International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Quasidifferentiable Problems in Optimal Control

Demyanov, V.F., Nikulina, V.N. and Shablinskaya, I.R.

IIASA Working Paper

WP-84-002

January 1984

NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

QUASIDIFFERENTIABLE PROBLEMS IN OPTIMAL CONTROL

V.F. Demyanov
V.N. Nikulina
I.R. Shablinskaya

January 1984
WP-84-2

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

QUASIDIFFERENTIABLE PROBLEMS IN OPTIMAL CONTROL

V.F. Demyanov, V.N. Nikulina, and I.R. Shablinskaya

1. INTRODUCTION

Nondifferentiability in control theory appears naturally on the right-hand side of the system of equations as well as in the functional (through saturation functions, by taking the modulus, etc.). In many cases both the system and the functional are described by *quasidifferentiable functions*, a class which is defined and investigated in [1-3].

This paper is concerned with the variations of trajectory caused by using different variations of the control for such quasidifferentiable right-hand sides. We consider five different types of control variations. Necessary conditions for an extremum of a quasidifferentiable functional are then stated.

The main intention of the authors is to draw the attention of specialists in control theory and its applications to a new class of problems which seems to be promising and practically oriented. A special case of this class of problems has already been discussed in [4].

1.1 Statement of the Problem

Let the object of study be governed by the following system of ordinary differential equations:

$$\dot{x}(t) = f(x(t),u(t),t) \tag{1}$$

$$x(0) = x_0 \in E_n \tag{2}$$

where $x = (x^{(1)},\ldots,x^{(n)})$, $u = (u^{(1)},\ldots,u^{(r)})$, $f = (f^{(1)},\ldots,f^{(n)})$, $t \in [0,T]$, and $T > 0$ is fixed.

We shall use N to denote the set of r-dimensional vector functions which are piece-wise continuous (right-hand continuous) on $[0,T]$. Let us set

$$U = \{u \in N \,|\, u(t) \in V \ \forall t \in [0,T]\}$$

where $V \subset E_r$ is a compact set. The set U is called the *class of admissible controls* and any $u \in U$ a *control*.

Functions $f^{(i)}$ are (i) defined on S (where $S \subset E_{n+r+1}$ is the set of all admissible x,u,t); (ii) continuous with respect to x and u; (iii) Lipschitzian with respect to x on S; (iv) piece-wise continuous with respect to x on S; and (v) quasidifferentiable with respect to x and S. (In Section 2.5 it will be assumed that the $f^{(i)}$'s are quasidifferentiable jointly with respect to x and u.)

Recall that the function F defined on E_n is quasidifferentiable at $x \in E_n$ if it is directionally differentiable and there exist convex compact sets $\underline{\partial}F(x) \subset E_n$ and $\overline{\partial}F(x) \subset E_n$ such that

$$\frac{\partial F(x)}{\partial g} \equiv \lim_{\alpha \to +0} \frac{1}{\alpha}[F(x+\alpha g)-F(x)] = \max_{v \in \underline{\partial}F(x)} (v,g) + \min_{w \in \overline{\partial}F(x)} (w,g) \quad \forall g \in E_n .$$

Let $x(t,u)$ denote the solution of system (1)-(2) for a chosen $u \in U$.

The problem is to minimize the functional

$$I(u) = \phi(x(T,u)) \tag{3}$$

subject to $u \in U$ where $\phi(x)$ is quasidifferentiable, finite and Lipschitzian on the set of admissible x.

Let $u^* \in U$ denote a u which minimizes I, i.e.,

$$I(u^*) = \min_{u \in U} I(u) \quad .$$

(We shall not consider here the problem of whether such a u exists or is unique.)

The pair of functions (x^*, u^*) where $x^*(t) = x(t,u^*)$ will be called an optimal process; $u^*(t)$ is known as an optimal control and $x^*(t)$ an optimal trajectory.

2. VARIATIONS OF A CONTROL

To derive necessary conditions for a minimum of (3) the following controls are generally used:

$$u_\varepsilon = u^* + \Delta u_\varepsilon \in U$$

where the function Δu_ε is called a *variation* of u^*.

We shall consider several variations of the control and the corresponding variations of the trajectory.

2.1 A Needle Variation (A Sharp Variation).

Let

$$\Delta u_\varepsilon(t) = \begin{cases} y - u^*(t) \,, & t \in [\theta, \theta+\varepsilon) \\ 0 \,, & t \notin [\theta, \theta+\varepsilon) \end{cases} \tag{4}$$

where $y \in V$, $\theta \in [0,T)$, $\varepsilon > 0$.

