ON THE NECESSARY CONDITION FOR OPTIMAL CONTROL OF NONLINEAR SYSTEMS*

By

HUBERT HALKIN

*in Stanford, California, U. S. A.***

ACKNOWLDGMENTS

The author wishes to express his sincere gratitude to Professor M. M. Schiffer of the Department of Mathematics of Stanford University. His constant guidance during the years 1960–1963 has been invaluable.

The author spent many months at the Department of Mathematics and the Center for Control Theory of the Research Institute for Advanced Studies, RIAS. Numerous discussions with Professors S. Lefschetz, J. P. LaSalle and R. E. Kalman have influenced him profoundly.

The interest of the author in the engineering aspects of the theory of optimal control was greatly enhanced by numerous conversations with Professor I. Flügge-Lotz of the Department of Aeronautics and Astronautics and Professor G. Franklin of the Department of Electrical Engineering, both at Stanford University.

The author would also like to thank Dr. L. W. Neustadt of Aerospace whose valuable comments on some of the author's previous papers were a motivating force in the development of this work.

Finally, Professor Baudouin Fraeijs de Veubeke of the University of Liège, who introduced the author to this interesting field of research, is sincerely acknowledged.

The author was the grateful recipient of a C.R.B. Graduate Fellowship (Perrin C. Galpin Fellow) of the Belgian American Educational Foundation during the years 1960–1962 and of a Stanford University Fellowship during the year 1962–1963.

---

* This work was supported in part by Office of Naval Research Contract Nonr-225(11) at Stanford University.

** Presently at Bell Telephone Laboratories, Whippany, N. J., U.S.A.

TABLE OF CONTENTS

INTRODUCTION

Words such as "system," "control," "optimal control," etc., have recently become very popular among a large group of engineers, particularly in aero-astronautics and electronics, as well as among many social science researchers in economics and psychology. The use of the same vocabulary in totally unrelated fields of study should not surprise us when we realize that they refer to problems which have the same mathematical structures. Among these mathematical structures, two are of particular importance: the *dynamical system* and the *control system* which are defined as follows:

A *dynamical system* is a pair (Y, R) where Y is an arbitrary space, called the event space, and where R is a binary relation on Y such that

(i) aRa for all $a \in Y$, i.e., R is reflexive

(ii) aRb and bRc implies aRc, i.e., R is transitive

(iii) aRb and bRa implies $a = b$, i.e., R is antisymmetric

(iv) aRb and aRc implies either bRc or cRb.

These properties of the relation R correspond to the usual concepts of causality and determinism. We may think of the formula aRb as meaning: the event b follows the event a. This correspondence will be seen clearly in the examples given later.

A *control system* is a triple $(Y, F, R(\cdot))$ where F is an arbitrary space, called the strategy space, and where $(Y, R(\alpha))$ is a dynamical system for every α in F. To clarify these ideas let us consider a particular type of dynamical system and the control system corresponding to it.

First we shall consider the dynamical system associated with a system of ordinary differential equations. In such a case the space Y is the Cartesian product of the real time axis T with elements t and of the n-dimensional Euclidean space X, called the state space, with elements $x = (x^1, \cdots, x^n)$. A function $f(x, t)$ from $X \times T$ into X is given. A binary relation R over $T \times X$ is then defined by

$$(t_1,x_1)R(t_2,x_2) \text{ if and only if } x(t_2;x_1,t_1)=x_2 \quad \text{and} \quad t_1 \leqq t_2$$

where $x(t;x^*,t^*)$ is the solution of the differential system

$$\dot{x}=f(x,t) \qquad \text{a.e. } t$$

satisfying the initial condition

$$x(t^*;x^*,t^*)=x^* .$$

We shall say that this binary relation R is generated over $T \times X$ by the differential system $\dot{x}=f(x,t)$. We easily see that the relation R satisfies the conditions (i) to (iv) given earlier, which implies that $(T \times X, R)$ is a dynamical system. More generally, the relation R could be generated by a system of difference equations, of differential-difference equations, of integral equations, etc. The relation R could even be defined explicitly by an appropriate subset of $T \times X \times T \times X$.

Let us now consider the control system associated with a class of systems of ordinary differential equations. The space Y is again the Cartesian product of the real time axis T and of the state space X. The strategy space F is given and for each $\alpha \in F$ a function $f(x,t;\alpha)$ from $X \times T$ into X is given. The binary relation $R(\alpha)$ is then generated over $T \times X$ by the differential system

$$\dot{x}=f(x,t;\alpha)$$

according to the previous definition. The triple $(T \times X, F, R(\cdot))$ is therefore a control system. We shall say that (t_2,x_2) is reachable from (t_1,x_1) if $(t_1,x_1)R(\alpha)(t_2,x_2)$ for some $\alpha \in F$.

Physics, as most descriptive sciences, is concerned with the study of dynamical systems, whereas engineering, economics and the other normative sciences are concerned with the study of control systems.

Classical mechanics offers a simple example of a dynamical system: the set Y in classical mechanics is the Cartesian product of the time axis and of the state space, and the relation R is generated by the laws of mechanics, which are given in most cases under the form of a set of differential equations.

However, the whole field of physics cannot be reduced to a scheme of such simplicity; no Laplace's observer could help. Quantum mechanics, for instance, is a dynamical system where the set Y is the Cartesian product of the time axis and of function spaces of probability distributions.

As an example of a control system, let us consider a rocket. A rocket is a mechanical system equipped with regulatory devices such as rudders, thrust modifiers, etc. The variables describing the position of these regulatory devices are called *control variables.* If these control variables are given functions of the time and of the state variables, i.e., position, velocity, etc., then we have a dynamical system. If, however, we are allowed to choose the functions describing the control variables in a certain given class of functions called the strategy space, then we have a control system.

It is not difficult to imagine many examples of control systems in other areas of engineering and economics. It should be noted that a control system could be stochastic: in many problems the set Y will be, as in quantum mechanics, the Cartesian product of the time axis and function spaces of probability distributions.

When dealing with a dynamical system the essential question which one should ask is: how does it behave, i.e., given $a \in Y$ which are the properties of the set $\{b : aRb\}$? This question has been the object of extensive studies, especially the theory of stability and oscillation of dynamical systems described by ordinary differential equations.

In the case of a control system a new type of question may be asked: what is the "best" element of the strategy space? For example: given a rocket and initial and terminal points in the state space, how should we choose the control variables so that the rocket will pass from the initial to the terminal point in the minimum amount of time?

Generally, we shall define an *optimal control problem* as follows:

Given

(1) a control system $(Y, F, R(\cdot))$

(2) a subset D of $Y \times Y$

(3) a real function g on $G = \{\langle a,b,f\rangle : (a,b,f) \in D \times F,\ aR(f)b\}$ where $\langle a,b,f\rangle = \{c : aR(f)c \text{ and } cR(f)b\}$.

Find an element $\langle a,b,f\rangle \in G$ such that $g(\langle a,b,f\rangle)$ is maximum. This optimal control problem is denoted by the quintuple $(Y,F,R(\cdot),D,g)$.

In the theory of optimal control, as in the theory of dynamical systems, it is possible to obtain very interesting results when we assume that the relation R is generated by the solutions of a system of differential equations.

In this work we shall consider *control systems* $(Y,F,R(\cdot))$ of the following type:

Y is the Cartesian product of the real time axis T with elements t and of an n-dimensional Euclidean space X with elenemts $x=(x^1,\cdots,x^n)$.

F is the class of measurable r-dimensional functions $[u]$ defined on T and taking their values in a given set Ω.

$R([u])$ for $[u]\in F$ is generated by the solutions of the system

$$\dot{x}=f(x,u(t),t) \qquad \text{a.e. } t$$

where $f(x,u,t)$ is a given function.

With this particular type of control system we shall associate an *optimal control problem* $(Y,F,R(\cdot),D,g)$ of the following type:

$D=A\times B$ where A is a set consisting of one point of Y and where B is the set consisting of one line of Y parallel to the axis x^n.

$g(\langle(x_a,t_a),(x_b,t_b),[u]\rangle)=x_b^n$.

In other words, we are given a point A in $X\times T$ and a line B in $X\times T$, parallel to the axis x^n; how can we find a function $[u]$ in a given class F such that starting at A and integrating the system $\dot{x}=f(x,u(t),t)$ we would end on B as far as possible in the positive direction on x^n? We call this problem the *fundamental problem of optimal control*. This fundamental problem is stated in greater detail in Section 1.

This work is principally devoted to the study of the necessary conditions for the solution of the fundamental problem of optimal control.

We want to stress here the fundamental difference between a classical problem of calculus of variations and a problem of optimal control. In an optimal control problem the set Ω may be quite arbitrary, and due to technological limitations it is very often a bounded and closed set: for instance, the thrust of a rocket can only vary on the closed interval $[0,m]$ where m

is the maximum available thrust. If a classical problem of calculus of variations is put in the form of an optimal control problem, the corresponding set Ω is always open. This explains why the classical techniques of the calculus of variations do not work for the general case of optimal control problems. Indeed, one of the most fundamental concepts of the calculus of variations is the concept of the arbitrary variation: you compare a nominal trajectory corresponding to the strategy f with the trajectory corresponding to a strategy $f + \delta$ and in calculus of variations this is always possible when δ is small enough. In optimal control this is not true anymore: if along a rocket trajectory the thrust has the maximum available value m at some time, it has no meaning to consider comparison trajectory where the thrust is augmented by a positive δ, however small this δ could be.

The most important result in the theory of optimal control is the "Maximum Principle of Pontryagin," a generalization of the Weierstrass E-test of the classical calculus of variations. In this work the "Maximum Principle" is obtained by a method fundamentally different from the method of Pontryagin and his associates, in particular we avoid some unresolved topological difficulties encountered in their reasoning. It should be remarked also that the assumptions we are making in the statement of our problem, in particular on the differentiability and boundedness of the function $f(x, u, t)$, are much weaker than the assumptions made by Pontryagin and his associates. In a previous publication, [12], using the same method, we obtained the same results for a more restrictive class of problems.

Any mathematical venture is made up of two parts: geometrical intuition and analytical machinery. From the chronological point of view the geometrical intuition always precedes the analytical manipulation in the formation of a theory and the first is of great help to understand the second. Unfortunately, this duality has a marked tendency to disappear and the role of geometrical intuition is barely noticeable in the final form of a theory. This work is no exception to this rule: the analytical machinery is easily seen. Besides classical results of the theory of ordinary differential equations, we use some extensions of the results of Lyapounov [16] and Blackwell [3] on the range of a vector integral and an application of Brouwer's Fixed Point Theorem. Unfortunately, the geometrical motivation is virtually absent from this work. For this reasom we shall make up for this deficiency in the introduction. More precisely, we shall generalize the concept of propagation and show that the fundamental problem of optimal control described above can be viewed as a problem of oqtimal propagation in an abstract space X.

A standard problem of classical propagation theory has the following structure: we are given a medium with a propagation law; the medium is a rest for $t < t_0$, we produce a certain perturbation at time t_0 and we want to predict what will happen for $t > t_0$.

This standard problem could be considered on two different levels. If we want to predict the intensity of the perturbation for every element in space-time as a function of the intensity of the initial perturbation at time t_0, we have what we call a quantitative propagation problem. In some circumstances, however, it is enough to predict which elements in space-time could possibly be perturbed as soon as we know which points are perturbed at the time t_0. This is what we call a qualitative propagation problem.

To solve a quantitative propagation problem we need the concept of intensity of a perturbation and a precise description of the space-time variation of this intensity which is usually given by a partial differential equation.

In this work we shall restrict our interest to qualitative propagation problems and consider the fundamental problem of optimal control as a generalization of the qualitative propagation problem.

To every element $(x,t) \in X \times T$ we shall associate the set

$$w(x,t) = \{f(x,u,t) : u \in \Omega\} .$$

The set $w(x,t)$ will be called the "wavelet" at the point x for the time t. The analogy with optics is clear: whenever a perturbation is produced at the point x at the time t then, in first approximation, all the points of the set

$$\{x + \alpha\, dt : \alpha \in w(x,t)\}$$

will be perturbed at the time $t + dt$.

If we write $x(t;[u])$ for the solution of

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

satisfying the initial condition $x(t_a; [u]) = x_a$ then

$$W(t) = \{x(t; [u]) : [u] \in F\}$$

is the set of points of X which at the time t could possibly be affected by a perturbation having taken place at $x = x_a$ at the time $t = t_a$. In other words, the boundary $\partial W(t)$ of the set $W(t)$ is the wavefront at the time t of a perturbation starting at $x = x_a$ at the time $t = t_a$.

We define a ray as a solution $x(t; [u])$ such that

$$x(t; [u]) \in \partial W(t) \qquad \text{for all } t \in [t_a, t_b].$$

We then have the following simple but fundamental property: if an element $[u]$ of F is optimal for the control problem, then $x(t; [u])$ is a ray of the propagation problem. The proof of this property is given in a previous paper [13] and may be summarized as follows: if $x(t_1; [u])$ is an interior point of $W(t_1)$ for some $t_1 \in [t_a, t_b]$ then $x(t_2; [u])$ is an interior point of $W(t_2)$ for all $t_2 \in [t_1, t_b]$ since the solutions of $\dot{x} = f(x, u(t), t)$ at the time t_2 depend continuously on the initial conditions at the time t_1; on the other hand, if $[u]$ is optimal then $x(t_b; [u])$ is a boundary point of $W(t_b)$ since otherwise there would be another $[\tilde{u}] \in F$ with

$$(x(t_b; [\tilde{u}]), t_b) \in B \quad \text{and} \quad x^n(t_b; [\tilde{u}]) > x^n(t_b; [u]),$$

contradicting the optimality of $[u]$; hence $x(t; [u]) \in \partial W(t)$ for all $t \in [t_a, t_b]$.

The optimal control problem is then reduced to the study of the rays of the abstract propagation problem. We may generalize to an abstract propagation problem the Huygens Principle and the associated Huygens construction. The basic facts about such a propagation may then be stated as follows: if a wavefront has a tangent plane at a point, then the wavelet leading to this point is entirely located on one side of this tangent plane. Consequently, we maximize the wavefront velocity, i.e., if p is the normal to the wavefront at this point then, along a ray passing through this point, the element $[u] \in F$ is such that the scalar product of p and $f(x, u(t), t)$ is maximum. This property is closely related to Pontryagin's Maximum Principle.

In this work we will give a precise analytical formulation of this scheme. To verify that an element $[v]$ of F is optimal, we adopt the point of view of an observer riding along the ray $x(t;[v])$ and making its observation in a moving frame of coordinates attached to the wavefront. For such an observer all the missed opportunities, i.e., the directions he could have followed but did not, are leading to points on one side of a hyperplane passing through the origin. This fact can be described analytically and leads to the mathematical formulation of the Maximum Principle of Pontryagin. The hyperplane mentioned above is the tangent hyperplane to the wavefront whenever such tangent hyperplane exists. It should be noted, however, that our derivation does not require the existence of such a tangent plane to the wavefront. In most intuitive derivations of Pontryagin's Maximum Principle the existence of the tangent plane is implicitly assumed: these derivations are very unsatisfactory since the real strength of the Maximum Principle of Pontyragin lies in its applicability to problems where this assumption cannot be made.

At the end of this introduction we want to compare the geometries of Finsler, Riemann, and Euclid with the geometry induced on an autonomous propagation space by the "wavelets" $w(x)$.

Let $F(x,y)$ be a real-valued function defined on $X \times X$. The function $F(x,y)$ induces in the space X a geometry for which the distance ds between two neighboring points x and $x + dx$ is given by

$$ds = F(x, dx) .$$

This geometry is called a Finsler geometry if

(i) $F(x, ky) = kF(x,y)$ for every $k > 0$ and all $(x,y) \in X \times X$.

(ii) $F(x,y) > 0$ if $y \neq 0$.

(iii) $F^2_{yy}(x,y)$ exists and is positive definite for all $(x,y) \in X \times X$.

A Riemannian geometry is a Finsler geometry such that

$$F^2(x,y) = \sum_{i,k=1}^{n} g_{ik}(x)\, y^i y^k$$

and a Euclidean geometry is a Riemannian geometry where

$$g_{ik}(x) = \delta_{ik}$$

or equivalently

$$F^2(x,y) = \sum_{i=1}^{n} (y^i)^2 .$$

To a Finsler geometry characterized by the function $F(x,y)$ on $X \times X$ we associate the set valued function $I(x)$ on X defined by

$$I(x) = \{y : F(x,y) \leqq 1\} .$$

The set $I(x)$ is called the *indicatrix* at the point x of the Finsler geometry on X. It follows from the conditions (i) to (iii) that for a Finsler geometry the set $I(x)$ has, with respect to the Euclidean norm, the following properties;

(α) $I(x)$ is closed and bounded.

(β) The origin is an interior point of $I(x)$.

(γ) $I(x)$ is strictly convex and has a continuously varying tangent hyperplane at each of its boundary points. In particular, $I(x)$ is an ellipsoid in the case of a Riemannian geometry and the unit sphere in the case of the Euclidean geometry.

Conversely, if we are given a space X and a set valued function $I(x)$ defined over X and satisfying the conditions (α), (β) and (γ) then there is a unique function $F(x,y)$ such that

$$I(x) = \{y : F(x,y) \leqq 1\} .$$

Moreover, this function $F(x,y)$ satisfies the conditions (i), (ii) and (iii) of the definition of a Finsler geometry.

From that follows that a Finsler geometry can be equivalently represented by the function $F(x,y)$ or by the indicatrix $I(x)$. We see immediately that the geodesics of the Finsler geometry characterized by the indicatrix $I(x)$ are the rays of the abstract autonomous propagation space characterized by the wavelets $w(x) = I(x)$.

More generally, we could start with a space X, an arbitrary set valued function $I(x)$ defined on X, and study the geometry induced on X by the indicatrix $I(x)$, i.e., the geometry for which the distance ds between two neighboring points x and $x + dx$ is the smallest nonnegative real number α such that

$$\frac{dx}{\alpha} \in I(x) .$$

In that case the rays of the abstract autonomous propagation space characterized by the wavelets $w(x)$ are the geodesics of the geometry induced on X by the indicatrix $I(x) = w(x)$.

We remark that the class of wavelets obtained by the definition

$$w(x) = \{f(x,u) : u \in \Omega\}$$

in the case of an abstract autonomous propagation space is much larger than the class of indicatrices defined by

$$I(x) = \{y : F(x,y) \leqq 1\}$$

in the case of a Finsler geometry.

For instance, we could obtain wavelets for which the origin is no more an interior point, which are not closed, which are not strictly convex or even with a lower dimension than the space itself. The geometry obtained by taking these wavelets as indicatrices can have some surprising properties: between two different points arbitrarily close to each other with respect to the Euclidean norm, we could have more than one geodesics or even no geodesics at all. Hence the geometry induced on the space X by the wavelets $w(x)$ is much more general than any Finsler geometry defined on the same space X.

A GUIDE TO THE READER

In Section 1 we give a precise statement of the fundamental problem of optimal control mentioned in the introduction. In Sections 2 and 3 we introduce some new concepts, perform some transformations, prove a few propositions in order to be able to state precisely the series of theorems given in Section 4. In this Section 4 are assembled all the results of this work: the necessary condition for the optimal control of a nonlinear dynamical system. In Sections 5 to 9 we establish some intermediary results on which are based the proofs of the theorems of Section 4. These proofs are given in Section 10.

Although this work is entirely devoted to the theory of the general nonlinear dynamical system defined in Section 1, we shall make frequent references to the theory of certain linear systems introduced at the end of Section 2. We do it for the following reasons:

(i) Many concepts and results which are necessary to the study of nonlinear systems but which are elaborate and difficult when dealing with these nonlinear systems become particularly clear when they are applied to the study of linear systems.

(ii) The methods developed here for nonlinear systems furnish a very simple theory for the linear systems.

SECTION 1

Statement of the Problem

In this section we shall give a precise formulation to the fundamental problem of optimal control described in the introduction.

We assume that we are given the following elements:

(i) a point

$$A = (x_a, t_a) = (x_a^1, x_a^2, \cdots, x_a^n, t_a) \in X \times T \tag{1.1}$$

where X, called the state space, is the Euclidean n-dimensional space with elements $x = (x^1, \cdots, x^n)$ and where T is the real line with elements t. T is usually interpreted as the time axis. The space $X \times T$ is called the event space.

