Control and Cybernetics

Vol. 19 (1990) No. 1-2

Generalized canonical formalism with applications

by

JÓZEF BARANOWICZ
ANDRZEJ NOWAKOWSKI

Institute of Mathematics
Łódź University
ul. S. Banacha 22
90-238 Łódź, Poland

The paper aims at giving a generalization of the classical Hamilton-Jacobi canonical formalism to the case when the Hamilton function H depends on the control function (parameter) u, $H(t, x, y) = \max_u H(t, x, y, u)$. The theory is illustrated in two fields: mechanics and optimal control theory. The theorem on Poincare-Cartan's integral invariant is generalized and the sufficient condition for a minimum in a control problem is given.

Introduction

Many modern questions of astronautics, mechanics and the theory of extremal problems are described by means of functions which depend on an additional control functions that plays the role of a parameter (see, for instance, [4], Ch. II. 2.).

The control function is, in general, only a measurable function in time (at best, piecewise continuous) and this, in turn, causes that the functions determining processes are only measurable, too. Examples of such problems can be found in astronautics, mechanics, optimization theory (see e.g. [7], intr. to part II). The fundamental function which then appears is the Hamilton function H which, in most cases, is not a continuous function in time.

In the classical theory of the Hamilton-Jacobi canonical formalism ([1], [3]) it is assumed that H is at least a C^1-function and that the flow trajectories, i.e. the integral curves of canonical equations, can never intersect. That is why the classical theory leaves out, except particular cases, the problems mentioned above. Therefore, it appears to be quite natural to look for the generalization of the classical formalism in such a way that it would be possible to apply it to most of the problems described by a control function.

What this paper presents is just an attempt of such a generalization.

In Section 1 we define an abstract Hamilton function $H(t, x, y) = \max H(t, x, y, u)$ and the 1-form $ydx - Hdt$ associated with it. Next, we note (Th. 1.1), under additional assumptions on H, that, as in the classical investigations, the characteristics of the form $ydx - Hdt$ are the pairs $(x(t), y(t))$ which satisfy the system of canonical equations (1.4), (1.5). This fact is the starting point for further considerations. Namely, omitting the additional assumptions on H, we consider some family of functions satisfying (1.4), (1.5) which depend on parameters σ, ρ. Next, we study the conditions this family should fulfil in order that the integral of $ydx - Hdt$ should possess all the basic properties such as: invariability (Theorem of Poincaré - Cartan), completeness (independence of integration path), the relationship with the activity function - the value function.

It becomes evident that we can omit most of the classical smoothness assumptions and we shall still have the above - mentioned properties. Moreover, the assumption about non-intersecting trajectories can be replaced by an essentially weaker one about the descriptivity of a suitable transformation.

The general results obtained in the paper are illustrated in two fields: mechanics and optimal control theory. The generalization of the theorem on Poincaré - Cartan's integral invariant is proved and some of its consequences are derived. Adapting the theory obtained to control problems, we prove the sufficient condition for a minimum of the functional in the classical optimization problem. This condition (Th. 6.1) generalizes the sufficient condition of Weierstrass from the calculus of variations ([7], Ch. I.).

The substantial advantage of the method we use here is the avoidance of multivalued functions that appear in a natural way if one allows the extremal trajectories to intersect (comp. [7], [6]). As a consequence, we obtain an essential simplification of ideas and calculations in many considerations.

1. Canonical differential equations

Let U be a Borel subset of the Euclidean space R^m. We shall be dealing with a measurable function $u(t) : [a, b] \to U$ where $[a, b]$ is any interval of R^1. In the sequel, $u(t)$ will appear as a parameter or *control function*. Along with the function u, we shall consider a function $\bar{u}$ (feedback function) where $\bar{u} = \bar{u}(t, x, y) : R \times R^n \times R^n \to U$.

We use H to denote a scalar function of variables t, x, y, u and suppose the following hypothesis satisfied:

the function $H(t, x, y, u)$ and its partial derivatives

(h_1) $H_x(t, x, y, u)$, $H_y(t, x, y, u)$ are continuous in the product $R \times R^n \times R^n \times U$.

We assume that the *feedback function* $\bar{u}$ satisfies the *maximum relation*

$$\mathrm{H}(t, x, y) = H(t, x, y, \bar{u}(t, x, y)) = \sup_{u \in U} H(t, x, y, u) \tag{1.1}$$

for any $\quad (t, x, y) \in R^{2n+1}$

The function $H = H(t, x, y)$ of (1.1) will be called the *Hamilton function.*

Suppose, for the use in Section 1 only, that $H(t, x, y)$ is a C^1-function in the space R^{2n+1}.

Let $(\tilde{t}, \tilde{x}, \tilde{y})$ be any point of R^{2n+1}. To motivate further considerations, we shall impose, for this section only, the requirement that at the point $(\tilde{t}, \tilde{x}, \tilde{y})$ the derivatives

$\dfrac{\partial}{\partial x} H(\tilde{t}, \tilde{x}, \tilde{y}, \bar{u}(t, x, y)), \qquad \dfrac{\partial}{\partial y} H(\tilde{t}, \tilde{x}, \tilde{y}, \bar{u}(t, x, y))$ are continuous and

satisfy the relationships

$$\begin{aligned} \frac{\partial H(t, x, y)}{\partial x} &= \frac{\partial}{\partial x} H(t, x, y, \bar{u}(\tilde{t}, \tilde{x}, \tilde{y})) + \frac{\partial}{\partial x} H(\tilde{t}, \tilde{x}, \tilde{y}, \bar{u}(t, x, y)), \\ \frac{\partial H(t, x, y)}{\partial y} &= \frac{\partial}{\partial y} H(t, x, y, \bar{u}(\tilde{t}, \tilde{x}, \tilde{y})) + \frac{\partial}{\partial y} H(\tilde{t}, \tilde{x}, \tilde{y}, \bar{u}(t, x, y)). \end{aligned} \tag{1.2}$$

Formulae (1.2) are to hold at the point $(\tilde{t}, \tilde{x}, \tilde{y})$.

Now, we can establish, in the space R^{2n+1}, a differential 1-form

$$\omega = y dx - H dt. \tag{1.3}$$

Since $H(t, x, y)$ is a C^1-function, there exists a differential 2-form $d\omega$ and we may speak about "*the field of directions of the rotation*" (see [1], § 44) of the form ω. Integral curves of this field are called "*rotational lines*" (comp. [1], § 44) or *characteristics* of the form ω.

We prove the following

Theorem 1.1 *Characteristics of form* (1.3) *in the* $(2n + 1)$*-dimensional Euclidean space of variables* t, x, y *have the description:* $x = x(t), y = y(t), t \in (\alpha, \beta)$. *The functions* $x(t), y(t), t \in (\alpha, \beta)$, *satisfy the canonical differential equations*

$$\dot{x} = H_y(t, x, y, \bar{u}(t, x, y)), \tag{1.4}$$

$$\dot{y} = -H_x(t, x, y, \bar{u}(t, x, y)) \tag{1.5}$$

where the dots over x *and* y *denote the derivatives in* t.

Proof. First, we note (by (1.1), (1.2)) that if we fix the first three variables in $H(t, x, y, \bar{u})$ by setting $t = \tilde{t}$, $x = \tilde{x}$, $y = \tilde{y}$, then

$$\frac{\partial}{\partial x} H(\tilde{t}, \tilde{x}, \tilde{y}, \bar{u}(t, x, y)) = 0,$$

$$\frac{\partial}{\partial y} H(\tilde{t}, \tilde{x}, \tilde{y}, \bar{u}(t, x, y)) = 0.$$

Hence the differential of form (1.3) has the form

$$d\omega = dy \wedge dx - H_x(t, x, y, \bar{u})\, dx \wedge dt - H_y(t, x, y, \bar{u}) dy \wedge dt.$$

The matrix of this 2-form in the variables y, x, t is equal to

$$A = \begin{bmatrix} 0 & -E & H_y \\ E & 0 & H_x \\ -H_y & -H_x & 0 \end{bmatrix}$$

where E is the unit matrix of degree n. The rank of A is $2n$, so the form $d\omega$ is non-singular. It is easy to check that the vector $(-H_x, H_y, 1)$ is the eigenvector of the matrix A with the eigenvalue zero for any fixed triplet (t, x, y). That vector defines the direction of the rotation of ω. Therefore the characteristics of ω have to satisfy equations (1.4) and (1.5), so the proof is now completed. ■

Let us note that from the above proof it follows that each pair $(x(t), y(t))$ satisfying (1.4), (1.5) is a characteristic of form (1.3).

