OPTIMIZATION THEORY FOR ORDINARY DIFFERENTIAL EQUATIONS

by

R. Timman*

1. *Introduction.*

In this paper an exposition will be given which is similar in structure to the classical calculus of variations for the determination of necessary conditions for the existence of a field of extremals for control problems, where the control variables or the state variables are subject to a set of inequality conditions.

If necessary conditions are derived for the existence of a field of extremals starting from a fixed point x_0^i for a value t_0 of the parameter t, it is relatively simple to derive Pontryagin's principle [Ref. 1].

Starting with a simple formulation of the necessary conditions for the minimum of a function of n variables x^i, which is the basic problem of non-linear programming, at one side and a rederivation of the classical Weierstrass condition at the other side, Pontryagin's principle for this never restricted case (existence of a field) is found. Here still variations are admitted which show discontinuities in the control variables for arbitrary values of the parameter t. It appears, however, that only discontinuities in the extremals will occur at a change of constraint conditions.

Hence it does not seem necessary to admit such discontinuities for arbitrary values of t, and the derivation of the corresponding conditions can be given for intervals, where a certain set of equality conditions are satisfied throughout this interval.

Then the derivation goes completely along the classical lines [compare e.g. Ref. 2].

For constraints on the state variables a similar reasoning immediately leads to results, which are a generalization of Pontryagin's result for one restriction [Ref. 1, chapter IV].

Finally it is shown that for time-optimal systems which are linear in the state and control variables, the solution requires the solution of a linear programming problem at each instant t. Here it is obvious that always discontinuities in the control variables occur, which correspond to a switch from one constraint to another.

It should be remarked that most of the results given here can be found in the recent work of Hestenes [Ref. 3].

2. *The minimum of a function of several variables with inequality constraints.*

As an introduction to the method which will be followed in the succeeding sections, we consider the minimization of a function $F(x^i)$ of n real variables x^i $(i = 1, \dots, n)$.

The function $F(x^i)$ is supposed to be of class C^2, i.e. the partial derivatives $F_{x^i x^j}$ exist everywhere and are continuous.

The problem is formulated as the requirement to determine a point $\hat{x}^i$, where $F(x^i)$ has a minimum value in the region R, determined by a set of inequality constraints

$$f^k(x^i) \leqq 0, \qquad k = 1, \dots, m.$$

where the $f^k(x^i)$ are also of class C^2.

* Technological University, Delft, the Netherlands.

In order to derive a set of necessary conditions, we assume that $\hat{x}^i$ is a minimum point, satisfying the conditions

$$f^k(\hat{x}^i) = 0, \qquad k = 1,\ldots,l \leqq n,$$

$$f^k(\hat{x}^i) < 0, \qquad k = l+1,\ldots,m.$$

Then for any variation δx^i satisfying

$$f^k(\hat{x}^i + \delta x^i) \leqq 0, \qquad k = 1,\ldots,m,$$

we must have

$$F(\hat{x}^i + \delta x^i) \geqq F(\hat{x}^i).$$

In order to transform these conditions we further assume that $F(x^i)$ is locally convex in a neighbourhood of $\hat{x}^i$, a property which is defined by the condition that for all variations in this neighbourhood

$$F(\hat{x}^i + \delta x^i) - F(\hat{x}^i) - F_{\hat{x}^j}\delta x^j \geqq 0.$$

Geometrically this means that the surface $F(x^i)$ in a n+1-dimensional space is everywhere above its tangent plane.

Then the requirement can be replaced by

$$F(\hat{x}^i + \delta x^i) - F(\hat{x}^i) \geqq F_{\hat{x}^j}\,\delta x^j \geqq 0,$$

for suppose that for a certain set of variables δx^j

$$F_{\hat{x}^j}\,\delta x^j \leqq 0,$$

we can, regarding the existence of continuous derivatives of second order, expand $F(\hat{x}^i + \delta x^i)$ into a Taylor series:

$$F(\hat{x}^i + \delta x^i) - F(\hat{x}^i) = F_{\hat{x}^j}\,\delta x^j + \tfrac{1}{2}\tilde{F}_{x^i x^j}\,\delta x^i \delta x^j,$$

where $\tilde{F}_{x^i x^j}$ denote the derivatives in a point of the neighbourhood.

Introducing a factor $\lambda > 0$, we then have

$$F(\hat{x}^i + \lambda\delta x^i) - F(\hat{x}^i) = F_{\hat{x}^j}\,\lambda\delta x^j + \tfrac{1}{2}\lambda^2\tilde{\tilde{F}}_{x^i x^j}\,\delta x^i \delta x^j$$

(where $\tilde{\tilde{F}}_{x^i x^j}$ may be different from the previous value $\tilde{F}_{x^i x^j}$).

By taking λ small enough, it is possible to arrange that the first term of the right-hand side exceeds the second one in absolute value. In this case

$$F(\hat{x}^i + \lambda\delta x^i) - F(\hat{x}^i) < 0$$

in contradiction to the assumption.

In a similar way it can be shown that the conditions

$$f^k(\hat{x}^i + \delta x^i) \leqq 0, \qquad k = 1,\ldots,l.$$

corresponding to the equality constraints

$$f^k(\hat{x}^i) = 0, \qquad k = 1,\ldots,l.$$

imply that

$$f^k_{\hat{x}^i}\,\delta x^i \leqq 0, \qquad k = 1,\dots,l.$$

On the other hand inequality constraints

$$f^k(\hat{x}^i) < 0, \qquad k = l+1,\dots,m,$$

do not imply an essential restriction on the δx^i, for, if $\max(\delta x^i)$ is sufficiently small, we always have

$$f^k(\hat{x}^i + \delta x^i) < 0, \qquad k = l+1,\dots,m,$$

because of the continuity of the $f^k(x^i)$.

From these considerations follows the necessary condition for a local minimum that

$$F_i\,\delta x^i \geqq 0, \qquad i = 1,\dots,n,$$

for all variations δx^i which satisfy

$$f^k_i\delta x^i \leqq 0, \qquad k = 1,\dots,l;\ i = 1,\dots,n.$$

Here the derivatives $F_{\hat{x}^i}$ and $f^k_{\hat{x}^i}$ are abbreviated as F_i and f^k_i. For a further treatment we introduce l slack variations $\delta z^k \geqq 0$, so that

$$f^k_i\delta x^i + \delta z^k = 0. \qquad k = 1,\dots,l;\ i = 1,\dots,n.$$

The constraints are supposed to be *independent*. This means that in the region considered the rank of the matrix f^k_i is l.

Then by rearranging the variables, we can always obtain that

$$\det(f^k_i) \neq 0, \qquad k = 1,\dots,l;\ i = 1,\dots,l.$$

Rewriting the constraint conditions as

$$f^k_i\delta x^i = -f^k_{l+h}\,\delta x^{l+h} - \delta z^k, \qquad i,\ k = 1,\dots,l;\ h = 1,\dots,n\text{-}l,$$

the first l δx^i can be expressed into the remaining δx^i and δz^k. For this expression is needed the reciprocal matrix γ^i_k of the f^k_j, defined by

$$\gamma^i_k f^k_j = \delta^i_j = 1, \quad \text{for} \quad i = j,$$
$$= 0, \quad \text{for} \quad i \neq j.$$

This gives

$$\delta x^i = -\gamma^i_k f^k_{l+h}\,\delta x^{l+h} - \gamma^i_k\delta z^k,$$

where now the δx^{l+h} are no longer subject to any restriction and the δz^k are $\geqq 0$. These values are substituted in the requirement

$$F_i\,\delta x^i + F_{l+h}\,\delta x^{l+h} \geqq 0, \qquad i = 1,\dots,l;\ h = 1,\dots,n\text{-}l.$$

yielding

$$(F_{l+h} - F_i\gamma^i_k f^k_{l+h})\delta x^{l+h} - F_i\gamma^i_k\delta z^k \geqq 0$$

for arbitrary δx^{l+h} and $\delta z^k \geqq 0$.

Introducing the vector

$$p_k = \gamma_k^i F_i, \quad i, k = 1, \dots, l,$$

we have

$$(F_{l+h} - p_k f_{l+h}^k)\, \delta x^{l+h} - p_k\, \delta z^k \geqq 0.$$

Since the δx^{l+h} are arbitrary, this leads to

$$F_{l+h} - p_k f_{l+h}^k = 0, \quad k = 1, \dots, l;\ h = 1, \dots, n-l,$$

and since the $\delta z^k \geqq 0$, the p_k must satisfy the restriction

$$p_k \leqq 0.$$

(This is essentially the Farkas lemma).

From the definition of the p_k we see that also

$$F_i - p_k f_i^k = F_i - \gamma_k^j F_j f_i^k = F_i - \delta_i^j F_j = F_i - F_i = 0, \quad i = 1, \dots, l.$$

This result justifies the Lagrange multiplier method:

If $F(x^i)$ has a minimum for $x^i = \hat{x}^i$, with equality for l of the constraints

$$f^k(\hat{x}^i) = 0, \quad k = 1, \dots, l,$$

there exist l multipliers $p_k \leqq 0$ $(k = 1, \dots, l)$, so that

$$F_i - p_k f_i^k = 0, \quad i = 1, \dots, n.$$

If the constraints are also locally convex, we now can easily derive the Kuhn-Tucker theorem by a consideration of the function

$$R(x^i, q_k) = F(x^i) - q_k f^k(x^i), \quad i = 1, \dots, n;\ k = 1, \dots, m.$$

Here the variables x^i satisfy the constraints

$$f^k(x^i) \leqq 0, \quad k = 1, \dots, m;\ 1 = 1, \dots, n,$$

while

$$q_k \leqq 0, \quad k = 1, \dots, m.$$

If $\hat{x}^i$ is the minimum point and $q_k = p_k$ for $k = 1, \dots, l$, and $q_k = 0$ for $k = l+1, \dots, m$, this theorem states that

$$R(x^i, p_k) \geqq R(\hat{x}^i, p_k) \geqq R(\hat{x}^i, q_k).$$

In fact, in the region, where F and f^k are all convex:

$$R(x^i, p_k) - R(\hat{x}^i, p_k) = F(x^i) - F(\hat{x}^i) - p_k \left\{ f^k(x^i) - f^k(\hat{x}^i) \right\}$$

$$\geqq F_i (x^i - \hat{x}^i) - p_k f_i^k (x^i - \hat{x}^i) = (F_i - p_k f_i^k)(x^i - \hat{x}^i) = 0.$$

On the other hand

$$R(\hat{x}^i, p_k) - R(\hat{x}^i, q_k) = F(\hat{x}^i) - p_k f^k(\hat{x}^i) - F(\hat{x}^i) + q_k f^k(\hat{x}^i) + q_{l+j} f^{l+j}(\hat{x}^i) =$$

$$= q_{l+j} f^{l+j}(\hat{x}^i) \geqq 0, \qquad k = 1, \ldots, l; \quad j = 1, \ldots, m-l.$$

The function $R(x^i, q_k)$ has a saddle point in $\hat{x}^i$, p_k, i.e. it has a minimum with respect to the x^i and a maximum with respect to the q_k.

