Journal of Engineering Mathematics Vol. II, No. 2, April 1968.
Wolters-Noordhoff Publishing - Groningen.
Printed in the Netherlands.

PONTRYAGIN'S MAXIMUM PRINCIPLE FOR A CONSTRAINED SYSTEM OF ORDINARY DIFFERENTIAL EQUATIONS

R. Nottrot

(D.S.M., Central Laboratory, Geleen, the Netherlands.)

SUMMARY

Recently TIMMAN [2] succeeded in setting up a theory of optimization by applying variational principles to problems of mathematical programming and control theory. These principles may be considered as basic when dealing with problems of optimization theory. In this paper we are concerned with a general problem of control theory: inequality-constraints for both the control-variables and the state-variables are taken into account. The point is to derive necessary conditions for the optimal control, which is such that the solution of a set of ordinary differential equations minimizes some given integral. Moreover end-conditions will be considered.

1. *Introduction*

In problems of control theory the state $x(t) = (x_1(t), x_2(t), \dots, x_n(t))$ of some system is determined by the choice of a set of control functions. The problem of finding the optimal control $u(t) = (u_1(t), u_2(t), \dots, u_m(t))$ that maximizes or minimizes some functional is the fundamental problem of optimal control theory. In some fashion this problem has been treated by Pontryagin [1] who derived the well known maximum principle, a set of necessary conditions for the optimal values of the control-variables. Recently Timman [2] succeeded in deriving the necessary conditions for an even more general problem by applying elementary variational methods to control problems.

In practical applications the control-variables $u_1, u_2, \dots, u_m$ are generally subject to a set of constraints:

$$\varphi_j(u,t) \leqslant 0, \qquad j = 1,2,\dots\dots, r \tag{1}$$

and also the state-variables $x_1, x_2, \dots\dots, x_n$ are subject to a number of constraints:

$$g_k(x,t) \leqslant 0, \qquad k = 1,2,\dots\dots, \nu. \tag{2}$$

In the following we shall be concerned with the problem of finding a control $u(t) = (u_1(t), u_2(t), \dots u_m(t))$ such that the arc $x(t) = (x_1(t), x_2(t), \dots x_n(t))$ from a fixed point (X_0, T_0), $X_0 = x(T_0)$ to a given point (X_1, T_1), $X_1 = x(T_1)$ in (x,t)-space, which arc satisfies a set of ordinary differential equations:

$$\dot{x}_i = \frac{dx_i}{dt} = f_i(x,u,t), \qquad i = 1,2,\dots, n, \tag{3}$$

minimizes the integral:

$$\int_{T_0}^{T_1} F(x,u,t)dt \tag{4}$$

Moreover the control-variables are subject to the constraints (1) and the state-variables are subject to the constraints (2). The variational approach as developed by Timman [2] is basic in the derivation of the necessary conditions for the optimal values of the control-variables. The formulation of the set of necessary conditions differs slightly from that given by Pontryagin [1], though both formulations are equivalent, as far as Pontryagin has been occupied with a similar problem.

2. *The set of necessary conditions.*

Let $u(t) = (u_1(t), u_2(t), \ldots\ldots, u_m(t))$ denote the optimal values of the control-variables and assume that the curve $x(t) = (x_1(t), x_2(t), \ldots\ldots x_n(t))$ from a fixed point (X_0, T_0), $X_0 = x(T_0)$ to some given point (X_1, T_1), $X_1 = x(T_1)$ in (x, t)-space satisfies the differential equations:

$$\dot{x}_i = f_i(x, u, t), \qquad i = 1, 2, \ldots, n, \tag{5}$$

subject to the conditions:

$$\varphi_j(u, t) \leqslant 0, \qquad j = 1, 2, \ldots, r, \tag{6}$$

$$g_k(x, t) \leqslant 0, \qquad k = 1, 2, \ldots, \nu \tag{7}$$

and minimizes the integral:

$$\int_{T_0}^{T_1} F(x, u, t)dt \tag{8}$$

as compared with all curves joining (X_0, T_0) to (X_1, T_1) and satisfying the same conditions (5), (6) and (7).

The functions F, f_i, φ_j and g_k are supposed to be of clase C^2 in (x, u, t)-space. The curve $x(t)$ is an *extremal* from (X_0, T_0) to the point (X_1, T_1). The conditions (7) define an admissible region R in (x, t)-space. It is assumed that there exists a field of such extremals from (X_0, T_0) to the points (x, t) in some neighbourhood of (X_1, T_1) in R; among all curves from (X_0, T_0) to a point (x, t) subject to the conditions (5), (6) and (7) an extremal has the characteristic property that it minimizes the integral

$$\int_{T_0}^{t} F(x, u, \tau)d\tau. \tag{9}$$

Let $J(x, t)$ denote this integral along an extremal from (X_0, T_0) to (x, t). Now consider the extremal $x(t)$ from (X_0, T_0) to (X_1, T_1). It is assumed that the right half open interval $[T_0, T_1)$ is the union of a number of such disjoint right half open sub-intervals that throughout the interior of a sub-interval certain constraints in (6) and (7) vanish whereas the other constraints are less than zero*. Let $[\tau_0, \tau_1)$ be such a sub-interval and suppose that for $\tau_0 < t < \tau_1$ the conditions

$$\varphi_j(u, t) = 0, \qquad j = 1, 2, \ldots, q \leqslant r, \tag{10}$$

$$g_k(x, t) = 0, \qquad k = 1, 2, \ldots, \mu \leqslant \nu, \tag{11}$$

whereas

* One might assume that $[T_0, T_1)$ is the union of a countable sequence of such intervals.