(ii) a line B in $X \times T$, parallel to the x^n-axis and determined by its projections x_b^i, $i = 1, \cdots, n-1$ and t_b on the other axis. More precisely, B is the set

$$(1.2) \qquad \{(x,t) : x^i = x_b^i \quad \text{for } i = 1, \cdots, n-1, \quad x^n \in R, \quad t = t_b\}$$

where R is the real line.

(iii) a set

$$(1.3) \qquad \Omega \subset U$$

where U, called the control space, is the Euclidean r-dimensional space with elements $u = (u^1, \cdots, u^r)$.

(iv) an n-dimensional vector valued function

$$(1.4) \qquad f(x,u,t) = (f^1(x,u,t), f^2(x,u,t), \cdots, f^n(x,u,t)) .$$

(v) the class F of all bounded measurable [r-dimensional] vector valued functions $[u] = \{(u(t),t) : t \in [t_a, t_b]\}$ satisfying the condition

$$(1.5) \qquad u(t) \in \Omega \quad \text{for all} \quad t \in [t_a, t_b] .$$

Given all these data we define E as the set of all [n-dimensional] vector valued functions $[x] = \{(x(t),t) : t \in [t_a, t_b]\}$ such that

(1.6) (1) $[x]$ is continuous and a.e. differentiable

(1.7) (2) $x(t_a) = x_a$

(3) there exists a $[v] \in F$ with the property

$$(1.8) \qquad \dot{x}(t) = f(x(t), v(t), t) \quad \text{a.e.} \quad t \in [t_a, t_b] .$$

(4) there exists an $\varepsilon > 0$ such that $f(x,u,t)$ and $f_x(x,u,t)$ are defined, measurable with respect to u anf t, uniformly equicontinuous with respect to x, and uniformly bounded for all

$$(1.9) \qquad (x,t,u) \in N([x],\varepsilon) \times \Omega^*$$

where

(1.10) $$N([x],\varepsilon)=\{(\bar{x},\bar{t}):|\bar{x}-x(t)|^2+|\bar{t}-t|^2\leqq\varepsilon^2,\quad t\in[t_a,t_b]\}$$

and where Ω^* is any bounded subset of Ω.

The fundamental problem of optimal control is then to find an element $[x]$ in E such that

(1.11) (α) $(x(t_b),t_b)\in B$

(β) for any $[\bar{x}]$ in E such that

(1.12) $$(\bar{x}(t_b),t_b)\in B$$

shall hold the relation

(1.13) $$\bar{x}^n(t_b)\leqq x^n(t_b)\ .$$

The problem as formulated above does not yet exhibit the characteristic structure of an optimal control problem: we have still to introduce the strategy space. One could be tempted to consider as strategy space for this problem the totality of the function space F introduced earlier. This can be done indeed but at great costs:we must make strong assumptions on the function (x,u,t) in order to insure for every $[u]$ in F the existence and uniqueness of the solution to the differential system

(1.14) $$\dot{x}=f(x,u(t),t)\quad\text{a.e.}\quad t\in[t_a,t_b]$$

with the initial conditions

(1.15) $$x(t_a)=x_a\ .$$

In this work we prefer to avoid making any further assumption on the function $f(x,u,t)$. Instead we shall restrict the strategy space to an appropriate subset F^* of F defined as the set of all $[v]\in F$ for which there exists an $[x]\in E$ with the property (1.8).

According to the following proposition, the set F^* has now all the properties of a strategy space in the sense of the introduction.

Proposition 1.1 *If* $[v] \in F^*$ *then there is a unique* $[x] \in E$ *satisfying the property* (1.8).

Proof of Proposition 1.1. By definition there exists at least one such $[x] \in E$; let us assume that there is another such $[\bar{x}] \in E$ with $[\bar{x}] \neq [x]$. Let $\tau = \sup_{t \in [t_a, t_b]} \{t : \bar{x}(\theta) = x(\theta) \text{ for } \theta \leqq t\}$ such a τ exists since $x(t_a) = \bar{x}(t_a) = x_a$. Moreover, $\bar{x}(\tau) = x(\tau)$ since $[\bar{x}]$ and $[x]$ are continuous and $\tau \neq t_b$ since $[\bar{x}] \neq [x]$. Let ε and $\bar{\varepsilon}$ be two positive real numbers corresponding to $[x]$ and $[\bar{x}]$ in the definition of E. Let $\varepsilon^* = \min\{\varepsilon, \bar{\varepsilon}\}$. By assumption the functions $f(x, v(t), t)$ and $f_x(x, v(t), t)$ are then bounded and measurable with respect to t in the ε^* neighborhood of the point $(x(\tau), \tau)$. Hence, from the theory of ordinary differential equations, there is a $\delta > 0$ such that $x(t) = \bar{x}(t)$ for $t \in [\tau, \tau + \delta]$. This contradicts the definition of τ and concludes the proof of Proposition 1.1.

Proposition 1.1 allows us to make the following definition: If $[v] \in F^*$, let $[x([v])] = \{(x(t; [v]), t) : t \in [t_a, t_b]\}$ be the unique element in E with the property (1.8).

We are now in a proper position to state the problem in terms of the strategy space F^* as follows:

Find an element $[v] \in F^*$ such that

(1.16) $\quad (\alpha)\ (x(t_b; [v]), t_b) \in B$

(β) for any $[w] \in F^*$ with the property

(1.17) $$(x(t_b; [w]), t_b) \in B$$

shall hold the relation

(1.18) $$x^n(t_b; [w]) \leqq x^n(t_b; [v]) .$$

The function $[v]$ satisfying the conditions (α) and (β) shall be called an optimal control function and the corresponding function $[x([v])]$ shall be called an optimal trajectory.

Remarks on the structure of the function $f(x, u, t)$

We should mention here two differences between the statement of the fundamental problem in optimal control given above and the fundamental problem treated by Pontryagin and his associates.

In our formulation we allow the function $f(x,u,t)$ to be dependent of the variable x^n to be maximized at time t_b. We are allowing this dependence for practical and esthetical reasons: to make the assumption that $f(x,u,t)$ is independent of x^n would lead to very little simplification of the subsequent developments but would nevertheless break the symmetry among the state variables. Moreover, many practical problems show a natural dependence of the differential equations on the variable to be maximized: in the classical problem of the maximization of the payload of a rocket, the evolution of the rocket depends on its mass at every intermediate instant of time.

Also, in contradistinction to Pontryagin's formulation, we do not require the differential system to be time independent.

However, by an appropriate introduction of new artificial variables we may transform our problem into the problem treated by Pontryagin and his associates. But the new problem obtained by this introduction of artificial variables is always degenerate in the following sense: any element $[u]$ of F^* satisfies Pontryagin's Maximum Principle. In order to obtain non-trivial necessary conditions for an element $[u]$ of F^* to be optimal we must, in such a case, use the so-called "transversality conditions". The methods developed in this work lead to a very clear geometric interpretation of this type of degeneracy which will be discussed in detail in Section 4.

Moreover, because of the assumption on the continuous dependence of the system of differential equations on the state variables, this transformation cannot be done if the time dependence of the differential equations is not continuous: this is to be contrasted with our very weak assumption on the time dependence of the differential equations: we require only measurability with respect to time.

We want to make another remark closely related to the introduction of new artificial variables and on the necessity to consider transversality conditions in such cases. The statement of the problem given here is made up of two parts: we define a *control system with initial conditions* by (1.1), (1.3), 1.4) and (1.5) and for this control system with initial conditions we define an *optimal control problem* by (1.2), (1.11), (1.12) and (1.13).

Most of the developments made in the followoing sections depend only on the *control system with initial conditions* but not on the particular *optimal control problem*. In fact, as we shall see in Section 4, our results are directly applicable to a large class of optimal control problems: this will allow us to dispense with the formal transformations required in order to apply the Maximum Principle of Pontryagin and with the consideration of transversality conditions which, after such transformations, are strictly needed if we want to obtain a nontrivial set of necessary conditions for an optimal solution.

SECTION 2

Comoving Coordinate Space along a Given Trajectory

In the introduction we wrote: "To verify that an element $[v]$ of F^* is optimal we adopt the point of view of an observer riding along the trajectory $[x([v])]$ and making his observations in a moving frame of coordinates attached to the wavefront." In this section we intend to carry out this scheme: for an arbitrary element $[v]$ of F^* we shall define a moving frame of coordinates $Y([v])$ by an appropriate transformation from $X \times [t_a, t_b]$ to $Y([v]) \times [t_a, t_b]$ and for each $[u]$ in F^* we shall study the trajectory $[y([u],[v])] = \{(y(t;[u],[v]),t) : t \in [t_a, t_b]\}$ which is the transformation in $Y([v]) \times [t_a, t_b]$ of the trajectory $[x([u])]$ in $X \times [t_a, t_b]$. The space $Y([v]) \times [t_a, t_b]$ is called the comoving coordinate space along the trajectory $[x([v])]$.

We introduce the space $Y([v]) \times [t_a, t_b]$ and the trajectories $[y([u],[v])]$ in that space for the following reasons: in the space $Y([v]) \times [t_a, t_b]$ there is a very natural way to associate with every trajectory $[y([u],[v])]$ an approximate trajectory

$$[\overset{+}{y}([u],[v])] = \{(\overset{+}{y}(t;[u],[v]),t) : t \in [t_a, t_b]\}$$

having a particularly simple structure. According to our previous analogy the trajectory $[\overset{+}{y}([u],[v])]$ could be considered as the most reasonable approximation of the trajectory $[y([u],[v])]$ made by the observer riding on the trajectory $[x([v])]$ and knowing the function $f(x,u,t)$ for only those values of x and t which are in the neighborhood of his own trajectory.

The consideration of these various types of trajectories will be of great help to derive the necessary conditions for the optimality of the trajectory $[x([v])]$; the guiding idea of this derivation, given in Section 4, could be summarized as follows: assuming that $[\overset{+}{y}([u],[v])]$ is the exact expression of $[y([u],[v])]$ we derive easily a set of necessary conditions for the optimality of $[x([v])]$, then we prove that our conclusions are still valid when $[\overset{+}{y}([u],[v])]$ is a close enough approximation of $[y([u],[v])]$.

After these commentaries we shall now proceed with the precise definitions of the entities mentioned above.

For any $[v]\in F^*$ we define an $n\times n$ matrix $D(t;[v])$ as follows:

$$D(t;[v]) = \frac{\partial f(x,v(t),t)}{\partial x}\bigg|_{x=x(t;[v])} \qquad t\in[t_a,t_b]\ . \tag{2.1}$$

More precisely, $D(t;[v])$ is the $n\times n$ matrix with elements $D^i_j(t;[v])$; $i,j=1,2,\cdots,n$; defined by

$$D^i_j(t;[v]) = \frac{\partial f^i(x,v(t),t)}{\partial x^j}\bigg|_{x=x(t;[v])} \qquad t\in[t_a,t_b]\ . \tag{2.2}$$

It is much more convenient to use these relations in the form (2.1) than in the form (2.2). Such a convention and its obvious generalizations will be used throughout this work.

From our assumptions we know that $D(t;[v])$ is bounded and measurable over $[t_a,t_b]$.

Let $G(t;[v])$ be an $n\times n$ matrix, continuous with respect to t, defined over $[t_a,t_b]$, satisfying the matrix differential equation

$$\dot{G}(t;[v]) = -G(t;[v])D(t;[v]) \text{ a.e. } t\in[t_a,t_b] \tag{2.3}$$

and such that

$$G(t_b;[v]) = I \tag{2.4}$$

where I is the identity $n \times n$ matrix.

Proposition 2.1. *The matrix valued function $G(t;[v])$ exists, is unique and bounded over $[t_a,t_b]$.*

Proposition (2.1) is an immediate consequence of the properties of $D(t;[v])$ and of the theory of ordinary differential equations.

We shall now introduce a Euclidean n-dimensional space $Y([v])$ with elements $y = (y^1, \cdots, y^n)$ by the mapping

$$\Phi([v]) : X \times [t_a,t_b] \to Y([v]) \times [t_a,t_b] \tag{2.5}$$

where

$$(y,t) = \Phi(x,t;[v]) \tag{2.6}$$

is defined by

$$y = G(t;[v])(x - x(t;[v])) . \tag{2.7}$$

Under the mapping $\Phi([v])$ the trajectory

$$[x([u])] = \{(x(t;[u]),t) : t \in [t_a,t_b]\} \text{ with } [u] \in F^* \tag{2.8}$$

will be transformed into the trajectory

$$[y([u],[v])] = \{(y(t;[u],[v]),t) : t \in [t_a,t_b]\} \tag{2.9}$$

defined by the relation

$$y(t;[u],[v]) = G(t;[v])(x(t;[u]) - x(t;[v])) \quad \text{for all } t \in [t_a,t_b] . \tag{2.10}$$

Proposition 2.2. *For every $[v]$ and $[u]$ in F^* the function $[y([u],[v])]$ exists, is unique and continuous.*

Proposition (2.2) follows directly from the relation (2.10) since we already know that $G(t;[v])$, $x(t;[u])$ and $x(t;[v])$ exist separately, are unique and continuous over $[t_a, t_b]$ for $[v]$ and $[u] \in F^*$.

Let us now define the approximate trajectory

$$[\overset{+}{y}([u],[v])] = \{(\overset{+}{y}(t;[u],[v]),t) : t \in [t_a, t_b]\} \tag{2.11}$$

by the relation

$$\overset{+}{y}(t;[u],[v]) = \int_{t_a}^{t} G(\tau;[v])(f(x(\tau;[v]),u(\tau),\tau) - f(x(\tau;[v]),v(\tau),\tau))d\tau \tag{2.12}$$

$$\text{for all } t \in [t_a, t_b] \ .$$

Proposition 2.3. *For every* $[v] \in F^*$ *and every* $[u] \in F$ *the function* $[\overset{+}{y}([u],[v])]$ *exists, is unique and continuous.*

Proposition (2.3) follows directly from the definition (2.12) since we already know that $G(t;[v])$, $f(x(t,[v]),u(t),t)$ and $f(x(t;[v]),v(t),t)$ are measurable and bounded over $[t_a, t_b]$ for all $[v] \in F^*$ and all $[u] \in F$.

As we mentioned at the beginning of this section, the trajectory $[\overset{+}{y}([u],[v])]$ can be considered as an approximation of the trajectory $[y([u],[v])]$. In Section 7 we shall define precisely in what sense the word "approximation" should be understood and what conclusions we may draw from it. We already see at this point that

$$\overset{+}{y}(t;[v],[v]) \equiv y(t;[v],[v]) \equiv 0 \qquad \text{all } t \in [t_a, t_b] \ . \tag{2.13}$$

The proximity of $[y([u],[v])]$ and $[\overset{+}{y}([u],[v])]$ is particularly apparent in the case of a particular class of linear systems, defined in the following paragraph, since we then have

$$[y([u],[v])] = [\overset{+}{y}([u],[v])] \qquad \text{for all } [u] \text{ and } [v] \in F \ . \tag{2.14}$$

Application to a linear system

We assume here that $f(x,u,t)$ has the form

$$f(x,u,t) = A(u,t)x + \phi(u,t) \tag{2.15}$$

or the form

$$f(x,u,t) = A(t)x + \phi(u,t) \ . \tag{2.16}$$

When the function $f(x,u,t)$ has the form (2.15) we shall speak of a linear system and when it has the form (2.16) we shall speak of a linear* system. From the definition it follows that a linear* system is a particular type of linear system. In the relations (2.15) and (2.16) the expressions $A(u,t)$ and $A(t)$ are $n \times n$ matrices, and $\phi(u,t)$ is an n-dimensional vector. We assume that $A(u,t)$, $A(t)$ and $\phi(u,t)$ are measurable with respect to their arguments and uniformly bounded over $\Omega^* \times [t_a,t_b]$ for any bounded set Ω^* subset of Ω.

In the case of a linear system we have

$$D(t;[v]) = A(v(t),t) \qquad \text{for all } t \in [t_a,t_b] \tag{2.17}$$

and $G(t;[v])$ is the continuous solution of the matrix differential equation

$$\dot{G}(t;[v]) = -\,G(t;[v])\,A(v(t),t) \qquad \text{a.e. } t \in [t_a,t_b] \tag{2.18}$$

with the terminal condition

$$G(t_b;[v]) = I \ . \tag{2.19}$$

Such a solution is usually written under the symbolic form

$$G(t;[v]) = \exp\left\{ \int_t^{t_b} A(v(\tau),\tau)d\tau \right\} \ . \tag{2.20}$$

We then have

$$y(t;[u],[v]) = G(t;[v])(x(t;[u]) - x(t;[v])) \qquad \text{for all } t \in [t_a,t_b] \tag{2.21}$$

and

$$\overset{+}{y}(t;[u],[v]) = \int_{t_a}^{t} G(\tau;[v])(A(u(\tau),\tau)x(\tau;[v]) + \phi(u(\tau),\tau) - A(v(\tau),\tau)x(\tau;[v]) - \phi(v(\tau),\tau))d\tau \qquad \text{for all } t \in [t_a,t_b] \ . \tag{2.22}$$

We shall now compare the two trajectories $[y([u],[v])]$ and $[\overset{+}{y}([u],[v])]$. We know already that

$$(2.23) \qquad y(t_a;[u],[v]) = \overset{+}{y}(t_a;[u],[v])$$

since from the relation (2.21) we have

$$(2.24) \qquad \begin{aligned} y(t_a;[u],[v]) &= G(t_a;[v])(x(t_a;[u]) - x(t_a;[v])) \\ &= G(t_a;[v])(x_a - x_a) = 0 \end{aligned}$$

and since from the relation (2.22) we have

$$(2.25) \qquad \overset{+}{y}(t_a;[u],[v]) = 0 .$$

Let us now consider $\overset{+}{y}{}^{\cdot}(t;[u],[v])$ and $\dot{y}(t;[u],[v])$. We have immediately

$$(2.26) \qquad \begin{aligned} \overset{+}{y}{}^{\cdot}(t;[u],[v]) = {} & G(t;[v])(A(u(t),t)x(t;[v]) + \phi(u(t),t) \\ & - A(v(t),t)x(t;[v]) - \phi(v(t),t)) \\ & \qquad\qquad \text{for a.e. } t \in [t_a, t_b] \end{aligned}$$

and

$$(2.27) \qquad \begin{aligned} & \dot{y}(t;[u],[v]) \\ &= (G(t;[v])(x(t;[u]) - x(t;[v])))^{\cdot} \\ &= \dot{G}(t;[v])(x(t;[u]) - x(t;[v])) + G(t;[v])(\dot{x}(t;[u]) - \dot{x}(t;[v])) \\ &= -G(t;[v])A(v(t),t)(x(t;[u]) - x(t;[v])) \\ &\quad + G(t;[v])(A(u(t),t)x(t;[u]) + \phi(u(t),t) - A(v(t),t)x(t;[v]) - \phi(v(t),t)) \\ &= G(t;[v])(A(u(t),t)x(t;[v]) + \phi(u(t),t) - A(v(t),t)x(t;[v]) + \phi(v(t),t)) \\ &\quad + G(t;[v])(A(u(t),t) - A(v(t),t))(x(t;[u]) - x(t;[v])) \\ &= G(t;[v])(A(u(t),t)x(t;[v]) + \phi(u(t),t) - A(v(t),t)x(t;[v]) - \phi(v(t),t)) \\ &\qquad + G(t;[v])(A(u(t),t) - A(v(t),t))G^{-1}(t;[v])\,y(t;[u],[v]) \\ &\qquad\qquad \text{a.e. } t \in [t_a, t_b] . \end{aligned}$$

The relations (2.26) and (2.27) imply

$$(2.28)\qquad \dot{y}(t;[u],[v]) - \overset{+}{\dot{y}}(t;[u],[v]) = G(t;[v])(A(u(t),t) - A(v(t),t))G^{-1}(t;[v])\,y(t;[u],[v]) \qquad \text{a.e. } t \in [t_a, t_b]\ .$$

We see immediately that in the case where

$$(2.29)\qquad A(u,t) = A(v,t) \qquad \text{a.e. } t \in [t_a, t_b]$$

we have

$$(2.30)\qquad \dot{y}(t;[u],[v]) = \overset{+}{\dot{y}}(t;[u],[v]) \qquad \text{a.e. } t \in [t_a, t_b]\ .$$

The relations (2.23) and (2.30) then implies

$$(2.31)\qquad y(t;[u],[v]) = \overset{+}{y}(t;[u],[v]) \qquad \text{all } t \in [t_a, t_b]\ .$$

We can then state the following result:

Proposition 2.4. *For a linear* system*

$$(2.32)\qquad [y([u],[v])] = [\overset{+}{y}([u],[v])] \qquad \textit{for all } [u] \textit{ and } [v] \in F.$$

On the other hand, we see that, even for a linear system, $[\overset{+}{y}([u],[v])]$ is in general different from $[y([u],[v])]$ and only an approximation of $[y([u],[v])]$ in a sense which will be defined in Section 7.