3. *The problem of the classical calculus of variations.*

As a preparation for the investigation of the minimization of integrals relating to control systems, some basic theorems of the classical calculus of variations will be rederived following a method which is also applicable to control problems.

Instead of a function $F(x^i)$ we now consider an integral

$$J = \int_{T_o}^{T_1} F(x^i(t), \dot{x}^i(t), t)\, dt, \qquad i = 1, \ldots, n,$$

where the function F is of class C^2 in the variables x^i, $\dot{x}^i$, t and the $x^i(t)$ are functions of a parameter t. The end points of the integral are fixed

for $t = T_o$: $x^i = X^i_o$,

for $t = T_1$: $x^i = X^i_1$.

The requirement is to find a set of functions $\hat{x}^i(t)$, for which the value of the integral J is a minimum compared to other curves $x^i(t)$ running through the same points.

In order to facilitate the later discussion, the integral is written into the form

$$J = \int_{T_o}^{T_1} F(x^i(t), u^i(t), t)\, dt,$$

where the x^i and u^i are related by

$$\dot{x}^i = u^i(t), \qquad i = 1, \ldots, n.$$

If the values of x^i for $t = T_o$ are fixed, the values of x^i are completely determined, if the values of $u^i(t)$ are given. We follow the method of section 2 regarding the equations as constraints. If a necessary condition is to be derived for the minimization problem, it is again supposed that

$$J = \int_{T_o}^{T_1} F(\hat{x}^i(t), \hat{u}^i(t), t)\, dt,$$

$$\dot{\hat{x}}^i = \hat{u}^i(t),$$

has a minimum value compared to other curves $x^i(t)$, $u^i(t)$.

Then for any other curve

$$u^i(t) = \hat{u}^i(t) + \delta u^i(t),$$

$$x^i(t) = \hat{x}^i(t) + \delta x^i(t),$$

we must have

$$\int_{T_0}^{T_1} F(\hat{x}^i + \delta x^i,\ \hat{u}^i + \delta u^i,\ t)\,dt \geqq \int_{T_0}^{T_1} F(\hat{x}^i,\ \hat{u}^i,\ t)\,dt.$$

Since the starting points are fixed, we have

$$\delta x^i(t) = \int_{T_0}^{t} \delta u^i(\tau)d\tau.$$

For a fixed end point the class of $\delta u^i(t)$ is restricted by the condition

$$\delta x^i(T_1) = \int_{T_0}^{T_1} \delta u^i(\tau)d\tau = 0.$$

Instead of restricting the variations δu^i to this class, we consider primarily general variations δu^i, which vanish for $t = T_0$ but otherwise are only restricted by the condition that the corresponding δx^i are small of order ϵ.

A variation δu^i is called an ϵ-variation, if it is integrable (in the Lebesgue sense) and if, for any bounded function $\theta(t)$, i.e. $|\theta(t)| \leqq M$ for $T_0 \leqq t \leqq T_1$ the integral

$$\max_i \left| \int_{T_0}^{T_1} \theta(t)\delta u^i(t)dt \right| \leqq \epsilon M.$$

A special class of these variations occurs, when $\delta u^i(t)$ is uniformly small, i.e.

$$\max_i \left| \delta u^i(t) \right| \leqq \frac{\epsilon}{T_1 - T_0},$$

but also locally, finite values of δu^i over a short interval can be admitted, e.g.

$$\begin{aligned} \delta u^i &= 0, & T_0 &\leqq t < T - \tfrac{1}{2}\epsilon \\ \delta u^i &= U^i, & T - \tfrac{1}{2}\epsilon &\leqq t \leqq T + \tfrac{1}{2}\epsilon \\ \delta u^i &= 0, & T + \tfrac{1}{2}\epsilon &< t \leqq T_1, \end{aligned}$$

where $\max U^i = 1$.

For these general variations the end point is not necessarily fixed, in particular for the variation quoted above $\delta x^i(T_1) = \epsilon U^i$.

For this type of variations necessary conditions will be derived for a case which occurs mostly in applications, viz. the existence of a field of extremals starting from the point X_0^i at $t = T_0$.

The field of extremals exists, if for every point X^i of a certain open region G in R^n and every T of an open interval a curve exists which joins X_0^i, T_0 with X^i, T, so that along this curve the integral

$$J(X^i, T) = \int_{T_0}^{T} F(x^i(t), u^i(t), t)dt; \quad u^i = \dot{x}^i$$

is a minimum compared to any other differentiable curve, for which $x^i(T_0) = X_0^i$ and $x^i(T) = X^i$. Moreover the minimum value $J(X^i, T)$ of the integral considered as a function of X^i and T is supposed to have partial derivatives of the first order with respect to the variables X^i and T.

Consider now an extremal of the field which reaches the point X_1^i at $t = T_1$. We first derive the classical Euler-Lagrange equations for variations along this curve, which are uniformly small.

If $\max_i |\delta u^i(t)| \leqq \frac{\epsilon}{T_1 - T_o}$,

we have

$$|\delta x^i| = \left| \int_{T_o}^{t} \delta u^i(\tau) d\tau \right| \leqq \epsilon.$$

The difference of the integral along $\hat{x}^i(t)$, $\hat{u}^i(t)$ and the varied curve

$$\int_{T_o}^{T_1} \left\{ F(\hat{x}^i + \delta x^i, \hat{u}^i + \delta u^i, t) - F(x^i, u^i, t) \right\} dt$$

can now be expanded into a Taylor series. This gives

$$\int_{T_o}^{T_1} (F_{\hat{x}^i} \delta x^i + F_{\hat{u}^i} \delta u^i) dt + 0(\epsilon^2) = \int_{T_o}^{T_1} \left\{ F_{\hat{x}^i} \int_{T_o}^{t} \delta u^i(\tau) d\tau + F_{\hat{u}^i} \delta u^i \right\} dt + 0(\epsilon^2).$$

The integral is now transformed by partial integration. We introduce a vector p_i, which satisfies

$$\dot{p}_i = F_{\hat{x}^i},$$

and find for uniformly small ϵ-variations δu^i that the difference is

$$\int_{T_o}^{T_1} \left\{ \dot{p}_i \int_{T_o}^{t} \delta u^i(\tau) d\tau + F_{\hat{u}^i} \delta u^i \right\} dt + 0(\epsilon^2) =$$

$$= \left[p_i \int_{T_o}^{t} \delta u^i(\tau) d\tau \right]_{T_o}^{T_1} + \int_{T_o}^{T_1} \left\{ - p_i + F_{\hat{u}^i} \right\} \delta u^i(t) dt + 0(\epsilon^2) =$$

$$= p_i \delta x^i(T_1) + \int_{T_o}^{T_1} \left\{ - p_i + F_{\hat{u}^i} \right\} \delta u^i(t) dt + 0(\epsilon^2).$$

Since we assumed that every point in a neighbourhood of X_1^i in the field could be reached by an extremal, we can make ϵ sufficiently small that $X_1^i + \delta x^i$ can be reached.

We regard further that $p_i(T_1)$ was arbitrary. If we choose as boundary values for p_i the values $J_{X_1^i}$, the term $p_i \delta x^i$ exactly equals the difference between the minimum values of the integrals from X_o^i to X_1^i and to $X_1^i + \delta x^i(T_1)$ (apart from terms of second order).

This means that in this case

$$\int_{T_o}^{T_1} \left\{ - p_i + F_{\hat{u}^i} \right\} \delta u^i(t) dt + 0(\epsilon^2) \geqq 0.$$

Since for sufficiently small ϵ the first term is dominant and δu^i is arbitrary, this leads to

$$p_i = F_{\hat{u}^i} .$$

Regarding the definition of p_i we see that

$$\dot{p}_i = \frac{d}{dt} F_{\hat{u}^i} = F_{\hat{x}^i} ,$$

which expresses the Euler-Lagrange equations.

4. The Weierstrass condition.

For general ϵ-variations the preceding consideration is not valid. If the integrand is developed into a Taylor series with respect to δu^i, the remainder terms are no longer of order ϵ^2 uniformly in the interval, however, this still applies to the expansion with respect to δx^i.

We again consider necessary conditions for the existence of a class of extremals which join the point X_0^i at $t = T_0$ to every point of a certain neighbourhood of the point X_1^i at values in an interval round T_1.