$$\varphi_j(u,t) < 0, \qquad j = q+1,\ q+2, \ldots, r, \tag{12}$$

$$g_k(x,t) < 0, \qquad k = \mu+1,\ \mu+2, \ldots, \nu \tag{13}$$

are satisfied along the extremal x(t). After introduction of slackvariables:

$$\varphi_j + z_j = 0, \qquad j = 1,2,\ldots, r,$$

$$g_k + v_k = 0, \qquad k = 1,2,\ldots, \nu$$

these conditions pass into

$$z_j = 0, \qquad j = 1,2,\ldots, q,$$

$$v_k = 0, \qquad k = 1,2,\ldots, \mu,$$

whereas

$$z_j > 0, \qquad j = q+1,\ q+2, \ldots, r,$$

$$v_k > 0, \qquad k = \mu+1,\ \mu+2, \ldots, \nu.$$

Later on it will be clear, that the constraints $\varphi_j \leqslant 0$, $j > q$ and $g_k \leqslant 0$, $k > \mu$ are not essential for $\tau_0 < t < \tau_1$.

We will derive a set of necessary conditions for the extremal x(t) joining (X_0, T_0) and (X_1, T_1) such that the control-functions are continuous functions of t on the interval (τ_0, τ_1). For the results that we have in view it is sufficient to consider only uniformly small variations of the control-variables on (τ_0, τ_1): $|\delta u_i(t)| < \epsilon$, $\tau_0 < t < \tau_1$. Then the corresponding variations of the state-variables are also uniformly small on (τ_0, τ_1). Terms of order ϵ^2 will be neglected.

Let us introduce the functions:

$$h_k = \sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i} f_i + \frac{\partial g_k}{\partial t} . \tag{14}$$

It is assumed that the matrix

$$\frac{\partial(h_1,\ h_2, \ldots h_\mu,\ \varphi_1,\ \varphi_2, \ldots, \varphi_q)}{\partial(u_1,\ u_2, \ldots\ldots\ldots, u_m)}$$

has rank $\mu + q$; Let us suppose that the matrix

$$\Delta = \frac{\partial(h_1,\ h_2, \ldots, h_\mu,\ \varphi_1,\ \varphi_2, \ldots, \varphi_q)}{\partial(u_1,\ u_2, \ldots\ldots\ldots, u_{\mu+q})}$$

is non-singular so that it has an inverse Γ with elements denoted by γ_{ij}.

Because of the $q + \mu$ constraints $\varphi_j = 0$, $j \leqslant q$ and $g_k = 0$, $k \leqslant \mu$ the variations $\delta u_1, \delta u_2, \ldots, \delta u_m$ are restricted. Therefore we shall eliminate $q + \mu$ variations $\delta u_1, \delta u_2, \ldots, \delta u_{\mu+q}$ so that we have $m - \mu - q$ arbitrary variations of the control-variables left.

From $g_k + v_k = 0$, $k \leqslant \mu$ it follows that

$$\delta g_k + \delta v_k = \sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i} \delta x_i + \delta v_k = 0, \qquad k = 1,2,\ldots\ldots, \mu,$$

where $\delta v_k \geqslant 0$. Differentiation with respect to t yields:

$$\sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i} \delta \dot{x}_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 g_k}{\partial x_i \partial x_j} f_j \delta x_i + \sum_{i=1}^{n} \frac{\partial^2 g_k}{\partial x_i \partial t} \delta x_i + \delta \dot{v}_k = 0,$$

$$k = 1, 2, \ldots, \mu.$$

Substituting for $\delta \dot{x}_i$ the expression

$$\delta \dot{x}_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} \delta x_j + \sum_{j=1}^{m} \frac{\partial f_i}{\partial u_j} \delta u_j$$

we find that

$$\sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i} \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} \delta x_j + \sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i} \sum_{j=1}^{m} \frac{\partial f_i}{\partial u_j} \delta u_j + \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 g_k}{\partial x_i \partial x_j} f_j \delta x_i +$$

$$+ \sum_{i=1}^{n} \frac{\partial^2 g_k}{\partial x_i \partial t} \delta x_i + \delta \dot{v}_k = 0.$$

Now

$$\frac{\partial h_k}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial^2 g_k}{\partial x_i \partial x_j} f_i + \sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i} \frac{\partial f_i}{\partial x_j} + \frac{\partial^2 g_k}{\partial x_j \partial t}.$$