The identity of $[y([u],[v])]$ and $[\overset{+}{y}([u],[v])]$ in the case of a linear* system is particularly helpful to obtain quickly, for a linear* system, the necessary conditions stated in Section 4 since Theorem III, the most difficult theorem of Section 4 is, as we shall see, trivially true in that case.

SECTION 3

Set of Reachable Events

In this section we shall introduce the important concept of the set of reachable events. Given a control system, we shall say that a point (x_β, t_β) in

$X \times T$ is reachable from the point (x_α, t_α) in $X \times T$ if $t_\beta \geqq t_\alpha$ and if there exists a control function $[u]$ in the strategy space F^* such that the solution of the system

$$\begin{cases} \dot{x} = f(x, u(t), t) & \text{a.e. } t \in [t_\alpha, t_\beta] \quad (3.1) \\ x(t_\alpha) = x_\alpha & \quad (3.2) \end{cases}$$

satisfies the terminal condition

$$x(t_\beta) = x_\beta \tag{3.3}$$

In other words, following the terminology used in the introduction, we say that (x_β, t_β) is reachable from (x_α, t_α) if and only if

$$(t_\alpha, x_\alpha) R([u])(t_\beta, x_\beta) \qquad \text{for some } [u] \text{ in } F^* . \tag{3.4}$$

We shall consider specially the set H, intersection by the hyperplane $t = t_b$ of the set of all events reachable from the initial event A by the trajectories $[x([u])]$ with $[u] \in F^*$. In the subsequent analytical developments we shall also use the set $H([v])$ which is the intersection by the hyperplane $t = t_b$ of the set of all events reachable from the initial event $y = 0$ by the trajectories $[y([u], [v])]$ with $[u] \in F^*$. Similarly, we shall consider the set $\overset{+}{H}([v])$ which is the intersection by the hyperplane $t = t_b$ of the set of all events reachable from the initial event $y = 0$ by the approximate trajectories $[\overset{+}{y}([u], [v])]$ with $[u] \in F$. According to our previous analogy, the set $\overset{+}{H}([v])$ may be considered as the most reasonable approximation of the set $H([v])$ made by an observer riding along the trajectory $[x([v])]$ but knowing the function $f(x, u, t)$ for only those values of x and t which are in the neighborhood of his own trajectory.

Formally, we then have

$$\begin{aligned} H &= \{x(t_b; [u]) : [u] \in F^*\} \\ H([v]) &= \{y(t_b; [u], [v]) : [u] \in F^*\} \qquad \text{for any } [v] \in F^* \\ \overset{+}{H}([v]) &= \{\overset{+}{y}(t_b; [u], [v]) : [u] \in F\} \qquad \text{for any } [v] \in F^* . \end{aligned}$$

We immediately have the relation

$$H([v]) = \{\alpha - x(t_b; [v]) : \alpha \in H\} .$$

The study of these sets and particularly of their boundaries will be made in the next section.

SECTION 4

Necessary Condition for the Optimal Control of a Dynamical System

In this section we give a series of seven theorems. These theorems summarize the whole content of this work. In the remaining sections we shall be concerned, directly or indirectly, with the proof of these theorems. More precisely, we shall establish in Sections 5 to 9 some preliminary results which will be used in Section 10 in the explicit proof of the seven theorems.

Theorem I. *If an element* $[v]$ *of* F^* *is optimal then the point* $x = x(t_b;[v])$ *is a boundary point of the set* H.

Theorem II. *If the point* $x = x(t_b;[v])$ *is a boundary point of the set* H *then the point* $y = 0$ *is a boundary point of the set* $H([v])$.

Theorem III. *If the point* $y = 0$ *is a boundary point of the set* $H([v])$ *then* $y = 0$ *is a boundary point of the set* $\overset{+}{H}([v])$.

Theorem IV. *If the point* $y=0$ *is a boundary point of the set* $\overset{+}{H}([v])$ *then there exists a nonzero constant vector* $\pi([v])$ *such that for all* $[u]$ *in* F:

$$\langle \pi([v]) \,|\, G(t;[v])(f(x(t;[v]),u(t),t) - f(x(t;[v]),v(t),t)) \rangle \leqq 0 \qquad \text{a.e. } t \in [t_a, t_b]\,. \tag{4.1}$$

By $\langle \alpha \,|\, \beta \rangle$ we mean the scalar product of α and β.

Theorem V. *If there is a nonzero constant vector* $\pi([v])$ *such that the condition* (4.1) *is satisfied for all* $[u]$ *in* F *then there is a vector* $p(t;[v])$ *continuous and nonidentically zero over* $[t_a, t_b]$ *such that*:

(4.2) (i) $p(t;[v]) = G^T(t;[v])\,\pi([v])$ al $t \in [t_a, t_b]$

(ii) *for all* $[u]$ *in* F

(4.3) $$\langle p(t;[v]) \,|\, (f(x(t;[v]),u(t),t) - f(x(t;[v]),v(t),t)) \rangle \leqq 0$$

a.e. $t \in [t_a, t_b]$

(4.4) (iii) $\dot{p}(t;[v]) = -D^T(t;[v])p(t;[v])$ a.e. $t \in [t_a, t_b]$.

The superscript T indicates the transposition of a matrix.

Theorem VI. *If the point $x = x(t_b;[v])$ is a boundary point of the set H then there exists a vector $p(t;[v])$, continuous and nonidentically zero on $[t_a, t_b]$, such that the conditions* (4.3) *and* (4.4) *are satisfied.*

Theorem VII. *If an element $[v]$ of F^* is optimal then there exists a vector $p(t;[v])$, continuous and nonidentically zero on $[t_a, t_b]$, such that the conditions* (4.3) *and* (4.4) *are satisfied.*

As we mentioned earlier, the demonstrations of these theroms are given in Section 10. The demonstrations of Theorems I, II, V, VI and VII are almost immediate. The proofs of Theorems III and IV are based on the results established in Sections 5 to 9.

If we define

(4.5) $$H(x,u,t,p) = \langle p \,|\, f(x,u,t) \rangle$$

then Theorem VII could be equivalently formulated as follows:

If an element $[v]$ of F^* is optimal then there exists a vector $p(t;[v])$, continuous and nonidentically zero on $[t_a, t_b]$ such that

(4.6) (i) $$\dot{x}(t;[v]) = \frac{\partial H(x(t;[v]),v(t),t,p)}{\partial p}\bigg|_{p=p(t;[v])} \quad \text{a.e. } t \in [t_a, t_b]$$

i.e.,

(4.7) $$\dot{x}(t,[v]) = f(x(t;[v]),v(t),t) \quad \text{a.e. } t \in [t_a, t_b]$$

(4.8) (ii) $$\dot{p}(t;[v]) = -\frac{\partial H(x,v(t),t,p(t;[v]))}{\partial x}\bigg|_{x=x(t;[v])} \quad \text{a.e. } t \in [t_a, t_b]$$

$$(4.9) \qquad \text{(iii)}\ H(x(t;[v]), v(t), t, p(t;[v])) \geqq H(x(t;[v]), u(t), t, p(t;[v]))$$

$$\text{for all } [u] \text{ in } F \text{ and a.e. } t \in [t_a, t_b] \ .$$

This equivalent formulation of Theorem VII is the well-known Maximum Principle of Pontryagin.

Let us make some comments on the logical structure of the series of theorems given earlier.

In Theorem I we associate two different notions: the concept of optimality for the particular optimal control problem under consideration and a topological property of the set H, which set depends only on the given control system with initial conditions but not on any particular optimal control problem.

In Theorems II to V we give a series of implications concerning certain properties of the sets H, $H([v])$ and $\overset{+}{H}([v])$.

In Threoem VI we give the combined result of all the implications contained in Theorems II to V.

In Theorem VII we use Theorem I as an intermediary in order to obtain from the topological results of Theorem VI the necessary condition for an optimal solution of the particular optimal control problem under consideration.

Theorem VI is the most important result in the theory of control systems. This theorem, we said earlier, depends only on the given control system with initial conditions but not on any particular optimal control problem. Hence, when a particular optimal control problem is given, we need only to verify that Theorem I is valid in order to derive from Theorem VI the appropriate necessary conditions for an optimal control. The verification of Theorem I is particularly simple in the case of the fundamental optimal control problem considered in this work but could also be easily done for a large class of different optimal control problems.

In contradistinction the method of Pontryagin and his associates is the following: when confronted with a particular optimal control problem they introduce new artificial variables which transform the control system itself, and hence the set H in such a way that, for the new control system, the particular optimal control problem has the form of the fundamental optimal control problem. Unfortunately, the new set H, obtained after introduction of these artificial variables has, with respect to the new event space, a lower dimensionality than before. In this case, Pontryagin's Maximum Principle, in the form given above, can be trivially satisfied for any control, not necessarily optimal, as we shall show in a later paragraph. For such case Pontryagin and his associates have given a stronger form of the Maximum Principle, including some auxiliary conditions similar to the transversality conditions in calculus of variations.

Finally, let us underline how closely these theorems correspond to the intuitive procedure stated in the introduction: "To verify that an element $[v]$ of F^* is optimal we adopt the point of view of an observer riding along the trajectory $[x([v])]$ and making his observations in a moving frame of coordinates attached to the wavefront. For such an observer all the missed opportunities, i.e., the directions he could have followed but did not, are leading to points on one side of a hyperplane passing through the origin."

By Theorem II we identify the moving frame of coordinates attached to the wavefront with the comoving coordinate system $Y([v]) \times [t_a, t_b]$ along the trajectory $[x([v])]$.

By Theorem III we show that for our purposes the set $\overset{+}{H}([v])$ is as good as the set $H([v])$. In our analogy the set $\overset{+}{H}([v])$ is the most reasonable approximation of the set $H([v])$ made by the observer riding along the trajectory $[x([v])]$ but knowing the function $f(x,u,t)$ for only those values of x and t which are in the neighborhood of his own trajectory. In other words, Theorem III states that the most reasonable approximation made by the moving observer is good enough as far as the derivation of necessary conditions is concerned.

In Theorem IV we identify the vector $\pi([v])$ with the normal to the hyperplane passing through the origin and such that all missed opportunities are directions leading to points located on one of its sides only.

Theorem V describes the same property as Theorem IV but from the point of view of an observer fixed in the space $X \times T$ instead of the moving observer considered earlier.

Remarks on the dimensionality of the set H

By construction the set H is a subset of the n-dimensional Euclidean space X. In Theorems II to VI we have derived some properties of the elements

$[u] \in F^*$ for which $x(t_b;[u])$ is a boundary point of the set H. In other words, we have given some necessary conditions on $[u]$ in order that $x(t;[u])$ be a boundary point of the set H. If the dimension of the set H is less than n then all the previous results become trivial since for any $[u] \in F^*$ the point $x(t_b;[u])$ will be a boundary point of the set H. This happens, for instance, when the set H is a subset of a sufficiently smooth $n-1$ dimensional manifold H^* nontangent to the line B' projection on X of the line B in $X \times T$. In such a case the set $H \cap B'$ has only isolated points and therefore all $[u] \in F^*$ with $(x(t_b;[u]), t_b) \in B$ are locally optimal. This explains why the necessary conditions for an optimal solution can be trivially satisfied in such a case for any control, not necessarily optimal. This example shows the need of stronger conditions. This need is partially satisfied by the consideration of the so-called "transversality conditions."

Application to linear systems.

We saw at the end of Section 2 that, for the linear* system (2.16) we have

(4.10) $$[\overset{+}{y}([u],[v])] = [y([u],[v])] \text{ for all } [u] \text{ and } [v] \text{ in } F \ .$$

In particular, this implies

(4.11) $$\overset{+}{y}(t_b;[u],[v]) = y(t_b;[u],[v]) \text{ for all } [u] \in F$$

i.e.,

(4.12) $$H([v]) = \overset{+}{H}([v]) \ .$$

The relation (4.12) simplifies greatly the derivation of the necessary condition for the optimal control of a linear* system. Indeed, Theorem III is trivially true for a linear* system and since none of the other theorems are particularly difficult, as we shall see in Section 10, we may now consider the entire theory for the optimal control of a linear* system as very simple.

In the case of the linear system (2.15), the relation (4.8) takes the simple form:

(4.13) $$\dot{p}(t;[v]) = - A^T(v(t),t)\, p(t;[v]) \ .$$

SECTION 5

Norms for the Space of Control Functions and for the Spaces of Trajectories

In this section we define various norms for the space F of control functions and for the spaces of trajectories. These norms will be extensively used in the remaining sections of this work.

Let us consider an arbitrary collection G of functions from $[t_a, t_b]$ to a Euclidean space. An element $\{(z(t), t) : t \in [t_a, t_b]\}$ in that collection will be denoted by $[z]$.

Let us define

(5.1) $$d([z],[\bar{z}]) = \operatorname*{ess\,sup}_{t \in [t_a, t_b]} |z(t) - \bar{z}(t)|$$

and

(5.2) $$\sigma([z],[\bar{z}]) = \mu(\{t : z(t) \neq \bar{z}(t) \text{ and } t \in [t_a, t_b]\})$$

for every $[z]$ and $[\bar{z}]$ in G.

By the symbol "ess sup" we mean the essential supremum, i.e.,

(5.3) $$\operatorname*{ess\,sup}_{t \in [t_a, t_b]} |z(t) - \bar{z}(t)| = \inf_{\alpha \in \mathscr{B}_0} \sup_{t \in \alpha} |z(t) - \bar{z}(t)|$$

where

(5.4) $$\mathscr{B}_0 = \{[t_a, t_b] \sim B : \mu(B) = 0\} .$$

It is easy to prove that $d(.,.)$ and $\sigma(.,.)$ are norms for a space of continuous functions and semi- norms for a space of measurable functions.

In this work we shall use $d(.,.)$ over the spaces of trajectories, continuous by definition, and we shall use $d(.,.)$ and $\sigma(.,.)$ over F, the space of control functions, measurable by definition.

If we define the equivalence relation $\doteq$ on F by

(5.5) $$[u] \doteq [v] \text{ if and only if } u(t) = v(t) \text{ for a.e. } t \in [t_a, t_b]$$

then $d(.,.)$ and $\sigma(.,.)$ are norms for the quotient space $\boldsymbol{F} = F/\doteq$.

In order to simplify the notations we shall talk of the set F even where we should talk strictly of the set $\bar{F}$ of equivalence classes of F under $\doteq$, and we shall simply write

$$[u] = [v] \tag{56}$$

even where we should write strictly

$$[u] \doteq [v] \ . \tag{5.7}$$

It should be stressed that the two norms $\sigma(.,.)$ and $d(.,.)$ are not equivalents they give rise to two completely different topologies on F.

To simplify the notation $d([z],0)$ shall be written $d([z])$ and similarly $\sigma([z],0)$ shall be written $\sigma([z])$.

A subset $\tilde{F}$ of F such that there exists a k with

$$d([u]) \leqq k \quad \text{for all} \quad [u] \in \tilde{F} \tag{5.8}$$

is called a d-bounded subset of F.

In particular, we shall denote by F_k the set of all elements in F such that $d([u]) \leqq k$, i.e.,

$$F_k = \{[u] : [u] \in F, d([u]) \leqq k\} \ . \tag{5.9}$$

Similarly, F_k^* will be the set

$$F_k^* = \{[u] : [u] \in F^*, d([u]) \leqq k\} \ . \tag{5.10}$$

SECTION 6

Curvilinear Coordinate Space along a Given Trajectory

In Section 2 we have associated to every element $[v]$ of the function space F^* a space $Y([v]) \times [t_a, t_b]$ called the comoving coordinate space along the trajectory $[x([v])]$. In this section we shall associate to every element $[v]$ of the function space F^* another space denoted by $Z([v]) \times [t_a, t_b]$ and called the curvilinear coordinate space along the trajectory $[x([v])]$. The consideration of the spaces $Z([v]) \times [t_a, t_b]$ is a very convenient tool in the study of the existence and uniform convergence of the trajectories corresponding to control functions in a d-bounded subset of F in terms of the norm σ defined in Section 5 over the space F of control functions.

In the case of the linear system introduced in the last paragraph of Section 2, we shall prove that $Z([v]) = Y([v])$ and we shall show that for a linear system the results given in this section take a much simpler form. Moreover, if the system is a linear* system we proved already in Section 2 that we have the relation $[y([u],[v])] = \overset{+}{y}([u],[v])]$. This will enable us to show that for a linear* system the results given in this section could be easily derived from the classical theory of nonhomogeneous linear differential equations. In the last paragraph of this section we shall consider briefly the case of linear and linear* systems.

If $[w] \in F^*$ and $(\tilde{x}, \tilde{t}) \in X \times [t_a, t_b]$ let

$$(6.1) \qquad [x([w], \tilde{x}, \tilde{t})] = \{(x(t; [w], \tilde{x}, \tilde{t}), t) : t \in [t_a, t_b]\}$$

be a continuous vector-valued solution of the differential equation

$$(6.2) \qquad \dot{x}(t; [w], \tilde{x}, \tilde{t}) = f(x(t; [w], \tilde{x}, \tilde{t}), w(t), t) \text{ a.e. } t \in [t_a, t_b]$$

satisfying the initial condition

$$(6.3) \qquad x(\tilde{t}; [w], \tilde{x}, \tilde{t}) = x(\tilde{t}; [w]) + \tilde{x} \ .$$

For every $[w] \in F^*$ we shall now introduce a Euclidean n-dimensional space $Z([w])$ with elements $z = (z^1, \cdots, z^n)$ by the mapping

$$(6.4) \qquad \Psi([w]) : X \times [t_a, t_b] \to Z([w]) \times [t_a, t_b]$$

for which

$$(6.5) \qquad (z, t) = \Psi(x, t; [w])$$

is determined by the relation

$$(6.6) \qquad z = x(t_b; [w], x - x(t; [w]), t) - x(t_b; [w])$$

or equivalently the inverse mapping

(6.7) $$\Psi^{-1}([w]) : Z([w]) \times [t_a, t_b] \to X \times [t_a, t_b]$$

for which

(6.8) $$(x,t) = \Psi^{-1}(z,t;[w])$$

is determined by the relation

(6.9) $$x = x(t;[w],z,t_b)\,.$$

In other words, the mapping $\Psi([w])$ associates to an element (x,t) of $X \times [t_a, t_b]$ an element (z,t) in $Z([w]) \times [t_a, t_b]$ determined as follows: z is the difference of the projections on the terminal hyperplane, i.e., the hyperplane $X \times \{t_b\}$, of the points (x,t) and $(x(t;[w]),t)$ of the hyperplane $X \times \{t\}$ where the projecting lines are the solutions of the differential system with the control function $[w]$. Conversely, the inverse mapping $\Psi^{-1}([w])$ associates to an element (z,t) in $Z([w]) \times [t_a, t_b]$ an element (x,t) in $X \times [t_a, t_b]$ determined as follows: x is the projection on the hyperplane $X \times \{t\}$ of the point $(x(t_b;[w]) + z, t_b)$ of the terminal hyperplane $X \times \{t_b\}$. The projecting lines are as before the solutions of the differential system with the control function $[w]$.

In the following propositions we shall prove some results concerning the existence, uniqueness and boundedness of the mappings $\Psi([w])$ and their inverses $\Psi^{-1}([w])$.