Suppose $\hat{x}^i(t)$, $\hat{u}^i(t)$ is the extremal which reaches X_1^i at $t = T_1$. The expression for a neighbouring extremal in rewritten

$$\int_{T_0}^{T_1} \left\{F(x^i(t), u^i(t), t) - F(\hat{x}^i(t), \hat{u}^i(t), t)\right\}dt =$$

$$= \int_{T_0}^{T_1} \left\{F(x^i, u^i, t) - F(x^i, \hat{u}^i, t) + F(x^i, \hat{u}^i, t) - F(\hat{x}^i, \hat{u}^i, t)\right\}dt$$

and use is made of the expansion

$$F(x^i, \hat{u}^i, t) - F(\hat{x}^i, \hat{u}^i, t) = F(\hat{x}^i + \delta x^i, \hat{u}^i, t) - F(\hat{x}^i, \hat{u}^i, t) = F_{\hat{x}^i}(\hat{x}^i, \hat{u}^i, t)\delta x^i + 0(\epsilon^2).$$

This leads to

$$\int_{T_0}^{T_1} \left\{F(x^i, u^i, t) - F(x^i, \hat{u}^i, t) + F_{\hat{x}^i}(\hat{x}^i, \hat{u}^i, t)\delta x^i\right\}dt + 0(\epsilon^2).$$

The first term is also of order ϵ. In fact, according to the mean value theorem

$$F(x^i, u^i, t) - F(x^i, \hat{u}^i, t) = \tilde{F}_{u^i}(u^i - \hat{u}^i) = \tilde{F}_{u^i}\delta u^i ,$$

where the derivative $\tilde{F}_{u^i}$ is bounded.

Using the fact that δu^i is an ϵ-variation the integral

$$\int_{T_0}^{T_1} \left\{F(x^i, u^i, t) - F(x^i, \hat{u}^i, t)\right\}dt$$

is of order ϵ, together with the second part of the integral.

As before the integral is transformed by partial integration. Defining the p_i by

$$\dot{p}_i = F_{\hat{x}^i}$$

the result is:

$$p_i(T_1)\delta x^i(T_1) + \int_{T_0}^{T_1} \left\{F(x^i, u^i, t) - F(x^i, \hat{u}^i, t) - p_i \delta u^i\right\} dt + 0(\epsilon^2).$$

The end values $p_i(T_1)$ are now fixed as the values of the partial derivatives of the minimum values of the integrals $J(X_1^i + \Delta X^i, T_1 + \Delta T)$ as functions of X_1^i and T_1:

$$p_i = J_{X^i}.$$

Then the integral passes into

$$J(X_1^i + \delta x^i, T_1) - J(X_1^i, T_1) + \int_{T_0}^{T_1} \left\{F(x^i, u^i, t) - F(x^i, \hat{u}^i, t) - p_i \delta u^i\right\} dt + 0(\epsilon^2).$$

With these values of p_i, which are now fixed by the end conditions the integral

$$\int_{T_0}^{T_1} \left\{F(x^i, u^i, t) - F(x^i, \hat{u}^i, t) - p_i(u^i - \hat{u}^i)\right\} dt + 0(\epsilon^2)$$

represents the difference between the integrals of the varied curve and the extremal to the point $X_1^i + \delta x^i$.

The condition then transforms into

$$\int_{T_0}^{T_1} \left\{F(x^i, u^i, t) - F(x^i, \hat{u}^i, t) - p_i(u^i - \hat{u}^i)\right\} dt + 0(\epsilon^2) \geqq 0.$$

If this is to be true, a necessary condition is that for every ϵ-variation δu^i the first order part

$$\int_{T_0}^{T_1} \left\{F(x^i, u^i, t) - F(x^i, \hat{u}^i, t) - p_i(u^i - \hat{u}^i)\right\} dt \geqq 0.$$

Now take for u^i the variation for arbitrary T in the interval

$$\delta u^i = 0, \quad T_0 \leqq t < T - \tfrac{1}{2}\epsilon,$$
$$\delta u^i = U^i, \quad T_0 - \tfrac{1}{2}\epsilon \leqq t \leqq T + \tfrac{1}{2}\epsilon,$$
$$\delta u^i = 0, \quad T + \tfrac{1}{2}\epsilon < t \leqq T_1.$$

This shows that

$$\int_{T-\frac{1}{2}\epsilon}^{T+\frac{1}{2}\epsilon} \left\{F(x^i, \hat{u}^i + U^i, t) - F(x^i, \hat{u}^i, t) - p_i U^i\right\} dt \geqq 0$$

for arbitrary small values of ϵ.

This is only possible, if for every T (except possibly in a set of measure zero)

$$F(x^i, \hat{u}^i + U^i, t) - F(x^i, \hat{u}^i, t) - p_i U^i \geqq 0.$$

If we use the Euler-Lagrange equations, thus $p_i = F_{u^i}$, we obtain the Weierstrass condition.

If the variations δu^i are subject to restrictions, the Euler-Lagrange equations are not necessarily satisfied, but the foregoing reasoning holds for all admissible δu^i.

5. The variational equations of control theory.

In control theory the fundamental problem is the minimization of an integral

$$\int_{T_0}^{T_1} F(x^i, u^k, t)dt, \qquad i = 1, \ldots, n; \; k = 1, \ldots, m,$$

where the state variables x^i depend on the control variables u^k through a set of differential equations

$$\dot{x}^i = f^i\{x^j(t), \; u^k(t), t\}.$$

The functions F, f^i are supposed to be of class C^2 in the variables x^i, u^k. If the initial point $x^i(T_0) = X_0^i$ is fixed, the trajectories are uniquely determined by a choice of the control variables $u^k(t)$.

In general these variables are subject to certain inequality constraints. The problem is to determine a set of control variables $\hat{u}^k(t)$, so that a point X_1^i is reached at a value $t = T_1$ and the integral along the trajectories $\hat{x}^i(t)$, determined by these control variables, has a minimum value compared to all other trajectories which join X_0^i, T_0 with X_1^i, T_1.

We shall derive necessary conditions corresponding to a more general situation.

Here it will be assumed that there exists a class of extremals which join the point X_0^i, T_0 with every point of a neighbourhood

$$X^i = X_1^i + \Delta X^i, \; T = T_1 + \Delta T \qquad \text{of the point } X_1^i, \; T_1.$$

Before deriving these necessary conditions for the minimization, we study the relation between ϵ-variations δu^k and the corresponding variation δx^k.

We recall that δu^k is an ϵ-variation, if for every bounded function $\theta(t)$ with $|\theta(t)| \leqq M$

$$\max_k \left| \int_{T_0}^{T_1} \theta(t) \, u^k(t)dt \right| \leqq \epsilon M.$$

The varied curve $x^i + \delta x^i$ is a solution of

$$\dot{x}^i + \delta\dot{x}^i = f^i\{x^j + \delta x^j, \; u^k + \delta u^k, \; t\}$$

and the variation δx^i satisfies

$$\delta\dot{x}^i = f^i\{x^j + \delta x^j, \; u^k + \delta u^k, \; t\} - f^i\{x^j, u^k, t\}.$$

We first show that, if δu^k is an ϵ-variation neglecting terms of second order, δx^i is uniformly small and is a solution of a non-homogeneous linear equation.

Hereto we expand the right-hand side in a Taylor expansion with respect to δx^i

$$\delta\dot{x}^i = f^i\left\{x^j + \delta x^j, u^k + \delta u^k, t\right\} - f^i\left\{x^j, u^k + \delta u^k, t\right\} + f^i\left\{x^j, u^k + \delta u^k, t\right\} - $$
$$- f^i\left\{x^j, u^k, t\right\} = f^i_{x^j}\delta x^j + f^i(x^j, u^k + \delta u^k, t) - f^i(x^j, u^k, t) + \tfrac{1}{2}\tilde{f}^i_{x^j x^l}\delta x^j \delta x^l$$

where the last term denotes the second derivatives in an intermediate point.

The equation will be transformed into an integral equation of Volterra type for δx^i by an investigation of the non-homogeneous linear equation

$$\dot{\xi}^i = f^i_j \xi^j + g^i,$$

where $f^i_j = f^i_{x^j}$.

The equation is solved by means of the solution of the adjoint equation

$$\dot{\psi}_j = - f^i_j \psi_i ,$$

remarking that for every solution ξ^i of the first equation and every solution ψ_i of the adjoint equation

$$\frac{d}{dt}(\xi^i \psi_i) = \dot{\xi}^i \psi_i + \xi^i \dot{\psi}_i = f^i_j \xi^j \psi_i + g^i \psi_i - \xi^i f_i \psi_j = g^i \psi_i ,$$

hence for the solution ξ^i of the non-homogeneous equation which is zero for $t = T_o$ and any solution ψ_i of the adjoint equation

$$\xi^i(t)\psi_i(t) = \int_{T_o}^{t} \psi_i(\tau) g^i(\tau) d\tau.$$

In order to solve ξ^i we remark that for any solution η^i of the homogeneous equation

$$\dot{\eta}^i = f^i_j \eta^j$$

and ψ_i we have

$$\frac{d}{dt}(\eta^i \psi_i) = 0.$$

Now take a set η^i_j which for $t = T_o$ passes into the unit vectors

$$\eta^i_j(T_o) = \delta^i_j$$

and for ψ^k_i also a set, for which

$$\psi^k_i(T_o) = \delta^k_i ,$$

then

$$\eta^i_j(t)\psi^k_i(t) = \delta^i_j \delta^k_i = \delta^k_j ,$$

and we see that the matrix $\psi^k_i(t)$ is the reciprocal of the matrix $\eta^i_j(t)$.