Thus we obtain the equations

$$\sum_{j=1}^{n} \frac{\partial h_k}{\partial x_j} \delta x_j + \sum_{j=1}^{m} \frac{\partial h_k}{\partial u_j} \delta u_j + \delta \dot{v}_k = 0, \quad k = 1, 2, \ldots, \mu. \tag{15}$$

From $\varphi_j + z_j = 0$, $j \leqslant q$ we have that

$$\delta \varphi_j + \delta z_j = \sum_{i=1}^{n} \frac{\partial \varphi_j}{\partial x_i} \delta x_i + \sum_{k=1}^{m} \frac{\partial \varphi_j}{\partial u_k} \delta u_k + \delta z_j = 0, \; j = 1, 2, \ldots, q, \tag{16}$$

where $\delta z_j \geqslant 0$. Solving $\delta u_1, \delta u_2, \ldots\ldots, \delta u_{\mu+q}$ from the equations (15) and (16) we obtain:

$$\delta u_i = - \sum_{k=1}^{n} A_{ik} \delta x_k - \sum_{k=\mu+q+1}^{m} B_{ik} \delta u_k - \sum_{j=1}^{\mu} \gamma_{ij} \delta \dot{v}_j - \sum_{j=1}^{q} \gamma_{i,\mu+j} \delta z_j, \tag{17}$$

$$j = 1, 2, \ldots\ldots, \mu+q,$$

where

$$A_{ik} = \sum_{j=1}^{\mu} \gamma_{ij} \frac{\partial h_j}{\partial x_k} + \sum_{j=1}^{q} \gamma_{i,u+j} \frac{\partial \varphi_j}{\partial x_k} = \sum_{j=1}^{\mu} \gamma_{ij} \frac{\partial h_j}{\partial x_k} \tag{18}$$

and

$$B_{ik} = \sum_{j=1}^{\mu} \gamma_{ij} \frac{\partial h_j}{\partial u_k} + \sum_{j=1}^{q} \gamma_{i,\mu+j} \frac{\partial \varphi_j}{\partial u_k}. \tag{19}$$

Now consider the difference

$$V = \int_{\tau_0}^{\tau_1} F(x+\delta x,\ u+\delta u,\ t)dt - \int_{\tau_0}^{\tau_1} F(x,u,t)dt$$

$$= \int_{\tau_0}^{\tau_1} \left\{ \sum_{i=1}^{n} \frac{\partial F}{\partial x_i} \delta x_i + \sum_{j=1}^{m} \frac{\partial F}{\partial u_j} \delta u_j \right\} dt,$$

where u(t) is the optimal control for the extremal x(t) joining (X_0, T_0) and (X_1, T_1). From (17) it follows that

$$V = \int_{\tau_0}^{\tau_1} \left\{ \sum_{k=1}^{n} \left[\frac{\partial F}{\partial x_k} - \sum_{j=1}^{\mu+q} \frac{\partial F}{\partial u_j} A_{jk} \right] \delta x_k + \sum_{k=\mu+q+1}^{m} \left[\frac{\partial F}{\partial u_k} - \sum_{j=1}^{\mu+q} \frac{\partial F}{\partial u_j} B_{jk} \right] \delta u_k \right.$$

$$\left. - \sum_{j=1}^{\mu+q} \frac{\partial F}{\partial u_j} \sum_{k=1}^{\mu} \gamma_{jk} \delta \dot{v}_k - \sum_{j=1}^{\mu+q} \frac{\partial F}{\partial u_j} \sum_{k=1}^{q} \gamma_{j,\mu+k} \delta z_k \right\} dt$$

and from

$$\delta \dot{x}_i = \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k} \delta x_k + \sum_{k=1}^{m} \frac{\partial f_i}{\partial u_k} \delta u_k$$

we find that

$$\delta \dot{x}_i = \sum_{k=1}^{n} \left[\frac{\partial f_i}{\partial x_k} - \sum_{j=1}^{\mu+q} \frac{\partial f_i}{\partial u_j} A_{jk} \right] \delta x_k + \sum_{k=\mu+q+1}^{m} \left[\frac{\partial f_i}{\partial u_k} - \sum_{j=1}^{\mu+q} \frac{\partial f_i}{\partial u_j} B_{jk} \right] \delta u_k$$

$$- \sum_{j=1}^{\mu+q} \frac{\partial f_i}{\partial u_j} \sum_{k=1}^{\mu} \gamma_{jk} \delta \dot{v}_k - \sum_{j=1}^{\mu+q} \frac{\partial f_i}{\partial u_j} \sum_{k=1}^{q} \gamma_{j,\mu+k} \delta z_k, \quad i = 1, 2, \ldots, n.$$