Definition. *If $[v] \in F^*$, k is a positive number and ε is the positive number associated to $[v]$ in the definition of E, let us define $F^*([v],k)$ to be the set*

(6.10) $$\left\{[w] : [w] \in F^*, d([w],[v]) \leqq k, d([x([w])],[x([v])]) \leqq \frac{\varepsilon}{2}\right\}.$$

It should be remarked that we do not know at this point if the set $F^*([v],k)$ contains any other elements besides $[v]$. In Proposition (6.10) we shall exhibit a large class of elements in F which also belong to $F^*([v],k)$ and a fortiori to F^*. The set $F^*([v],k)$ plays a very important role in this work: the necessary condition for the optimality of the element $[v]$ of F^* will be derived from the comparison of $[v]$ with only those elements in F^* which also belong to $F^*([v],k)$.

Proposition 6.1. *If $[w] \in F^*([v],k)$ and ε is the positive number associated to $[v]$ in the definition of E then $\varepsilon/2$ can be used as the positive number associated to $[w]$ in the definition of E.*

Proof of Proposition 6.1. From the definition of E we know that $f(x, u, t)$ and $f_x(x, u, t)$ satisfies certain conditions for all

$$(x,t,u) \in N([x([v])],\varepsilon) \times \Omega^*$$

where Ω^* is any bounded subset of Ω. But $d([x([w])],[x([v])] \leqq \varepsilon/2$ implies $N([x([w])],\varepsilon/2) \subset N([x([v])],\varepsilon)$ hence $f(x,u,t)$ and $f_x(x,u,t)$ will a fortiori satisfy the same conditions for all $(x,t,u) \in N([([w])],\varepsilon/2) \times \Omega^*$. This concludes the proof of Proposition (6.1).

Proposition 6.2. *If $[v] \in F^*$, ε is the positive number associated to $[v]$ in the definition of E and k is an arbitrary positive number, then there exists an $M > 0$ such that for all $|\tilde{x}| \leqq \varepsilon/2M$, all $\tilde{t} \in [t_a, t_b]$ and all $[w] \in F^*([v],k)$ we have*

(6.11) (i) $[x([w],\tilde{x},\tilde{t})]$ *and* $\left[\dfrac{\partial x([w],\tilde{x},\tilde{t})}{\partial \tilde{x}}\right]$ *exist and are unique*

(6.12) (ii) $d([x([w],\tilde{x},\tilde{t})],[x([w])]) \leqq |\tilde{x}|M$.

Proof of Proposition 6.2. From our assumption we know that

(6.13) (i) $[x([w])]$ exists

(6.14) (ii) $f(x,w(t),t)$ and $\dfrac{\partial f(x,w(t),t)}{\partial x}$ exist and are uniformly

bounded for all $(x,t) \in N\left([x([w])],\dfrac{\varepsilon}{2}\right)$.

By definition $[x([w])]$ and $[x([w],\tilde{x},\tilde{t})]$ are solutions of the *same* differential equation

$$(6.15) \qquad \dot{x} = f(x,w(t),t) \qquad \text{a.e. } t \in [t_a,t_b]$$

but corresponding to different initial values. Hence from the theorem on the continuous dependence of the solution of a differential equation, we know that there exists a constant $M([w])$ such that for all $|\tilde{x}| \leqq \frac{\varepsilon}{2M([w])}$ and all $\tilde{t} \in [t_a, t_b]$ the functions $[x([w], \tilde{x}, \tilde{t})]$ and $\left[\frac{\partial x([w], \tilde{x}, \tilde{t})}{\partial \tilde{x}}\right]$ exist and are unique. It remains to show that we can find a constant M such that $M([w]) \leqq M$ for all $[w] \in F^*([v], k)$.

Let $\left|\frac{\partial f(x,u,t)}{\partial x}\right|$ denote the norm of the matrix $\frac{\partial f(x,u,t)}{\partial x}$. This norm is defined as usual by the relation

(6.16) $$\left|\frac{\partial f(x,u,t)}{\partial x}\right| = \sup_{|y| \leqq 1} \left|\frac{\partial f(x,u,t)}{\partial x} y\right| .$$

Let

(6.17) $$R = \sup \left|\frac{\partial f(x,u,t)}{\partial x}\right|$$

over all values of $(x,t) \in N([x([v])], \varepsilon)$ and all values of u with $|u| \leqq d([v]) + k$.

Let

(6.18) $$M = 2e^{R(t_b - t_a)} .$$

We shall now show that this constant M satisfies our requirements.

Let $[\tau_1, \tau_2]$ be the supremum of all closed intervals $[\theta_1, \theta_2] \subset [t_a, t_b]$ such that

(6.19) (i) $x(t; [w], \tilde{x}, \tilde{t})$ exists and is unique on $[\theta_1, \theta_2]$

(6.20) (ii) $|x(t; [w], \tilde{x}, \tilde{t}) - x(t; [w])| \leqq |\tilde{x}| M$ for all $t \in [\theta_1, \theta_2]$.

The previous definition makes sense since the class of closed intervals satisfying the conditions (6.19) and (6.20) is not empty: it contains the closed interval $[\tilde{t}, \tilde{t}]$.

We then have

$$\frac{d}{dt}|x(t;[w],\tilde{x},\tilde{t}) - x(t;[w])| \leqq R\,|x(t;[w],\tilde{x},\tilde{t}) - x(t;[w])| \tag{6.21}$$

for a.e. $t \in [\tau_1, \tau_2]$

and

$$|x(\tilde{t};[w],\tilde{x},\tilde{t}) - x(\tilde{t};[w])| = |\tilde{x}|. \tag{6.22}$$

This implies

$$|x(t;[w],\tilde{x},\tilde{t}) - x(t;[w])| \leqq |\tilde{x}|\,e^{R(t_b - t_a)} = |\tilde{x}|\frac{M}{2}. \tag{6.23}$$

Hence $[\tau_1, \tau_2] = [t_a, t_b]$ because otherwise there would be a closed interval $[\theta_1^*, \theta_2^*]$ with $[\tau_1, \tau_2] \subsetneqq [\theta_1^*, \theta_2^*] \subset [t_a, t_b]$ for which the conditions (6.19) and (6.10) would be satisfied, contradicting the definition of $[\tau_1, \tau_2]$.

The relation (6.23) now becomes

$$d([x([w],\tilde{x},\tilde{t})],[x([w])]) \leqq |\tilde{x}|\,\frac{M}{2}. \tag{6.24}$$

This concludes the proof of Proposition (6.2).

Propposition (6.2) leads immediatelly to the following result:

Proposition 6.3. *If $[v] \in F^*$, ε is the positive number associated to $[v]$ in the definition of E, k is an arbitrary positive number, and M is the positive constant introduced in Proposition (6.2), then for every $[w] \in F^*([v],k)$ the mapping $\Psi([w]): X \times [t_a, t_b] \to Z([w]) \times [t_a, t\ ]$ described by the relation (6.6) is well defined for all (x,t) in $X \times [t_a, t_b]$ such that $|x - x(t;[w])| \leqq \frac{\varepsilon}{2M}$ and the resulting (z,t) in $Z([w]) \times [t_a, t_b]$ is such that $|z| \leqq M|x - x(t;[w])|$. Conversely, the mapping $\Psi^{-1}([w]): Z([w]) \times [t_a, t_b] \to X \times [t_a, t_b]$ described by the relation (6.9) is well defined for all (z,t) in $Z([w]) \times [t_a, t_b]$ such that $|z| \leqq \frac{\varepsilon}{2M}$ and the resulting (x,t) in $X \times [t_a, t_b]$ is such that $|x - x(t;[w])| \leqq M|z|$.*

Convention 1. We shall write $x(t;[w],\tilde{x})$ for $x(t;[w],\tilde{x},t_b)$.

2. In the remaining part of this work $[v]$ will always be an element in F^*, ε will be the positive number associated to $[v]$ in the definition of E, k will be an arbitrary positive number and M will be the positive number introduced in Proposition (6.2).

Definition. *For* $[w] \in F^*([v],k)$ *and* $|z| \leqq \varepsilon/M$ *let*

$$D(t;[w],z) = \left.\frac{\partial f(x,w(t),t)}{\partial x}\right|_{x=x(t;[w],z)} \quad \text{for all } t \in [t_a,t_b] \,. \tag{6.25}$$

Proposition 6.4. *The matrix valued function* $D(t;[w],z)$ *is measurable with respect to* t *and uniformly bounded for all* $t \in [t_a,t_b]$, *all* $[w] \in F^*([v],k)$ *and all* $|z| \leqq \varepsilon/2M$.

Proof of Proposition 6.4. From $[w] \in F^*([v],k)$ and $|z| \leqq \varepsilon/2M$ we know that

$$\text{(i)} \quad (x(t;[w],z),t) \in N([x([v])],\varepsilon) \quad \text{for all } t \in [t_a,t_b] \tag{6.26}$$

$$\text{(ii)} \quad |w(t)| \leqq d([v]) + k \qquad \text{for all } t \in [t_a,t_b] \tag{6.27}$$

hence

$$|D(t;[w],z)| \leqq R \qquad \text{for all } t \in [t_a,t_b] \tag{6.28}$$

where R is the positive constant introduced in the proof of Proposition (6.2). This concludes the proof of Proposition (6.4).

Definition. *For* $[w] \in F^*([v],k)$, $|z| \leqq \varepsilon/2M$ *let* $G(t;[w],z)$ *be the continuous solution of the matrix differential equation*

$$\dot{G}(t;[w],z) = -\,G(t;[w],z)\,D(t;[w],z) \quad \text{a.e. } t \in [t_a,t_b] \tag{6.29}$$

with the terminal condition

$$G(t_b;[w],z) = I \tag{6.30}$$

where I *is the identity matrix.*

Proposition 6.5. *The matrix $G(t;[w],z)$ exists, is uniformly continuous with respect to t and uniformly bounded for all $t \in [t_a, t_b]$, all $[w] \in F^*([v],k)$ and all $|z| \leqq \varepsilon/2M$.*

Proof of Proposition 6.5. Immediate from Proposition (6.4) and the theory of linear differential equations with bounded measurable coefficients. Moreover,

$$|G(t;[w],z)| \leqq e^{N(t_b-t_a)} < M \tag{6.31}$$

for all $t \in [t_a, t_b]$, all $[w] \in F^*([v],k)$ and all $|z| \leqq \varepsilon/2M$, where M is the positive constant introduced in the proof of Proposition (6.2). This concludes the proof of Proposition (6.5).

Proposition 6.6. *The matrix $G^{-1}(t;[w],z)$, inverse of the matrix $G(t;[w],z)$ exists, is uniformly continuous with respect to t and uniformly bounded for all $t \in [t_a, t_b]$, all $[w] \in F^*([v],k)$ and all $|z| \leqq \varepsilon/2M$.*

Proof of Proposition 6.6. Let $G^*(t;[w],z)$ be the continuous solution of the matrix differential equation

$$\dot{G}^*(t;[w],z) = D(t;[w],z)\,G^*(t;[w],z) \quad \text{a.e. } t \in [t_a, t_b] \tag{6.32}$$

with the terminal condition

$$G^*(t_b;[w],z) = I \ . \tag{6.33}$$

For the same reasons as in Proposition (6.5), we know that the matrix $G^*(t;[w],z)$ exists, is uniformly continuous with respect to t and uniformily bounded for all $t \in [t_a, t_b]$, all $[w] \in F^*([v],k)$ and all $|z| \leqq \varepsilon/2M$. We will now prove that

$$G(t;[w],z)\,G^*(t;[w],z) = I \quad \text{for all} \quad t \in [t_a, t_b] \tag{6.34}$$

i.e.,

$$G^*(t;[w],z) = G^{-1}(t;[w],z) \quad \text{for all} \quad t \in [t_a, t_b] \ . \tag{6.35}$$

We already know that

$$G(t_b;[w],z)\,G^*(t_b;[w],z) = II = I \tag{6.36}$$

and that $G(t;[w],z)\,G^*(t;[w],z)$ is continuous over $[t_a,t_b]$. It remains to verify that

$$(G(t;[w],z)\,G^*(t;[w],z))^{\cdot} = 0 \text{ a.e. } t\in[t_a,t_b] \ . \tag{6.37}$$

This last relation is immediate since

$$\begin{aligned} &\dot G(t;[w],z)\,G^*(t;[w],z) + G(t;[w],z)\,\dot G^*(t;[w],z) \\ &= -G(t;[w],z)D(t;[w],z)G^*(t;[w],z) + G(t;[w],z)D(t;[w],z)G^*(t;[w],z) \\ &= 0 \quad \text{a.e.} \quad t\in[t_a,t_b] \ . \end{aligned} \tag{6.38}$$

This concludes the proof of Proposition (6.6).

Proposition 6.7. *For all* $t\in[t_a,t_b]$, *all* $[w]\in F^*([v],k)$ *and all* $|z|\leqq\varepsilon/M$ *we have*

$$G^{-1}(t;[w],z) = \frac{\partial x(t;[w],z)}{\partial x} \ . \tag{6.39}$$

Proof of Proposition 6.7. Let

$$\Delta(t) = G^{-1}(t;[w],z) - \frac{\partial x(t;[w],z)}{\partial z} \tag{6.40}$$

we have

$$\Delta(t_b) = I - I = 0 \tag{6.41}$$

and

$$\begin{aligned} \dot\Delta(t) &= D(t;[w],z)\,G^{-1}(t;[w],z) - \frac{\partial}{\partial z}f(x(t;[w],z),w(t),t) \\ &= D(t;[w],z)\,G^{-1}(t;[w],z) - D(t;[w],z)\,\frac{\partial x(t;[w],z)}{\partial z} \\ &= D(t;[w],z)\,\Delta(t) \quad \text{a.e.} \quad t\in[t_a,t_b] \ . \end{aligned} \tag{6.42}$$

From the relations (6.41) and (6.42), we obtain

(6.43) $$\Delta(t) = 0 \quad \text{all} \quad t \in [t_a, t_b] \,.$$

This concludes the proof of Proposition (6.7).

Under the mapping

(6.44) $$\Psi([w]) : X \times [t_a, t_b] \to Z([w]) \times [t_a, t_b]$$

the trajectory

(6.45) $$[x([u])] = \{(x(t;[u]), t) : t \in [t_a, t_b]\} \text{ with } [u] \in F^*$$

will be transformed into the trajectory

(6.46) $$[z([u],[w])] = \{(z(t;[u],[v]), t) : t \in [t_a, t_b]\}$$

acccording to the relation

(6.47) $$z(t;[u],[w]) = x(t_b;[w], x(t;[u]) - x(t;[w]), t) - x(t_b;[w]) \quad \text{for all} \quad t \in [t_a, t_b] \,.$$

Conversely, under the mapping

(6.48) $$\Psi^{-1}([w]) : Z([w]) \times [t_a, t_b] \to X \times [t_a, t_b]$$

the trajectory $[z([u],[w])]$ will be transformed into the trajectory $[x([u])]$ according to the relation

(6.49) $$x(t;[u]) = x((t;[v], z(t;[u],[v])) \quad \text{all} \quad t \in [t_a, t_b] \,.$$

In other words, the mapping $\Psi([w])$ associates to a trajectory $[x([u])]$ in $X \times [t_a, t_b]$ a trajectory $[z([u],[w])]$ in $Z([w]) \times [t_a, t_b]$ determined as follows: $z(t;[u],[w])$ is the difference of the projections on the terminal hyperplane $X \times \{t_b\}$ of the points $(x(t;[u]), t)$ and $(x(t;[w]), t)$ of the hyperplane $X \times \{t\}$ where the projecting lines are the solutions of the differential system with the control function $[w]$. Conversely, the inverse mapping $\Psi^{-1}([w])$ associates to a trajectory $[z([u],[w])]$ in $Z([w]) \times [t_a, t_b]$ a trajectory $[x([u])]$ in

$X \times [t_a, t_b]$ determined as follows: $x(t;[u])$ is the projection on the hyperplane $X \times \{t\}$ of the point $(x(t_b;[w]) + x(t_b;[u]), t_b)$ of the terminal hyperplane $X \times \{t_b\}$. The projecting lines are as before the solutions of the differential system with control function $[w]$.

Proposition 6.8. *If* $[u]$ *and* $[w] \in F^*([v],k)$ *then*

(6.50) (i) $$\dot{z}(t;[u],[w]) = G(t;[w],z(t;[u],[w]))\,(f(x(t;[u]),u(t),t) - f(x(t;[u]),w(t),t))$$ *for* a.e. $t \in [t_a, t_b]$

(6.51) (ii) $$z(t_a;[u],[w]) = 0 \ .$$

Proof of Proposition 6.8. If $[u]$ and $[w] \in F^*([v],k)$ then $[z([u],[w])]$ exists and we have the relation

$$x(t;[u]) = x(t;[w],z(t;[u],[w])) \text{ for all } t \in [t_a, t_b] \ . \tag{6.52}$$

By differentiation of (6.52) with respect to t we obtain

(6.53) $$f(x(t;[u]),u(t),t) = f(x(t;[u]),w(t),t) + G^{-1}(t;[w],z(t;[u],[v]))\,\dot{z}(t;[u],[w])$$ for a.e. $t \in [t_a, t_b]$

since

$$\frac{\partial x(t;[w],z)}{\partial z} = G^{-1}(t;[w],z) \tag{6.54}$$

as was proved in Proposition (6.7). From the relation (6.53), we obtain the relation (6.50).

If we let $t = t_a$ in relation (6.47), we have

$$\begin{aligned} z(t_a;[u],[w]) &= x(t_b;[w],0,t_a) - x(t_b;[w]) \\ &= x(t_b;[w]) - x(t_b;[w]) = 0 \ . \end{aligned} \tag{6.55}$$

This concludes the proof of Proposition (6.8).

Proposition 6.9. *There are positive constants P and Q such that for all* $[w_1] \in F^*([v], k)$ *and all* $[w_2] \in F$ *with*

(6.56) (i) $d([w_1],[w_2]) \leqq k$

(6.57) (ii) $\sigma([w_1],[w_2]) \leqq Q$

we have

(6.58) (i) $[w_2] \in F^*$

(6.59) (ii) $d([x([w_1])],[x([w_2])]) \leqq \frac{\varepsilon}{2}$

(6.60) (iii) $d([x([w_1])],[x([w_2])]) \leqq P\sigma([w_1],[w_2])$.

Proof of Proposition 6.9. Let

$$L = \sup|f(x,u,t)| \tag{6.61}$$

over all $(x,t) \in N([x([v])],\varepsilon)$, all $|u| \leqq d([v]) + 2k$. Let M be the positive constant introduced in Proposition (6.2). We shall prove that the relations (6.58), (6.59) and (6.60) are valid when the constants P and Q are determined by

(6.62
$$P = 2M^2L$$

$$Q = \frac{\varepsilon}{8M^3L}. \tag{6.63}$$

Let τ be supremum of all times $\theta \in [t_a, t_b]$ such that

(6.64) (i) $x(t;[w_2])$ exists and is unique on $[t_a, \theta]$

(6.65) (ii) $|x(t;[w_2]) - x(t;[w_1])| \leqq \frac{\varepsilon}{2M}$ for all $t \in [t_a, \theta]$

(6.66) (iii) $z(t;[w_2],[w_1])$ exists and is unique on $[t_a, \theta]$

(6.67) (iv) $|z(t;[w_2],[w_1])| \leqq \frac{\varepsilon}{2M}$ for all $t \in [t_a, \theta]$.

This definition makes sense since the set of all times θ such that the relations (6.64) to (6.67) are satisfied on $[t_a, \theta]$ is not empty: it contains the time t_a.