Since

$$\xi^i(t)\psi^k_i(t) = \int_{T_o}^{t} \psi^k_i(\tau) g^i(\tau) d\tau,$$

we have

$$\xi^i(t)\psi_i^k(t)\eta_k^j(t) = \eta_k^j(t)\int_{T_o}^{t} \psi_i^k(\tau)g^i(\tau)d\tau$$

or

$$\xi^j(t) = \eta_k^j(t)\int_{T_o}^{t} \psi_i^k(\tau)g^i(\tau)d\tau.$$

In this way we transform the equation for δx^i into

$$\delta x^i(t) = \eta_k^j(t)\int_{T_o}^{t} \psi_i^k(\tau)\left[f^i(u^k+\delta u^k) - f^i(u^k)\right]d\tau + \tfrac{1}{2}\eta_k^j(t)\int_{T_o}^{t} \psi_i^k(\tau)\tilde{f}^i_{x^j x^l}\,\delta x^j(\tau)\delta x^l(\tau)d\tau.$$

Now $\psi_i^k(t)$ and $\eta_k^j(t)$ are bounded and $f^i(u^k+\delta u^k) - f^i(u^k) = \tilde{f}^i_{u^k}\delta u^k$ from the mean value theorem. Application of successive approximations gives for $\delta x^i(t)$ a series. The first term is of order ϵ, since δu^k is an ϵ-variation.

Then the successive additions which result from the last term in the integral equations are of order ϵ^2.

Hence we can write

$$\delta x^i(t) = \eta_k^j(t)\int_{T_o}^{t} \psi_i^k(\tau)\left[f^i(u^k+\delta u^k) - f^i(u^k)\right]d\tau + 0(\epsilon^2)$$

and to the first order of approximation δx^i satisfies the non-homogeneous linear equations

$$\delta\dot{x}^i(t) = f^i_j\delta x^j(t) + f^i(u^k+\delta u^k) - f^i(u^k).$$

6. *The maximum principle.*

The necessary condition for the existence of an extremal which is imbedded in the above-mentioned field of extremals can now be derived. If $\hat{x}^i(t)$, $\hat{u}^i(t)$ is the extremal from X_o^i, T_o to X_1^i, T_1 we consider an ϵ-variation δu^i which gives another curve $x^i(t)$:

$$x^i(t) = \hat{x}^i(t) + \delta x^i(t), \quad T_o \leqq t \leqq T_1,$$

with

$$u^i(t) = \hat{u}^i(t) + \delta u^i(t), \quad T_o \leqq t \leqq T_1.$$

where $u^i(t)$ is supposed to be completely inside the admissible domain of the control variables.

The value of the integral along $\hat{x}^i(t)$, $\hat{u}^k(t)$ is $J(X_1^i, T_1)$ and the difference between the value of the integral along the curve $x^i(t)$ and $J(X_1^i, T_1)$ is

$$\int_{T_o}^{T_1} \left\{F(x^i(t),\, u^k(t), t) - F(\hat{x}^i(t),\, \hat{u}^k(t), t)\right\}dt.$$

Since δx^i is of order ϵ, we again expand into a Taylor series with respect to the δx^i

$$\int_{T_o}^{T_1} \left\{F(x^i, u^k, t) - F(x^i, \hat{u}^k, t) + F(\hat{x}^i + \delta x^i, \hat{u}^k, t) - F(\hat{x}^i, \hat{u}^k, t)\right\} dt =$$

$$= \int_{T_o}^{T_1} \left\{F(x^i, u^k, t) - F(x^i, \hat{u}^k, t) + F_{\hat{x}^i}\, \delta x^i\right\} dt + 0(\epsilon^2),$$

where δx^i is a solution of the equation

$$\delta \dot{x}^i = f^i_{x^j} \delta x^j + f^i(x^j, u^k, t) - f^i(x^j, \hat{u}^k, t).$$

Introducing a vector $p_i(t)$ we remark that

$$\frac{d}{dt}(p_i \delta x^i) = \dot{p}_i \delta x^i + p_i \delta \dot{x}^i = \dot{p}_i \delta x^i + p_j f^j_{\hat{x}^i} \delta x^i + p_i f^i(u^k) - p_i f^i(\hat{u}^k).$$

If the vector p_i is to satisfy the equation

$$\dot{p}_i = - f^j_{\hat{x}^i} p_j + F_{\hat{x}^i}$$

the integral transforms into

$$\int_{T_o}^{T_1} \left\{F(x^i, u^k, t) - F(x^i, \hat{u}^k, t) + \frac{d}{dt}(p_i \delta x^i) - p_i f^i(u^k) + p_i f^i(\hat{u}^k)\right\} dt + 0(\epsilon^2)$$

or

$$p_i(T_1)\delta x^i(T_1) + \int_{T_o}^{T_1} \left\{F(x^i, u^k, t) - F(x^i, \hat{u}^k, t) - p_i f^i(u^k) + p_i f^i(\hat{u}^k)\right\} dt + 0(\epsilon^2).$$

We supposed the value $J(X^i, T)$ to be a function of class C^1 in a neighbourhood of X^i_1, T_1.

If now p_i is fixed by the conditions that at $t = T_1$

$$p_i = J_{X^i_1},$$

the term

$$p_i(T_1)\delta x^i(T_1) = J_{X^i_1} \delta x^i(T_1) = J(X^i_1 + \delta x^i, T_1) - J(X^i_1, T_1) + 0(\epsilon^2),$$

and the integral

$$\int_{T_o}^{T_1} \left\{F(x^i, u^k, t) - F(x^i, \hat{u}^k, t) - p_i f^i(u^k) + p_i f^i(\hat{u}^k)\right\} dt + 0(\epsilon^2)$$

is the difference of the value of the integral along the curve $x^i(t)$ and the value along the extremal with the same end point (at $t = T_1$). If this actually is to be an extremal, we must have

$$\int_{T_o}^{T_1} \left\{F(x^i, u^k, t) - F(x^i, \hat{u}^k, t) - p_i f^i(u^k) + p_i f^i(\hat{u}^k)\right\} dt \geqq 0$$

for all ϵ-variations δu^i.

Taking a special variation

$$\delta u^i = 0, \quad T_o \leqq t < T - \tfrac{1}{2}\epsilon,$$
$$\delta u^i = U^i, \quad T - \tfrac{1}{2}\epsilon \leqq t \leqq T + \tfrac{1}{2}\epsilon,$$
$$\delta u^i = 0, \quad T + \tfrac{1}{2}\epsilon < t \leqq T_1,$$

we infer as in section 4 that

$$F(x^i, u^k, t) - F(x^i, \hat{u}^k, t) - p_i f^i(u^k) + p_i f^i(\hat{u}^k) \geqq 0$$

in nearly every point of the interval for all admissible values of u^k.

In this way a maximum principle is derived, which is similar to Pontryagin's principle:

In order that $\hat{x}^i, \hat{u}^k$ gives an extremal from X_o^i, T_o to X_1^i, T_1 which is embedded in a field of extremals starting from the same point, it is necessary that there exists a solution p_i of the equations

$$\dot{p}_i = -f^j_{\hat{x}^i} p_j + F_{\hat{x}^i},$$

for which for every value of t on the interval $T_o \leqq t \leqq T_1$:

$$H(x^i, u^k, t) = -F(x^i, u^k, t) + p_j f^j(x^i, u^k, t) \leqq H(x^i, \hat{u}^k, t).$$

Since this principle is to hold for every value of ϵ and $\delta x^i = x^i - \hat{x}^i$ is of order ϵ, we must also have for points on the extremal

$$H(\hat{x}^i, u^k, t) \leqq H(\hat{x}^i, \hat{u}^k, t).$$

7. *Specification of the maximum principle for inequality constraints on the control variables.*

In the preceding section the minimum condition of the integral is reduced to a maximum condition of the control variables u^i on points $\hat{x}^i(t)$ of the extremal.

This means that for a fixed value of t we have a simple optimization problem for functions of a number of variables u^k

$$H(u^k) \leqq H(\hat{u}^k),$$

where, since x^i, p_i and t are fixed, we only have

$$H = -F(u^k) + p_i f^i(u^k).$$

The constraints are expressed in the form of a number (μ) of inequalities

$$g^\kappa\left\{u^k(t), t\right\} \leqq 0, \quad \kappa = 1, \ldots, \mu,$$

where the functions g^κ are of class C^2.

$$g^\kappa(u^k) \leqq 0, \qquad \kappa = 1, \ldots, \mu.$$

We introduce slack variables z^κ

$$g^\kappa(u^k) + z^\kappa = 0, \quad \kappa = 1, \ldots, \mu.$$

Then the z^κ are subject to the conditions

$$z^\kappa \geqq 0, \qquad \kappa = 1, \ldots, \mu.$$

We now suppose that in the maximum point $z^{\kappa} = 0$ for $\kappa = 1, \ldots, \nu < m$, and $z^{\kappa} > 0$ for $\kappa = \nu + 1, \ldots, \mu$ and impose the condition that for a small variation $\delta u^k = u^k - \hat{u}^k$, which does not leave the admissible region

$$\delta H \leqq 0.$$

The constraints give conditions on the δu^k which follow from

$$g^{\kappa}_{\hat{u}^k} \cdot \delta u^k + \delta z^{\kappa} = 0, \quad \kappa = 1, \ldots, \mu,$$

but for $\kappa = 1, \ldots, \nu$ the δz^{κ} are limited

$$\delta z^{\kappa} \geqq 0, \qquad \kappa = 1, \ldots, \nu,$$

for $\kappa = \nu + 1, \ldots, \mu$ for sufficiently small δu^k we always have

$$z^{\kappa} + \delta z^{\kappa} > 0.$$

Hence the δu^k are only restricted by

$$g^{\kappa}_k \delta u^k = -\delta z^{\kappa}, \qquad \kappa = 1, \ldots, \nu,$$

where g^{κ}_k is an abbreviation for $g^{\kappa}_{\hat{u}^k}$.