Let us introduce a set of functions $\psi_i(t)$ of t on (τ_0, τ_1) such that

$$\frac{\partial F}{\partial x_k} - \sum_{j=1}^{\mu+q} \frac{\partial F}{\partial u_j} A_{jk} = \frac{d\psi_k}{dt} + \sum_{i=1}^{n} \psi_i \left[\frac{\partial f_i}{\partial x_k} - \sum_{j=1}^{\mu+q} \frac{\partial f_i}{\partial u_j} A_{jk} \right], \quad k = 1, 2, \ldots, n. \tag{20}$$

Then it follows from

$$\dot{\psi}_i \delta x_i = \frac{d}{dt}(\psi_i \delta x_i) - \psi_i \delta \dot{x}_i$$

that

$$V = \int_{\tau_0}^{\tau_1} \left\{ \frac{d}{dt} \sum_{i=1}^{n} \psi_i \delta x_i + \sum_{k=\mu+q+1}^{m} \left[\frac{\partial F}{\partial u_k} - \sum_{j=1}^{\mu+q} \frac{\partial F}{\partial u_j} B_{jk} - \sum_{i=1}^{n} \psi_i \left(\frac{\partial f_i}{\partial u_k} - \sum_{j=1}^{\mu+q} \frac{\partial f_i}{\partial u_j} B_{jk} \right) \right] \delta u_k \right.$$

$$- \sum_{k=1}^{\mu} \sum_{j=1}^{\mu+q} \left(\frac{\partial F}{\partial u_j} \gamma_{jk} - \sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_j} \cdot \gamma_{jk} \right) \delta v_j - \sum_{k=1}^{q} \sum_{j=1}^{\mu+q} \left(\frac{\partial F}{\partial u_j} \gamma_{j,\mu+k} + \right.$$

$$\left. \left. - \sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_j} \gamma_{j,\mu+k} \right) \delta z_k \right\} dt.$$

Putting

$$\lambda_k = \sum_{j=1}^{\mu+q} \gamma_{j,\mu+k} \left(\sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_j} - \frac{\partial F}{\partial u_j} \right), \tag{21}$$

and

$$\alpha_k = \sum_{j=1}^{\mu+q} \gamma_{jk} \left(\sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_j} - \frac{\partial F}{\partial u_j} \right), \tag{22}$$

we find from

$$\sum_{j=1}^{\mu+q} \left(\sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_j} B_{jk} - \frac{\partial F}{\partial u_j} B_{jk} \right) = \sum_{j=1}^{\mu+q} \left(\sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_j} - \frac{\partial F}{\partial u_j} \right) \left(\sum_{l=1}^{\mu} \gamma_{jl} \frac{\partial h_l}{\partial u_k} + \sum_{l=1}^{q} \gamma_{j,\mu+l} \frac{\partial \varphi_l}{\partial u_k} \right) = \sum_{l=1}^{\mu} \alpha_l \frac{\partial h_l}{\partial u_k} + \sum_{l=1}^{q} \lambda_l \frac{\partial \varphi_l}{\partial u_k}$$

that

$$V = \left[\sum_{i=1}^{n} \psi_i \, \delta x_i \right]_{\tau_1}^{\tau_0} + \int_{\tau_0}^{\tau_1} \left\{ \sum_{k=\mu+q+1}^{m} \left(\frac{\partial F}{\partial u_k} - \sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_k} + \sum_{j=1}^{\mu} \alpha_j \frac{\partial h_j}{\partial u_k} + \sum_{j=1}^{q} \lambda_j \frac{\partial \varphi_j}{\partial u_k} \right) \delta u_k + \sum_{k=1}^{\mu} \alpha_k \delta \dot{v}_k + \sum_{k=1}^{q} \lambda_k \delta z_k \right\} dt.$$

Since

$$\sum_{j=1}^{\mu+q} \left(\sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_j} A_{jk} - \frac{\partial F}{\partial u_j} A_{jk} \right) = \sum_{j=1}^{\mu+q} \left(\sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_j} - \frac{\partial F}{\partial u_j} \right) \left(\sum_{l=1}^{\mu} \gamma_{jl} \frac{\partial h_l}{\partial x_k} \right) = \sum_{l=1}^{\mu} \alpha_l \frac{\partial h_l}{\partial x_k}$$

the functions $\psi_i(t)$ satisfy the differential equations*:

$$\dot{\psi}_k = \frac{\partial F}{\partial x_k} - \sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial x_k} + \sum_{l=1}^{\mu} \alpha_l \frac{\partial h_l}{\partial x_k}, \quad i = 1, 2, \ldots\ldots, n. \tag{23}$$

By virtue of

$$\sum_{j=1}^{\mu} \gamma_{lj} \frac{\partial h_j}{\partial u_k} + \sum_{j=1}^{q} \gamma_{l,\mu+j} \frac{\partial \varphi_j}{\partial u_k} = \delta_{lk} = \begin{cases} 0, & l \neq k \\ 1, & l = k \end{cases}$$

we have for $k = 1, 2, \ldots\ldots \mu+q$ the equations

* This set of differential equations is identical with the system of adjoint equations that Pontryagin introduced in his derivation of the maximum-principle.