From (6.65) and (6.61) we have

$$(6.68)\quad |f(x(t;[w_2]), w_2(t), t) - f(x(t;[w_2]), w_1(t), t)| \leqq 2L \text{ for all } t \in [t_a, \tau] .$$

From the relation (6.67) and Proposition (6.5), we have

$$(6.69)\quad |G(t;[w_1], z(t;[w_2],[w_1]))| \leqq M \text{ for all } t \in [t_a, \tau] .$$

By definition, we have

$$(6.70)\quad \mu(\{t: f(x(t;[w_2]), w_2(t), t) - f(x(t;[w_2]), w_1(t), t) \neq 0,\ t \in [t_a, \tau]\}) \leqq \sigma([w_1],[w_2]) .$$

From the relations (6.68), (6.69) and (6.70), and from the differential equation (6.50) with initial condition (6.51), we obtain

$$(6.71)\quad |z(t;[w_2],[w_1])| \leqq 2LM\sigma([w_1],[w_2]) \text{ for all } t \in [t_a, \tau] .$$

With the help of relations (6.57) and (6.63), the inequality (6.71) becomes

$$(6.72)\quad |z(t;[w_2],[w_1])| \leqq \frac{\varepsilon}{4M^2} \text{ for all } t \in [t_a, \tau] .$$

Applying Proposition (6.3) to relations (6.71) and (6.71), we obtain the two relations:

$$(6.73)\quad |x(t;[w_2]) - x(t;[w_1])| \leqq 2LM^2\sigma([w_1],[w_2]) \text{ for all } t \in [t_a, \tau]$$

$$(6.74)\quad |x(t;[w_2]) - x(t;[w_1])| \leqq \frac{\varepsilon}{4M} \text{ for all } t \in [t_a, \tau] .$$

From (6.72) and (6.74) we conclude that

$$(6.75)\quad \tau = t_b$$

because otherwise there would be a $\theta^* \in (\tau, t_b]$ for which all the conditions (6.64) to (6.67) would be satisfied, in contradiction with the definition of τ. Hence

(6.76) $$[w_2] \in F^*$$

and the expression (6.74) can now be written

(6.77) $$|x(t;[w_2]) - x(t;[w_1])| \leqq \frac{\varepsilon}{4M} \quad \text{for all} \quad t \in [t_a, t_b]$$

i.e.,

(6.78) $$d([x([w_2])],[x([w_1])]) \leqq \frac{\varepsilon}{4M}$$

and a fortiori

(6.79) $$d([x([w_2])],[x([w_1])]) \leqq \frac{\varepsilon}{2} .$$

From relations (6.62), (6.73) and (6.75), we have

(6.80) $$|x(t;[w_2]) - x(t;[w_1])| \leqq P\sigma([w_1],[w_2]) \text{ for all } t \in [t_a, t_b]$$

i.e.,

(6.81) $$d([x([w_2])],[x([w_1])]) \leqq P\sigma([w_1],[w_2]) .$$

Relations (6.76), (6.79) and (6.81) are the required relations (6.58),(6.59) and (6.60). This concludes the proof of Proposition (6.9).

Proposition 6.10. *If Q is the positive constant introduced in Proposition (6.9), then for all $[u] \in F$ with $d([u],[v]) \leqq k$ and $\sigma([u],[v]) \leqq Q$, we have $[u] \in F^*([v],k)$.*

Proof of Proposition 6.10. By applying Proposition (6.9) to $[w_1] = [v]$ and $[w_2] = [u]$, we obtain

(6.82) (i) $[u] \in F^*$

(6.83) (ii) $d([x([v])],[x([u])]) \leqq \frac{\varepsilon}{2}$.

From relations (6.82) and (6.83) and from the assumption $d([u],[v]) \leqq k$ we then have

(6.84) $$[u] \in F^*([v],k) .$$

This concludes the proof of Proposition (6.10).

Proposition 6.11. *If Q and P are the positive constants introduced in Proposition* (6.9), *then for all $[u_1]$ and $[u_2] \in F$ with $d([u_1],[v]) \leqq k$, $d[(u_2],[v]) \leqq k$, $\sigma([u_1],[v]) \leqq Q$ and $\sigma([u_2],[v]) \leqq Q$, we have $d([x([u_1])],[x([u_2])]) \leqq P\sigma([u_1],[u_2])$.*

Proof of Proposition 6.11. By applying Proposition (6.10) to $[u]=[u_1]$ we obtain $[u_1] \in F^*([v],k)$, and by applying Proposition (6.9) to $[w_1]=[u_2]$ and $[w_2]=[u_2]$ we obtain $d([x([u_1])],[x([u_2])]) \leqq P\sigma([u_1],[u_2])$. This concludes the proof of Proposition (6.11).

Proposition 6.12. *If Q is the positive constant introduced in Proposition* (6.9), *then there is a positive number K such that for all $[u_1]$ and $[u_2] \in F$ with $d([u_1],[v]) \leqq k$, $d([u_2],[v]) \leqq k$, $\sigma([u_1],[v]) \leqq Q$ and $\sigma([u_2],[v]) \leqq Q$, we have*

$$d([y([u_1],[v])],[y([u_2],[v])]) \leqq K\sigma([u_1],[u_2]) \,. \tag{6.85}$$

Proof of Proposition 6.12. By definition (see (2.8)), we have

$$y(t;[u],[v]) = G(t;[v])(x(t;[u]) - x(t;[v])) \qquad \text{for all } t \in [t_a,t_b] \,. \tag{6.86}$$

From Proposition (6.5), we have

$$|G(t;[v])| = |G(t;[v],0)| \leqq M \qquad \text{for all } t \in [t_a,t_b] \,. \tag{6.87}$$

From relation (6.86), we have

$$y(t;[u_1],[v]) - y(t;[u_2],[v]) = G(t;[v])(x(t;[u_1]) - x(t;[u_2])) \qquad \text{for all } t \in [t_a,t_b] \,. \tag{6.88}$$

From relations (6.87) and (6.88), we have

$$d([y([u_1],[v])],[y([u_2],[v])]) \leqq Md([x([u_1])],[x([u_2])]) \,. \tag{6.89}$$

But by Proposition (6.11), we have

$$d([x([u_1])],[x([u_2])]) \leqq P\sigma([u_1],[u_2]) \,. \tag{6.90}$$

From (6.89) and (6.90), we then have

$$(6.91)\qquad d([y([u_1],[v])],[y([u_2],[v])]) \leqq PM\sigma([u_1],[u_2]) = K\sigma([u_1],[u_2])\,.$$

This concludes the proof of Proposition (6.12).

All the resullts of this section which we shall need later on can be summarized as follows:

Proposition 6.13. *If* $[v]\in F^*$, *ε is the positive number associated to* $[v]$ *in the definition of E, k is an arbitrary positive number, then there exist two positive numbers K and Q such that for all* $[u_1]$ *and* $[u_2]\in F$ *with*

$$(6.92)\qquad \text{(i)}\quad d([u_1],[v]) \text{ and } d([u_2],[v]) \leqq k$$

$$(6.93)\qquad \text{(ii)}\quad \sigma([u_1],[v]) \text{ and } \sigma([u_2],[v]) \leqq Q$$

we have

$$(6.94)\qquad \text{(i)}\quad [u_1] \text{ and } [u_2]\in F^*$$

$$(6.95)\qquad \text{(ii)}\quad d([y([u_1],[v])],[y([u_2],[v])]) \leqq K\sigma([u_1],[u_2])\,.$$

Application to a linear system.

We shall assume that $f(x,u,t)$ has the particular form given in relation (2.15), namely

$$(6.96)\qquad f(x,u,t) = A(u,t)x + \phi(u,t)\,.$$

The results obtained in this section in the case of a nonlinear system can a fortiori be applied to a linear system. We want to show here how these results could be obtained directly in the case of a linear system.

Since for any $[u]\in F$ the coefficients of Equation (6.96) are measurable and bounded, we know that a solution $[x([u])]$ will always exist, i.e., $[u]\in F^*$. In other words, we have

$$(6.97)\qquad F = F^*\,.$$

Hence the question of existence of a solution for a particular $[u] \in F$ which was the main difficulty in the treatment of the general nonlinear system, given in the beginning of this section, is trivially solved in the case of a linear system. To complete this direct study of a linear system, we shall perform some algebraic manipulations and prove the existence of uniform bounds for some constants associated to each element of F.

The study of a linear system is particularly simple because the matrix

$$D(t;[w],z) = \frac{\partial f(x,w(t),t)}{\partial x}\bigg|_{x=x(t;[w],z)} = A(w(t),t) \qquad \text{for all } t \in [t_a,t_b] \tag{6.98}$$

is independent of z. From there follows that the matrix $G(t;[w],z)$ will also be independent of z and we shall write $G(t;[w])$ instead of $G(t;[w],z)$. Proposition (6.7) can now be stated

$$G^{-1}(t;[w]) = \frac{\partial x(t;[w],z)}{\partial z} \tag{6.99}$$

and from the relation

$$x = x(t;[w],z) \tag{6.100}$$

defining the mapping

$$\Psi^{-1}([w]) : Z([w]) \times [t_a,t_b] \to X \times [t_a,t_b] \tag{6.101}$$

it follows that

$$x = x(t;[w],0) + G^{-1}(t;[w])z\ . \tag{6.102}$$

But by definition

$$x(t;[w],0) = x(t;[w]) \tag{6.103}$$

hence

$$x = x(t;[w]) + G^{-1}(t;[w])z \tag{6.104}$$

i.e.,

$$z = G(t;[w])(x - x(t;[w])) . \tag{6.105}$$

If we compare relation (6.104) with the relation

$$y = G(t;[w])(x - x(t;[w]) \tag{6.106}$$

defining the mapping (see Section 2)

$$\Phi([w]) : X \times [t_a, t_b] \to Y([w]) \times [t_a, t_b] \tag{6.107}$$

we obtain

$$\Psi([w]) \equiv \Phi([w]) \tag{6.108}$$

i.e.,

$$Z([w]) \times [t_a, t_b] = Y([w]) \times [t_a, t_b] \tag{6.109}$$

and in particular

$$[z([u],[w])] = [y([u],[w])] . \tag{6.110}$$

For a linear system Propositions (6.9), (6.10), (6.11) and (6.12) are combined into the following result:

Proposition 6.14. *For any positive number k, there is a number P such that for all $[u]$ and $[w] \in F_k$ we have*

$$d([y([u],([w])]) \leqq P\sigma([u],[w]) . \tag{6.111}$$

Proof of Proposition 6.14. For any positive number k there exist uniform bounds for each of the expressions on the right side of the differential equation (2.26), hence there exists a P such that

$$|\dot{y}(t;[u],[w])| \leqq P \tag{6.112}$$

for a.e. $t \in [t_a, t_b]$, all $[u]$ and $[w] \in F_k$.

Moreover,

$$y(t_a;[u],[w]) = 0 \tag{6.113}$$

and

$$\dot{y}(t;[u],[w]) = 0 \tag{6.114}$$

for all $t \in [t_a, t_b]$ such that $u(t) = w(t)$.

From relations (6.112), (6.113) and (6.114) we immediately have

$$d([y([u],[w])]) \leqq P\sigma([u],[w]) \quad . \tag{6.115}$$

This completes the proof of Proposition (6.14).

Application to a linear* system.

If we assume that the system is linear*, i.e., if the function $f(x,u,t)$ has the particular form given in relation (2.16):

$$f(x,u,t) = A(t)x + \phi(u,t) \tag{6.116}$$

the treatment given above for linear systems can be further simplified, since the matrices $D(t;[w])$ and $G(t;[w])$ are independent of $[w]$ in the case of a linear* system. We shall write $G(t)$ instead of $G(t;[w])$.

The differential equation for $[y([u],[w])]$ becomes

$$\dot{y}(t;[u],[w]) = G(t)(\phi(u(t),t) - \phi(w(t),t)) \quad \text{for a.e. } t \in [t_a, t_b] \tag{6.117}$$

and may be directly integrated to give

$$y(t;[u],[w]) = \int_{t_a}^{t} G(\tau)\,(\phi(u(\tau),\tau) - \phi(w(\tau),\tau))\,d\tau \quad \text{all } t \in [t_a, t_b] \tag{6.118}$$

which implies

$$x(t;[u]) - x(t;[w]) = G^{-1}(t)\int_{t_a}^{t} G(\tau)(\phi(u(\tau),\tau) - \phi(w(\tau),\tau))\, d\tau \quad \text{all } t \in [t_a, t_b] \ . \tag{6.119}$$

The results of Proposition (6.17) may then be immediately read off from relation (6.118).

Relation (6.119) could have been immediately derived from the theory of nonhomogeneous linear differential equations. Indeed, we have

$$(x(t;[u]) - x(t;[w]))^{\cdot} = A(t)(x(t,[u]) - x(t;[w])) + \phi(u(t),t) - \phi(w(t),t) \quad \text{for a. e. } t \in [t_a, t_b] \tag{6.120}$$

and relation (6.119) is the well-known solution of the differential equation (6.120) for the initial condition

$$x(t_a;[u]) - x(t_a;[w]) = 0 \ . \tag{6.121}$$

SECTION 7

Approximation of the Comparison Trajectories in the Comoving Space along a Given Trajectory

In $Y([v]) \times [t_a, t_b]$, the comoving coordinate space along the trajectory $[x([v])]$, we have for every $[u] \in F^*$ a trajectory $[y([u],[v])]$ which is the image of the trajectory $[x([u])]$ and a trajectory $[\overset{+}{y}([u],[v])]$ which is a certain approximation of the trajectory $[y([u],[v])]$. In this section we shall study the properties of this approximation. More precisely, this section will be devoted to the proof of Proposition (7.3) in which we give an upper bound for the uniform distance between $[y([u],[v])]$ and $[\overset{+}{y}([u],[v])]$ as a function of the distance between the two control functions $[u]$ and $[v]$ when this distance is measured with respect to the norm σ.

We remind the reader that the results of this section are highly trivial when the system is linear*, since in that case we already know that $[y([u],[v])] \equiv [\overset{+}{y}([u],[v])]$ (see Section 2). When the system is linear the situation is not trivial but nevertheless very simple as is shown in a paragraph at the end of this section.

Proposition 7.1. *Let $F(t,x)$ be an n-dimensional vector-valued function defined for all $t \in [t_a, t_b]$ and all n-dimensional vectors x with $|x| \leqq \eta$, where η is a fixed positive number, such that*

(7.1) (i) *$F(t,x)$ is measurable with respect to t for all $|x| \leqq \eta$, uniformly equicontinuous wit respect to x and uniformly bounded for all $t \in [t_a, t_b]$ and all $|x| \leqq \eta$*

$$(7.2) \qquad \text{(ii)} \quad \lim_{|x| \to 0} \frac{|F(t,x)|}{|x|} = 0 \qquad \textit{uniformly for } t \in [t_a, t_b]$$

then there exists a function $G(r)$ defined, continuous and nondecreasing over $[0, \eta]$ such that

$$(7.3) \qquad \text{(i)} \quad \lim_{r \to 0} \frac{G(r)}{r} = 0$$

$$(7.4) \qquad \text{(ii)} \quad \int_{t_a}^{t_b} |F(t, a(t))| \, dt \leqq G(d([a]))$$

for all bounded measurable n-dimensional vector -valued functions $[a]$ such that $d([a]) \leqq \eta$.

Proof of Proposition 7.1. Let

$$(7.5) \qquad G(r) = (t_b - t_a) \sup_{\substack{|x| \leqq r \\ t \in [t_a, t_b]}} |F(t,x)| \qquad \text{for all } r \in [0, \eta] \; .$$

By construction $G(r)$ is continuous and nondecreasing over $[0, \eta]$. We also have

$$(7.6) \qquad \lim_{r \to 0} \frac{G(r)}{r} = 0$$

since we have assumed that

$$(7.7) \qquad \lim_{|x| \to 0} \frac{|F(t,x)|}{|x|} = 0 \quad \text{uniformly for } t \in [t_a, t_b] \; .$$

Moreover,

$$(7.8) \qquad \int_{t_a}^{t_b} |F(t,a(t))|\,dt \leqq (t_b - t_a) \sup_{t\in[t_a,t_b]} |F(t,a(t))|$$

$$\leqq (t_b - t_a) \sup_{\substack{t\in[t_a,t_b]\\ |x|\leqq d([a])}} |F(t,x)|$$

$$\leqq G(d([a])) .$$

This concludes the proof of Proposition (7.1).

Proposition 7.2. *Let $K(t,x,u)$ be an n-dimensional vector-valued function defined for all $t\in[t_a,t_b]$, all n-dimensional vectors x with $|x| \leqq \eta$, all r-dimensional vectors u with $|u| \leqq \delta$, where η and δ are fixed positive numbers, such that*

(7.9) (i) *$K(t,x,u)$ is measurable with respect to t and u, uniformly equicontinuous with respect to x and uniformly bounded for all $t\in[t_a,t_b]$, all $|x|\leqq\eta$ and all $|u|\leqq\delta$*

(7.10) (ii) $K(t,0,u)=0$ *all $t\in[t_a,t_b]$, all $|u|\leqq\delta$*

(7.11) (iii) $K(t,x,0)=0$ *all $t\in[t_a,t_b]$, all $|x|\leqq\eta$*

then there exists a function $H(r)$ defined, continuous and nondecreasing over $[0,\eta]$ such that

(7.12) (i) $H(0)=0$

$$(7.13) \quad \text{(ii)} \int_{t_a}^{t_b} |K(t,x(t),u(t))|\,dt \leqq H(d([x]))\,\sigma([u]) .$$

Proof of Proposition 7.2. Let

$$(7.14) \qquad H(r) = \sup_{\substack{t\in[t_a,t_b]\\ |x|\leqq r\\ |u|\leqq\delta}} |K(t,x,u)| .$$

Then $H(r)$ is continuous and nondercreasing over $[0,\eta]$ since $K(t,x,u)$ is uniformly equicontinuous with respect to x for all $t \in [t_a, t_b]$ and all $|u| \leqq \delta$. We also have

$$(7.15) \qquad H(0) = 0$$

since we have assumed $K(t,0,u) = 0$. Moreover,

$$(7.16) \qquad \int_{t_a}^{t_b} |K(t,x(t),u(t))|\,dt \leqq \int_{\substack{t_a \\ u(t)\neq 0}}^{t_b} |K(t,x(t),u(t))|\,dt$$

$$\leqq \int_{\substack{t_a \\ u(t)\neq 0}}^{t_b} \sup_{\substack{t \in [t_a,t_b] \\ |r| \leqq d([x]) \\ |v| \leqq \delta}} |K(t,r,v)|\,dt$$

$$\leqq \int_{\substack{t_a \\ u(t)\neq 0}}^{t_b} H(d([x]))dt$$

$$\leqq H(d([x]))\sigma([u]) \ .$$

This concludes the proof of Proposition (7.2).