In what follows we shall only consider functions $x(t)$, $y(t)$ which satisfy canonical differential equations (1.4), (1.5), but with essentially weaker assumptions imposed on the function H.

2. Canonical flow

From now on, we shall only suppose the function H to satisfy (h_1). All further assumptions will concern some special family of functions. We want to stress that upon the feedback function $\bar{u}(t, x, y)$ no requirements are imposed.

Each function $x(t)$, $t \in (t_1, t_2)$, $(t_1, t_2) \subset [a, b]$, that satisfies (1.4) will be termed a *trajectory* (comp. [7], p. 263) and each pair of functions $x(t)$, $y(t)$, $t \in (t_1, t_2)$, satisfying (1.4), (1.5) a *canonical pair* (comp. [7], p. 265). A trajectory and a canonical pair will usually be associated with the corresponding control function $u(t) = \bar{u}(t, x(t), y(t))$ and, for brevity, we shall speak of the trajectory

$$(x(t), \ u(t)) \tag{2.1}$$

and the canonical pair

$$x(t),\ \ y(t),\ \ u(t). \tag{2.2}$$

Consider extended real-valued functions $t^-(\sigma)$, $t^+(\sigma)$ where $-\infty \leqslant t^-(\sigma) < t^+(\sigma) \leqslant +\infty$, $\sigma \in G$ and G is an open Euclidean set. The points σ at which $t^-(\sigma) \neq +\infty$ or $t^+(\sigma) \neq +\infty$ are to form an open set and the two functions are to be C^1-functions, except for the case of $t^-(\sigma)$ or $t^+(\sigma)$ at any point where the values $-\infty$ or $+\infty$ are taken.

We further suppose that a set R of points ρ is also open in a Euclidean space and we put $\tilde{G} = G \times R$. By Z we shall denote the set of (t, σ) for which $\sigma \in G$ and t is subject to the condition

$$-\infty < t^-(\sigma) \leqslant t \leqslant t^+(\sigma) < +\infty, \tag{2.3}$$

and let Z^* stand for the set of (t, σ, ρ) for which $(\sigma, \rho) \in \tilde{G}$ and t satisfies (2.3).

Having introduced the above sets, we consider a family F of canonical pairs given by

$$x(t, \sigma, \rho),\ \ y(t, \sigma, \rho),\ \ u(t, \sigma, \rho) \qquad (t, \sigma, \rho) \in Z^*. \tag{2.4}$$

Here σ, ρ are "labels" which distinguish a member of the family, i.e. σ, ρ remain constant on a member of F and this member then corresponds to the interval $t^-(\sigma) \leqslant t \leqslant t^+(\sigma)$. We use D to denote the set of triples (t, x, y) where $x = x(t, \sigma, \rho)$, $y = y(t, \sigma, \rho)$, $(t, \sigma, \rho) \in Z^*$.

We say that the map $\varphi: Z^* \to D$ is *descriptive* if, for each point $(t, \sigma, \rho) \in Z^*$, the following condition is satisfied: for each rectifiable curve $C \subset D$ (i.e. with a finite length) with the initial point $\varphi(t, \sigma, \rho)$, there exists a rectifiable curve $\Gamma \subset Z^*$ with the initial point (t, σ, ρ) such that each sufficiently small arc $\tilde{C} \subset C$ with the same initial point as C is the image under the map φ of a sufficiently small arc $\tilde{\Gamma} \subset \Gamma$ starting from (t, σ, ρ) (comp. [7], p.266)

Let us set

$$f(t, x, y) = H_y(t, x, y, \bar{u}(t, x, y)), \tag{2.5}$$

$$L(t, x, y) = yf(t, x, y) - H(t, x, y, \bar{u}(t, x, y)). \tag{2.6}$$

Finally, when $(t, \sigma, \rho) \in Z^*$, we shall write

$$\tilde{L}(t, \sigma, \rho),\ \ \tilde{f}(t, \sigma, \rho),\ \int_t^{t+(\delta)} \tilde{L}_\sigma(\tau, \sigma, \rho)\, d\tau$$

instead of the expressions

$$L(t, x(t, \sigma, \rho), y(t, \sigma, \rho)), \quad f(t, x(t, \sigma, \rho), y(t, \sigma, \rho)),$$

$$\int_t^{t+(\delta)} \tilde{L}_\sigma(\tau, \sigma, \rho)\, d\tau + \tilde{L}(t^+(\sigma), \sigma, \rho)\, t_\sigma^+(\sigma).$$

We now suppose the following hypotheses satisfied:

(h_2) the function $\tilde{L}(t, \sigma, \rho)$ is continuous in Z^*;
there exist continuous derivatives $\tilde{L}_\rho(t, \sigma, \rho)$, $\tilde{L}_\sigma(t, \sigma, \rho)$, $\tilde{f}_\rho(t, \sigma, \rho)$, $\tilde{f}_\sigma(t, \sigma, \rho)$, $\tilde{L}_{\sigma\rho}(t, \sigma, \rho)$ in Z^*;
for each fixed $(t, x, y) \in D$ there exist derivatives

$\frac{\partial}{\partial \rho} H(t, x, y, u(t, \sigma, \rho))$, $\frac{\partial}{\partial \sigma} H(t, x, y, u(t, \sigma, \rho))$, satisfying the relations

$$\frac{\partial \tilde{L}}{\partial \rho} = y\tilde{f}_\rho - H_x x_\rho - \frac{\partial}{\partial \rho} H(t, x, y, u(t, \sigma, \rho)),$$

$$\frac{\partial \tilde{L}}{\partial \sigma} = y\tilde{f}_\sigma - H_x x_\sigma - \frac{\partial}{\partial \sigma} H(t, x, y, u(t, \sigma, \rho)),$$

for any (t, x, y) where $x = x(t, \sigma, \rho)$, $y = y(t, \sigma, \rho)$ and $(t, \sigma, \rho) \in Z^*$;

(h_3) the function $y(t, \sigma, \rho)$ is continuous in the set Z^*

(h_4) for (t, σ, ρ) in Z^*, the derivatives $x_t(t, \sigma, \rho)$, $x_\sigma(t, \sigma, \rho)$, $x_\rho(t, \sigma, \rho)$ are continuous;

(h_5) the map $Z^* \to D$ defined by $(t, \sigma, \rho) \to (t, x(t, \sigma, \rho), y(t, \sigma, \rho))$ is descriptive.

The above hypotheses are of fundamental importance in our further considerations. If all of them are satisfied, the family F is called a *canonical flow*. In the classical mechanics the R^{2n+1} - space of (t, x, y) is called a phase space and the family F a phase flow (see [1], § 16).