$$\frac{\partial F}{\partial u_k} - \sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_k} + \sum_{j=1}^{\mu} \alpha_j \frac{\partial h_j}{\partial u_k} + \sum_{j=1}^{q} \lambda_j \frac{\partial \varphi_j}{\partial u_k} =$$

$$= \frac{\partial F}{\partial u_k} - \sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_k} + \sum_{j=1}^{\mu} \sum_{l=1}^{\mu+q} \gamma_{lj} \left(\sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_l} - \frac{\partial F}{\partial u_l} \right) \frac{\partial h_j}{\partial u_k} + \tag{24}$$

$$+ \sum_{j=1}^{q} \sum_{l=1}^{\mu+q} \gamma_{l,\mu+j} \left(\sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_l} - \frac{\partial F}{\partial u_l} \right) \frac{\partial \varphi_j}{\partial u_k} = 0.$$

Now suppose the functions α_k to be differentiable with respect to t on (τ_0, τ_1). Later on this assumption will be justified. Setting:

$$\mu_k = - \dot{\alpha}_k \tag{25}$$

we find from

$$- \frac{d}{dt} \left(\sum_{k=1}^{\mu} \alpha_k \delta g_k \right) = \frac{d}{dt} \left(\sum_{k=1}^{\mu} \alpha_k \delta v_k \right) = \sum_{k=1}^{\mu} \dot{\alpha}_k \delta v_k + \sum_{k=1}^{\mu} \alpha_k \delta \dot{v}_k =$$

$$= - \sum_{k=1}^{\mu} \mu_k \delta v_k + \sum_{k=1}^{\mu} \alpha_k \delta \dot{v}_k$$

that

$$V = \left[\sum_{i=1}^{n} \left(\psi_i - \sum_{k=1}^{\mu} \alpha_k \frac{\partial g_k}{\partial x_i} \right) \delta x_i \right]_{\tau_0}^{\tau_1} +$$

$$+ \int_{\tau_0}^{\tau_1} \left\{ \sum_{k=\mu+q+1}^{m} \left(\frac{\partial F}{\partial u_k} - \sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_k} + \sum_{j=1}^{\mu} \alpha_j \frac{\partial h_j}{\partial u_k} + \sum_{j=1}^{q} \lambda_j \frac{\partial \varphi_j}{\partial u_k} \right) \delta u_k + \right.$$

$$\left. + \sum_{k=1}^{\mu} \mu_k \delta v_k + \sum_{k=1}^{q} \lambda_k \delta z_k \right\} dt. \tag{26}$$

Introducing the functions

$$p_i = \psi_i - \sum_{j=1}^{\mu} \alpha_j \frac{\partial g_j}{\partial x_i}, \qquad i = 1, 2, \ldots\ldots, n \tag{27}$$

we obtain from (26) the expression:

$$V = \left[\sum_{i=1}^{n} p_i \delta x_i \right]_{\tau_0}^{\tau_1} + \int_{\tau_0}^{\tau_1} \left\{ \sum_{k=\mu+q+1}^{m} \left(\frac{\partial F}{\partial u_k} - \sum_{i=1}^{n} p_i \frac{\partial f_i}{\partial u_k} + \sum_{j=1}^{q} \lambda_j \frac{\partial \varphi_j}{\partial u_k} \right) \delta u_k + \right.$$

$$\left. + \sum_{k=1}^{\mu} \mu_k \delta v_k + \sum_{k=1}^{q} \lambda_k \delta z_k \right\} dt. \tag{28}$$

From

$$\dot{\psi}_i = \frac{\partial F}{\partial x_i} - \sum_{k=1}^{n} \psi_k \frac{\partial f_k}{\partial x_i} + \sum_{l=1}^{\mu} \alpha_l \frac{\partial h_l}{\partial x_i}$$

$$= \frac{\partial F}{\partial x_i} - \sum_{k=1}^{n} \psi_k \frac{\partial f_k}{\partial x_i} + \sum_{l=1}^{\mu} \alpha_l \left[\sum_{k=1}^{n} \frac{\partial^2 g_l}{\partial x_i \partial x_k} f_k + \sum_{k=1}^{n} \frac{\partial g_l}{\partial x_k} \frac{\partial f_k}{\partial x_i} + \frac{\partial^2 g_l}{\partial x_i \partial t} \right]$$

$$= \frac{\partial F}{\partial x_i} - \sum_{k=1}^{n} \psi_k \frac{\partial f_k}{\partial x_i} + \sum_{l=1}^{\mu} \alpha_l \left[\frac{d}{dt} \left(\frac{\partial g_l}{\partial x_i} \right) + \sum_{k=1}^{n} \frac{\partial g_l}{\partial x_k} \frac{\partial f_k}{\partial x_i} \right]$$

$$= \frac{\partial F}{\partial x_i} - \sum_{k=1}^{n} \left(\psi_k - \sum_{l=1}^{\mu} \alpha_l \frac{\partial g_l}{\partial x_k} \right) \frac{\partial f_k}{\partial x_i} + \sum_{l=1}^{\mu} \frac{d}{dt} \left(\alpha_l \frac{\partial g_l}{\partial x_i} \right) + \sum_{l=1}^{\mu} \mu_l \frac{\partial g_l}{\partial x_i}$$

it is seen that the functions $p_i(t)$ satisfy the adjoint differential equations

$$\dot{p}_i = \frac{\partial F}{\partial x_i} - \sum_{k=1}^{n} p_k \frac{\partial f_k}{\partial x_i} + \sum_{j=1}^{\mu} \mu_j \frac{\partial g_j}{\partial x_i}, \qquad i = 1, 2, \ldots, n. \tag{29}$$

