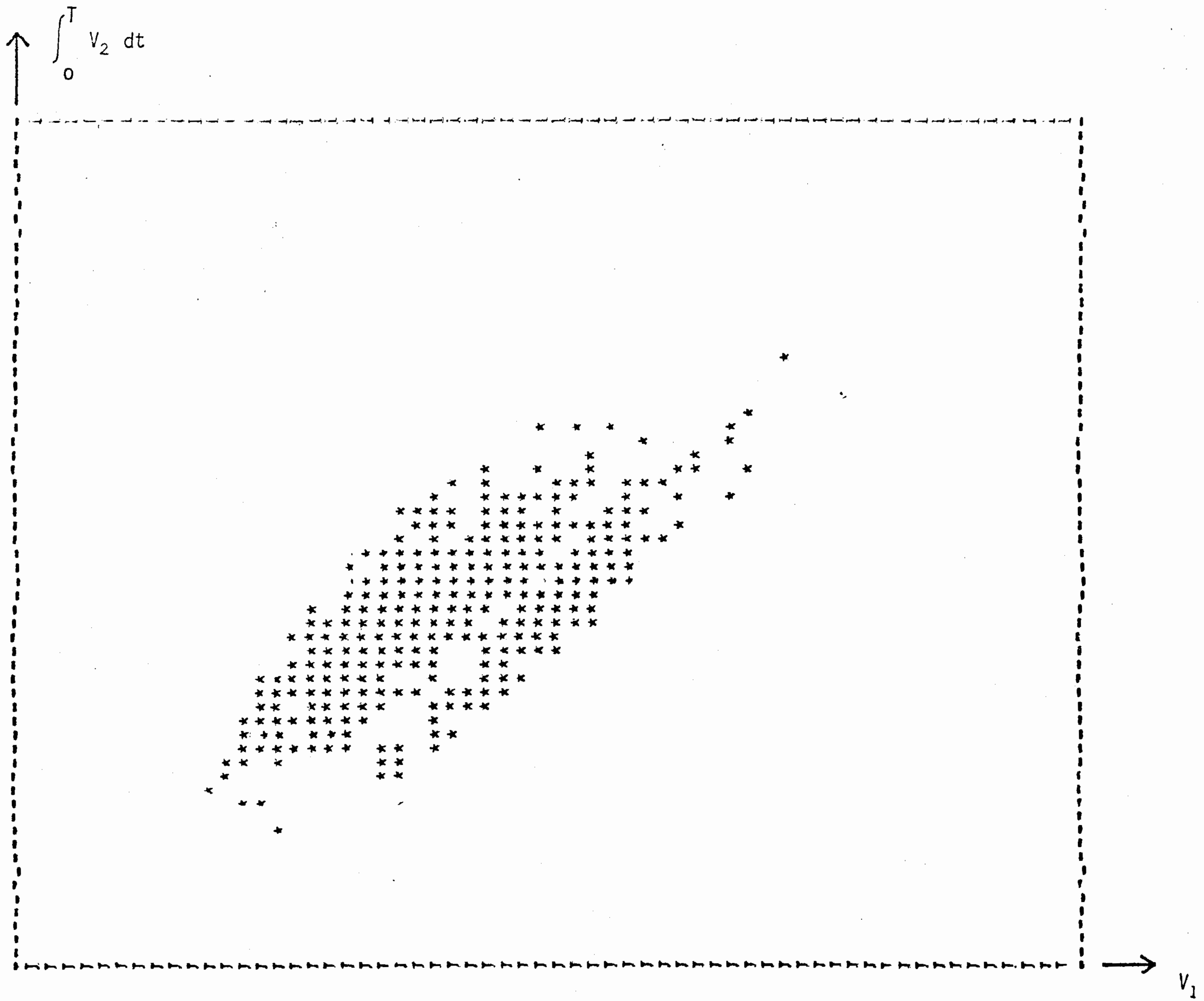

Fig 11

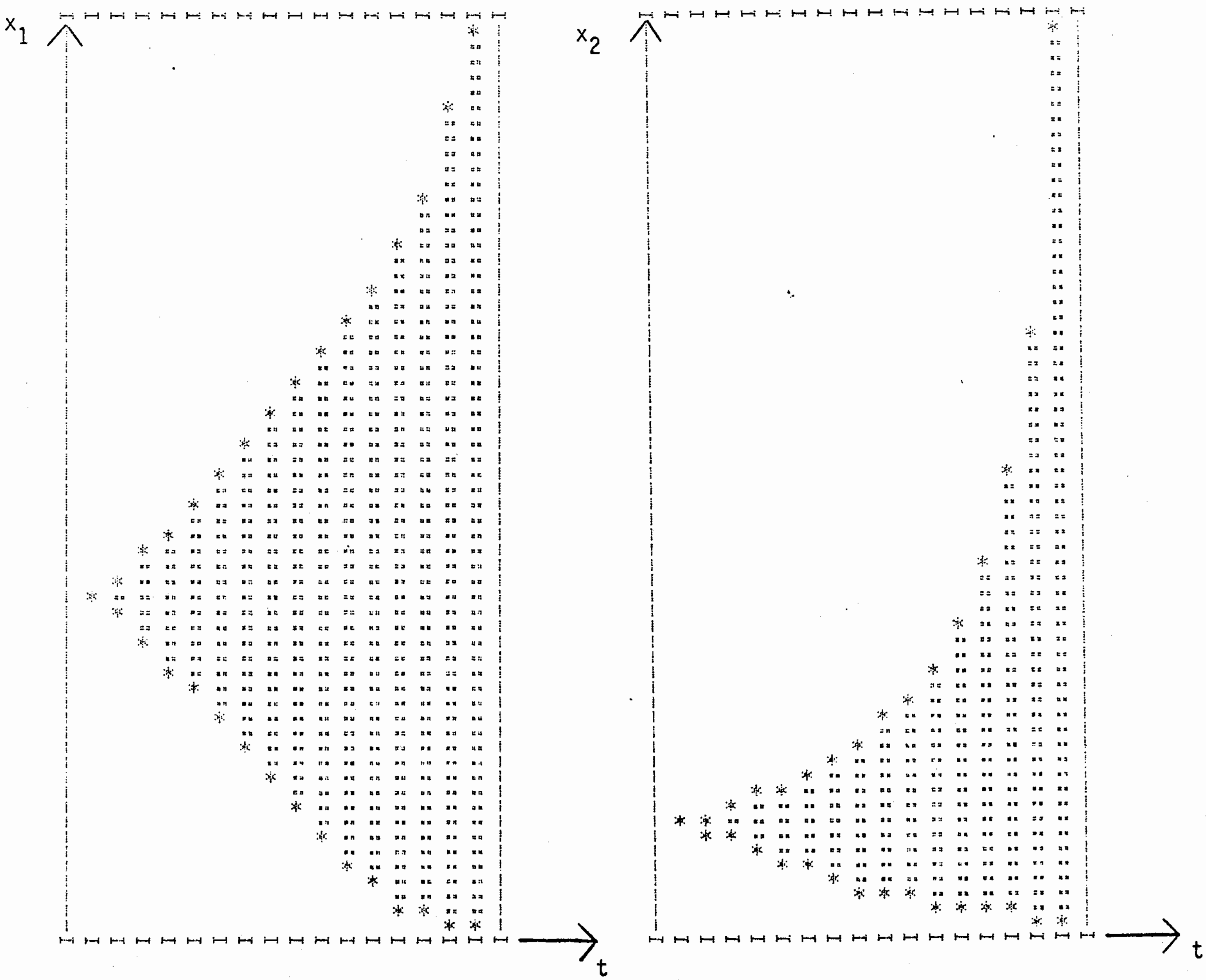