3. Preliminary lemmas

Suppose we are given a canonical flow as described in Section 2.
LEMMA 3.1. *Let C be a rectifiable curve situated, together with its end points, in D. Then, the function $H(t, x, y)$ defined by (1.1) is bounded and Borel measurable along C.*

P r o o f. By (h_5), to each point of C, including its end points, there corresponds a neighbourhood (in the topology of C induced from R^{2n+1}) that is the image of a curve Γ, $\Gamma \subset Z^*$. From Borel's covering theorem it follows that we can suppose C so small that the corresponding (t, σ, ρ)-curve is just the Γ. By (2.6) and (h_2), (h_3) the function $\tilde{H}(t, \sigma, \rho) = H(t, x(t, \sigma, \rho), y(t, \sigma, \rho))$ is continuous in Z^* and thus, bounded on Γ. To each point $(t, x, y) \in C$ we can now attach the first point (t, σ, ρ) of Γ at which $x(t, \sigma, \rho) = x, y(t, \sigma, \rho) = y$. Putting these values into $\tilde{H}(t, \sigma, \rho)$, we obtain a Borel measurable function $\bar{H}(t, x, y)$ defined on C. This function is also bounded. But, of course, $\bar{H}(t, x, y) = H(t, x, y)$ on C; this completes the proof.

For $t^-(\sigma) \leqslant t \leqslant t^+(\sigma)$ and $(\sigma, \rho) \in \tilde{G}$, we set

$$S(t, x(t, \sigma, \rho), y(t, \sigma, \rho)) = \int_t^{t^+(\sigma)} L(\tau, x(\tau, \sigma, \rho), y(\tau, \sigma, \rho))d\tau.$$

We shall prove the following

LEMMA 3.2 *Let $B \subset Z^*$ be any simply connected closed domain and let Γ denote any rectifiable curve in B with (t_0, σ_0, ρ_0) as the initial point and (t_1, σ_1, ρ_1) as the terminal one. Then*

$$\int_\Gamma \tilde{L}\,dt - \left(\int_t^{t^+(\sigma)} \bar{L}_\sigma d\tau\right) d\sigma - \left(\int_t^{t^+(\sigma)} \tilde{L}_\rho d\tau\right) d\rho =$$

$$S(t_0, x(t_0, \sigma_0, \rho_0), y(t_0, \sigma_0, \rho_0,)) - S(t_1, x(t_1, \sigma_1, \rho_1), y(t_1, \sigma_1, \rho_1,)).$$

P r o o f. Let us note that, by (h_2), the expression

$$\tilde{L}\,dt - \left(\int_t^{t^+(\sigma)} \tilde{L}_\sigma d\tau\right) d\sigma - \left(\int_t^{t^+(\sigma)} \tilde{L}_\rho d\tau\right) d\rho$$

is an exact differential in the variables (t, σ, ρ) of some function $Q(t, \sigma, \rho)$, $(t, \sigma, \rho) \in B$. Assuming $Q(t_0, \sigma_0, \rho_0) = 0$, we see that the function $Q(t, \sigma, \rho)$ has the form

$$Q(t, \sigma, \rho) = \left(\int_{t_0}^{t^+(\sigma_0)} \tilde{L}(\tau_0, \sigma_0, \rho_0) d\tau - \left(\int_t^{t^+(\sigma)} \tilde{L}(\tau, \sigma, \rho) d\tau.\right.\right.$$

From the above formula we get

$$\int_\Gamma \tilde{L}\,dt - \left(\int_t^{t^+(\sigma)} \bar{L}_\sigma d\tau\right) d\sigma - \left(\int_t^{t^+(\sigma)} \tilde{L}_\rho d\tau\right) d\rho =$$

$$= Q(t_1, \sigma_1, \rho_1) - Q(t_0, \sigma_0, \rho_0) =$$

$$= S(t_0, x(t_0, \sigma_0, \rho_0), y(t_0, \sigma_0, \rho_0)) - S(t_1, x(t_1, \sigma_1, \rho_1), y(t_1, \sigma_1, \rho_1)),$$

as asserted. ■

The next lemma appears to be a restatement, under our weakened assumptions, of the well-known theorem of Malus from geometrical optics (see [7], § 26).

LEMMA 3.3. *On each canonical pair of the canonical flow F the expressions*

$$yx_\sigma + \int_t^{t^+(\sigma)} \bar{L}_\sigma d\tau, \qquad yx_\rho + \int_t^{t^+(\rho)} \tilde{L}_\rho d\tau, \tag{3.1}$$

are constant.

In this statement x and y stand for the functions $x(t, \sigma, \rho)$, $y(t, \sigma, \rho)$ and the property is to hold for $(t, \sigma, \rho) \in Z^*$.

P r o o f. We shall show that expressions (3.1) depend only on (σ, ρ), i.e. do not depend on t.

Let $(\hat{t}, \hat{\sigma}, \hat{\rho})$ be any point of the set Z^* and $\hat{x}(t)$, $\hat{y}(t)$, $\hat{u}(t)$ the corresponding values of the functions $x(t, \hat{\sigma}, \hat{\rho})$, $y(t, \hat{\sigma}, \hat{\rho})$, $u(t, \hat{\sigma}, \hat{\rho})$; $\hat{t} \leqslant t < t^+(\hat{\sigma})$. Let further α, β stand for any coordinates of $\sigma \in G$, $\rho \in R$, respectively, and $a(t)$, $b(t)$ for the values of

$$\frac{\partial}{\partial \alpha} H(t, \hat{x}(t), \hat{y}(t), u(t, \sigma, \rho)), \quad \frac{\partial}{\partial \beta} H(t, \hat{x}(t), \hat{y}(t), u(t, \sigma, \rho))$$

at $(t, \hat{\sigma}, \hat{\rho})$, $\hat{t} \leqslant t < t^+(\hat{\sigma})$, accordingly. By integrating with respect to t, then differentiating with respect to α and again differentiating with respect to t, from (1.4), taking into account (2.5), we obtain the following relations calculated at the point $(t, \hat{\sigma}, \hat{\rho})$, $\hat{t} \leqslant t < t^+(\hat{\sigma})$,

$$x_\alpha(t, \sigma, \rho) - x_\alpha(\hat{t}, \sigma, \rho) = \int_{\hat{t}}^{t} \tilde{f}_\alpha(\tau, \sigma, \rho) d\tau;$$

$$x_\beta(t, \sigma, \rho) - x_\beta(\hat{t}, \sigma, \rho) = \int_{\hat{t}}^{t} \tilde{f}_\beta(\tau, \sigma, \rho) d\tau,$$

$$\frac{\partial}{\partial t} x_\alpha(t, \sigma, \rho) = \tilde{f}_\alpha(t, \sigma, \rho), \quad \frac{\partial}{\partial t} x_\beta(t, \sigma, \rho) = \tilde{f}_\beta(t, \sigma, \rho). \tag{3.2}$$

From equation (1.5) we see that at the same point $(t, \hat{\sigma}, \hat{\rho})$ the equalities

$$x_\alpha(t, \sigma, \rho) \frac{\partial}{\partial t} \hat{y}(t) = - x_\alpha(t, \sigma, \rho) H_x(t, \hat{x}(t), \hat{y}(t), \hat{u}(t)),$$

$$\tag{3.3}$$

$$x_\beta(t, \sigma, \rho) \frac{\partial}{\partial t} \hat{y}(t) = - x_\beta(t, \sigma, \rho) H_x(t, \hat{x}(t), \hat{y}(t), \hat{u}(t))$$

hold.