If the $g^{\kappa}(u^k)$ are independent, the rank of the matrix g^{κ}_k is ν and we can rearrange the variables u^k, so that

$$\det(g^{\kappa}_k) \neq 0, \qquad \kappa = 1, \ldots, \nu; \; k = 1, \ldots, \nu.$$

Introducing the inverse matrix γ^l_{κ} of g^{κ}_k, defined by

$$\gamma^l_{\kappa} g^{\kappa}_k = \delta^l_k = 1, \quad \text{for} \quad l = k,$$
$$= 0, \quad \text{for} \quad l \neq k,$$

we can express $\delta u^k (k = 1, \ldots, \nu)$ into $\delta u^{\nu+h}$ $(h = 1, \ldots, \mu - \nu)$ and δz^{κ}.

$$\gamma^l_{\kappa} g^{\kappa}_k \delta u^k = -\gamma^l_{\kappa} \delta z^{\kappa}, \quad l = 1, \ldots, \nu,$$

or

$$\delta u^l = -\gamma^l_{\kappa} g^{\kappa}_{\nu+h} \delta u^{\nu+h} - \gamma^l_{\kappa} \delta z^{\kappa}, \quad l = 1, \ldots, \nu.$$

Here the variations in $\delta u^{\nu+h}$ are arbitrary (if sufficiently small), but the δz^{κ} are subject to the restriction $\delta z^{\kappa} \geqq 0$.

Inserting into the expression for δH

$$\delta H = -F_{u_j} \delta u^j + p_i f^i_{u^j} \delta u^j - F_{u^{\nu+h}} \delta u^{\nu+h} + p_i f^i_{u^{\nu+h}} \delta u^{\nu+h}; \; j = 1, \ldots, \nu,$$

we obtain

$$\delta H = -(F_{u^{\nu+h}} + \gamma^j_{\kappa} g^{\kappa}_{\nu+h} F_{u^j}) \delta u^{\nu+h} + p_i (f^i_{u^{\nu+h}} + \gamma^j_{\kappa} g^{\kappa}_{\nu+h} f^i_{u^j}) \delta u^{\nu+h} - (-F_{u^j} + p_i f^i_{u^j}) \gamma^j_{\kappa} \delta z^{\kappa} \leqq 0.$$

The condition is simplified when introducing parameters λ_{κ}

$$\lambda_{\kappa} = (-F_{u^j} + p_i f^i_{u^j}) \gamma^j_{\kappa}$$

and takes the form

$$(-F_{u^{\nu+h}} + p_i f^i_{u^{\nu+h}} + \lambda_{\kappa} g^{\kappa}_{\nu+h}) \delta u^{\nu+h} - \lambda_{\kappa} \delta z^{\kappa} \geqq 0$$

for arbitrary $\delta u^{\nu+h}$ and $\delta z^{\kappa} \geqq 0$.

This gives the necessary conditions

$$- F_{u^{\nu+h}} + p_i f^i_{u^{\nu+h}} + \lambda_\kappa g^\kappa_{u^{\nu+h}} = 0,$$

$$\lambda_\kappa \leqq 0,$$

while from the definition of λ_κ it follows that also

$$- F_{u^j} + p_i f^i_{u^j} + g^\kappa_{u^j}\lambda_\kappa = 0, \quad j = 1, \ldots, \nu.$$

Hence the maximum principle requires the existence of ν multipliers $\lambda_\kappa \leqq 0$, which satisfy the m conditions

$$- F_{u^j} + p_i f^i_{u^j} + \lambda_\kappa g^\kappa_{u^j} = 0.$$

In this way we have expressed a set of necessary conditions in the form of a set of equations for the functions $x^i(t)$, $p_i(t)$, $u^k(t)$, $\lambda_\kappa(t)$:

$$\dot{x}^i = f^i(x^j, u^k, t),$$

$$\dot{p}_i = - f^j_{x^i} p_j + F_{x^i},$$

$$g^\kappa(u^k, t) = 0,$$

$$- F_{u^j} + p_i f^i_{u^j} + \lambda_\kappa g^\kappa_{u^j} = 0,$$

$$\lambda_\kappa \leqq 0.$$

Now it is possible to locate discontinuities in the extremals. As long as $\lambda_\kappa < 0$ in a certain interval $\tau_0 \leqq t \leqq \tau_1$ the variables satisfy a set of algebraic and differential equations and hence x^i, u^k, λ_κ and p_i are differentiable. But as soon as λ_κ passes through zero, the conditions are no longer satisfied. Here the assumption that $z^\kappa = 0$ ($\kappa = 1, \ldots, \nu$) must be altered. This means that the set of equality constraints changes. This gives rise to a possible jump in the u^i, which results in a discontinuity in the derivatives of $x^i(t)$.

In the interval $\tau_0 \leqq t \leqq \tau_1$ it is not necessary to consider variations δu^k, which show this jump, only uniformly small δu^k suffice to derive the necessary conditions, shown here.

8. *Direct derivation of the differential equations with general inequality restrictions on the control variables.*

Once it is shown to be possible to locate discontinuities in $\hat{u}^k$, it is also possible to give a direct derivation of the set of differential equations, which are necessary conditions for the functions $\hat{u}^k(t)$ to give a minimum value to the integral. With this aim in view we specify the constraints in an even more general form containing both state and control variables

$$g^\kappa(x^i(t), u^k(t), t) \leqq 0, \quad \kappa = 1, \ldots, \mu,$$

where we suppose that the functions g^κ considered as functions of u^k are independent, i.e. for every combination of ν equations the matrix

$$(g^\kappa_{u^k}) \quad \kappa = 1, \ldots, \nu \leqq m$$

has the rank ν.

We suppose that in a subinterval $\tau_o \leqq t \leqq \tau_1$ of $T_o \leqq t \leqq T_1$ the u^k, x^i satisfy the conditions

$$g^\varkappa(x^i, u^k, t) = 0, \quad \text{for } \varkappa = 1, \dots, \nu \leqq m$$
$$g^\varkappa(x^i, u^k, t) < 0, \quad \text{for } \varkappa = \nu + 1, \dots, \mu.$$

Then again introducing the slack variables $z^\varkappa(t)$ by

$$g^\varkappa + z^\varkappa = 0$$

we have

$$z^\varkappa = 0 \quad \text{for} \quad \varkappa = 1, \dots, \nu,$$
$$z^\varkappa > 0 \quad \text{for} \quad \varkappa = \nu + 1, \dots, \mu.$$

In the interval $\tau_o \leqq t \leqq \tau_1$ we introduce uniformly small ϵ-variations δu^k.

Then the δx^i satisfy the differential equations (neglecting terms of order ϵ^2)

$$\delta \dot{x}^i = f^i_{\hat{x}^j} \delta x^j + f^i_{\hat{u}^k} \delta u^k, \qquad i = 1, \dots, n.$$

According to the constraints we have in every point

$$g^\varkappa_{\hat{x}^i} \delta x^i + g^\varkappa_{\hat{u}^k} \delta u^k + \delta z^\varkappa = 0,$$

with $\delta z^\varkappa \geqq 0$ for $\varkappa = 1, \dots, \nu \leqq m$.

The conditions for $\varkappa = \nu + 1, \dots, \mu$ do not give restrictions on the δx^i and δu^k, if they are sufficiently small.

The difference in value between the integral along the varied curve and the extremal is

$$\int_{\tau_o}^{\tau_1} \left\{ F(\hat{x}^i + \delta x^i, \hat{u}^k + \delta u^k, t) - F(\hat{x}^i, \hat{u}^k, t) \right\} dt =$$
$$= \int_{\tau_o}^{\tau_1} \left\{ F_{\hat{x}^i} \delta x^i + F_{\hat{u}^k} \delta u^k \right\} dt + 0(\epsilon^2).$$

Now we have $n+\nu$ relations between the variations δu^k, δx^i and $\delta z^\varkappa$. According to our assumption we can rearrange the u^k, so that

$$\det(g^\varkappa_{u^k}) \neq 0, \qquad \varkappa = 1, \dots, \nu; \ k = 1, \dots, \nu.$$

Introducing the inverse matrix $\gamma^l_\varkappa$

$$\gamma^l_\varkappa g^\varkappa_{u^k} = \delta^l_k,$$

we have

$$\delta u^l = -\gamma^l_\varkappa g^\varkappa_{\hat{u}^{\nu+h}} \delta u^{\nu+h} - \gamma^l_\varkappa g^\varkappa_{\hat{x}^i} \delta x^i - \gamma^l_\varkappa \delta z^\varkappa,$$

and the equations for $\delta \dot{x}^i$ take the form:

$$\delta \dot{x}^i = (f^i_{\hat{x}^j} - f^i_{\hat{u}^l} \gamma^l_\varkappa g^\varkappa_{\hat{x}^j}) \delta x^j + \left\{ f^i_{\hat{u}^{\nu+h}} - f^i_{\hat{u}^l} \gamma^l_\varkappa g^\varkappa_{\hat{u}^{\nu+h}} \right\} \delta u^{\nu+h} - f^i_{\hat{u}^l} \gamma^l_\varkappa \delta z^\varkappa.$$
$$i, j = 1, \dots, n; \ l, \varkappa = 1, \dots, \nu; \ h = 1, \dots, m-\nu,$$

while the integral transforms into

$$\int_{\tau_0}^{\tau_1}\left\{(F_{\dot{x}^i} - F_{\hat{u}^l}\,\gamma^l_\kappa g^\kappa_{\dot{x}^i})\delta x^i + (F_{\hat{u}^{\nu+h}} - \gamma^l_\kappa g^\kappa_{\hat{u}^{\nu+h}} F_{\hat{u}^l})\delta u^{\nu+h} - F_{\hat{u}^l}\,\gamma^l_\kappa \delta z^\kappa\right\} dt.$$

For further reduction of this integral we introduce a vector p_i, for which

$$\frac{d}{dt}(p_i\,\delta x^i) = \dot{p}_i \delta x^i + p_i \delta \dot{x}^i =$$
$$= \dot{p}_i\,\delta x^i + p_j(f^j_{\dot{x}^i} - f^j_{\hat{u}^l}\gamma^l_\kappa\, g^\kappa_{\dot{x}^i})\,\delta x^i + p_j\left\{f^j_{\hat{u}^{\nu+h}} - f^j_{\hat{u}^l}\gamma^l_\kappa\, g^\kappa_{\hat{u}^{\nu+h}}\right\}\delta u^{\nu+h} - p_i f^i_{\hat{u}^l}\gamma^l_\kappa\,\delta z^\kappa.$$