Fig 12

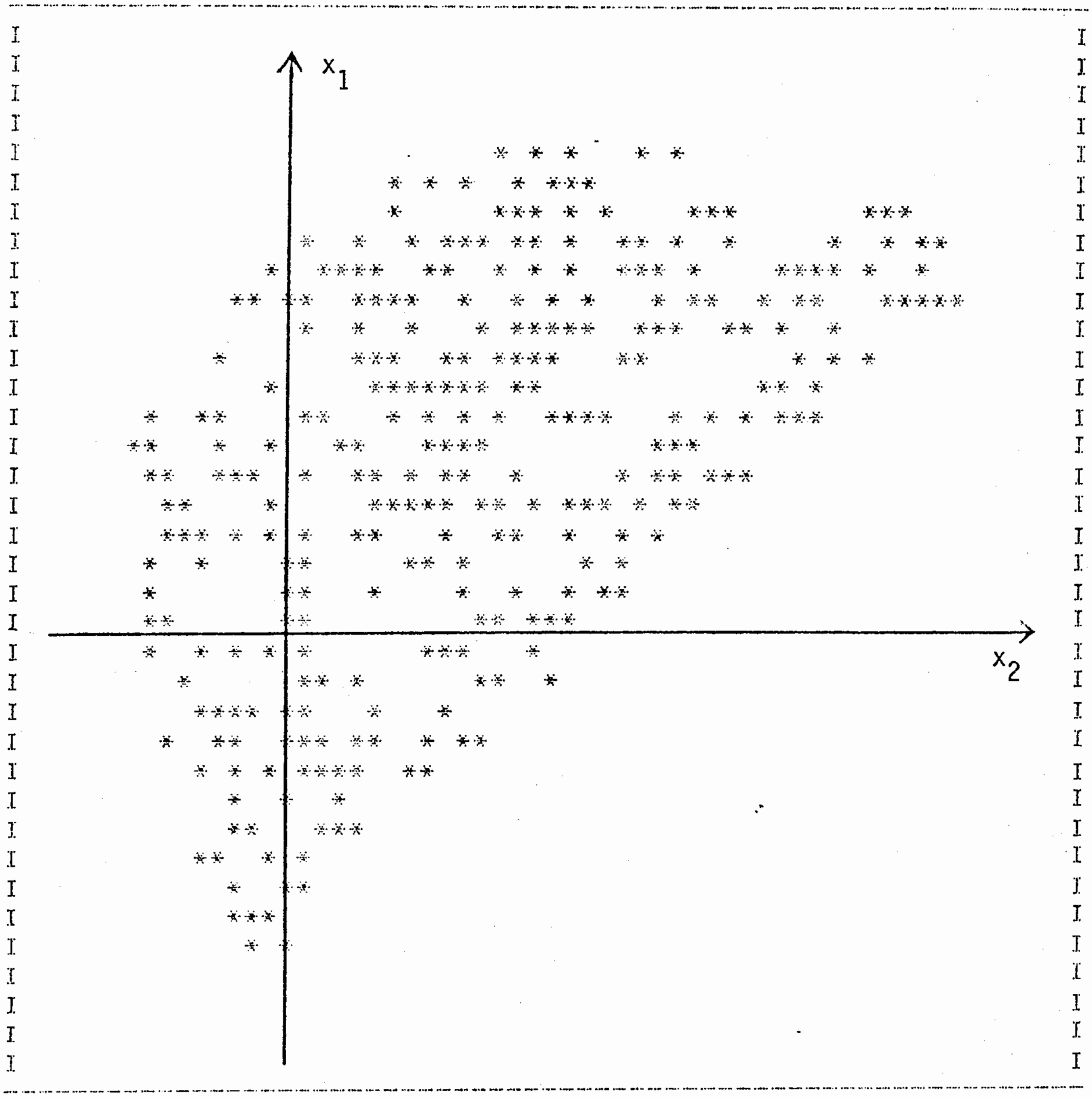

Fig 13