Multiplying relations (3.2) by $\hat{y}(t)$ and adding the results to (3.3), respectively, we obtain - with the use of (2.5), (2.6) - at this very point

$$\frac{\partial}{\partial t}\{\hat{y}x_\alpha\} - \tilde{L}_\alpha = a(t), \frac{\partial}{\partial t}\{\hat{y}x_\beta\} - \tilde{L}_\beta = b(t),\ \ \hat{t} \leqslant t < t^+(\hat{\sigma}). \tag{3.4}$$

Since G and R are open sets, the maximum relation (1.1) implies

$$a(t) = 0, \quad b(t) = 0, \quad \hat{t} \leqslant t < t^+(\hat{\sigma}).$$

Integration of (3.4) over the interval $[\hat{t}, t^+(\hat{\sigma})]$ yields

$$\begin{aligned} \hat{y}x_\alpha + \int_{\hat{t}}^{t^+(\hat{\sigma})} \tilde{L}_\alpha dt &= \hat{y}(t^+(\hat{\sigma}))x_\alpha + \tilde{L}(t^+(\hat{\sigma}), \hat{\sigma}, \hat{\rho})t_\alpha{}^+(\sigma), \\ \hat{y}x_\beta + \int_{\hat{t}}^{t^+(\hat{\sigma})} \tilde{L}_\rho dt &= \hat{y}(t^+(\hat{\sigma}))x_\beta, \end{aligned} \tag{3.5}$$

where the left - hand sides are calculated at $(\hat{t}, \hat{\sigma}, \hat{\rho})$ and the right - hand ones at $(t^+(\hat{\sigma}), \hat{\sigma}, \hat{\rho})$. Now, we see that the right - hand sides of (3.5) depend only on (σ, ρ) and do not depend on t. Hence, in view of the free choice of the point $(\hat{t}, \hat{\sigma}, \hat{\rho}) \in Z^*$ and the coordinates α, β we infer that expressions (3.1) take constant values on each fixed member of the canonical flow F. The proof of Lemma 3.3 is now completed. ∎

4. Theorem on Poincaré - Cartan's integral invariant

For each point $(t_0, x_0, y_0) \in D$, we consider the integral

$$S(t_0, x_0, y_0) = \int_{t_0}^{t^+} L(\tau, x(\tau), y(\tau))d\tau \tag{4.1}$$

"along" a canonical pair $x(t), y(t), t \in [t^-, t^+], t^- \leqslant t_0 \leqslant t^+$, of the flow F such that $x(t_0) = x_0$, $y(t_0) = y_0$, where t^-, t^+ are the points from the loci $t = t^-(\sigma)$, $t = t^+(\sigma)$, $\sigma \in G$, respectively. Of course, through each point $(t_0, x_0, y_0) \in D$ there may pass more than one member of F. Therefore, an additional principle is formulated to be satisfied by the canonical flow F:

(h_6) If, through any point $(t_0, x_0, y_0) \in D$, there pass two members of F, then the values of integral (4.1) along each of them are the same.

Thus the function $S(t, x, y)$ of (4.1) is uniquely defined in D. This function in mechanics is called *the activity function,* while in optimal control theory - *the value function.*

Let C be any rectifiable curve lying in D with the description

$$t = \bar{t}(s), \quad x = \bar{x}(s), \quad y = \bar{y}(s), \qquad 0 \leqslant s \leqslant \zeta,$$

where s is the arc length parameter. On the curve C we define the curvilinear integral

$$\int_C y\,dx - H(t, x, y)\,dt = \int_0^{\zeta} [y(s)\frac{dx}{ds} - H(\bar{t}(s), \bar{x}(s), \bar{y}(s))\frac{dt}{ds}\,]ds \quad (4.2)$$

That this integral is well defined follows from Lemma 3.1.

We say that two curves C_1, C_2 in D *cut the same subfamily* $F(C_1, C_2)$ of the canonical flow F if, for each point of C_1, there existpoints of C_2 and, a member of $F(C_1, C_2)$ which passes through these points and, conversely; for each point of C_2, there is a point of C_1 and a member of $F(C_1, C_2)$ which passes through these points.

We now come to our main result on some integral invariant.

THEOREM 4.1. *Let two closed rectifiable curves C_1, C_2 lying in D and cutting the same subfamily be given. Let further Γ_1, Γ_2 denote rectifiable curves of (t, σ, ρ)-space such that C_1, C_2 are their images under the map $(t, \sigma, \rho) \to (t, x(t, \sigma, \rho), y(t, \sigma, \rho))$. We also assume that Γ_1, Γ_2 are included in some simply connected sets contained in Z^* and that the sets of values of Γ_1 and Γ_2 in coordinates σ, ρ coincide.*

Then the curvilinear integrals of the form $\omega = y\,dx - H\,dt$ along C_1 and C_2 are equal, i.e.

$$\int_{C_1} y\,dx - H\,dt = \int_{C_2} y\,dx - H\,dt\,. \tag{4.3}$$

P r o o f. In view of (1.4), (2.5) and (2.6), we have

$$\int_{C_1} y\,dx - H\,dt = \int_{\Gamma_1} y x_\sigma d\sigma + y x_\rho d\rho + \bar{L}\,dt =$$

$$= \int_{\Gamma_1} \{y x_\sigma + (\int_t^{t^+(\sigma)} \bar{L}_\sigma d\tau)\} d\sigma + \{y x_\rho + (\int_t^{t^+(\sigma)} \bar{L}_\rho)\} d\rho +$$

$$+ \int_{\Gamma_1} \bar{L}\,dt - (\int_t^{t^+(\sigma)} \bar{L}_\sigma d\tau)\,d\sigma - (\int_t^{t^+(\sigma)} \bar{L}_\rho d\tau)\,d\rho\,.$$

Since C_1 is closed, Lemma 3.2 now implies

$$\int_{C_1} \omega = \int_{\Gamma_1} \{ y x_\sigma + (\int_t^{t^+(\sigma)} \bar{L}_\sigma d\tau) \} d\sigma + \{ y x_\rho + (\int_t^{t^+(\sigma)} \bar{L}_\rho d\tau) \} d\rho \quad (4.4)$$

Analogously, we get

$$\int_{C_2} \omega = \int_{\Gamma_2} \{ y x_\sigma + (\int_t^{t^+(\sigma)} \bar{L}_\sigma d\tau) \} d\sigma + \{ y x_\rho + (\int_t^{t^+(\sigma)} \bar{L}_\rho d\tau) \} d\rho \quad (4.5)$$

By the assumptions of the theorem, the sets of values of Γ_1 and Γ_2 in coordinates σ, ρ coincide; hence, from Lemma 3.3 we conclude that the right-hand sides of (4.4) and (4.5) have the same value. The proof is now completed. ■

REMARK 4.1. The above theorem generalizes the well-known theorems of this type (cf. [1], § 44) not only because here weaker assumptions on H are imposed but mainly, because the members of F may intersect. The form $\omega = y\,dx - H\,dt$ is called Poincaré-Cartan's integral invariant.

Consider the particular case when the curves C_1, C_2 lie in the hyperplanes $t = \text{const}$. Along such curves $\int \omega = \int y\,dx$. Let

$$D^{t_0} = \{ (t, x, y) \in D : t = t_0 \},$$

$$D^{t_1} = \{ (t, x, y) \in D : t = t_1 \}$$

and consider the map $g_{t_0}^{t_1} : D^{t_0} \to D^{t_1}$ defined by the canonical flow F when t varies from t_0 to t_1, $t^-(\sigma) \leqslant t_0 \leqslant t_1 \leqslant t^+(\sigma)$, $\sigma \in G$. We note that this map may be multivalued.

We now assume, in addition, that the map $(t, \sigma, \rho) \to (t, x(t, \sigma, \rho), y(t, \sigma, \rho))$ is such that any closed curve $C \subset D_{t_0}$ is the image under this map of some rectifiable closed curve Γ of (t_0, σ, ρ)-space included in some simply connected set. We also suppose $y(t_1, \sigma, \rho)$ to be a C^1-function for $(\sigma, \rho) \in \tilde{G}$.

Let C_{t_0} be any closed rectifiable curve in D^{t_0}. As its image under the mapping $g_{t_0}^{t_1}$ we understand the rectifiable curve C_{t_1} obtained in the following way. We take Γ_0 whose image is C_{t_0}. To the curve Γ_0 there corresponds a rectifiable curve Γ_1 such that $(t_1, \sigma, \rho) \in \Gamma_1$ if $(t_0, \sigma, \rho) \in \Gamma_0$. Now, C_{t_1} is defined as the image of Γ_1 under the map $(t_1, \sigma, \rho) \to (t_1, x(t_1, \sigma, \rho), y(t_1, \sigma, \rho))$. In the sequel, the image of C_{t_0} under the map $g_{t_0}^{t_1}$ will always be understood in the above described way.