Proposition 7.3. *If $[v] \in F^*$ and if Q is the positive constant introduced in Proposition (6.9), then there exists a function $g(r)$ defined, continuous and nondecreasing over $[0,Q]$ such that*

$$(7.17) \qquad \text{(i)} \lim_{r\to 0} \frac{g(r)}{r} = 0$$

(7.18) (ii) *for all $[u] \in F^*([v],k)$ with $\sigma([u],[v]) \leqq Q$, we have*

$$d([y([u],[v])],[\overset{+}{y}([u],[v])]) \leqq g(\sigma([u],[v])) \ .$$

Proof of Proposition 7.3. By definition, we have

$$(7.19) \qquad y(t;[u],[v]) = G(t;[v])(x(t;[u]) - x(t;[v])) \text{ all } t \in [t_a, t_b] \ .$$

By differentiation with respect to t of relation (7.19), we obtain

$$
\begin{aligned}
(7.20)\qquad & \dot{y}(t;[u],[v]) \\
&= \dot{G}(t;[v])(x(t;[u]) - x(t;[v])) \\
&\quad + G(t;[v])(\dot{x}(t;[u]) - \dot{x}(t;[v])) \\
&= - G(t;[v])D(t;[v])(x(t;[u]) - x(t;[v])) \\
&\quad + G(t;[v])(f(x(t;[u]),u(t),t) - f(x(t;[v]),v(t),t)) \\
&= G(t;[v])(f(x(t;[v]),u(t),t) - f(x(t;[v]),v(t),t)) \\
&\quad + F(t;x(t;[u]) - x(t;[v])) \\
&\quad + K(t;x(t;[u]) - x(t;[v]),u(t) - v(t)) \text{ a.e. } t \in [t_a,t_b]
\end{aligned}
$$

where the functions F and K are defined by

$$
\begin{aligned}
(7.21)\quad F(t;x) = &- G(t;[v])D(t;[v])x + G(t;[v])(f(x(t;[v]) + x,v(t),t) \\
&- f(x(t;[v]),v(t),t))
\end{aligned}
$$

$$
\begin{aligned}
(7.22)\quad K(t;x,s) = G(t;[v])(&f(x(t;[v]) + x,v(t) + s,t) - f(x(t;[v]),v(t) + s,t) \\
&+ f(x(t;[v]),v(t),t) - f(x(t;[v]) + x,v(t),t)) \; .
\end{aligned}
$$

By definition we have

$$
(7.23)\qquad \dot{\overset{+}{y}}(t;[u],[v]) = G(t;[v])(f(x(t;[v]),u(t),t) - f(x(t;[v]),v(t),t)) \quad \text{for a.e. } t \in [t_a,t_b] \; .
$$

From relations (7.20) and (7.23), we obtain

$$
\begin{aligned}
(7.24)\qquad & \dot{y}(t;[u],[v]) - \dot{\overset{+}{y}}(t;[u],[v]) \\
&= F(t;x(t;[u]) - x(t;[v])) + K(t;x(t;[u]) - x(t;[v]),u(t) - v(t)) \; .
\end{aligned}
$$

By definition we also have

$$
(7.25)\qquad y(t_a;[u],[v]) = \overset{+}{y}(t_a;[u],[v]) = 0 \; .
$$

Hence, from relations (7.24) and 7.25), we may write:

(7.26) $d([y([u],[v])],[\overset{+}{y}([u],[v])])$

$$\leqq \int_{t_a}^{t_b} |F(t;x(t;[u])-x(t;[v]))|\,dt + \int_{t_a}^{t_b} |K(t;x(t;[u])-x(t;[v]),u(t)-v(t))|\,dt.$$

Let us first estimate the integral

$$\int_{t}^{t_b} |F(t;x(t;[u]) - x(t;[v]))|\,dt \; . \tag{7.27}$$

From the definition of $F(t;x)$ we have

$$F(t;0) = 0 \tag{7.28}$$

and

$$F_x(t;0) = -\,G(t;[v])D(t;[v]) + G(t;[v])D(t;[v]) = 0 \; . \tag{7.29}$$

From the relations (7.28) and (7.29), we then have

$$\lim_{|x|\to 0} \frac{|F(t,x)|}{|x|} = 0 \; . \tag{7.30}$$

By Proposition (7.1) there then exists a function $G(r)$ defined, continuous and nondecreasing over $[0,\varepsilon/2]$ such that

$$\text{(i)} \lim_{r\to 0} \frac{G(r)}{r} = 0 \tag{7.31}$$

$$\text{(ii)} \int_{t_a}^{t_b} |F(t;x(t;[u]) - x(t;[v]))|\,dt \leqq G(d([x([u])],[x([v])])) \; . \tag{7.32}$$

From Proposition (6.11) we know that

$$d([x([u])],[x([v])]) \leqq P\sigma([u],[v]) \; . \tag{7.33}$$

If we define $G_1(r) = G(Pr)$, relations (7.32) and (7.33) may then be written

$$\int_{t_a}^{t_b} |F(t;x(t;[u]) - x(t;[v]))|\,dt \leqq G_1(\sigma([u],[v])) \; . \tag{7.34}$$

By definition $G_1(r)$ is continuous, nondecreasing over $[0, Q]$ and such that

7.35) $$\lim_{r \to 0} \frac{G_1(r)}{r} = 0 .$$

Let us now estimate the integral

(7.36) $$\int_{t_a}^{t_b} | K(t; x(t; [u]) - x(t; [v]), u(t) - v(t)) | \, dt .$$

From the definition of $K(t, r, s)$ we have

(7.37) $K(t, x, 0) = 0$ for all $t \in [t_a, t_b]$ and all x with $|x| \leqq \frac{\varepsilon}{2}$

and

(7.38) $K(t, 0, s) = 0$ for all $t \in [t_a, t_b]$ and all s with $|s| \leqq k$.

Hence by Proposition (7.2), there is a function $H(r)$ defined, continuous and nondecreasing over $[0, \varepsilon/2]$ such that

(7.39) (i) $H(0) = 0$

(7.40) (ii) $$\int_{t_a}^{t_b} | K(t; x(t; [u]) - x(t; [y]), u(t) - v(t)) | \, dt \leqq H(d([x([u])], [x([v])])) \sigma ([u], [v]) .$$

From Proposition (6.11) ,we know that

(7.41) $$d([x([u])], [x([v])]) \leqq P\sigma([u], [v]) .$$

If we define $G_2(r) = rH(Pr)$ then relations (7.40) and (7.41) may be written:

(7.42) $$\int_{t}^{t_b} |K(t; x(t; [u]) - x(t; [v]), u(t) - v(t)) | \, dt \leqq G_2(\sigma([u], [v])) .$$

By definition $G_2(r)$ is continuous, nondecreasing over $[0, Q]$ and such that

(7.43) $$\lim_{r \to 0} \frac{G_2(r)}{r} = 0 .$$

Combining relations (7.26), (7.34) and (7.42), we obtain

(7.44) $$d([y([u],[v])],[\overset{+}{y}([u],[v])]) \leqq G_1(\sigma([u],[v])) + G_2(\sigma([u],[v])) \ .$$

We define

(7.45) $$g(r) = G_1(r) + G_2(r) \ .$$

From (7.35) and (7.43) it follows that the function $g(r)$ is continuous. nondecreasing over $[0,Q]$ and such that

7.46) $$\lim_{r\to 0} \frac{g(r)}{r} = 0$$

Moreover, combining relations (7.44) and (7.45), we obtain

(7.47) $$d([y([u],[v])],[\overset{+}{y}([u],[v])]) \leqq g(\sigma([u],[v])) \ .$$

This concludes the proof of Proposition (7.3).

Application to a linear system.

We shall again consider the linear system

(7.48) $$f(x,u,t) = A(u,t)x + \phi(u,t)$$

introduced in Section 2 and show how the results of this section could be obtained directly for a linear system.

In Section 2 we have derived the following relations:

(7.49) $$\dot{y}(t;[u],[v]) - \dot{\overset{+}{y}}(t;[u],[v])$$
$$= G(t;[v])(A(u(t),t) - A(v(t),t))G^{-1}(t;[v])\, y(t;[u],[v])$$
$$\text{for a.e. } t \in [t_a, t_b]$$

(7.50) $$y(t_a;[u],[v]) = \overset{+}{y}(t_a;[u],[v]) = 0 \ .$$

We know that the coefficient $G(t;[v])(A(u(t),t) - A(v(t),t))G^{-1}(t;[v])$, occuring in the right side of Equation (7.49), is uniformly bounded by some constant V for all $[u]$ and $[v] \in F_k$, and we have proved in Proposition(6.14) that there exists a positive constant P such that

(7.51) $$d([y([u],[w])]) \leqq P\sigma([u],[w]) \ .$$

From relations (7.49) and (7.51), it follows that

(7.52) $$|\dot{y}(t;[u],[v]) - \overset{+}{\dot{y}}(t;[u],[v])| \leqq VP\sigma([u],[v]) \quad \text{for a.e. } t \in [t_a, t_b] \ .$$

Moreover, we immediately see from relation (7.49) that

(7.53) $$\dot{y}(t;[u],[v]) - \overset{+}{\dot{y}}(t;[u],[v]) = 0 \quad \text{for all } t \in [t_a,t_b] \text{ such that } u(t) = v(t) \ .$$

If we write

(7.54) $$P^* = PV$$

then from relations (7.50), (7.52), (7.53) and (7.54), we immediately obtain the following result:

Proposition 7.4. *If k is a positive real number, then there is a positive constant P^* such that*

(7.55) $$d([y([u],[v])],[\overset{+}{y}([u],[v])]) \leqq P^*(\sigma([u],[v]))^2 \quad \textit{for all } [u] \textit{ and } [v] \in F_k \ .$$

In the case of a linear system, Proposition (7.4) implies Proposition (7.3) since the function $g(r) = P^*r^2$ satisfies the conditions of Proposition (7.3): P^*r^2 is continuous, nondecreasing for positive r and $\lim_{r \to 0} \frac{P^*r^2}{r} = 0$.

SECTION 8

The Range of a Vector Integral over Borel Sets

In this section we shall derive some properties of the set $\overset{+}{H}([v])$. In Proposition (8.11) we shall prove that the set $\overset{+}{H}([v])$ is convex. This result will be used later in the proof of Theorem IV. Let $\Gamma : F \to \overset{+}{H}([v])$ be the mapping which maps an element $[u]$ of F into the element $\overset{+}{y}(t_b;[u],[v])$ of $\overset{+}{H}([v])$. In Proposition (8.13) we shall prove that if 0 is an interior point of the set $\overset{+}{H}([v])$ then there is a subset $F([v])$ of F such that 0 is also an interior point of the image of $F([v])$ under the mapping Γ and such that the restriction to $F([v])$ of the mapping Γ has a continuous inverse. We shall need Proposition (8.13) in the proof of Theorem III.

We shall assume that the reader knows the basic elements of the theory of measure which can be found in Halmos' book, "Measure Theory".

First let us recall some classical notations and definitions:

1. If A is a set, $\mathscr{A}$ a σ-algebra of subsets of A, μ a non-negative measure defined on $\mathscr{A}$ with $\mu(A) < +\infty$, then $(A,\mathscr{A},\mu)$ is called a *measure σ-algebra.*

2. An element $B \in \mathscr{A}$ is called an *atom* of the measure σ-algebra $(A,\mathscr{A},\mu)$ if $\mu(B) \neq 0$ and if $D \subset B$ with $D \in \mathscr{A}$ implies either

$$\mu(D) = 0 \tag{8.1}$$

or

$$\mu(D) = \mu(B) \ . \tag{8.2}$$

3. A measure σ-algebra $(A,\mathscr{A},\mu)$ is called *nonatomic* if it has no atom.

Proposition 8.1. *If $(A,\mathscr{A},\mu)$ is a nonatomic measure σ-algebra, $B \in \mathscr{A}$, $\mathscr{A}_B = \{D : D \in \mathscr{A},\ D \subset B\}$ then $(B,\mathscr{A}_B,\mu)$ is a nonatomic measure σ-algebra.*

Proof of Proposition 8.1. Let $\overline{\mathscr{A}_B} \supset \mathscr{A}_B$ be the minimal σ-algebra containing $\mathscr{A}_B$. We have by construction $\overline{\mathscr{A}_B} \subset \mathscr{A}$, moreover, $D \in \overline{\mathscr{A}_B}$ implies $D \subset B$, hence $\overline{\mathscr{A}_B} \subset \mathscr{A}_B$. In other words, $\overline{\mathscr{A}_B} = \mathscr{A}_B$ and $(B,\mathscr{A}_B,\mu)$ is a measure σ-algebra. Finally, $(B,\mathscr{A}_B,\mu)$ is nonatomic since any atom of $(B,\mathscr{A}_B,\mu)$ is also an atom of $(A,\mathscr{A},\mu)$. This concludes the proof of Proposition (8.1).

Proposition 8.2. *If $(A,\mathscr{A},\mu)$ is a nonatomic measure σ-algebra, $\mathscr{A}_\alpha = \{B : B \in \mathscr{A}, \mu(B) = \alpha\mu(A)\}$, then there exists a nest $\mathscr{N} \subset \mathscr{A}$ such that $\mathscr{N} \cap \mathscr{A}_\alpha \neq \emptyset$ for all $\alpha \in [0,1]$.*

Proof of Proposition 8.2. Let $\mathscr{N}$ be a maximal nest in $\mathscr{A}$. Such a nest exists by the Hausdorff Maximal Principle. Let $\mathscr{N}_\alpha = \mathscr{A} \cap \mathscr{A}_\alpha$,

$\mathcal{N}_\alpha^- = \bigcup_{0 \leqq a \leqq \alpha} \mathcal{N}_a$, $\mathcal{N}_\alpha^+ = \bigcup_{\alpha \leqq a \leqq 1} \mathcal{N}_a$, then $\mathcal{N}_\alpha^- \neq \emptyset$ since $\emptyset \in \mathcal{N}_0$, similarly $\mathcal{N}_\alpha^+ \neq \emptyset$ since $A \in \mathcal{N}_1$. Let $N_\alpha^+ = \bigcap_{B \in \mathcal{N}_\alpha^+} B$, $N_\alpha^- = \bigcup_{B \in \mathcal{N}_\alpha^-} B$. Since $\mathcal{N}$ is a maximal nest over the σ-algebra $\mathscr{A}$, then $N_\alpha^- \in \mathcal{N}$, $N_\alpha^+ \in \mathcal{N}$, $\mu(N_\alpha^-) \leqq \alpha$, $\mu(N_\alpha^+) \geqq \alpha$, $N_\alpha^- \subset N_\alpha^+$ and there is no $N \in \mathscr{A}$ such that $N \neq N_\alpha^-$, $N \neq N_\alpha^+$, $N_\alpha^- \subset N \subset N_\alpha^+$. Moreover, we have $\mu(N_\alpha^-) = \mu(N_\alpha^+) = \alpha$. Otherwise, from the nonatomicity of $(A, \mathscr{A}, \mu)$, there would be a subset K of $N_\alpha^+ \sim N_\alpha^-$ such that $\mu(K) \neq 0$ and $\mu(K) \neq \mu(N_\alpha^+ \sim N_\alpha^-) = \mu(N_\alpha^+) - \mu(N_\alpha^-)$, i.e., such that $N_\alpha^- \cup K \neq N_\alpha^-$, $N_\alpha^- \cup K \neq N_\alpha^+$ and $N_\alpha^- \subset N_\alpha^- \cup K \subset N_\alpha^+$ which contradicts our previous results. For every $\alpha \in [0,1]$ we have exhibited elements N_α^- and N_α^+ in $\mathcal{N}_\alpha = \mathcal{N} \cap \mathscr{A}_\alpha$. This concludes the proof of Proposition (8.2).

Proposition 8.3. *If $(A, \mathscr{A}, \mu)$ is a nonatomic measure σ-algebra, then there exists a set $\mathscr{D} = \{D_\alpha : \alpha \in [0,1]\}$ such that*

(8.3) (i) $D_\alpha \in \mathscr{A}$ *all* $\alpha \in [0,1]$

(8.4) (ii) $\mu(D_\alpha) = \alpha\mu(A)$

(8.5) (iii) $D_{\alpha_1} \subset D_{\alpha_2}$ *if and only if* $\alpha_1 \leqq \alpha_2$.

Proof of Proposition 8.3. For every $\alpha \in [0,1]$ let D_α be an element of the nonempty set $\mathcal{N}_\alpha$. Such D_α exists by the axiom of choice. The conditions (i), (ii) and (iii) are then satisfied by construction. This concludes the proof of Proposition (8.3).

Proposition 8.4. *If $(A, \mathscr{A}, \mu)$ is a nonatomic measure σ-algebra, $\{a_i : i = 1, \cdots, k\}$ a finite set of nonnegative real numbers, then there exists a set $\{A_i : i = 1, \cdots, k\}$ such that*

(8.6) (i) $A_i \in \mathscr{A}$ *for* $i = 1, \cdots, k$

(8.7) (ii) $A_i \cap A_j = \emptyset$ *for* $i = 1, \cdots, k; j = 1, \cdots, k$ *and* $i \neq j$

(8.8) (iii) $\bigcup_{i=1}^{k} A_i = A$

(8.9) (iv) $\mu(A_i) = \dfrac{a_i}{\sum_{j=1}^{k} a_j} \mu(A)$.

Proof of Proposition 8.4. Let

(8.10) $$\alpha_i = \frac{a_i}{\sum_{j=i}^{k}} \quad \text{for } i = 1, \cdots, k \; .$$

By Proposition (8.3) there is a $A_1 \in \mathscr{A}$ such that

(8.11) $$\mu(A_1) = \alpha_1 \mu(A) \; .$$

Let

(8.12) $$A^{(1)} = A \sim A_1$$

and

(8.13) $$\mathscr{A}_1 = \mathscr{A}_{A^{(1)}} \; .$$

By Proposition (8.1), $(A^{(1)}, \mathscr{A}_1, \mu)$ is a nonatomic measure σ-algebra and by Proposition(8.2) there is a $A_2 \in \mathscr{A}_1$ such that

(8.14) $$\mu(A_2) = \alpha_2 \mu(A^{(1)}) \; .$$

Let

(8.15) $$A^{(2)} = A^{(1)} \sim A_2$$

and

(8.16) $$\mathscr{A}_2 = \mathscr{A}_{A^{(2)}} \; .$$

By proposition (8.1), $(A^{(2)}, \mathscr{A}_2, \mu)$ is a nonatomic measure σ-algebra and by Proposition (8.2) there is a $A_3 \in \mathscr{A}_2$ such that

(8.17) $$\mu(A_3) = \alpha_3 \mu(A^{(2)}) \; .$$

By repeating $k-1$ times the same process, we obtain a set $\{A_i : i = 1, \cdots, k\}$. It is a trivial matter to verify that the set $\{A_i : i = 1, \cdots, k\}$ satisfies conditions (i) to (iv). This concludes the proof of Proposition (8.4).

Proposition 8.5. *If $(A,\mathscr{B},\mu)$ is a nonatomic measure σ-algebra, f a bounded and $\mathscr{B}$ measurable function over A, there is a set $D\in\mathscr{B}$ with $\mu(D)=\frac{1}{2}\mu(A), \int_D f\,d\mu=\frac{1}{2}\int_A f\,d\mu$.*

Proof of Proposition 8.5. From Proposition (8.2) there is a subset $\mathscr{D}=\{D_\alpha:\alpha\in[0,1]\}$ of $\mathscr{B}$ such that

(i) $\mu(D_\alpha)=\alpha\mu(A)$

(ii) $D_{\alpha_1}\subset D_{\alpha_2}$ if and only if $\alpha_1\leqq\alpha_2$.

Let $B_\alpha=D_\alpha\sim D_{\alpha-1/2}$ all $\alpha\in[\frac{1}{2},1]$, hence $B_\alpha\in\mathscr{B}$ and

$$\mu(B_\alpha)=\mu(D_\alpha)-\mu(D_{\alpha-1/2})=\alpha-(\alpha-\tfrac{1}{2})=\tfrac{1}{2}\ .$$

We shall assume temporarily that $\int_A f\,d\mu\neq 0$.
Let

$$\Phi(s)=\frac{\int_S f\,d\mu}{\int_A f\,d\mu}\ .$$

We then have

$$\Phi(B_1)+\Phi(B_{1/2})=1 \tag{8.18}$$

since $(\Phi(D_1)-\Phi(D_{1/2}))=(\Phi(D_{1/2})-\Phi(D_0))=(1-\frac{1}{2})+(\frac{1}{2}-0)=1$. Moreover,

$$\Phi(B_\alpha)\text{ is continuous over }\alpha\in[\tfrac{1}{2},1] \tag{8.19}$$

since

$$\begin{aligned}|\Phi(B_{\alpha_1})-\Phi(B_{\alpha_2})| &\leqq |(\Phi(D_{\alpha_1})-\Phi(D_{\alpha_1-1/2}))-(\Phi(D_{\alpha_2})-\Phi(D_{\alpha_2-1/2}))|\\ &\leqq |\Phi(D_{\alpha_1})-\Phi(D_{\alpha_2})|+|\Phi(D_{\alpha_1-1/2})-\Phi(D_{\alpha_2-1/2})|\\ &\leqq M|\mu(D_{\alpha_1})-\mu(D_{\alpha_2})|+M|\mu(D_{\alpha_1-1/2})-\mu(D_{\alpha_2-1/2})|\\ &\leqq M|\alpha_1-\alpha_2|+M|(\alpha_1-\tfrac{1}{2})-(\alpha_2-\tfrac{1}{2})|=2M|\alpha_1-\alpha_2|\,.\end{aligned}$$

From (8.19) we know that $\Phi(B_{[1/2,1]})=\{\Phi(B_\alpha):\alpha\in[\frac{1}{2},1]\}$ is a segment. Moreover, $\Phi(B_{[1/2,1]})$ contains $\Phi(B_1)$ and $\Phi(B_{1/2})=1-\Phi(B_1)$, from (8.18), hence $\Phi(B_{[1/2,1]})$ contains

$$\frac{\Phi(B_1) + \Phi(B_{1/2})}{2} = \frac{\Phi(B_1) + (1 - \Phi(B_1))}{2} = \frac{1}{2}$$

since a segment is a convex set.