In order to use this result for simplification of the integral the p_i are defined as a solution of the differential equation

$$\dot{p}_i + (f^j_{\dot{x}^i} - f^j_{\hat{u}^l}\gamma^l_\kappa g^\kappa_{x^i})p_j = F_{\dot{x}^i} - F_{\hat{u}^l}\,\gamma^l_\kappa\, g^\kappa_{\dot{x}^i},$$

and the integral transforms into

$$\left[p_i\delta x^i\right]_{\tau_0}^{\tau_1} + \int_{\tau_0}^{\tau_1}\Big[\left\{F_{\hat{u}^{\nu+h}} - \gamma^l_\kappa\, g^\kappa_{\hat{u}^{\nu+h}}\, F_{\hat{u}^l} - p_j\,(f^j_{\hat{u}^{\nu+h}} - f^j_{\hat{u}^l}\gamma^l_\kappa\, g^\kappa_{\hat{u}^{\nu+h}})\right\}\delta u^{\nu+h} +$$
$$+ p_i f^i_{\hat{u}^l}\gamma^l_\kappa\,\delta z^\kappa - F_{\hat{u}^l}\,\gamma^l_\kappa\,\delta z^\kappa\Big]\, dt.$$

The multipliers λ_κ are introduced as

$$\lambda_\kappa = -F_{\hat{u}^l}\,\gamma^l_\kappa + p_i f^i_{\hat{u}^l}\gamma^l_\kappa$$

and the reduced form of the integral is

$$\left[p_i\,\delta x^i\right]_{\tau_0}^{\tau_1} + \int_{\tau_0}^{\tau_1}\left\{(F_{\hat{u}^{\nu+h}} - p_j\, f^j_{\hat{u}^{\nu+h}} + \lambda_\kappa\, g^\kappa_{\hat{u}^{\nu+h}})\delta u^{\nu+h} + \lambda_\kappa\,\delta z^\kappa\right\}dt.$$

Since now we have only considered uniformly small variations δu^k, it is sufficient to consider variations δu^k for which $\delta x^i(\tau_0) = \delta x^i(\tau_1) = 0$. In this case we have the condition that

$$\int_{\tau_0}^{\tau_1}\left\{(F_{\hat{u}^{\nu+h}} - p_j f^j_{\hat{u}^{\nu+h}} + \lambda_\kappa\, g^\kappa_{\hat{u}^{\nu+h}})\delta u^{\nu+h} + \lambda_\kappa\,\delta z^\kappa\right\} dt \geqq 0$$

for arbitrary $\delta u^{\nu+h}$ and $\delta z^\kappa \geqq 0$.

Using the fundamental lemma of the calculus of variations this leads to

$$F_{\hat{u}^{\nu+h}} - p_j f^j_{\hat{u}^{\nu+h}} + \lambda_\kappa\, g^\kappa_{\hat{u}^{\nu+h}} = 0$$

together with

$$\lambda_\kappa \geqq 0.$$

From the definition of λ_κ we also see that

$$-F_{\hat{u}^l} + p_j f^j_{\hat{u}^l} + \lambda_\kappa\, g^\kappa_{\hat{u}^l} = 0$$

and we obtain for the determination of the extremals the following equations

$$\dot{x}^i = f^i(x^j, u^k, t), \qquad i = 1, \ldots, n,$$

$$\dot{p}_i = - f^j_{\hat{x}^i} p_j + F_{\hat{x}^i} + \lambda_\kappa g^\kappa_{\hat{x}^i},$$

$$0 = - F_{\hat{u}^k} + p_j f^j_{\hat{u}^k} + \lambda_\kappa g^\kappa_{\hat{u}^k}, \qquad k = 1, \ldots, m,$$

$$g^\kappa(x^i, u^k, t) = 0, \qquad k = 1, \ldots, \nu$$

$$\lambda_\kappa \geqq 0.$$

9. The Weierstrass-Erdmann conditions and the equation of Hamilton-Jacobi.

With the restricted variations $\delta x^i(\tau_0) = \delta x^i(\tau_1) = 0$ we derived the differential equations which are satisfied along the extremal. We repeat that we establish a set of necessary conditions for the existence of a field of extremals, where $\hat{x}^i(t)$, $\hat{u}^k(t)$ gives a single extremal, extending from $X^i_0(T_0)$ to $X^i_1(T_1)$.

We now consider the general variations $\delta x^i(t)$ and moreover we extend the interval $\tau_0 \leqq t \leqq \tau_1$ to $\tau_0 + \delta\tau_0 \leqq t \leqq \tau_1 + \delta\tau_1$.

Then along the arc of the variation curve the variation of the integral is

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

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

$$+ \int_{\tau_0}^{\tau_1} \left\{F(\hat{x}^i + \delta x^i, \ \hat{u}^k + \delta u^k, \ t) - F(\hat{x}^i, \hat{u}^k, t)\right\} dt =$$

$$= - F(x^i_0, u^k_0, \tau_0)\delta\tau_0 + F(x^i_1, u^k_1, \tau_1)\delta\tau_1 + p_i(\tau_1)\delta x^i_1 -$$

$$- p_i(\tau_0)\delta x^i_0 + \int_{\tau_0}^{\tau_1} \lambda_\kappa \delta z^\kappa dt + 0(\epsilon^2),$$

where use is made of the reduction of section 6.

The total displacements of the end points are

$$\Delta x^i_1 = \delta x^i_1 + f^i_1 \delta\tau_1,$$

$$\Delta x^i_0 = \delta x^i_0 + f^i_0 \delta\tau_0$$

and the variation of the integral is

$$\left\{F(x^i_1, u^k_1, \tau_1) - p_i f^i_1\right\}\delta\tau_1 + p_i(\tau_1)\Delta x^i_1 - \left\{F(x^i_0, u^k_0, \tau_0) - p_i f^i_0\right\}\Delta\tau_0 - p_i(\tau_0)\Delta x^i_0 =$$

$$= - H_1 . \Delta\tau_1 + p_i(\tau_1)\Delta x^i_1 + H_0 . \Delta\tau_0 - p_i(\tau_0)\Delta x^i_0 + \int_{\tau_0}^{\tau_1} \lambda_\kappa \delta z^\kappa dt.$$

We now remark that this expression is only the contribution of the arc between $\tau_0 + \delta\tau_0$ and $\tau_1 + \delta\tau_1$ of the varied curve. Moreover, the variations $\delta\tau_0$, $\delta\tau_1$ are arbitrary (although small) and also the variations δx^i_0 and δx^i_0 can be chosen arbitrarily.

In fact, the variations δx^i satisfy a set of linear equations where the δu^k are functions, only subjected to the restrictions

$$g^{\kappa}_{x^i}\delta x^i + g^{\kappa}_{u^k}\delta u^k + \delta z^{\kappa} = 0, \qquad \kappa = 1,\dots,\nu \leqq n$$

$$\delta z^{\kappa} \geqq 0.$$

Starting with $\delta x^i(\tau_0)$ it is always possible to find a set of functions δu^k so that end values $\delta x^i(\tau_1)$ are reached by superposition of a set of n special functions δu^k which satisfy the restriction.

Now for two intervals, adjacent, so that

$$\tau_1^- = \tau_0^+, \ \tau_1^- + \delta\tau_1^- = \tau_0^+ + \delta\tau_0^+,$$

$$x_1^{i-} = x_0^{i+}, \ x_1^{i-} + \delta x_1^{i-} = x_0^{i+} + \delta x_0^{i+}$$

the contribution of the common endpoint to the variation of the integral is

$$(- H_1^- + H_0^+)\Delta\tau + (p_{i_1}^- - p_{i_0}^+)\Delta x^i.$$

Since $\Delta\tau$ and Δx^i are arbitrary, we must have

$$H_0^+ = H_1^-, \ p_{i_0}^+ = p_{i_1}^-,$$

i.e. the p_i and H must be continuous at the junction of the intervals. At the final end point T_1 we have

$$\Delta J = - H(T_1)\Delta T_1 + p_i(T_1)\Delta X ,$$

which shows that not only, as was used in the definition of p_i

$$J_{X^i} = p_i(T_1),$$

but also

$$J_T = - H(T_1).$$

Since the p_i are defined by differential equations and they are continuous at the junction of intervals, this gives the only boundary condition.

It is possible to introduce along the extremal the variables x^i and p_i as canonical variables. In this case u^k and λ_κ are considered as functions of x^i and p_i, defined by the $m + \nu$ equations

$$0 = - F_{u^k} + p_j f^j_{u^k} + \lambda_\kappa g^\kappa_{u^k}, \qquad k = 1,\dots,m$$

$$g^\kappa(x^i, u^k, t) = 0. \qquad \kappa = 1,\dots,\nu.$$

We introduce the Hamiltonian in the form

$$H = - F + p_i f^i + \lambda_\kappa g^\kappa.$$

Along the part $\tau_0 \leqq t \leqq \tau_1$ $g^\kappa = 0$, at a switching point where for a value of κ $\lambda_\kappa = 0$ we keep this value for λ_κ and go to another constraint $g^{\kappa'} = 0$ which starts with $\lambda_{\kappa'} = 0$.