Theorem 4.1 implies the following, self-evident but essential

COROLLARY 4.1. *Let* $g_{t_0}^{t_1}$, C_{t_0}, C_{t_1} *be as described above. Then*

$$\int_{C_{t_0}} y\,dx = \int_{C_{t_1}} y\,dx.$$

REMARK 4.2. The form $y\,dx$ is termed Poincare's conditional invariant. This form has a simple geometric interpretation. Assume that there exists a two-dimensional oriented chain γ_{t_0} with C_{t_0} as its boundary contained in D_{t_0} and, analogously, a chain γ_{t_1} with C_{t_1} as its boundary, $C_{t_1} \subset D_{t_1}$. The curve C_{t_1} is obtained from C_{t_0} by using the map $g_{t_0}^{t_1}$. Now, we have

$$\int_{\gamma_{t_0}} dy \wedge dx = \int_{\gamma_{t_1}} dy \wedge dx. \tag{4.6}$$

Indeed, Stokes' theorem implies

$$\int_{C_{t_0}} y\,dx = \int_{\gamma_{t_0}} dy \wedge dx, \qquad \int_{C_{t_1}} y\,dx = \int_{\gamma_{t_1}} dy \wedge dx;$$

hence Corollary 4.1 gives (4.6).

We say that a mapping $g : R^{2n} \to R^{2n}$ is *canonical* if

$$\int_{C} y\,dx = \int_{gC} y\,dx$$

where C, $C \subset R^{2n}$, denotes any rectifiable curve and gC stands for a rectifiable curve being the image of C under the mapping g.

From Corollary 4.1 we immediately have the following

REMARK 4.3. The map $g_{t_0}^{t} : D_{t_0} \to D_t$, $t^-(\sigma) \leqslant t_0 < t \leqslant t^+(\sigma)$, $\sigma \in G$, t_0-fixed, defines in D the canonical map ($g_{t_0}^{t}$, D_{t_0}, D_t are understood as in Corollary 4.1).

In this way one can further study consequences of Theorem 4.1; we shall not do this here. We come back to form (1.3) and to functions (1.1) and (2.6).

5. Properties of the activity function, Hamilton - Jacobi - Bellman equation

In this section we give more exact relations between the functions L, H and the form ω defined in Sections 1 and 2. These relations have miscellaneous applications, e.g. in mechanics, in optimal control theory.

We put

$$Z^{*+} = \{(t, \sigma, \rho) : t = t^{+}(\sigma), \ \ \sigma \in G, \ \ \rho \in R\}$$

(for G and R see Section 2) and, next,

$$D^{+} = \{(t,x,y): x = x(t,\sigma,\rho), y = y(t,\sigma,\rho), (t,\sigma,\rho) \in Z^{*+}\}. \quad (5.1)$$

From now on, in place of hypothesis (h_5), we shall only suppose that the map $Z^* \setminus Z^{*+} \to D \setminus D^{+}$, defined by $(t, \sigma, \rho) \to (t, x(t, \sigma, \rho), y(t, \sigma, \rho))$, is descriptive.

A subset $\dot{D}$ of D is called *an exact set* if, for each rectifiable curve C, $C \subset \dot{D}$, with end points (t_1, x_1, y_1), (t_2, x_2, y_2), the equality

$$\int_C y\,dx - H\,dt = S(t_1, x_1, y_1) - S(t_2, x_2, y_2) \tag{5.2}$$

holds.

(h_7) In what follows we suppose D^{+} to be an exact set. We shall, at the same time, assume the integrand on the right of (4.2) to be Lebesgue integrable.

If (h_7) is satisfied, then, of course, the right-hand side of (5.2) is equal to zero for $C \subset D^{+}$.

LEMMA 5.1. *The quantities*

$$yx_\sigma + \tilde{L}\,t_\sigma^{+}, \ \ yx_\rho$$

are identically zero in the set Z^{*+}.

P r o o f. Let (t_0, σ_0, ρ_0) be any point of the set Z^{*+} and Γ - any sufficiently small rectifiable curve in Z^{*+} which corresponds, by setting $t = t^{+}(\sigma)$, to a small segment parallel to one of the σ-axes, with the initial point (σ_0, ρ_0). Let C, $C \subset D^{+}$, stand for the image of Γ under the map $(t, \sigma, \rho) \to (t, x(t, \sigma, \rho), y(t, \sigma, \rho))$. Since D^{+} is exact, we have

$$0 = \int_C y\,dx - H\,dt = \int_\Gamma y\,x_\sigma\,d\sigma + \tilde{L}\,dt = \int_\Gamma (y\,x_\sigma + \tilde{L}\,t_\sigma^{+})\,d\sigma\,.$$

Hence, taking into account that Γ was chosen arbitrarily, and that $y\,x_\sigma + \tilde{L}t_\sigma^{+}$ is a continuous function in Z^{*+}, we infer that $y\,x_\sigma + \tilde{L}t_\sigma^{+}$ is equal to zero at $(t^{+}(\sigma_0), \sigma_0, \rho_0)$, i.e. at any point of Z^{*+}.

The proof concerning $y\,x_\rho$ is analogous. ∎

COROLLARY 5.1. *The expressions*

$$y x_\sigma + \int_t^{t^+(\sigma)} \bar{L}_\sigma d\tau , \qquad y x_\rho + \int_t^{t^+(\sigma)} \tilde{L}_\rho d\tau$$

vanish in Z^*.

P r o o f. These expressions are continuous functions in Z^* by (h_2)-(h_4). Lemma 3.3 now implies that the expressions are constant on each member of F, therefore the assertion follows from the above lemma.

Let C be any rectifiable curve contained in $D \setminus D^+$ with parametric description

$$t = \bar{t}(s), \quad x = \bar{x}(s), \quad y = \bar{y}(s), \quad 0 \leqslant s \leqslant \zeta,$$

where s is the arc length parameter. Then function (4.1) restricted to the curve C has the form $S(\bar{t}(s), \bar{x}(s), \bar{y}(s))$, i.e. is the function of the variable s for $s \in [0, \zeta]$. ■

The next theorem will show that form (1.3), introduced in the artificial way in Section 1, appears in the natural way in mechanics or optimal control theory. Namely, we prove the following important.

contr121-Control and Cybernetics-skł-G.P. korekta DM

THEOREM 5.1 *The function* $s \to (S(\bar{t}(s), \bar{x}(s), \bar{y}(s)), s \in [0, \zeta]$, *is absolutely continuous and, for almost all s in* $[0, \zeta]$, *the relation*

$$\frac{d}{ds} S(\bar{t}(s), \bar{x}(s), \bar{y}(s)) = -[\bar{y}(s)\frac{dx}{ds} - H(\bar{t}(s), \bar{x}(s), \bar{y}(s))\frac{dt}{ds}] \qquad (5.3)$$

holds.

Proof. First, we prove the absolute continuity of $S(\bar{t}(s), \bar{x}(s), \bar{y}(s))$ as a function of s. The function $(\frac{dt}{ds}, \frac{dx}{ds})$, defined for almost all $s \in [0, \zeta]$, is integrable over $[0, \zeta]$.