Since

$$\frac{\partial F}{\partial u_k} - \sum_{i=1}^{n} \psi_i \frac{\partial f_i}{\partial u_k} + \sum_{j=1}^{\mu} \alpha_j \frac{\partial h_j}{\partial u_k} + \sum_{j=1}^{q} \lambda_j \frac{\partial \varphi_j}{\partial u_k} = \frac{\partial F}{\partial u_k} - \sum_{i=1}^{n} p_i \frac{\partial f_i}{\partial u_k} + \sum_{j=1}^{q} \lambda_j \frac{\partial \varphi_j}{\partial u_k}$$

it follows from (24) that for $k \leqslant q+\mu$

$$\frac{\partial F}{\partial u_k} - \sum_{i=1}^{n} p_i \frac{\partial f_i}{\partial u_k} + \sum_{j=1}^{q} \lambda_j \frac{\partial \varphi_j}{\partial u_k} = 0. \tag{30}$$

For $q < j \leqslant r$ we set $\lambda_j = 0$ and for $\mu < k \leqslant \nu$ we set $\mu_k = 0$. Assuming the functions $p_1(t)$, $p_2(t)$, ..., $p(t)$ to be continuous on $[T_0, T_1]$ we fix the solution of the differential equations (29) by imposing the endcondition, that at the point (X_1, T_1):

$$p_i = \frac{\partial J}{\partial x_i} \tag{31}$$

provided these derivatives exist in R.

The constraints $\varphi_j \leqslant 0$, $j > q$ and $g_k \leqslant 0$, $k > \mu$ can be ignored on (τ_0, τ_1); if t is some point of (τ_0, τ_1) then there exists a neighbourhood $(t-\Delta, t+\Delta)$ such that $\varphi_j(u + \delta u, t) \leqslant 0$, $j > q$ and $g_k(x + \delta x, t) \leqslant 0$, $k > \mu$ on $(t-\Delta, t+\Delta)$ for sufficiently small variations of the control-variables.

Now the functions $p_i(t)$ are continuous on $[T_0, T_1]$ and

$$\sum p_i(T_1)\delta x_i(T_1) = \delta J(X_1, T_1)$$

Thus, since

$$\int_{T_1}^{T_0} \left\{ F(x+\delta x, u+\delta u, t) - F(x, u, t) \right\} dt \geqslant \delta J(X_1, T_1)$$

and since the variations δu_k, $k > q + \mu$, δz_k and δv_k in (28) are arbitrary except for the condition $\delta z_k \geqslant 0$ and $\delta v_k \geqslant 0$ we obtain from the fundamental lemma of the calculus of variations (taking the variations $\delta\mu_k = 0$, $k > q+\mu$ outside a sufficiently small neighbourhood $(t - \Delta, t + \Delta)$ of $t \in (\tau_0, \tau_1)$) that also for $k = \mu+q+1, \mu+q+2, \ldots, m$

$$\frac{\partial F}{\partial u_k} - \sum_{i=1}^{n} p_i \frac{\partial f_i}{\partial u_k} + \sum_{j=1}^{q} \lambda_j \frac{\partial \varphi_j}{\partial u_k} = 0 \tag{32}$$

on (τ_0, τ_1) and moreover that

$$\lambda_k \geqslant 0, \qquad k = 1, 2, \ldots, q$$

and

$$\mu_k \geqslant 0, \qquad k = 1, 2, \ldots, \mu.$$

So we have on a sub-interval (τ_0, τ_1) of $[T_0, T_1]$, which is such that throughout (τ_0, τ_1) certain constraints vanish whereas the remaining constraints are less than zero, the following *necessary conditions* for the extremal joining (X_0, T_0) with (X_1, T_1):

$$\left.\begin{aligned}
&\frac{\partial F}{\partial u_j} - \sum_{i=1}^{n} p_i \frac{\partial f_i}{\partial u_j} + \sum_{k=1}^{r} \lambda_k \frac{\partial \varphi_k}{\partial u_j} = 0, \quad j = 1, 2, \ldots, m, \\
&\lambda_k \varphi_k = 0;\ \lambda_k \geqslant 0, \quad k = 1, 2, \ldots, r, \\
&\dot{p}_i = \frac{\partial F}{\partial x_i} - \sum_{j=1}^{n} p_j \frac{\partial f_j}{\partial x_i} + \sum_{k=1}^{\nu} \mu_k \frac{\partial g_k}{\partial x_i}, \quad i = 1, 2, \ldots, n, \\
&\mu_k g_k = 0;\ \mu_k \geqslant 0, \quad k = 1, 2, \ldots, \nu, \\
&\dot{x}_i = f_i, \quad i = 1, 2, \ldots, n.
\end{aligned}\right\} \tag{33}$$