We wish to find

$$h(t) \equiv (h^{(1)}(t),\dots,h^{(n)}(t)) = \lim_{\varepsilon\to+0} \frac{1}{\varepsilon}[x(t,u_\varepsilon)-x(t,u^*)]. \quad (5)$$

where the vector function h is the variation of the trajectory x^* caused by variation of the control u^*.

It is clear that $h(t) = 0 \quad \forall t \in [0,\theta)$.

For $t > \theta$ we have

$$x_\varepsilon(t) \equiv x(t,u_\varepsilon) = x_0 + \int_0^\theta f(x^*(\tau),u^*(\tau),\tau)d\tau + \int_\theta^{\theta+\varepsilon} f(x_\varepsilon(\tau),y,\tau)d\tau + \int_{\theta+\varepsilon}^t f(x_\varepsilon(\tau),u^*(\tau),\tau)d\tau.$$

Invoking (5), and taking the limit as $\varepsilon \to +0$, we obtain

$$h^{(i)}(t) = f^{(i)}(x^*(\theta),y,\theta) - f^{(i)}(x^*(\theta),u^*(\theta),\theta) + \int_\theta^t \left[\max_{v\in\underline{\partial} f^{(i)}(\tau)} (v,h(\tau)) + \min_{w\in\overline{\partial} f^{(i)}(\tau)} (w,h(\tau)) \right] d\tau \quad \forall i\in 1:n \quad (6)$$

where $\underline{\partial} f^{(i)}(\tau) = \underline{\partial} f_x^{(i)}(x^*(\tau),u^*(\tau),\tau)$ and $\overline{\partial} f^{(i)}(\tau) = \overline{\partial} f_x^{(i)}(x^*(\tau), u^*(\tau),\tau)$ are respectively a subdifferential and a superdifferential of $f^{(i)}$ with respect to x.

For every τ the sets $\underline{\partial} f^{(i)}(\tau) \subset E_n$ and $\overline{\partial} f^{(i)}(\tau) \subset E_n$ are convex and compact.

Let us now rewrite system (6) in the following shorter form:

$$h(t) = f(x^*(\theta),y,\theta) - f(x^*(\theta),u^*(\theta),\theta) + \int_\theta^t \left[\max_{v\in\underline{\partial} f(\tau)} (v,h(\tau)) + \min_{w\in\overline{\partial} f(\tau)} (w,h(\tau)) \right] d\tau \quad (7)$$

where

$$\underline{\partial} f(\tau) = [\underline{\partial} f^{(1)}(\tau),\dots,\underline{\partial} f^{(n)}(\tau)], \quad \overline{\partial} f(\tau) = [\overline{\partial} f^{(1)}(\tau),\dots,\overline{\partial} f^{(n)}(\tau)] .$$

Suppose that all mappings $\underline{\partial} f^{(i)}$ and $\overline{\partial} f^{(i)}$ are piece-wise continuous on $[0,T]$. Then it follows from (6) that

$$\dot{h}^{(i)}(t) = \max_{v \in \underline{\partial} f^{(i)}(t)} (v, h(t)) + \min_{w \in \overline{\partial} f^{(i)}(t)} (w, h(t))$$

$$h^{(i)}(\theta) = f^{(i)}(x^*(\theta), y, \theta) - f^{(i)}(x^*(\theta), u^*(\theta), \theta) \quad \forall i \in 1:n .$$

We can again rewrite this system in a shorter form:

$$\dot{h}(t) = \max_{v \in \underline{\partial} f(t)} (v, h(t)) + \min_{w \in \overline{\partial} f(t)} (w, h(t)) \tag{8}$$

$$h(\theta) = f(x^*(\theta), y, \theta) - f(x^*(\theta), u^*(\theta), \theta) \quad . \tag{9}$$

If $\underline{\partial} f(t)$ and $\overline{\partial} f(t)$ are piece-wise continuous mappings then a solution to (8)-(9) exists and is unique for any fixed $y \in V$ and $\theta \in [0,T)$. Here, $h(t)$ depends on θ and y.