Hence we may write

$$\Big(H_{x^i} \Big)_{p_i = \text{const.}} = \Big(H_{x^i} \Big)_{p_i, \lambda^\kappa, u^k = \text{const.}} + H_{u^k} \frac{\partial u^k}{\partial x^i} + H_{\lambda_\kappa} \frac{\partial \lambda_\kappa}{\partial x^i},$$

$$\left(H_{p_i}\right)_{x^i = \text{const.}} = \left(H_{x^i}\right)_{x^i, \lambda_\kappa, u^k = \text{const.}} + H_{u_k}\frac{\partial u^k}{\partial p_i} + H_{\lambda_\kappa}\frac{\partial \lambda_\kappa}{\partial p_i}.$$

But $H_{u^k} = -F_{u^k} + p_j f^j_{u^k} + \lambda_\kappa g^\kappa_{u^k} = 0$

as well as

$$H_{\lambda_\kappa} = g^\kappa = 0,$$

and the equations take the canonical form

$$\dot{x}^i = H_{p_i},$$
$$\dot{p}_i = -H_{x^i}.$$

Since H can be considered as a function of x^i and p_i only the value of the integral satisfies the partial differential equation

$$J_T + H(X^i, J_{X^i}, T) = 0,$$

where, however, the dependence of H on J_{X^i} is implicit.

In explicit form we obtain a set of partial differential equations

$$J_T - F(X^i, u^k, T) + f^i(X^i, u^k, T).J_{X^i} = 0$$
$$g^\kappa(X^i, u^k, T) = 0 \qquad \kappa = 1, \ldots, \nu,$$
$$-F_{u^k} + f^i_{u^k} J_{X^i} + \lambda_\kappa g^\kappa_{u^k} = 0 \qquad k = 1, \ldots, m.$$

These partial differential equations are the equations of Hamilton-Jacobi, or the eiconal equations.

This differential equation is only linear in appearance. Putting $J_{X^i} = p_i$ we have

$$J_T - F(X^i, u^k, T) + f^i p_i = 0$$

where the u^k depend on x^i and p_i from

$$g^\kappa(X^i, u^k, T) = 0$$
$$-F_{u^k} + f^i_{u^k} p_i + \lambda_\kappa g^\kappa_{u^k} = 0.$$

The equations of Charpit-Lagrange, which give the characteristics for the eiconal equation

$$\frac{dt}{1} = \frac{dx^i}{f^i + (-F_{u^k} + f^i_{u^k} p_i)\dfrac{\partial u^k}{\partial p_i}} = \frac{dp_i}{-F_{x^i} + f^j_{x^i} p_j + (-F_{u^k} + f^i_{u^k} p_i)\dfrac{\partial u^k}{\partial x^i}} =$$

using the last equation

$$\frac{dt}{1} = \frac{dx^i}{f^i - \lambda_\kappa g^\kappa_{u^k}\dfrac{\partial u^k}{\partial p_i}} = \frac{dp_i}{-F_{x^i} + f^j_{x^i} p_j - \lambda_\kappa g^\kappa_{u^k}\dfrac{\partial u^k}{\partial x^i}}.$$

From

$$g^{\kappa}(X^i, u^k, T) = 0$$

we find

$$g^{\kappa}_{u^k} \cdot \frac{\partial u^k}{\partial p_i} = 0$$

and

$$g^{\kappa}_{X^i} + g^{\kappa}_{u^k} \frac{\partial u^k}{\partial X^i} = 0$$

which shows that the characteristic equations can be written as

$$\frac{dt}{1} = \frac{dx^i}{f^i} = \frac{dp_i}{-F_{x^i} + f^j_{x^i} p_j + \lambda_\kappa g^{\kappa}_{x^i}}.$$

i.e. here again the canonical equations are the characteristics of the eiconal equation.

At a switching point, where $\lambda_\kappa = 0$, the last equation shows that the derivatives J_{X^i} are continuous, as they are expected to be.

10. Constraints in the state variables only.

If the constraints contain the state variables only, the matrix $g^{\kappa}_{u^k}$ is completely degenerate and the preceding conditions are no longer valid. The constraints are supposed to have the form

$$g^{\kappa}(x^i, t) \leqq 0, \qquad \kappa = 1, \ldots, \mu,$$

which, after introduction of the slack variables z^{κ} passes into

$$g^{\kappa}(x^i, t) + z^{\kappa} = 0,$$

where

$$z^{\kappa} \geqq 0.$$

We suppose that in the interval $\tau_0 \leqq t \leqq \tau_1$,

$$z^{\kappa} = 0, \qquad \kappa = 1, \ldots, \nu,$$

and

$$z^{\kappa} > 0, \qquad \kappa = \nu + 1, \ldots, \mu.$$

Then the variations δx^i at each point are limited by

$$g^{\kappa}_{x^i} \delta x^i + \delta z^{\kappa} = 0 \qquad \kappa = 1, \ldots, \nu,$$

where $\delta z^{\kappa} \geqq 0$. This leads to ν multipliers $\lambda_\kappa \geqq 0$ so that we have to determine the u^k in such a way that the integral

$$\int_{\tau_0}^{\tau_1} \left\{ F(x^i, u^k, t) - \lambda_\kappa g^{\kappa}(x^i, t) \right\} dt$$

attains a minimum value.

After the introduction of the appropriate λ_κ the variations δx^i are no longer restricted. This means that we can now study the effect of a variation δu^k on the integral. The corresponding variation δx^i follows from the differential equations

$$\delta \dot{x}^i = f^i_{x^j}\,\delta x^j + f^i_{u^k}\,\delta u^k$$

and the variation in the integral is

$$\int_{\tau_0}^{\tau_1} \left\{F_{x^i}\,\delta x^i - F_{u^k}\,\delta u^k - \lambda_\kappa g^\kappa_{x^i}\,\delta x^i\right\} dt$$

where $\lambda_\kappa \geqq 0$ for $\kappa = 1, \ldots, \nu$.

Introduce as before conjugate variables p_i, we have

$$\frac{d}{dt}(p_i\,\delta x^i) = \dot{p}_i\,\delta x^i + p_i\,\delta\dot{x}^i = \dot{p}_i\,\delta x^i + p_j f^j_{x^i}\,\delta x^i + p_j f^j_{u^k}\,\delta u^k.$$

Hence, putting

$$\dot{p}_i = -p_j f^j_{x^i} + F_{x^i} - \lambda_\kappa g^\kappa_{x^i}$$

the variation of the integral takes the form

$$\int_{\tau_0}^{\tau_1} \left\{\frac{d}{dt}(p_i\,\delta x^i) + (F_{u^k} - p_j f^j_{u^k})\delta u^k\right\} dt = \Big[p_i\,\delta x^i\Big]_{\tau_0}^{\tau_1} + \int_{\tau_0}^{\tau_1} (F_{u^k} - p_j f^j_{u^k})\delta u^k\,dt.$$

This leads, since the δu^k are free, to the condition

$$F_{u^k} - p_j f^j_{u^k} = 0.$$

We obtain the system of equations for the determination of p_i, x^i and u^k

$$\dot{x}^i = f^i(x^j, u^k, t),$$

$$\dot{p}_i = -p_j f^j_{x^i} + F_{x^i} - \lambda_\kappa g^\kappa_{x^i},$$

$$0 = -p_j f^j_{u^k} + F_{u^k},$$

where $\lambda_\kappa \geqq 0$.

If for a value of τ_0 for certain values α of κ, λ_α goes through zero, the condition $g^\alpha = 0$ ceases to apply and the value of λ_α remains zero.

In order to derive the Weierstrass-Erdmann conditions, we consider the variation along an interval, extending from $\tau_0 + \delta\tau_0$ to $\tau_1 + \delta\tau_1$ and remarking that

$$(\Delta x^i)_1 = (\delta x^i)_1 + f^i_1\,\delta\tau_1, \quad (\Delta x^i)_0 = (\delta x^i)_0 + f^i_0\,\delta\tau_0,$$

the variation takes the form

$$-(F\delta\tau)_0 + (F\delta\tau)_1 + \int_{\tau_0}^{\tau_1} (F_{x^i}\,\delta x^i + F_{u^k}\,\delta u^k)\,dt =$$

$$-(F + p_i f^i)_1\,\delta\tau_1 + (-F + p_i f^i)_0\,\delta\tau_0 + (p_i \Delta x^i)_1 - (p_i \Delta x^i)_0.$$

Now for the total variation Δx^i, Δt we must have

$$g^\kappa_{x^i}\Delta x^i + g^\kappa_t \Delta t + \Delta z^\kappa = 0$$

with $\Delta z^\kappa \geqq 0$, or

$$- g^\kappa_{x^i}\delta x^i - (g^\kappa_t + f^i g^\kappa_{x^i})\delta t = \delta z^\kappa - (g^\kappa_t + f^i g^\kappa_{x^i})\delta t \geqq 0.$$

Now we already have $\delta z^\kappa \geqq 0$, in order that $\Delta z^\kappa \geqq 0$ for positive and negative values of δt we must have

$$\frac{dg^\kappa}{dt} \equiv g^\kappa_t + g^\kappa_{x^i} f^i = 0,$$

i.e. at the endpoints the trajectory must be tangent to the surfaces expressing the boundary conditions.