So, let $s_0 \in [0, \zeta]$ be any point such that $(\frac{dt}{ds}, \frac{dx}{ds})$ is approximately continuous at it (almost all points in $[0, \zeta]$ are such). We introduce the notations

$$\hat{t} = \bar{t}(s_0), \quad \hat{x} = \bar{x}(s_0), \quad \hat{\dot{t}} = \frac{d\bar{t}(s_0)}{ds},$$

$$\hat{\dot{x}} = \frac{d\bar{x}(s_0)}{ds}, \quad \hat{y} = \bar{y}(s_0), \quad \hat{H} = H(\hat{t}, \hat{x}, \hat{y}).$$

Let $(\hat{t}, \hat{\sigma}, \hat{\rho})$ be any point of $Z^* \setminus Z^{*+}$ for which $x(\hat{t}, \hat{\sigma}, \hat{\rho}) = \hat{x}, y(\hat{t}, \hat{\sigma}, \hat{\rho}) = \hat{y}$.

Denote by Γ a rectifiable curve in $Z^* \setminus Z^{*+}$ such that small arcs of the curve C, issuing from $(\hat{t}, \hat{x}, \hat{y})$, are, in accordance with (h_5), the images under the map $(t, \sigma, \rho) \to (t, x(t, \sigma, \rho), y(t, \sigma, \rho))$ of small arcs γ of Γ issuing from the point $(\hat{t}, \hat{\sigma}, \hat{\rho})$. Let us now find a parametric description of the curve

$$t = \tilde{t}(\lambda), \quad \sigma = \tilde{\sigma}(\lambda), \quad \rho = \tilde{\rho}(\lambda), \quad 0 \leqslant \lambda \leqslant \eta,$$

such that the point $(\hat{t}, \hat{\sigma}, \hat{\rho}) \in \Gamma$ corresponds to zero, while λ is the arc length parameter. We can then define a continuous increasing function $s = s(\lambda)$, $\lambda \in [0, \eta]$, such that $s(0) = s_0$, which gives rise to the corresponding arc length along C, i.e. which satisfies in any subinterval $[0, \eta_1] \subset [0, \eta]$ the relations

$$\bar{t}(s(\lambda)) = \tilde{t}(\lambda), \quad \bar{x}(s(\lambda)) = x(\tilde{t}(\lambda), \tilde{\sigma}(\lambda), \tilde{\rho}(\lambda)),$$

$$\bar{y}(s(\lambda)) = y(\tilde{t}(\lambda), \tilde{\sigma}(\lambda), \tilde{\rho}(\lambda)).$$

Let now γ be a sufficiently small arc of Γ issuing from the point $(\hat{t}, \hat{\sigma}, \hat{\rho})$, described above, defined in the interval $K = [0, \eta_1]$. Denote by Δs and ΔS the corresponding differences in s and in $S(t, x, y)$ at the ends of a small arc of C issuing from the point $(\hat{t}, \hat{x}, \hat{y})$, being the image of γ. Let further

$$\Delta \tilde{S} = \int_\gamma \bar{L}\, dt - \Big(\int_t^{t^+(\sigma)} \tilde{L}_\sigma d\tau\Big) d\sigma - \Big(\int_t^{t^+(\sigma)} \tilde{L}_\rho d\tau\Big) d\rho$$

(comp. Lemma 3.2). We have $\Delta S = -\Delta \tilde{S}$. On the other hand, by Corollary 5.1, along γ we obtain

$$y x_\sigma + \int_t^{t^+(\sigma)} \bar{L}_\sigma d\tau = 0, \qquad y x_\rho + \int_t^{t^+(\sigma)} \tilde{L}_\rho d\tau = 0.$$

Hence, using (2.5) and (2.6)

$$\Delta \tilde{S} = \int_\gamma (y x_t - \tilde{H})\, dt + y x_\sigma d\sigma + y x_\rho d\rho = \int_K \Big(y \frac{dx}{ds} - \tilde{H} \frac{dt}{ds}\Big) ds(\lambda) \qquad (5.4)$$

where $-\tilde{H} = \tilde{L} - y\tilde{f}$. By the assumptions imposed upon the functions $\tilde{L}, \tilde{f}, y$ in Z^*, we get the boundedness of $\tilde{H}$ in K and next, from (5.4), the uniform boundedness of $\dfrac{\Delta S}{\Delta s}$ as $\Delta s \to 0$. This means that S is a locally Lipschitz function of the variable s and, hence, absolutely continuous in s.

To prove (5.3), we note that the approximate continuity of $(\frac{dt}{ds}, \frac{dx}{ds})$ at s_0 implies that, given $\varepsilon > 0$, there exists a closed set B of values of s such that, for each sufficiently small interval $B_\delta = \{s : s_0 \leqslant s \leqslant s_0 + \delta\}$, the inequalities

(i) $\quad |(\frac{dt}{ds}, \frac{dx}{ds}) - (\hat{t}, \hat{x})| < \varepsilon \qquad s \in B \cap B_\delta,$

(ii) $\quad \text{meas}\,(B_\delta \setminus B) < \varepsilon\delta$

hold ($|\cdot|$ stands here for the Euclidean norm).

Denote by Λ the set of those values of λ for which $s(\lambda) \in B$. In view of the above considerations, we can regard the expression $y\frac{dx}{ds} - \tilde{H}\frac{dt}{ds}$ as a function of λ in K. Thus, if we set $g = g(\lambda) = (y\frac{dx}{ds} - \tilde{H}\frac{dt}{ds}) - (\hat{y}\hat{x} - \hat{H}\hat{t})$, then we have

$$\frac{\Delta\tilde{S}}{\Delta s} - (\hat{y}\hat{x} - \hat{H}\hat{t}) = \frac{1}{\Delta s}\int_K g(\lambda)\,ds(\lambda) = \tag{5.5}$$
$$= \frac{1}{\Delta s}\int_{K\cap\Lambda} g(\lambda)\,ds(\lambda) + \frac{1}{\Delta s}\int_{K\setminus\Lambda} g(\lambda)\,ds(\lambda).$$

For a sufficiently small segment K, g is bounded in $K \setminus \Lambda$ and this set has $s(\lambda)$-measure less than $\varepsilon\delta$, by (ii) above with $\delta = \Delta s$. The $s(\lambda)$-measure of the set $K \cap \Lambda$ is at most δ. From (i) and the continuity of $\tilde{H}$ and y it now follows that the absolute value $|g|$ of the integrand in (5.5) does not exceed in the set $K \cap \Lambda$ a fixed multiplier of ε. This means that the last two terms in (5.5) cannot exceed certain fixed multipliers of an arbitrarily small positive ε. The proof of the theorem is now completed. ∎

From Theorem 5.1 we derive the following (also important) corollary.

COROLLARY 5.2. *The set D is exact, i.e. equality (5.2) holds for each rectrifiable curve C of D.*
P r o o f. Since D^+ is an exact set, we can confine ourselves to show that the set $D \setminus D^+$ is exact. By integrating (5.3), in view of Theorem 5.1, we obtain (5.2). ∎

COROLLARY 5.3. *Let C be any rectifiable curve in D with the arc length parametric description*

$$t = t_0, \quad x = x_0, \quad y = \bar{y}(s), \qquad 0 \leqslant s \leqslant \zeta.$$

Then, for all s, $0 \leqslant s \leqslant \zeta$, *we have*

$$S(t_0, x_0, \bar{y}(s)) = S(t_0, x_0, \bar{y}(0)).$$

P r o o f. First, we note that if the curve C lies in D^+, we have the assertion at once; hence we assume C to be included in $D \setminus D^+$. By (5.3), we conclude that

$$\frac{d}{ds} S(t_0, x_0, \bar{y}(s)) = 0 \qquad \text{a.e. in } [0, \zeta].$$

From Theorem 5.1 follows the absolute continuity of $S(t_0, x_0, \bar{y}(s))$ in $[0, \zeta]$. Thus $S(t_0, x_0, \bar{y}(s))$ is a constant function equal to $S(t_0, x_0, \bar{y}(0))$, ■ as required.