2.2 A Multiple Needle Variation (Needle Variations at Several Points)

Let

$$\Delta u_\varepsilon(t) = \begin{cases} y_i - u^*(t), & t \in [\theta_i, \theta_i + \varepsilon \ell_i) \quad \forall i \in 1:r \ , \\ 0 \ , & t \notin \bigcup_{i \in 1:r} [\theta_i, \theta_i + \varepsilon \ell_i) \end{cases}$$

where $y_i \in V$, $\theta_i \in [0,T)$, $\ell_i \geq 0$, $r > 0$, and r is an arbitrary (but fixed) natural number.

It is clear that $x_\varepsilon(t) = x^*(t)$ for $t \leq \theta_1$. If $t > \theta_1$ then we have

$$x_\varepsilon(t) \equiv x(t, u_\varepsilon) = x_0 + \int_0^{\theta_1} f(x^*(\tau), u^*(\tau), \tau) d\tau + \int_{\theta_1}^{\theta_1 + \varepsilon \ell_1} f(x_\varepsilon(\tau), y_1, \tau) d\tau$$

$$+ \int_{\theta_1 + \varepsilon \ell_1}^{\theta_2} f(x_\varepsilon(\tau), u^*(\tau), \tau) d\tau + \int_{\theta_2}^{\theta_2 + \varepsilon \ell_2} f(x_\varepsilon(\tau), y_2, \tau) d\tau + \ldots +$$

$$+ \int_{\theta_{r(t)}}^{\theta_{r(t)} + \varepsilon \ell_{r(t)}} f(x_\varepsilon(\tau), y_{r(t)}, \tau) d\tau + \int_{\theta_{r(t)} + \varepsilon \ell_{r(t)}}^{t} f(x_\varepsilon(\tau), u^*(\tau), \tau) d\tau \tag{10}$$

where $r(t) \in 1:r$ is such that

$$\theta_{r(t)} < t \leq \theta_{r(t)+1} \quad . \tag{11}$$

If $r(t) = r$ then $r+1 = T$.

Without loss of generality we can assume that $t > \theta_{r(t)} + \varepsilon \ell_{r(t)}$.

From (10) it follows that

$$h(t) \equiv \lim_{\varepsilon \to +0} \frac{1}{\varepsilon} [x_\varepsilon(t) - x^*(t)] = \ell_1 [f(x^*(\theta_1), y_1, \theta_1) - f(x^*(\theta_1), u^*(\theta_1), \theta_1)]$$

$$+ \int_{\theta_1}^{\theta_2} \max_{v \in \underline{\partial} f(\tau)} (v, h(\tau)) + \min_{w \in \overline{\partial} f(\tau)} (w, h(\tau))] d\tau$$

$$+ \ell_2 [f(x^*(\theta_2), y_2, \theta_2) - f(x^*(\theta_2), u^*(\theta_2), \theta_2)]$$

$$+ \int_{\theta_2}^{\theta_3} \max_{v \in \underline{\partial} f(\tau)} (v, h(\tau)) + \min_{w \in \overline{\partial} f(\tau)} (w, h(\tau))] d\tau + \ldots +$$

$$+ \ell_{r(t)} [f(x^*(\theta_{r(t)}), y_{r(t)}, \theta_{r(t)}) - f(x^*(\theta_{r(t)}), u^*(\theta_{r(t)}, \theta_{r(t)})]$$

$$+ \int_{\theta_{r(t)}}^{t} \max_{v \in \underline{\partial} f(\tau)} (v, h(\tau)) + \min_{w \in \overline{\partial} f(\tau)} (w, h(\tau))] d\tau \quad . \tag{12}$$

Now let us introduce the functions

$$h_0(t) = 0 \qquad \forall t \in [0,T]$$

$$h_i(t) = 0 \qquad \forall t < \theta_i$$

while for $t > \theta_i$ the function $h_i(t)$ satisfies the differential equation

$$\dot{h}_i(t) = \max_{v \in \underline{\partial} f(t)} (v, h_i(t)) + \min_{w \in \overline{\partial} f(t)} (w, h_i(t)) \tag{13}$$

with initial condition

$$h_i(\theta_i) = h_{i-1}(\theta_i) + h_i\left[f(x^*(\theta_i), y_i, \theta_i) - f(x^*(\theta_i), u^*(\theta_i), \theta_i)\right] \quad . \tag{14}$$

From (12) it is clear that $h(t) = h_{r(t)}(t)$.

Thus, $h(t)$ (which depends on $\{y_i\}$, $\{\theta_i\}$, and $\{1_i\}$) is a piecewise continuous function satisfying the system of differential equations (8) (or, equivalently (13)) with several "jumps" as indicated by (14).