At the junction of the two intervals, the p_i and $H = - F + p_i f^i + \lambda_\kappa g^\kappa$ are continuous. Here we have derived conditions for the minimum problem for the integral

$$J = \int_{\tau_0}^{\tau_1} (F - \lambda_\kappa g^\kappa)dt = \int_{\tau_0}^{\tau_1} F dt - \int_{\tau_0}^{\tau_1} \lambda_\kappa g^\kappa dt.$$

Introducing a factor μ_κ by

$$\dot{\mu}_\kappa = \lambda_\kappa$$

and remarking that along the extremal $\frac{dg^\kappa}{dt} = 0$, we can also write

$$J = \int_{\tau_0}^{\tau_1} F dt - \int_{\tau_0}^{\tau_1} g^\kappa d\mu_\kappa = \int_{\tau_0}^{\tau_1} F dt - \Big[\mu_\kappa g^\kappa\Big]_{\tau_0}^{\tau_1} .$$

The Hamilton-Jacobi equations refer to this function. Hence

$$p_i = J_{x^i}, \quad J_t = - H.$$

If, instead of J we want to consider the original integral

$$J^* = \int_{\tau_0}^{\tau_1} F dt = J + \Big[\mu_\kappa g^\kappa\Big]_{\tau_1} - \Big[\mu_\kappa g^\kappa\Big]_{\tau_0} ,$$

we derive the modification of the Hamilton-Jacobi equations by the consideration of a variation, which extends from $\tau_0 + \delta\tau_0$ to $\tau_1 + \delta\tau_1$, and where the variation starts at

$$\Delta x^i_0 = \delta x^i_0 + (f^i\delta\tau)_0, \quad \Delta x^i_1 = \delta x^i_1 + (f^i\delta\tau)_1 .$$

Then
$$\Delta J = \Big[J_{x^i}\Delta x^i + J_T\,\delta T\Big]_0^1 = \Big[p_i\Delta x^i + J_T\,\delta t\Big]_0^1 =$$
$$= \Big[J^*_{x^i}\Delta x^i + J^*_t\,\delta t\Big]_0^1 - \Big[\mu_\kappa(g^\kappa_{x^i}\Delta x^i + g^\kappa_t\Delta t)\Big]_0^1 .$$

From this result we see that

$$J^*_{x^i} = p_i + \mu_\kappa g^\kappa_{x^i} .$$
$$J^*_t = - H + \mu_\kappa g^\kappa_t .$$

At a junction, where the trajectory leaves the boundary condition $g^{\alpha} = 0$ the μ_{κ} jumps to zero and we have a jump in the values of $J^*_{x^i}$ and J^*_t, which is proportional to $g^{\alpha}_{x^i}$ and g^{α}_t with the factor μ_{κ}.

11. *The problem of Mayer.*

In many applications the object function is not an integral, but only a function of the end values X^i_1, T_1 of the variables x^i, t: $\Phi(X^i_1, T_1)$. In this case obviously the endpoint cannot be fixed. Hence we can only consider the problem, where we have a set of end conditions

$$\varphi^{\rho}(X^i_1, T_1) = 0, \qquad \rho = 1, \ldots, l < n,$$

which leave a certain amount of freedom.

The differential equations again have the form

$$\dot{x}^i = f^i(x^j(t), u^k(t), t), \qquad i = 1, \ldots, n,$$

with constraints

$$g^{\kappa}(x^i, u^k, t) \leqq 0, \qquad \kappa = 1, \ldots, \mu.$$

As before the g^{κ} are assumed to be independent.

The fundamental problem here again is to establish a set of necessary conditions for the existence of a field of extremals, starting from a fixed point X^i_0, T_0 to each point of a neighbourhood of points which satisfy the end conditions. Suppose we have an extremal $\hat{u}(t), \hat{x}(t)$ which satisfies

$$\dot{\hat{x}}^i = f^i(\hat{x}^j, \hat{u}^k, t)$$

with

$$g^{\kappa}(\hat{x}^i, \hat{u}^k, t) + z^{\kappa} = 0, \qquad \kappa = 1, \ldots, \mu,$$

$$z^{\kappa} = 0, \qquad \kappa = 1, \ldots, \nu,$$

$$\varphi^{\rho}\left\{\hat{x}^i(T_1), T_1\right\} = 0, \qquad \rho = 1, \ldots, l,$$

and $\Phi\left\{\hat{x}^i(T_1), T_1\right\}$ minimum.

We consider a neighbouring curve with $\hat{u}^k + \delta u^k$ as control variables extending from $T = T_0$ to $T = T_1$ where δu^k is uniformly small.

Then

$$\delta x^i = x^i(u^k, t) - x^i(\hat{u}^k, t), \quad T_0 \leqq t \leqq T_1$$

is the solution of the equation

$$\delta \dot{x}^i = f^i_{x^j} \delta x^j + f^i_{u^k} \delta u^k$$

with initial condition $\delta x^i(T_0) = 0$.

$$\Delta x^i = \delta x^i(T_1) + f^i(X^i_1, u^k_1, T_1)\delta T.$$

In order to find an expression for $\delta x^i(T_1)$ in terms of δu^k we again introduce the variables p_i as solutions of the equation

$$\dot{p}_i = - f^j_{\hat{x}^i} p_j + \lambda_{\kappa} g^{\kappa}_{\hat{x}^i}.$$

Then

$$\frac{d}{dt}(p_i\delta x^i) = \dot{p}_i\delta x^i + p_i\delta\dot{x}^i = p_i f^i_{x^j}\delta x^j + p_i f^i_{u^k}\delta u^k - f^j_{\hat{x}^i} p_j \delta x^i + \lambda_\kappa g^\kappa_{\hat{x}^i}\delta x^i.$$

But

$$g^\kappa_{x^i}\delta x^i + g^\kappa_{u^k}\delta u^k + \delta z^k = 0,$$

which yields

$$\frac{d}{dt}(p_i\delta x^i) = p_i f^i_{u^k}\delta u^k - \lambda_\kappa g^\kappa_{u^k}\delta u^k - \lambda_\kappa\delta z^k$$

or

$$\Big[p_i\delta x^i\Big]_{T_o}^{T_1} = \int_{T_o}^{T_1}(p_i f^i_{u^k} - \lambda_\kappa g^\kappa_{u^k})\delta u^k dt - \int_{T_o}^{T_1}\lambda_\kappa\delta z^\kappa dt.$$

The end conditions $\varphi^\rho(X^i_1, T_1) = 0$ are taken into account by the introduction of Lagrange multipliers μ_ρ, we seek the minimum of

$$\Phi - \mu_\rho\varphi^\rho,$$

i.e. the condition gives:

$$\left[\Phi_{x^i} - \mu_\rho\varphi^\rho_{x^i}\right]\delta x^i \geqq 0.$$

Choosing now the end conditions for p_i equal to

$$p_i = \Phi_{X^i} - \mu_\rho\varphi^\rho_{X^i}$$

we have

$$\int_{T_o}^{T_1}\left\{p_i f^i_{u^k} - \lambda_\kappa g^\kappa_{u^k}\right\}\delta u^k dt - \int_{T_o}^{T_1}\lambda_\kappa\delta z^\kappa dt \geqq 0.$$

Since the δu^k are arbitrary, we have

$$p_i f^i_{u^k} - \lambda_\kappa g^\kappa_{u^k} = 0$$

and

$$\lambda_\kappa \leqq 0.$$

This gives the following set of necessary conditions:

$$\begin{aligned}
&\dot{\hat{x}}^i = f^i(\hat{x}^j, u^k, t) && i = 1,\dots,n\\
&\dot{p}_i = -f^j_{\hat{x}^i}p_j + \lambda_\kappa g^\kappa_{\hat{x}^i} &&\\
&p_i f^i_{u^k} - \lambda_\kappa g^\kappa_{u^k} = 0 && k = 1,\dots,m\\
&g^\kappa(x^i, u^k, t) = 0 && \kappa = 1,\dots,\nu,
\end{aligned}$$

where for $t = T_1$ the end conditions for p_i are determined from

$$\begin{aligned}
&p_i(T_1) = \Phi_{X^i} - \mu_\rho\varphi^\rho_{X^i} && \rho = 1,\dots,1\\
&\varphi^\rho(X^i, T_1) = 0 &&
\end{aligned}$$

12. Time optimal problems, which are linear in the state and control variables.

If the x^i have to satisfy a set of linear differential equations

$$\dot{x}^i = a^i_j x^j + c^i_k u^k + d^i, \qquad i = 1,\dots,n, \quad k = 1,\dots,m,$$

where the u^k are restricted by the conditions

$$g^\kappa_k u^k + h^\kappa \leqq 0, \qquad \kappa = 1,\dots,\mu,$$

and it is required to find those values of u^k, for which a curve, starting at a point X^i_0 for $t = 0$ reaches the origin $X^i_1 = 0$ for a minimum value of T_1, the conditions of the preceding section require the solution of the set of adjoint equations

$$\dot{p}_i = - a^j_i p_j,$$

where the u_i have to be chosen such as to *maximize*

$$H = -1 + p_i a^i_j x^j + p_i c^i_k u^k + p_i d^i$$

with the constraints

$$g^\kappa_k u^k + h^\kappa \leqq 0, \qquad \kappa = 1,\dots,\mu.$$

This means that at each time the values of u^k are determined by a linear programming problem.

For a maximum the λ_κ which are determined from the conditions

$$p_i c^i_k + \lambda_\kappa g^\kappa_k = 0 \qquad k = 1,\dots,m$$

have to satisfy $\lambda_\kappa \leqq 0$.

Apparently the λ_κ are the dual variables of the linear programming problem and the well-known results apply that always exactly m equality conditions on the u^k are satisfied. For changing values of p_i the cost function changes. Hence in this case we have a case of parametric linear programming.

It is necessary here that the number of conditions μ is greater than m, the number of parameters u^k.

If the endpoint is the point $X^i_1 = 0$, it is necessary that this endpoint is always reached along a curve, where the u^k are completely determined by a set of m equality conditions. Hence, we have a set of fixed trajectories through the origin, along which the origin can be reached, each corresponding to a vertex of the Simplex formed by the hyperplanes.

Apparently, only in the linear case these simple conditions apply.

References.

1. L.S.Pontryagin, V.G.Boltyanskii, R.S.Gamkrelidze, E.F.Mishchenko — The Mathematical Theory of Optimal Processes. Wiley and Sons, N.Y., 1962/Pergamon Press, London, 1964/Oldenbourg Verlag, München, Wien, 1964.
2. I.M.Gelfand, S.V.Fomin — Calculus of Variations. Translation by R.A.Silverman, Prentice-Hall Inc., New Jersey, 1963.
3. M.R.Hestenes — Calculus of Variations and Optimal Control Theory, Princeton University Press, New York, 1966.

[Received May 15, 1967]