Corollaries 5.2 and 5.3 immediately imply

COROLLARY 5.4. *Let* (t_1, x_1, y_1), (t_2, x_2, y_2) *be any points of D. Suppose the sets*

$$P_1 = \{(t, x, y) : t = t_1, \quad x = x_1, \quad y \in R^n\} \cap D,$$

$$P_2 = \{(t, x, y) : t = t_2, \quad x = x_2, \quad y \in R^n\} \cap D.$$

are connected. If C is any rectifiable curve in D with its end points in P_1 *and* P_2, *respectively, then*

$$\int_C y\,dx - H\,dt = S(t_1, x_1, y_1) - S(t_2, x_2, y_2).$$

REMARK 5.1. The word "connected" in this paper is understood as follows:

"A set B is connected if, for any two points of B, there exists a rectifiable curve in B which joints them".

In the special case when the points (t_1, x_1, y_1), (t_2, x_2, y_2) are joined by a curve γ which belongs to the canonical flow F, Corollary 5.4 is known in mechanics as *the principle of the least activity*. The curve γ is then called the *extremal curve* of the integral $\int y\,dx - H\,dt$.

Denote by D_x a set of (t, x)-space covered by graphs of trajectories $x(t)$ such that $x(t)$, $y(t)$, $u(t)$ is a canonical pair and member of the flow F at the same time. Assume also that, for each $(t_0, x_0) \in D_x$, the set

$$P_0 = \{(t, x, y,) : t = t_0, \quad x = x_0, \quad y \in R^n\} \cap D$$

is connected. Then, by Corollary 5.3, the function

$$\bar{S}(t_0, x_0) = S(t_0, x_0, y) \tag{5.6}$$

is uniquely determined in D_x, where $S(t_0, x_0, y)$, $(t_0, x_0, y) \in P_0$, is defined by (4.1). Suppose further that D_x is an open set and that there are continuous derivatives $\bar{S}_t$, $\bar{S}_x$. Then from Theorem 5.1 we infer directly that the function $\bar{S}(t, x)$ satisfies in D_x the Hamilton - Jacobi - Bellman differential equation

$$-\bar{S}_t + H(t, x, -\bar{S}_x) = 0.$$

Indeed, in classical mechanics the function $\bar{S}(t, x)$ is called the activity function, thus our function $S(t, x, y)$, $(t, x, y) \in D$ is its generalization.

REMARK 5.2. Considering the particular case when the function $H(t,x,y,u)$ does not depend on the function u, i.e. $H(t,x,y) = H(t,x,y,u)$, we see that all we have done so far forms the generalization of the classical theory of the canonical formalism (see [1], Ch.9).

6. Adaptation to optimal control theory

Let, as in Section 1, U denote a Borel subset of R^m and let $u(t) : [a,b] \to U$ be a measurable function which will now be called *a control* or *an admissible control.*

Consider a vector function f, $f = f(t,x,u)$, where $f : [a,b] \times R^n \times R^m \to R^n$ and a scalar function L defined in $[a,b] \times R^n \times R^m$. We assume these functions and derivatives $f_x(t,x,u)$, $L_x(t,x,u)$ continuous.

An admissible trajectory $x(t)$ corresponding to a control $u(t)$ is an absolutely continuous function $x : [a,b] \to R^n$ satisfying

$$\dot{x}(t) = f(t, x(t), u(t)) \tag{6.1}$$

for almost all t in $[a,b]$.

We shall consider the following problem:

"Find a minimum of the functional

$$I(x,u) = \int_a^b L(t, x(t), u(t))\, dt \tag{6.2}$$

over all admissible trajectories $x(t)$ and corresponding controls $u(t)$ such that $x(a) = c$, $x(b) = d$, where c, d are fixed points of R^n".

To solve this problem, we put

$$\begin{aligned} &H(t,x,y,u) = yf(t,x,u) - L(t,x,u); \\ &(t,x,y,u) \in [a,b] \times R^n \times R^n \times U. \end{aligned} \tag{6.3}$$

An admissible pair of functions (x, u) will be said *to satisfy the maximum principle* (comp. [2], Cor.2) if there exists an absolutely continuous conjugate function $y : [a, b] \to R^n$ such that

$$\begin{aligned} &\text{(i) } \dot{y}(t) = -H_x(t, x(t), y(t), u(t)) \\ &\text{(ii) } H(t, x(t), y(t), u(t)) = \sup\{H(t, x(t), y(t), u : u \in U\}. \end{aligned} \tag{6.4}$$

As in Section 2, we shall call each pair $(x(t), u(t))$, $t \in [a, b]$, satisfying the maximum principle a trajectory. We term a canonical pair a trio of functions $(x(t), y(t), u(t))$, $t \in [a, b]$, such that $(x(t), u(t))$ defines a trajectory and $y(t)$ is the corresponding conjugate function which satisfies (6.4). Thus, we can further consider a canonical flow F, if it exists, i.e. a family of canonical pairs which satisfies (h_1)-(h_7). Of course, all these assumptions do not exclude the case when, at some point of D, the canonical pairs intersect. Having this in mind, with each point (t, x, y) of D we associate the set

$$U(t, x, y) \tag{6.5}$$

of values of such controls u at the point t for which the corresponding members of the flow F take, at the point t, the values x, y, i.e. the set of values of the function $u(t, \sigma, \rho)$ at those points (t, σ, ρ) of Z^* where $x(t, \sigma, \rho), y(t, \sigma, \rho)$ take the given values x, y, respectively. Naturally, (6.5) may be a multivalued function. We shall call it *the generalized feedback control*. By an *admissible feedback control* we shall mean a single-valued function $\bar{u}(t, x, y)$, $\bar{u}(t, x, y) \in U(t, x, y)$ for $(t, x, y) \in D$, such that at each point (t, x, y) the point $(t, x, y, \bar{u}(t, x, y))$ lies on a canonical pair of our flow F. It is clear that at any point (t, x, y) of D we have

$$H(t, x, y) = H(t, x, y, \bar{u}(t, x, y)) = \sup\{H(t, x, y, u) : u \in U\} \tag{6.6}$$

for all admissible feedback controls $\bar{u}(t, x, y)$.

It is evident that if we take for $H(t, x, y, \bar{u})$ from Section 1 function (6.3) with $u = \bar{u}(t, x, y)$, then our functions f and L from (6.1) and (6.2), also with $u = \bar{u}(t, x, y)$, are the same as the functions f and L computed from (2.5) and (2.6).

Now, we have the following lemma which is quite similar to Lemma 3.1.

LEMMA 6.1. *Let C be any rectifiable curve situated, together with its end points, in $D \setminus D^+$. Then, there exists along C an admissible feedback control $\bar{u}_m(t, x, y)$ such that the function*

$$H(t, x, y) = H(t, x, y, \bar{u}_m(t, x, y)),$$

given by (6.6), *is bounded and Borel measurable.*

Having done all that, we are now in a position to apply the results of Section 5. Indeed, we have defined the functions L, f, H and the sets D, Z^*, D^+, Z^{*+}, D_x as well as the function $S(t,x,y)$ in D. Therefore, we are able to prove the following

THEOREM 6.1. (*Sufficiency Theorem*). *Assume that a canonical flow F exists and that the set*

$$P_a = \{(t,x,y): t = a,\ x = c,\ y \in R^n\} \cap D$$

is connected. If an admissible trajectory $\bar{x}(t), \bar{x}(a) = c, \bar{x}(b) = d$, under control $\bar{u}(t), t \in [a,b]$, is a trajectory of the flow F, then, for this pair $(\bar{x}, \bar{u})$, functional (6.2) *attains its minimum relative to all admissible pairs (x, u) such that:*

(i) $x(a) = c, \quad x(b) = d;$

(ii) *graphs of the trajectories $x(t)$ lie in the set D_x;*

(iii) *there exists a rectifiable function $y(t)$, $t \in [a,b]$, such that the triplets $(t, x(t), y(t)) \in D$ for $t \in [a,b]$.*