2.3 A Bundle of Variations

Let

$$\Delta u_\varepsilon(t) = \begin{cases} y_i - u^*(t), & t \in [\theta_i, \theta_i + \varepsilon \ell_i) \\ 0 \;, & t \notin [\theta, \theta+\varepsilon) \end{cases}$$

where $y_i \in V$, $\ell_i > 0$, $\sum_{i=1}^{r} \ell_i = 1$, $\theta_1 = \theta$, $\theta_{i+1} = \theta_i + \varepsilon\ell_i$, $\theta_r + \varepsilon\ell_r = \theta + \varepsilon$

and r is an arbitrary natural number.

It is not difficult to check that

$$h(t) = 0 \qquad \forall t \in [0, \theta) \quad .$$

For $t \geqslant \theta$, we have

$$h(t) = \sum_{i=1}^{r} \ell_i [f(x^*(\theta), y_i, \theta) - f(x^*(\theta), u^*(\theta), \theta)] +$$

$$+ \int_\theta^t \left[\max_{v \in \underline{\partial} f(\tau)} (v, h(\tau)) + \min_{w \in \overline{\partial} f(\tau)} (w, h(\tau))\right] d\tau \quad . \tag{15}$$

The variation of trajectory $h(t)$ satisfies the system of ordinary differential equations (8) with initial condition

$$h(\theta) = \sum_{i=1}^{r} \ell_i [f(x^*(\theta), y_i, \theta) - f(x^*(\theta), u^*(\theta), \theta)] \quad .$$

The vectro function h(t) depends here on $\{y_i\}$, $\{\ell_i\}$ and θ.

2.4 A Multiple Bundle of Variations (A Bundle of Variations at Several Points).

Take

$$\Delta_\varepsilon(t) = \begin{cases} y_{ij} - u^*(t), & t \in [\theta_i + \varepsilon \sum_{k=0}^{j-1} \ell_{ik}, \theta_i + \varepsilon \sum_{k=0}^{j} \ell_{ik}) \\ & \forall j \in 1:M_i \quad \forall i \in 1:N \\ 0, & t \notin \bigcup_{i \in 1:N} [\theta_i, \theta_i + \varepsilon \ell_i) \end{cases}$$

where $\varepsilon>0$, $\theta_i \in [0,T]$, $y_{ij} \in V$, $\ell_{ij} \in V$, $\ell_{ij} \geqslant 0$, and $\ell_{i0} = 0$ for all $i \in 1:N$, $j \in 1:M_i$ where, $\sum_{j=1}^{M_i} \ell_{ij} = 1$ and M_i and N are natural numbers.

Consider the functions

$$h_0(t) = 0 \qquad \forall t \in [0,T]$$

$$h_i(t) = 0 \qquad \forall t<\theta_i$$

while for $t>\theta_i$ the function $h_i(t)$ satisfies the differential equation (13) with initial conditions

$$h_i(\theta_i) = h_{i-1}(\theta_i) + \sum_{k=0}^{Mi} \ell_{ik}\left[f(x^*(\theta_i),y_{ik},\theta_i) - f(x^*(\theta_i),u^*(\theta_i),\theta_i)\right] .$$

It is now possible to show that

$$h(t) = h_{r(t)}(t)$$

where r(t) was defined in (11). The function h(t) depends on $\{y_{ij}\}$, $\{\theta_i\}$, and $\{\ell_{ij}\}$.

2.5 A Classical Variation

Suppose in addition to the above assumptions that the set U is convex and f is quasidifferentiable jointly in x and u, i.e.,

$$\frac{\partial f(x,u,t)}{\partial [h,q]} \equiv \lim_{\alpha\to+0} \frac{1}{\alpha} [f(x+\alpha h,u+\alpha q,t)-f(x,u,t)] =$$

$$= \max_{[v_1,v_2]\in \underline{\partial} f_{x,u}(t)} [(v_1,h)+(v_2,q)] + \min_{[w_1,w_2]\in \overline{\partial} f_{x,u}(t)} [(w_1,h)+(w_2,q)].$$

Now let

$$\Delta u_\varepsilon(t) = \varepsilon(u(t)-u^*(t)) \equiv \varepsilon q(t), \quad u\in U \quad .$$

Proceeding as above we find that h(t) satisfies the system of ordinary differential equations

$$\dot{h}(t) = \max_{[v_1,v_2]\in \underline{\partial} f_{x,u}(t)} [(v_1,h(t))+(v_2,q(t))] +$$

$$+ \min_{[w_1,w_2]\in \overline{\partial} f_{x,u}(t)} [(w_1,h(t))+(w_2,q(t))]$$

with initial condition

$$h(0) = 0 \quad .$$

Here $\underline{\partial} f_{x,u}(t) \subset E_{n+r}$ and $\overline{\partial} f_{x,u}(t) \subset E_{n+r}$ are convex compact sets.