Obviously x, u and p are differentiable with respect to t on (τ_0, τ_1). So the assumed differentiability of α_k in (25) is justified. If we introduce the Hamiltonian

$$H = -F + \sum_{i=1}^{n} p_i f_i - \sum_{j=1}^{r} \lambda_j \varphi_j - \sum_{k=1}^{\nu} \mu_k g_k, \tag{34}$$

then the necessary conditions can be expressed in the canonical form:

$$\left.\begin{aligned}
&\frac{\partial H}{\partial u_j} = 0, \qquad j = 1, 2, \ldots, m, \\
&\frac{\partial H}{\partial x_i} = -\dot{p}_i, \quad i = 1, 2, \ldots, n, \\
&\frac{\partial H}{\partial p_i} = \dot{x}_i, \qquad i = 1, 2, \ldots, n.
\end{aligned}\right\} \tag{35}$$

From these relations it is seen that the Hamiltonian satisfies the equation:

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}$$

on an interval (τ_0, τ_1).

Along the extremal x(t) joining (X_0, T_0) and (X_1, T_1) we have from (28) and (32) that

$$\delta \int_{\tau_2}^{\tau_1} F(x,u,t)dt = \left[\sum_{i=1}^{n} p_i \delta x_i \right]_{\tau_0}^{\tau_1} + \int_{\tau_0}^{\tau_1} \left(\sum_{k=1}^{r} \lambda_k \delta z_k + \sum_{k=1}^{\nu} \mu_k \delta v_k \right) dt.$$

Hence

$$\delta \int_{T_0}^{t} \left\{ F(x,u,\tau) + \sum_{k=1}^{r} \lambda_k(\tau)\varphi_k(u,\tau) + \sum_{k=1}^{\nu} \mu_k(\tau) g_k(x,\tau) \right\} d\tau = \sum_{i=1}^{n} p_i(t)\delta x_i(t). \qquad (36)$$

If we consider only variations of the control-variables for which $\delta x_i(T_1) = 0$ then it follows from this relation that

$$\delta \int_{T_0}^{T_1} \left\{ F + \sum_{k=1}^{r} \lambda_k \varphi_k + \sum_{k=1}^{\nu} \mu_k g_k \right\} dt = 0.$$

Therefore, among all curves joining (X_0, T_0) to (X_1, T_1) and satisfying the differential equations (5) the extremal x(t) minimizes the integral

$$\int_{T_0}^{T_1} \left\{ F + \sum_{k=1}^{r} \lambda_k \varphi_k + \sum_{k=1}^{\nu} \mu_k g_k \right\} dt. \qquad (37)$$

Note that the constraints (6) and (7) are omitted. Let $J^*(x,t)$ denote the integral

$$\int_{T_0}^{t} \left\{ F + \sum_{k=1}^{r} \lambda_k \varphi_k + \sum_{k=1}^{\nu} \mu_k g_k \right\} d\tau$$

from (X_0, T_0) to some point (x, t) of the extremal x(t). Since x(t) minimizes this integral as compared with all curves joining (X_0, T_0) to (x, t) and satisfying the differential equations (5) we have that

$$\delta \int_{T_0}^{t} \left(F + \sum_{k=1}^{r} \lambda_k \varphi_k + \sum_{k=1}^{\nu} \mu_k g_k \right) d\tau = \delta J^*(x,t) + \int_{T_0}^{t} \left\{ \sum_{j=1}^{m} \left[\frac{\partial (F + \Sigma \lambda_k \varphi_k + \Sigma \mu_k g_k)}{\partial u_j} - \sum_{i=1}^{n} p_i \frac{\partial f_i}{\partial u_j} \right] \delta u_j \right\} d\tau,$$

where

$$\frac{\partial (F + \Sigma \lambda_k \varphi_k + \Sigma \mu_k g_k)}{\partial u_j} - \Sigma p_i \frac{\partial f_i}{\partial u_j} = 0.$$

Consequently we find from (36) that

$$\Sigma p_i(t)\delta x_i(t) = \delta J^*(x,t). \qquad (38)$$

Since the extremal x(t) minimizes the integral (37) with the constraints (6) and (7) omitted it follows that the condition

$$\frac{\partial H}{\partial u_j} = 0$$

corresponds to a maximum of the Hamiltonian (34) (cf. Timman, [2], section 6), which expresses the maximum-principle.