P r o o f. Making use of Corollary 5.3 and of (5.6), we have

$$\bar{S}(a,c) = S(a,c,y_a) = \int_a^b L(t,\bar{x}(t),\bar{u}(t))\,dt \tag{6.7}$$

for all $y_a \in P_a$. From Corollary 5.2 we conclude that the set D is exact, which implies the equality

$$\int_C y\,dx - H\,dt = S(a,c,y_a) \tag{6.8}$$

for each rectifiable curve C, $C \subset D$, with ends $(a,c,y_a) \in P$, $(b,d,y_b) \in D^+$. In particular, (6.8) holds for all curves $(x(t), y(t))$ that satisfy conditions (i)-(iii) of the theorem. Comparing (6.7) and (6.8) and taking into account that $x(t)$ satisfies (6.1), we obtain

$$I(\bar{x},\bar{u}) = \int_a^b L(t,\bar{x}(t),\bar{u}(t))\,dt =$$

$$= \int_a^b [y(t)f(t,x(t),u(t)) - H(t,x(t),y(t),u(t,x(t),y(t)))]\,dt$$

for some admissible feedback control $\bar{u}(t,x,y)$. Since

$$I(x,u) = \int_a^b L(t,x(t),u(t))\,dt,$$

we have

$$I(x,u)-I(\bar{x},\bar{u}) = \tag{6.9}$$
$$= \int_a^b \{H[t,x(t),y(t),\bar{u}(t,x(t),y(t))] - H(t,x(t),y(t),u(t,))\}\,dt.$$

By virtue of relation (6.6), the inequality

$$I(x,u)-I(\bar{x},\bar{u}) \geqslant 0$$

holds, as asserted. This completes the proof of our sufficiency theorem. ∎

REMARK 6.1. Let us note that formula (6.9) generalizes Theorem 3.1 of [5]. In particular, the integrand in (6.9) is a generalization of the Weierstrass excess function.

REMARK 6.2. Theorem 6.1 generalizes the classical sufficiency theorem of Weierstrass from the calculus of variations. Indeed, let $U = R^n$ and $f(t,x,u) = u$, i.e. $x = u$. Suppose the classical assumptions fulfilled: the function $L(t,x,u)$ is a C^2-function; the family of extremal trajectories $x(t,\sigma,\rho)$ (trajectories of our canonical flow) is independent of the parameter ρ, i.e. has the form $x(t,\sigma)$, $(t,\sigma) \in Z$ and $x(t,\sigma)$ is a C^2-function in Z, $x_\sigma \neq 0$ in $Z \setminus Z^+$ $(Z^+ = \{(t,\sigma) : t = t^+(\sigma), \sigma \in G\})$; the set $Z \setminus Z^+$ is simply connected and through each point of $D_x \setminus D_x^+$ $(D_x^+ = \{(t^+(\sigma), x(t^+(\sigma),\sigma)) : \sigma \in G\})$ there passes one and only one trajectory x of the family $x(t,\sigma)$. We notice that in this case the map $(t,\sigma) \to (t, x(t,\sigma)) : Z \setminus Z^+ \to D_x \setminus D_x^+$ has the inverse C^1-map $(t,x) \to (t,\sigma(t,x))$, which implies that it is descriptive. We set $y(t,\sigma) = L_u(t,x(t,\sigma),x_t(t,\sigma))$. Of course, $y(t,\sigma)$ is a C^1-function in Z and the map $(t,\sigma) \to (t,x(t,\sigma))$, $y(t,\sigma))$ is descriptive in $Z \setminus Z^+$. The set G is connected, so the set P_a from Theorem 6.1 is connected. If the E-Weierstrass function is non-negative, then condition (6.4) (ii) is fulfilled. The family $(x(t,\sigma), y(t,\sigma))$, $\sigma \in G$, forms a canonical flow. As the function $\bar{y}(t,\sigma) = y(t,\sigma(t,x))$ is a C^1-function in $D_x \setminus D_x^+$, we have assumption (iii) of Theorem 6.1 satisfied, too. Hence we can finally conclude that if the classical assumptions of Weierstrass' theorem are satisfied, then all the assumptions of Theorem 6.1 are also satisfied.

Now, we give a simple example from the classical calculus of variations to show the value of the above theory and, in particular, of Theorem 6.1.

EXAMPLE. Consider the optimal problem

$$\text{minimize } \frac{1}{2}\int_0^\pi (-x^2(t) + y^2(t))\,dt \tag{6.10}$$

in the space of absolutely continuous functions with $x(0) = 0, x(\pi) = 0$. It is known that this problem cannot be solved by any classical method (the field theory of extremals or conjugate points) and by a generalization of the field theory described in [6], either.

The function $H(t,x,y)$ is now equal to

$$H(t,x,y) = \frac{1}{2}\ x^2 + \frac{1}{2}\ y^2.$$

The canonical triplets satisfy

$$\dot{x}(t) = y(t), \qquad \dot{y}(t) = -x(t).$$

We choose the family

$$x(t,\sigma) = \sigma \sin t, \ \ y(t,\sigma) = \sigma \cos t, \ \ t \in [0,\pi], \sigma \in R.$$

Hypotheses (h_2)-(h_5) are satisfied ($D = \{(t,x,y) : x = x(t,\sigma), y = y(t,\sigma), t \in [0,\pi], \sigma \in R\}$). The set $P_0 = \{(0,0,y) : y \in R\}$ is connected.

Therefore, by Theorem 6.1, $\bar{x}(t) \equiv 0$, $t \in [0,\pi]$, affords (6.10) the global minimum.

We notice that in [6] it is required that the map $(t,\sigma) \to (t, x(t,\sigma))$, $t \in [0,\pi]$, $\sigma \in R$, be descriptive, which is not true in our example, while in the theory described here we require that $(t,\sigma) \to (t, x(t,\sigma), y(t,\sigma))$ have to be descriptive and this holds in the above example.

References

[1] ARNOLD V.J. Mathematical methods of classical mechanics. Springer-Verlag, New York 1978.

[2] CLARK F.H. The Maximum Principle under Minimal Hypotheses. *SIAM J. Contr.*, **14** (1976), 1078-1091.

[3] COURANT R., HILBERT D. Methods of mathematical physics. Interscience, 1962.

[4] FLEMING W.H., RISHEL R. Deterministic and Stochastic Optimal Control. Springer-Verlag, 1975.

[5] MAYNE D.Q. Sufficient conditions for a control to be a strong minimum. *J. Opt. Theory Appl.* **3** (1977), 339-351.

[6] NOWAKOWSKI A. Sufficient conditions for a strong relative minimum in an optimal control problem. *J. Opt. Theory Appl.* **50** (1986), 129-147.

[7] YOUNG L.C. Lectures on the calculus of variations and optimal control theory. W.B. Sounders Company, Philadelphia-London-Toronto, 1969.

Received, September 1988.

Uogólniony formalizm kanoniczny z zastosowaniami

Celem pracy jest uogólnienie klasycznego formalizmu Hamiltona - Jacobiego na przypadek, w którym funkcja Hamiltona H zależy od sterowania (parametru) u,

$$H(t,x,y) = \max_u H(t,x,y,u)$$

Podana teoria jest zilustrowana dwoma przykładami zastosowań: w mechanice i teorii sterowania optymalnego. Uogólniono twierdzenie dotyczące całkowego niezmiennika Poincare - Cartana oraz podano warunki wystarczające dla istnienia minimum w pewnym zadaniu sterowania.

Обобщеный канонический формализм с премененими

Целью работы является обобщение классического формализма Гамильтона-Якоби для случая, когда функция Гамильтона H зависит от управления параметра u,

$$H(t, x, y) = \max u \, H(t, x, y, u)$$

Приведенная теория иллюстрируется двумя примерами применений: в механике и в теории оптимального управления. Обобщена теорема, касающаяся интегральной постоянной Пуанкаре-Картана, а также даны достаточные условия существования минимума для нокоторой задачи управления.