Let $\bar{\alpha} \in [\frac{1}{2}, 1]$ be such that $\Phi(B_{\bar{\alpha}}) = \frac{1}{2}$ then $D = B_{\bar{\alpha}}$ is the requested set.

If $\int_A f\,d\mu = 0$ let $\Phi(s) = \int_S f\,d\mu$ then

(8.20) $$\Phi(B_1) + \Phi(B_{1/2}) = 0$$

and

(8.21) $$\Phi(B_\alpha) \text{ is continuous for } \alpha \in [\tfrac{1}{2}, 1]$$

for the same reasons as (8.18) and (8.19).

From(8.20) and (8.1), we conclude as before that there is a $\bar{\alpha} \in [\frac{1}{2}, 1]$ such that $\Phi(B_{\bar{\alpha}}) = 0$. This ends the proof of Proposition (8.5).

Proposition 8.6. *If $(A, \mathscr{B}, \mu)$ is a nonatomic measure σ-algebra, f a uniformly bounded and $\mathscr{B}$ measurable function over A, $\mathscr{A}$ the minimal σ-algebra over $\mathscr{D} = \{D_i : i = 0, 1, 2, \cdots,\}$ where the sets D_i have the following properties:*

(8.22) $$D_0 = A$$

(8.23) $$D_n \in \mathscr{B}$$

(8.24) $$D_{2n+1} \cap D_{2n+2} = \emptyset$$

(8.25) $$D_{2n+1} \cup D_{2n+2} = D_n$$

(8.26) $$\mu(D_{2n+1}) = \mu(D_{2n+2}) = \tfrac{1}{2}\mu(D_n)$$

(8.27) $$\int_{D_{2n+1}} f\,d\mu = \int_{D_{2n+2}} f\,d\mu = \tfrac{1}{2}\int_{D} f\,d\mu$$

then

(8.28) $$\int_D f\,d\mu = \frac{\mu(D)}{\mu(A)} \int_A f\,d\mu \qquad \textit{for all } D \in \mathscr{A}.$$

Proof of Proposition 8.6. Let

$$\lambda_1(D) = \frac{\mu(D)}{\mu(A)} \int_A f d\mu \tag{8.29}$$

$$\lambda_2(D) = \int_D f d\mu \, . \tag{8.30}$$

We then have to prove that $\lambda_1(D) = \lambda_2(D)$ for all $D \in \mathscr{A}$. We may assume without loss of generality that $\int_A f d\mu \geqq 0$. Indeed, the case $\int_A f d\mu < 0$ can be reduced to the case $\int_A f d\mu \geqq 0$ by introducing the function $f^* = -f$.

1. *The proposition is true for each* D_k *in* $\mathscr{D}$.

If k is a nonnegative integer then there is a unique sequence of nonnegative integers $k_0, k_1, \cdots, k_n$ with

$$k_0 = 0 \tag{8.31}$$

$$k_{i+1} = 2k_i + 1 \text{ or } k_{i+1} = 2k_i + 2, \quad i = 0,\ 1, \cdots, n-1 \tag{8.32}$$

$$k_n = k \tag{8.33}$$

hence

$$\int_{D_{k_{i+1}}} f d\mu = \frac{1}{2} \int_{D_{k_i}} f d\mu \tag{8.34}$$

and

$$\mu(D_{k_{i+1}}) = \frac{1}{2}\, \mu(D_{k_i}) \qquad \text{for all } i = 0, \cdots, n-1 \, .$$

Relations (8.34) and (8.35) imply

$$\int_{D_{k_{i+1}}} f d\mu = \frac{\mu(D_{k_{i+1}})}{\mu(D_{k_i})} \int_{D_{k_i}} f d\mu \qquad \text{all } i = 0, \cdots, n-1 \tag{8.36}$$

i.e.,

$$\int_{D_{k_n}} f d\mu = \frac{\mu(D_{k_n})}{\mu(D_{k_0})} \int_{D_{k_0}} f d\mu \tag{8.37}$$

or

$$\int_{D_k} d\mu = \frac{\mu(D_k)}{\mu(A)} \int_A f\,d\mu \qquad \text{all } D_k \in \mathscr{D} \tag{8.38}$$

and

$$\lambda_2(D_k) = \lambda_1(D_k \qquad \text{all } D_k \in \mathscr{D}\,. \tag{8.39}$$

Hence the proposition is true for each $D_k \in \mathscr{D}$.

2. *The proposition is true for each D in the minimal algebra $\mathscr{F}$ over $\mathscr{D}$.* If

$$D \in \mathscr{F} \tag{8.40}$$

then

$$D = \bigcup_{i=1}^{k} D_{n_i} \tag{8.41}$$

with

$$D_{n_i} \cap D_{n_j} = \emptyset \qquad \text{for } i \neq j\;. \tag{8.42}$$

From definition (8.29) and relations (8.41) and (8.42), we obtain

$$\lambda_1(D) = \sum_{i=1}^{k} \lambda_1(D_{n_i})\;. \tag{8.43}$$

Similarly, from definition (8.30) and relations (8.41) and (8.42), we obtain

$$\lambda_2(D) = \sum_{i=1}^{k} \lambda_2(D_{n_i})\;. \tag{8.44}$$

From relations (8.39), (8.43) and (8.44), it follows that

$$\lambda_1(D) = \lambda_2(D)\;. \tag{8.45}$$

Hence the proposition is true for each $D \in \mathscr{F}$.

3. *The proposition is true for each $D \in \mathscr{S}$.*
The position is the following:

(a) λ_1 and λ_2 are two set-valued functions defined over the σ-algebra $\mathscr{B}$.

(b) λ_1 is a measure over $\mathscr{B}$ since μ is a measure over $\mathscr{B}$.

(c) $\lambda_1(D) = \lambda_2(D)$ for all $D \in \mathscr{F}$, where $\mathscr{F}$ is a subalgebra of $\mathscr{B}$.

Since $\mathscr{A}$, the σ-algebra generated by the algebra $\mathscr{F}$, is by construction a subalgebra of the σ-algebra $\mathscr{B}$, it follows from the theorem on the uniqueness of the extension of a measure that $\lambda_1(D) = \lambda_2(D)$ for each $D \in \mathscr{A}$. (See Halmos, Measure Theory, page 54). This concludes the proof of Proposition (8.6).

Proposition 8.7. *If $(A,\mathscr{B},\mu)$ is a nonatomic measure σ-algebra, f a bounded and $\mathscr{B}$ measurable function over A, then there is a nonatomic measure σ-algebra $(A,\mathscr{A},\mu)$ with $\mathscr{A} \subset \mathscr{B}$ such that*

$$\int_D f d\mu = \frac{\mu(D)}{\mu(A)} \int_A f d\mu \qquad \text{for all } D \in \mathscr{A}\ . \tag{8.46}$$

Proof of Proposition 8.7. Let us call $D_0 = A$ and $\mathscr{B}_{D_0} = \mathscr{B}$. Let P_n be the following precedure:

If $(D_n, \mathscr{B}_{D_n}, \mu)$ is a nonatomic measure σ-algebra, then by Propoition (6.1) there are sets D_{2n+1} and $D_{2n+2} \in \mathscr{B}_{D_n}$ such that

$$D_{2n+1} \cap D_{2n+2} = \emptyset \tag{8.47}$$

$$D_{2n+1} \cup D_{2n+2} = D_n \tag{8.48}$$

$$\mu(D_{2n+1}) = \mu(D_{2n+2}) = \frac{1}{2}\, \mu(D_n) \tag{8.49}$$

$$\int_{D_{2n+1}} f d\mu = \int_{D_{2n+2}} f d\mu = \frac{1}{2} \int_{D_n} f d\mu \tag{8.50}$$

and $(D_{2n+1}, \mathscr{B}_{D_{2n+1}}, \mu)$ and $(D_{2n+2} \mathscr{B}_{D_{2n+2}}, \mu)$ are nonatomic measure σ-algebras by Proposition (8.1).

Let us apply recurrently P_n to $(D_n, \mathscr{B}_{D_n}, \mu)$ for $n = 0, 1, 2, \cdots$. This is possible since D_0 and $\mathscr{B}_{D_0}$ are given. Let $\mathscr{A}$ be the minimal σ-algebra generated by $\{D_i : i = 0, 1, 2, \cdots\}$. By construction, $(A,\mathscr{A},\mu)$ is a nonatomic measure σ-algebra and $\mathscr{A} \subset \mathscr{B}$. Moreover, $\mathscr{A}$ satisfies the assumptions of Proposition (8.6), hence

$$\int_D f\,d\mu = \frac{\mu(D)}{\mu(A)}\int_A f\,d\mu \qquad \text{for all } D\in\mathscr{A}\ . \tag{8.51}$$

This concludes the proof of Proposition (8.7).

Proposition 8.8. *If* $(A,\mathscr{B},\mu)$ *is a nonatomic measure* σ*-algebra,* $f=(f^1,\cdots,f^n)$ *an* n*-dimensional vector-valued, bounded and* $\mathscr{B}$ *measurable function over* A*, there is a nonatomic measure* σ*-algebra* $(A,\mathscr{A},\mu)$*,* $\mathscr{A}\subset\mathscr{B}$ *such that*

$$\int_D f\,d\mu = \frac{\mu(D)}{\mu(A)}\int_A f\,d\mu \qquad \textit{for all } D\in\mathscr{A}\ . \tag{8.52}$$

Proof of Proposition 8.8. Let us apply Proposition (8.7) to f^1 over $(A,\mathscr{B},\mu)$ and let $(A,\mathscr{A}_1,\mu)$, $\mathscr{A}_1\subset\mathscr{B}$ be the nonatomic measure σ-algebra so obtained. More generally, let us apply for $i=1,2,\cdots,n-1$, Proposition (8.7) to f^{i+1} over $(A,\mathscr{A}_i,\mu)$ and let $(A,\mathscr{A}_{i+1},\mu)$, $\mathscr{A}_{i+1}\subset\mathscr{A}_i$ be the nonatomic measure σ-algebra so obtained.

Let $\mathscr{A}=\mathscr{A}_n$. We then have

$$\int_D f^i d\mu = \frac{\mu(D)}{\mu(A)}\int_A f^i d\mu, \qquad i=1,\cdots,n \ \text{ for all } \ D\in\mathscr{A} \tag{8.53}$$

i.e.,

$$\int_D f\,d\mu = \frac{\mu(D)}{\mu(A)}\int_A f\,d\mu \qquad \text{for all } D\in\mathscr{A}\ . \tag{8.54}$$

This concludes the proof of Proposition (8.8).

Proposition 8.9. (Lyapounov's Theorem). *If* $(A,\mathscr{B},\mu)$ *is a nonatomic measure* σ*-algebra,* $f=(f^1,f^2,\cdots,f^n)$ *an* n*-dimensional vector-valued, bounded and* $\mathscr{B}$ *measurable function over* A*, then*

$$R=\left\{\int_B f\,d\mu : B\in\mathscr{B}\right\} \tag{8.55}$$

is convex.

Proof of Propositiion 8.9. It is enough to prove that if b_1 and $b_2\in R$, then

$$\{\alpha b_1 + (1-\alpha)b_2 : \alpha \in [0,1]\} \subset R . \tag{8.56}$$

If b_1 and $b_2 \in R$, then there are B_1 and $B_2 \in \mathscr{B}$ such that

$$\int_{B_i} f d\mu = b_i \qquad \text{for } i = 1,2 . \tag{8.57}$$

Let

$$f_i = f\chi(B_i) \qquad \text{for } i = 1,2 \tag{8.58}$$

and

$$f^* = (f_1^1, f_1^2, \cdots, f_1^n, f_2^1, \cdots, f_2^n) . \tag{8.59}$$

If we apply Proposition (8.8) to f^* over $(A,\mathscr{B},\mu)$, we obtain a nonatomic measure σ-algebra $(A,\mathscr{A},\mu)$, $\mathscr{A} \subset \mathscr{B}$. Let $\mathscr{D} = \{D_\alpha : \alpha \in [0,1]\}$ be a subset of $\mathscr{A}$ such that

$$\mu(D_\alpha) = \alpha\mu(A) \qquad \text{all } \alpha \in [0,1] . \tag{8.60}$$

$\mathscr{D}$ exists by Proposition (8.3). We then have

$$R \supset \left\{\int_X f d\mu : X = [D_\alpha \cap B_1] \cup [(A \sim D_\alpha) \cap B_2],\ \alpha \in [0,1]\right\} \tag{8.61}$$

$$= \left\{\int_{D_\alpha} f_1 d\mu + \int_{A\sim D_\alpha} f_2 d\mu : \alpha \in [0,1]\right\} \tag{8.62}$$

$$= \left\{\frac{\mu(D_\alpha)}{\mu(A)}\int_A f_1 d\mu + \frac{\mu(A\sim D_\alpha)}{\mu(A)}\int_A f_2 d\mu : \alpha \in [0,1]\right\} \tag{8.63}$$

$$= \{\alpha b_1 + (1-\alpha)b_2 : \alpha \in [0,1]\} .$$

This concludes the proof of Proposition (8.9).

Proposition 8.10. *If $(A,\mathscr{B},\mu)$ is a nonatomic finite measure σ-algebra X is an n-dimensional Euclidean space, S is a class of bounded $\mathscr{B}$ measurable functions from A to X such that f and $g \in S$ implies*

(8.65) $$f\chi(B) + g\chi(A \sim B) \in S \qquad \text{for all } B \in \mathscr{B}$$

then

(8.66) $$L(S) = \left\{ \int_A f \, d\mu : f \in S \right\}$$

is convex.

Proof of Proposition 8.10. Let f and $g \in S$. We shall prove that there is a set $L(f,g)$ such that

(8.67) (i) $L(f,g)$ is convex

(8.68) (ii) $L(f,g) \subset L(S)$

(8.69) (iii) $\int_A f \, d\mu$ and $\int_A g \, d\mu \in L(f,g)$.

The existence of such a set $L(f,g)$ is a sufficient condition for the convexity of the set $L(S)$.

Let

(8.70) $$L(f,g) = \left\{ \int_A (f\chi(B) + g\chi(A \sim B)d\mu : B \in \mathscr{B} \right\}$$

and

(8.71) $$L^*(f,g) = \left\{ -\int_A g \, d\mu + \alpha : \alpha \in L(f,g) \right\} .$$

We may write $L^*(f,g) = \{\int_A (f-g)\chi(B) \, d\mu : B \in \mathscr{B}\}$, hence $L^*(f,g)$ is convex by Proposition (8.9). The convexity of $L(f,g)$ then follows from the convexity of $L^*(f,g)$. This proves relation (i).

Relation (ii) is an immediate consequence of the definitions of $L(S)$ and $L(f,g)$. Moreover, for $B = A \in \mathscr{B}$ we have $\int_A f \, d\mu \in L(f,g)$ and for $B = \emptyset \in \mathscr{B}$ we have $\int_A g \, d\mu \in L(f,g)$. This proves relation (iii) and concludes the proof of Proposition (8.10).

Proposition 8.11. *The set* $\overset{+}{H}([v])$ *is convex.*

Proof of Proposition 8.11. From the definition of $\overset{+}{H}([v])$ we have

(8.72) $$\overset{+}{H}([v]) = \left\{ \int_{t}^{t_b} \phi(t, u(t); [v]) dt : [u] \in F \right\}$$

where

(8.73) $$\phi(t, u(t); [v]) = G(t; [v])(f(x(t; [v]), u(t), t) - f(x(t; [v]), v(t), t)) \ .$$

From the definition of the class F we know that if $[u_1]$ and $[u_2] \in F$ and if B is a measurable subset of $[t_a, t_b]$ there exists an element $[u] \in F$ such that

(8.74) $$\begin{aligned} u(t) &= u_1(t) \qquad \text{for all } t \in B \\ &= u_2(t) \qquad \text{for all } t \in [t_a, t_b] \sim B \ . \end{aligned}$$

Hence the class S^* of all functions $\phi(t, u(t); [v])$ with $[u] \in F$ satisfies condition (8.65), and we may apply Proposition (8.10). This concludes the proof of Proposition (8.11).

Notations. If f and g are elements of the class S introduced in Proposition (8.10), we shall write

(8.75) $$\sigma(f, g) = \mu(\{t : f(t) \neq g(t), t \in A\})$$

and

(8.76) $$d(f, g) = \operatorname*{ess\,sup}_{t \in A} |f(t) - g(t)| \ .$$

The remarks made in Section 5 also apply to these norms.

Proposition 8.12. *Under the assumptions of Proposition* (8.10), *if* $\int_A f_0 \, d\mu$ *is an interior point of* $L(S)$, *then there is a subseet* $S(f_0)$ *of* S *and two positive constants* m *and* k *such that*

(8.77) (i) $d(f_0, g) \leqq k$ *for all* g *in* $S(f_0)$

(8.78) (ii) *the mapping* $S(f_0) \to L(S(f_0))$ *is one-to-one*

(8.79) (iii) $\int_A f_0 \, d\mu$ *is an interior point of* $L(S(f_0))$

(8.80) (iv) $\sigma(g,h) \leqq m \left| \int_A g\,d\mu - \int_A h\,d\mu \right|$ *for all* g *and* $h \in S(f_0)$.

Proof of Proposition 8.12.

Notations. If A is a set, then int A is the set of interior points of A, co A is the convex hull of A, and int co A is the set of interior points of the convex hull of A.

If $\int_A f_0 d\mu \in \operatorname{int} L(S)$, then there are functions $f_1, f_2, \cdots, f_{n+1}$ in S such that $l_0 \in \operatorname{int}\operatorname{co} M$ where $M = \{l_i : i = 0, 1, 2, \cdots, n+1\}$ with $l_i = \int_A f_i d\mu$ for $i = 0, 1, \cdots, n+1$. Let $M_i = (M \sim \{l_i\})$ for $i = 0.1, 2, \cdots, n+1$. For every $l \in \operatorname{co} M$ let $\lambda(l) = \{\lambda_i(l) : i = 0, 1, \cdots, n+1\}$ be defined by the following rules:

(i) k is the smallest integer such that $l \in \operatorname{co} M_k$

(ii) $\lambda_k(l) = 0$

(iii) $\lambda(l) \sim \{\lambda_k(l)\}$ are the barycentric coordinates of l with respect to M_k.

By construction there exists a positive constant m such that

$$\sum_{i=0}^{n+1} |\lambda_i(l') - \lambda_i(l'')| \leqq m\,|l' - l''| \text{ for all } l' \text{ and } l'' \text{ in co } M . \tag{8.81}$$

Let

$$\phi = \{f_0^1, f_0^2, \cdots, f_0^n, f_1^1, \cdots, f_1^n, f_2^1, \cdots, f_{n+1}^n\} . \tag{8.82}$$

Let $(A, \mathscr{A}, \mu)$ with $\mathscr{A} \subset \mathscr{B}$ be a nonatomic measure σ-algebra such that

$$\int_D \phi\,d\mu = \frac{\mu(D)}{\mu(A)} \int_A \phi\,d\mu \qquad \text{for all } D \in \mathscr{A} . \tag{8.83}$$

Such an $\mathscr{A}$ exists by Proposition (8.8).

Let $\{D_i : i = 0, 1, \cdots, n+1\}$ be such that

(8.84) (i) $D_i \in \mathscr{A}$

(8.85) (ii) $D_i \cap D_j = \emptyset$ if $i \neq j$

(8.86) (iii) $\bigcup_{i=0}^{n+1} D_i = A$

(8.87) (iv) $\mu(D_i) = \dfrac{1}{n+2}\mu(A)$.

Such a set $\{D_i : i = 0, 1, 2, \cdots, n + 1\}$ exists by Proposition (8.4).

Let $\mathscr{D}_i$ be a nest $\{D_i^\alpha : \alpha \in [0,1]\} \subset \mathscr{A}_{D_i}$ such that

(8.88) (i) $\alpha < \alpha'$ implies $D_i^\alpha \subset D^{\alpha'}$

(8.89) (ii) $D_i^1 = D_i$

(8.90) (iii) $\mu(D_i^\alpha) = \alpha\mu(D_i)$.