3. *Endconditions*

Thusfar we assumed the endpoint (X_1, T_1) to be some fixed point in (x, t)-space. In applications however it often happens that a number of *endconditions* for the endpoint (X, T):

$$\chi_k(X, T) = 0, \qquad k = 1, 2, \ldots\ldots, s, \tag{39}$$

where X = x(T) has to be satisfied. It is then the point to minimize the function J*(X, T) subject to the conditions:

$$\chi_k(X, T) = 0, \qquad k = 1, 2, \ldots\ldots, s.$$

So the Lagrangian function

$$L = J^*(X, T) - \sum_{k=1}^{s} \beta_k \chi_k(X, T), \tag{40}$$

where $\beta_1, \beta_2, \ldots\ldots, \beta_s$ are s multipliers, has to be minimized. For the minimum of this function we have the necessary conditions:

$$\frac{\partial L}{\partial X_i} = \frac{\partial J^*}{\partial X_i} - \sum_{k=1}^{s} \beta_k \frac{\partial \chi_k}{\partial X_i} = 0 \tag{41}$$

and

$$\frac{\partial L}{\partial T} = \frac{\partial J^*}{\partial T} - \sum_{k=1}^{s} \beta_k \frac{\partial \chi_k}{\partial T} = 0, \tag{42}$$

the transversality conditions at the endpoint (X, T) of the extremal x(t) joining (X_0, T_0) with the manifold defined by the equations (39).

Let us consider the case that the endcondition

$$T = T_1$$

is prescribed, which condition corresponds to a variable endpoint X in x-space. From the transversality-conditions we find that

$$\frac{\partial J^*}{\partial X_i} = 0$$

at the endpoint (X, T). From (38) we have that at the endpoint (X, T) of the extremal joining (X_0, T_0) to (X, T)

$$\delta J^*(X, T) = \Sigma p_i(T) \delta X_i(T)$$

where apparently $\delta J^*(X, T) = 0$. Therefore we can impose the endcondition

$$p_i(T_1) = 0 \tag{43}$$

for the system of adjoint differential equations in case of a problem with free endpoint X.

Let us now consider the case that the endcondition

$$X = X_1 \tag{44}$$

is prescribed, which condition corresponds to a fixed endpoint X with T variable. From the transversality-condition we find that

$$\frac{\partial J^*}{\partial T} = 0$$

at the endpoint (X, T). Now we have from (38) that for a displacement $(\Delta X, \delta T)$ of the endpoint (X, T):

$$\Delta J^* = \delta J^* + \left(F + \sum_{j=1}^{r} \lambda_j \varphi_j + \sum_{k=1}^{\nu} \mu_k g_k\right) \delta T$$

$$= \sum_{i=1}^{n} p_i \delta X_i + \left(F + \sum_{j=1}^{r} \lambda_j \varphi_j + \sum_{j=1}^{\nu} \mu_k g_k\right) \delta T,$$

where $\delta X_i = \Delta X_i - f_i \delta T$. Hence

$$\Delta J^* = \Sigma p_i \Delta X_i + H \delta T,$$

where δT is an arbitrary variation. Thus we obtain the endcondition

$$H = 0 \tag{45}$$

for a problem with fixed endpoint X.

4. *Final remarks*

If constraints of the form

$$\varphi_j(u, x, t) \leqslant 0$$

for the control-variables are prescribed then the same conditions as (35) hold for the extremal joining (X_0, T_0) and (X_1, T_1). The derivation of these conditions is quite analogous to that given in section 2.

When maximizing the Hamiltonian H at some point t of $[T_0, T_1]$ obviously the term $\sum_{k=1}^{\nu} \mu_k g_k$ does not matter.

The state-variables are found by means of the differential equations (5), subject to the conditions (7).

In practical applications the constraints often have the form:

$$a_j \leqslant u_j \leqslant b_j, \qquad j = 1, 2, \ldots\ldots, m \tag{46}$$

for the control-variables and

$$\alpha_k \leqslant x_k \leqslant \beta_k, \qquad k = 1, 2, \ldots\ldots, n \tag{47}$$

for the state-variables. In numerical computations suchlike constraints are easy to deal with, In case $x_k = \alpha_k$ or $x_k = \beta_k$ in (47) $f_k = 0$ which implies that the Hamiltonian does not contain p_k.

As long as the constraints (7) are all less than zero the functions $p_i(t)$ are determined from the system (29) of adjoint differential equations, where $\mu_j = 0$. As soon as some of these constraints vanish the relations (27) may be used, where α_j is found from (22) and the functions $\psi_i(t)$ are solutions of the system (23) of differential equations. The unknown initial value $p(T_0)$ which must be chosen such that the endconditions are satisfied means a fairly complicated feature of the numerical procedure.

REFERENCES

1. L.S. Pontryagin, V.G. Boltyanskii, R.S. Gamkrelidze, E.F. Mishchenko. The Mathematical Theory of Optimal Processes, Pergamon Press, London, 1964.
2. R. Timman. Optimization Theory for Ordinary Differential Equations, Journal of engineering mathematics, I, 3, 1967, p. 159-186.

[Received February 7, 1968]