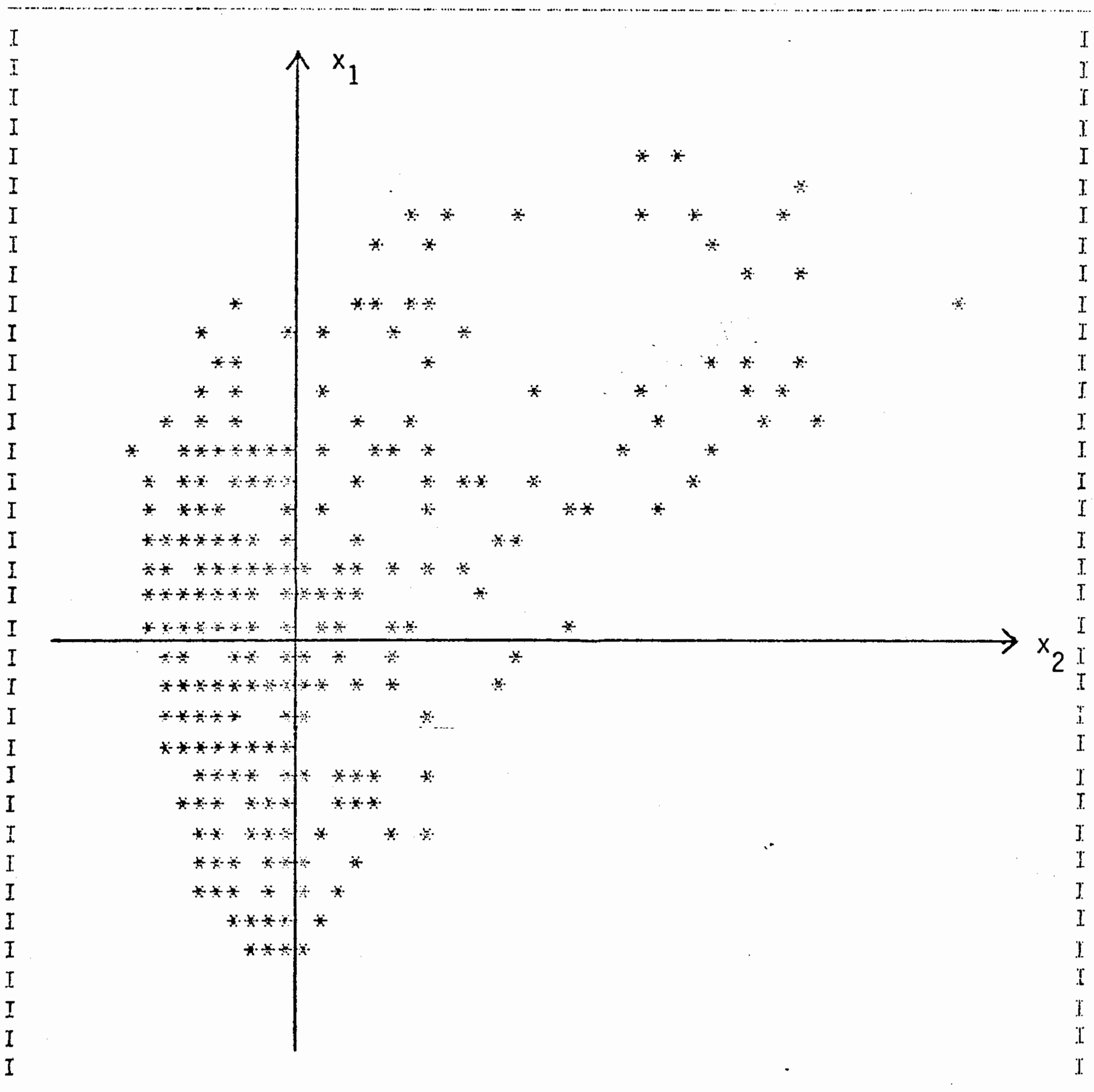

Fig 14

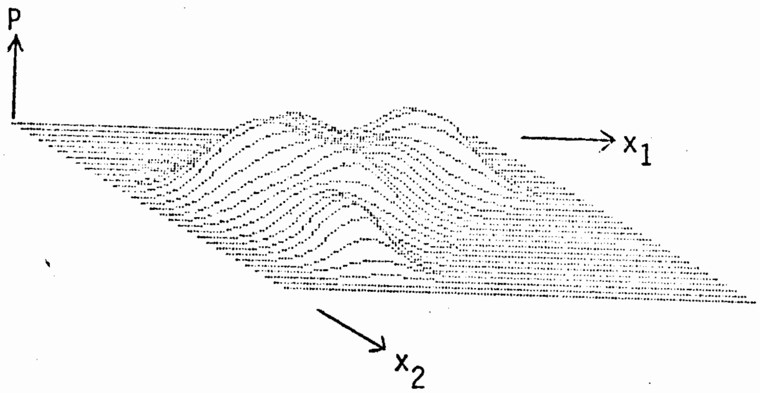