Such a nest $\mathscr{D}_i$ exists for all $i = 0, 1, 2, \cdots, n + 1$ by Proposition (8.3).

We shall use the following convention: if α is a positive number, the set α co M is defined by

(8.91) $$\alpha \text{ co } M = \{l_0 + \alpha(l - l_0) : l \in \text{co } M\} .$$

Let

(8.92) $$f_{(l)} = f_0\chi\left(A \sim \bigcup_{i=1}^{n+1} D_i^{\lambda_i((n+2)l)}\right) + \sum_{i=1}^{n+1} f_i\chi\left(D_i^{\lambda_i((n+2)l)}\right)$$

$$\text{for all } l \in \frac{1}{n+2} \text{co } M .$$

By construction, we have:

(8.93) $$\sigma(f_{(l')}, f_{(l'')}) = \sum_{i=0}^{n+1} |\lambda_i(l') - \lambda_i(l'')|$$

$$\text{for all } l' \text{ and } l'' \in \frac{1}{n+2} \text{co } M .$$

Let

(8.94) $$S(f_0) = \left\{f_{(l)} : l \in \frac{1}{n+2} \text{co } M\right\} .$$

We have $S(f_0) \subset S$ since $f_{(l)} \in S$ for all $l \in \frac{1}{n+2}$ co M because $\mathscr{A}_{D_i} \subset \mathscr{A} \subset \mathscr{B}$ for all $i = 0, 1, 2, \cdots, n + 1$ by Proposition (8.2).

We also have, by construction,

(8.95) $$\int_A f_{(l)} d\mu = l \quad \text{for all } l \in \frac{1}{n+2} \text{co } M$$

i.e.,

$$L(S(f_0)) = \frac{1}{n+2} \operatorname{co} M \ . \tag{8.96}$$

But since

$$l_0 \in \operatorname{int} \frac{1}{n+2} \operatorname{co} M \tag{8.97}$$

then relation (8.96) implies

$$l_0 \in \operatorname{int} L(S(f_0)) \ . \tag{8.98}$$

Relation (8.98) is the required property (8.79). Property (8.78) is satisfied by virtue of relations (8.94) and (8.95). Property (8.77) follows from the fact that the class $S(f_0)$ has been constructed from a finite number of bounded mesurable functions. From relations (8.81) and (8.93), we have

$$\sigma(f_{(l')}, f_{(l'')}) \leqq m \,|\, l' - l'' \,| \qquad \text{for all } l' \text{ and } l'' \in \frac{1}{n+2} \operatorname{co} M \ . \tag{8.99}$$

Relation (8.99) is equivalent to property (8.80) by virtue of definition (8.94) This concludes the proof of Proposition (8.12).

Proposition 8.13. *If* 0 *is an interior point of* $\overset{+}{H}([v])$ *then there is a subset* $F([v])$ *of* F *and two constants* m *and* k *such that*

(8.100) (i) $d([u],[v]) \leqq k$ *for all* $[u] \in F([v])$

(8.101) (ii) *the mapping*: $F([v]) \to \{\overset{+}{y}(t_b; [u],[v]) : [u] \in F([v])\}$ *is one-to-one*

(8.102) (iii) 0 *is an interior point of* $\{\overset{+}{y}(t_b; [u],[v]) : [u] \in F([v])\}$

(8.103) (iv) $\sigma([u_1],[u_2]) \leqq m \,|\, \overset{+}{y}(t_b; [u_1],[v]) - \overset{+}{y}(t_b; [u_1],[v]) \,|$
for all $[u_1]$ and $[u_2] \in F([v])$.

Proof of Proposition 8.13. Let S^* be the class of functions introduced in Proposition (8.11). By replacing S by S^* in the proof of Proposition (8.12), we immediately obtain Proposition (8.13).

SECTION 9

An Application of Brouwer's Fixed Point Theorem

In this section we prove a single proposition which will play a fundamental role in the proof of Theorem III, given in Section 10.

Proposition 9.1. *If f is a continuous mapping of a ball $S^\eta = \{x : x \in E, |x| \leqq \eta\}$, with $\eta > 0$, of the Euclidean space E into E such that there exists a function $g(r)$ defined, continuous and nondecreasing over $[0, \eta]$ and having the following properties:*

$$(9.1) \qquad \text{(i)} \quad \lim_{r \to 0} \frac{g(r)}{r} = 0$$

$$(9.2) \qquad \text{(ii)} \quad |a - f(a)| \leqq g(|a|) \qquad \textit{for all } a \in S^\eta$$

then 0 is an interior point of the set $f(S^\eta)$ image of the set S^η through the mapping f.

Proof of Proposition 9.1. Let $\rho \in (0, \eta]$ be such that $g(\rho) < \frac{\rho}{2}$, i.e., $\frac{2g(\rho)}{\rho} < 1$. Such a ρ exists since $\frac{2g(\rho)}{\rho}$ is continuous over $(0, \eta]$ and $\lim_{\rho \to 0} \frac{2g(\rho)}{\rho} = 0$ by assumption.

Let $h_z(x) = z + x - f(x)$. If $z \in S^{\rho/2}$ then h_z maps S^ρ into itself since for $x \in S^\rho$ we have

$$(9.3) \quad |h_z(x)| \leqq |z| + |x - f(x)| \leqq \frac{\rho}{2} + g(|x|) \leqq \frac{\rho}{2} + g(\rho) \leqq \frac{\rho}{2} + \frac{\rho}{2} = \rho \,.$$

Moreover, the continuity of the mapping f implies the continuity of the mapping h_z, hence by Brouwer's Fixed Point Theorem, there exists an x_1 such that

$$(9.4) \qquad h_z(x_1) = x_1$$

i.e.,

$$(9.5) \qquad z = f(x_1) \,.$$

For all $z \in S^{\rho/2}$ there exists an x_1 such that relation (9.5) holds, hence

$$S^{\rho/2} \subset f(S^{\rho}) \subset f(S^{\eta}) . \tag{9.6}$$

The point 0 is then an interior point of the set $f(S^{\eta})$. This concludes the proof of Proposition (9.1).

Remark. A weaker form of this proposition, corresponding to $g(r) = Mr^2$ for some $0 < M < +\infty$, has been introduced without proof in a previous publication [12]. An elegant proof for the case $g(r) = Mr^2$ has been communicated to the author by Dr. G. S. Jones of RIAS. The proof given here for the larger class of functions g described in the statement of Proposition (9.1) is a generalization of Dr. G. S. Jones' proof.

SECTION 10

Proofs of the Theorems of Section 4

Proof of Theorem I. If $x = x(t_b;[v])$ is an interior point of the set H, there is an $\varepsilon > 0$ such that

$$\xi = (x^1(t_b;[v]), x^2(t_b;[v]), \cdots, x^{n-1}(t_b;[v]), x^n(t_b;[v]) + \varepsilon) \tag{10.1}$$

is also a point of the set H, hence there is a $[u]$ in F^* such that

$$x(t_b;[u]) = \xi . \tag{10.2}$$

We then have

$$x^n(t_b;[u]) = x^n(t_b;[v]) + \varepsilon > x^n(t_b;[v]) . \tag{10.3}$$

But by construction we also have

$$(\xi, t_b) \in B . \tag{10.4}$$

Relations (10.2), (10.3) and (10.4) contradict the assumed optimality of the element $[v]$ of F^*, hence $x = x(t_b;[v])$ is a boundary point of the set H. This concludes the proof of Theorem I.

Proof of Theorem II. The mapping $H \to H([v])$ is defined by

$$y = G(t_b;[v])(x - x(t_b;[v])) . \tag{10.5}$$

But $G(t_b;[v])$ is the identity matrix, hence $H([v])$ is a simple translation of H. This translation conserves the topological properties of the points in H; in particular to a boundary point of H corresponds a boundary point of $H([v])$ and conversely.

Proof of Theorem III. Let us assume that 0 is interior to $\overset{+}{H}([v])$ and show that 0 is then interior to H($[v]$).

If 0 is interior to $\overset{+}{H}([v])$ then by Proposition (8.13) we know that there exists a subset $F([v])$ of F and two constants m and k such that

(10.6) (i) 0 is interior to $\{\overset{+}{y}(t_b;[u],[v]):[u]\in F([v])\}$

(10.7) (ii) the mapping $F([v])\to\{\overset{+}{y}(t_b;[u],[v]):[u]\in F([v])\}$ is one-to-one

(10.8) (iii) $d([u],[v])\leqq k$ for all $[u]\in F([v])$

(10.9) (iv) $\sigma([u_1],[u_2])\leqq m\,|\overset{+}{y}(t_b;[u_1],[v])-\overset{+}{y}(t_b;[u_2],[v])|$ for all $[u_1]$ and $[u_2]\in F([v])$.

Let $F([v],Q)$ be the subset of $F([v])$ defined by

$$F([v],Q)=\{[u]:[u]\in F([v]),\sigma([u],[v])\leqq Q\} \quad (10.10)$$

or all $[u]\in F([v])$ such that

$$|\overset{+}{y}(t_b;[u],[v])|\leqq\frac{Q}{m} \quad (10.11)$$

we have by applying (10.9),

$$\sigma([u],[v])\leqq Q \quad (10.12)$$

which implies

$$[u]\in F([v],Q)\ . \quad (10.13)$$

We may then write

(10.14) (i) 0 is interior to $\{ \overset{+}{y}(t_b;[u],[v]) : [u] \in F([v],Q)\}$

(10.15) (ii) the mapping $F([v],Q) \to \{ \overset{+}{y}(t_b;[u],[v]) : [u] \in F([v],Q)\}$ is one-to one

(10.16) (iii) $d([u],[v]) \leqq k$ for all $[u] \in F([v],Q)$

(10.17) (iv) $\sigma([u_1],[u_2]) \leqq m \,| \overset{+}{y}(t_b;[u_1],[v]) - \overset{+}{y}(t_b;[u_2],[v])|$
for all $[u_1]$ and $[u_2] \in F([v],Q)$.

From Proposition (6.13) we then have

(10.18) $$F([v],Q) \subset F^*([v],k) \subset F^*$$

i.e.,

(10.19) $y(t_b;[u],[v])$ exists for all $[u] \in F([v],Q)$.

We also know that

(10.20) (i) the mapping from $\{ \overset{+}{y}(t_b;[u],[v]) : [u] \in F([v],Q)\}$ to $F([v],Q)$ is continuous (see relation (10.17))

(10.21) (ii) the mapping from $F([v],Q)$ to $\{y(t_b;[u],[v]) : [u] \in F([v],Q)\}$ is continuous (see Proposition (6.13)).

(10.22) Hence the mapping from $\{ \overset{+}{y}(t_b;[u],[v]) : [u] \in F([v],Q)\}$ to $\{y(t_b;[u],[v]) : [u] \in F([v],Q)\}$ is continuous.

From proposition (7.13) we know that there exists a function $g(r)$, defined, continuous and nondecreasing over $[0,Q]$ such that

(10.23) (i) $\lim_{r\to 0} \frac{g(r)}{r} = 0$

(10.24) (ii) $| \overset{+}{y}(t_b;[u],[v]) - y(t_b;[u],[v])| \leqq g(\sigma([u],[v]))$
for all $[u] \in F([v],Q)$.

We define the function $G(r)$ by the relation

$$G(r) = g(mr) \ . \tag{10.25}$$

From relations (10.17), (10.23) and (10.24) we know that there exists a function $G(r)$, defined, continuous, nondecreasing over $[0, Q/m]$ and such that

$$\text{(i)} \lim_{r \to 0} \frac{G(r)}{r} = 0 \tag{10.26}$$

$$\text{(ii)} \ |\overset{+}{y}(t_b;[u],[v]) - y(t_b;[u],[v])| \leqq G(|\overset{+}{y}(t_b;[u],[v])|) \ . \tag{10.27}$$

We apply Proposition (9.1) to relations (10.14), (10.22), (10.26) and (10.27) and we obtain

$$0 \text{ is interior to } \{y(t_b;[u],[v]) : [u] \in F([v], Q)\} \tag{10.28}$$

and a fortiori

$$0 \text{ is interior to } H([v]) \ . \tag{10.29}$$

This concludes the proof of Theorem III .

Proof of Theorem IV. By Proposition (8.11), the set $\overset{+}{H}([v])$ is convex, hence if $y = 0$ is a boundary point of $\overset{+}{H}([v])$ there exists a hyperplane

$$\langle \pi([v]) \mid y \rangle = 0 \tag{10.30}$$

such that

$$\langle \pi([v]) \mid y \rangle \leqq 0 \qquad \text{for all} \quad y \in \overset{+}{H}([v]) \ . \tag{10.31}$$

Let us assume that there is a $[u] \in F$ such that

$$\langle \pi([v]) \mid G(t;[v])(f(x(t;[v]), u(t), t) - f(x(t;[v]), v(t), t)) \rangle \geqq \varepsilon > 0 \tag{10.32}$$

for $t \in E$, where $E \in \mathscr{B}$ with $\mu(E) > 0$, then by introducing the vector-valued function

$$u^*(t) = v(t) + \chi(E)(u(t) - v(t)) \text{ all } t \in [t_a, t_b] \tag{10.33}$$

we obtain

$$\langle \pi([v]) \mid \overset{+}{y}(t_b; [u^*], [v]) \rangle \geqq \varepsilon\mu(E) > 0 \; . \tag{10.34}$$

We also have

$$\overset{+}{y}(t_b; [u^*], [v]) \in \overset{+}{H}([v]) \tag{10.35}$$

since $[u^*] \in F$. Relations (10.31), (10.34) and (10.35) are contradictory. This concludes the proof of Theorem IV.

Proof of Theorem V. Let the vector $p(t;[v])$ be defined by relation (4.2). This vector ($p(t;[v])$) is nonidentically zero and continuous over $[t_a, t_b]$ since $G(t;[v])$ and $G^{-1}(t;[v])$ exist and are bounded over $[t_a, t_b]$. Relation (4.1) may be written under the form:

$$\langle \pi([v]) \mid G(t;[v])(f(x(t;[v]), u(t), t) - f(x(t;[v]), v(t), t)) \rangle \leqq 0 \tag{10.36}$$
$$\text{for all } [u] \text{ in } F \text{ and a.e. } t \in [t_a, t_b]$$

or, from the definition of a transposed matrix

$$\langle G^T(t;[v])\pi([v]) \mid f(x(t;[v]), u(t), t) - f(x(t;[v]), v(t), t) \rangle \leqq 0 \tag{10.37}$$
$$\text{for all } [u] \text{ in } F \text{ and a.e. } t \in [t_a, t_b]$$

i.e., from relation (4.2),

$$\langle p(t;[v]) \mid f(x(t;[v]), u(t), t) - f(x(t;[v]), v(t) t) \rangle \leqq 0 \tag{10.38}$$
$$\text{for all } [u] \text{ in } F \text{ and a.e. } t \in [t_a, t_b] \; .$$

This proves relation (4.3). By differentiation with respect to t of relation (4.2), we obtain

$$\begin{aligned} \dot{p}(t;[v]) &= (G^T(t;[v]))^{\cdot}\pi([v]) \\ &= (\dot{G}(t;[v]))^T\pi([v]) \text{ a.e. } t \in [t_a, t_b] \end{aligned} \tag{10.39}$$

but by definition (see relation (2.3)), we have

(10.40) $$\dot{G}(t;[v]) = -G(t;[v])D(t;[v]) \text{ a.e. } t \in [t_a, t_b]$$

hence

(10.41) $$\begin{aligned} \dot{p}(t;[v]) &= (-G(t;[v])D(t;[v]))^T\pi([v]) \\ &= -D^T(t;[v])G^T(t;[v])\pi([v]) \\ &= -D^T(t;[v])p(t;[v]) . \end{aligned}$$

This proves the relation (4.4) and concludes the proof of Theorem V.

Proof of Theorem VI. This theorem is just a logical conclusion of Theorem II, III, IV and V: if the point $x(t_b;[v])$ is a boundary point of the set H, then the point $y = 0$ is a boundary point of the set $H([v])$ (see Theorem II), then the point $y = 0$ is a boundary point of the set $\overset{+}{H}([v])$ (see Theorem III), then there exists a non-zero constant vector $\pi([v])$ such that condition (4.1) is satisfied for all $[u]$ in F (see Theorem IV), then there exists a vector $p(t;[v])$ continuous, nonidentically zero on $[t_a, t_b]$ and satisfying conditions (4.3) and (4.4)(see Theorem V). This concludes the proof of Theorem VI.

Proof of Therem VII. This theorem is just a logical conclusion of Theorems I and VI: if an element $[v]$ of F^* is optimal, then the point $x(t_b;[v])$ is a boundary point of the set H (see Theorem I), then there exists a vector $p(t;[v])$ continuous, nonidentically zero on $[t_a, t_b]$ and satisfying conditions (4.3) and (4.4)(see Theorem VI). This concludes the proof of Theorem VII.

REFERENCES

1. M. Aoki, Minimal effort control systems with an upper bound of the control time, IEEE *Transactions on Automatic Control*, No. 1, 1963, pp. 60–61.

2. L. D. Berkovitz, An optimum thrust control problem, *J. Math. Anal. Appl.*, Vol. 3, 1961, pp. 122–132.

3. D. Blackwell, The range of a vector measure, *Proc. Amer. Math. Soc.*, Vol. 2, 1951, pp. 390–395.

4. P. Finsler, Über Kurven und Flächen in Allgemeinen Räumen, Birkhäuser, 1961.

5. I. Flügge-Lotz, Discontinuous automatic control, Princeton University Press, Princeton, N. J., 1953.

6. I. Flügge-Lotz, and H. Halkin, Pontryagin's maximum principle and optimal control, Technical Report No. 130, Department of Engineering Mechanics, Stanford University, 1961.

7. B. Fraeijs de Veubeke, Méthodes variationnelles et performances optimales en aéronautique, *Bull. Soc. Math. de Belgique*, Vol. 8, 1956, pp. 136-157.

8. A. T. Fuller, Bibliography of optimum nonlinear control of determinate and stochastic-definite systems, *J. Electronics Control*, Vol. 13, 1962, pp. 589-611.

9. J. Hadamard, Le principe de Huyghens, *Bull. Soc. Math. de France*, Vol. 52, 1924, pp. 610–640.

10. H. Halkin, Trajectoires optimales, Engineering Dissertation, Université de Liège 1960.

11. H. Halkin, Nondegenerate variational problems and the principle of optimal evoution, *Proceedings of the Symposium on Vehicle Systems Optimization, Institute for the Aerospace Sciences*, New York, 1962, pp. 43–44.

12. H. Halkin, Lyapounov's theorem on the range of a vector measure and Pontryagin's maximum principle, *Arch. Rational Mech. Anal.*, Vol. 10, 1962, pp. 296–304.

13. H. Halkin, The principle of optimal evolution, in "Nonlinear differential equations and nonlinear mechanics", J. P. LaSalle and S. Lefschetz ed., Academic Press, New York, 1936, pp. 284–302.

14. R. E. Kalman, Canonical structure of linear dynamical systems, *Proc. Nat. Acad. Sci.*, Vol. 48, 1962, pp. 596–600.

15. J. P. LaSalle, The time optimal control problem, in "Contributions to the theory of nonlinear oscillations", Vol. V, Princeton University Press, 1960, pp. 1–24.

16. A. Lyapounov, Sur les fonctions vecteurs complètement additives, *Bull. Acad. Sci. U. R. S. S.*, Vol. 4, 1940, pp. 465–478. [In Russian with a French Résumé].

17. E. J. McShane, On multipliers for Lagrange problems, *Amer. J. Math.*, Vol. 61, 1939, pp. 809–819.

18. L. W. Neustadt, The existence of optimal controls in the absence of convexity conditions, to appear in *J. Math. Anal. Appl.*

19. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, The mathematical theory of optimal processes, English translation, L. W. Neustadt ed., Interscience, New York, 1962.

20. P. A. Samuelson, Efficient paths of capital accumulation in terms of the calculus of variations, in "Mathematical Methods in the social sciences", Standford University Press, 1959.

Department of Mathematics
Stanford University
Stanford, California, U. S. A.
and
Bell Telephone Laboratories
Whippany, N.J., U.S.A.

(Received June 10, 1963).