Fig 15

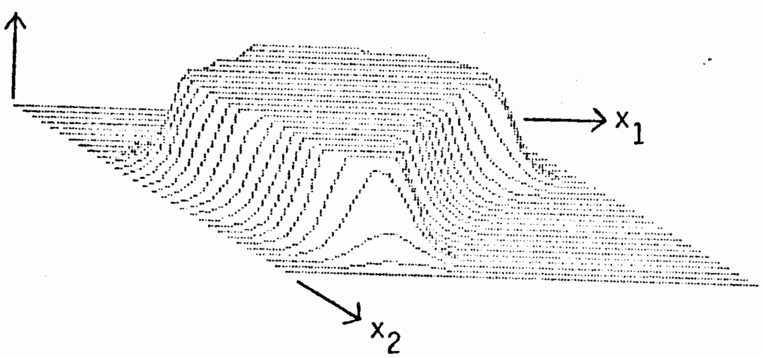

Fig 16

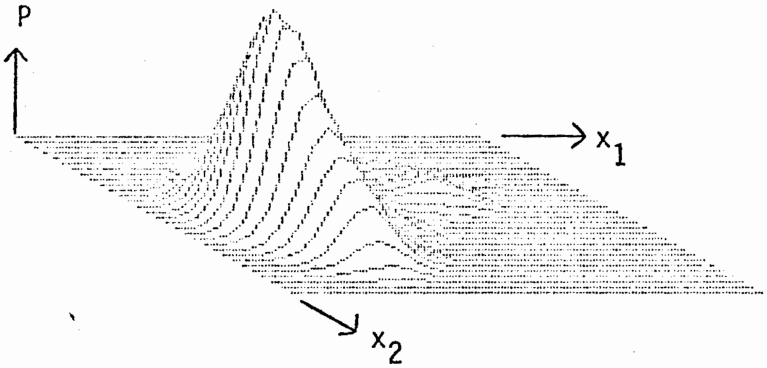

Fig 17

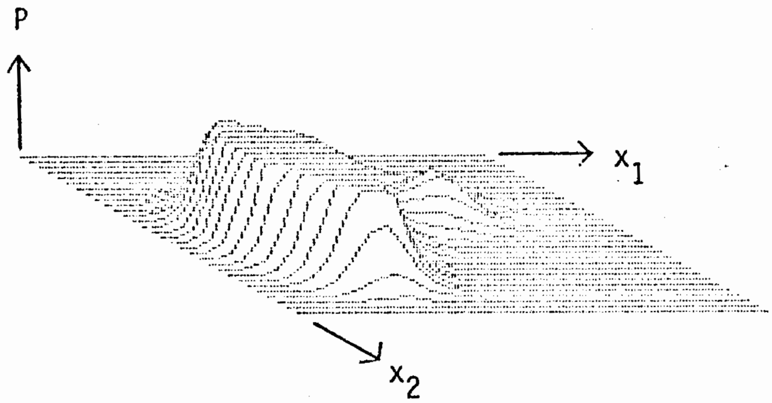

Fig 18