Thus for all of the five control variations considered here we obtain

$$x_\varepsilon(t) = x^*(t) + \varepsilon h(t) + o(\varepsilon)$$

where h(t) satisfies a particular system of equations, depending on the control variation chosen.

3. NECESSARY OPTIMALITY CONDITIONS

Since ϕ is quasidifferentiable and Lipschitzian we have

$$J(u_\varepsilon) = \phi(x^*(T)+\varepsilon h(T)+o(\varepsilon)) = \phi(x^*(T)) + \varepsilon \frac{\partial \phi(x^*(T))}{\partial h(T)} + o(\varepsilon)$$

and therefore the following necessary condition holds:

Theorem 1. If $u^* \in U$ is an optimal control then

$$\frac{\partial \phi(x^*(T))}{\partial h(T)} = \max_{v \in \underline{\partial}\phi(x^*(T))} (v,h(T)) + \min_{w \in \overline{\partial}\phi(x^*(T))} (w,h(T)) \geqslant 0 \qquad (16)$$

for all admissible variations of trajectory $h(T)$.

It is possible to obtain different necessary conditions by considering different types of control variations. Suppose, for example, that f is smooth with respect to x and that we choose a needle variation.

Introduce the functions $\Psi_v(t)$, $\Psi_w(t)$:

$$\frac{d\Psi_v(t)}{dt} = - \frac{\partial f^T(x^*(t),u^*(t),t)}{\partial x} \Psi_v(t), \quad t \leqslant T$$

$$\Psi_v(T) = v, \quad v \in \underline{\partial}\phi(x^*(T))$$

$$\frac{d\Psi_w(t)}{dt} = - \frac{\partial f^T(x^*(t),u^*(t),t)}{\partial x} \Psi_w(t), \quad t \leqslant T$$

$$\Psi_w(T) = w, \quad w \in \overline{\partial}\phi(x^*(T))$$

where $\frac{\partial f}{\partial x}$ is the derivative of f with respect to x. Then the following result can be obtained from (16):

Theorem 2. For a control $u^* \in U$ to be optimal it is necessary that

$$\min_{y \in V} \left[\max_{v \in \underline{\partial}\phi(x^*(T))} \Delta_y H(x^*,u^*,\Psi_v,\theta) + \min_{w \in \overline{\partial}\phi(x^*(T))} \Delta_y H(x^*,u^*,\Psi_w,\theta) \right] =$$

$$= 0 \quad \forall \theta \in (0,T) \tag{17}$$

where

$$H(x,u,\Psi,\theta) = (f(x(\theta),u(\theta),\theta), \Psi(\theta))$$

$$\Delta_y H(x^*,u^*,\Psi,\theta) = H(x^*,y,\Psi,\theta) - H(x^*,u^*,\Psi,\theta) \quad .$$

Condition (17) is a generalization of the Pontryagin maximum principle [5].

Remark 1. If f is not smooth we cannot rewrite (16) in a compact form such as (17). In the case where there are only a finite number of points at which $\underline{\partial}f$ and $\overline{\partial}f$ are not singletons, the function h can again be presented in a shortened form, allowing (16) to be varified comparatively easily.

Remark 2. The most interesting case arises when there exists a set of nonzero measure for which $\underline{\partial}f(t)$ and $\overline{\partial}f(t)$ are not singletons. This introduces the problem of the so-called "sliding modes" - a very important area for further study.

Remark 3. Different control variations are associated with different necessary conditions (the case in which f - but not ϕ - is smooth has been discussed in [4]). Note that if both f and ϕ are smooth the more complicated controls (a multiple needle variation, bundle of variations, multiple bundle of variations) are useless since we obtain the same necessary condition as in the case of a needle variation. If ϕ is non-smooth, however, these conditions differ, yielding necessary conditions of differing complexity. Some illustrative examples are described in [4].

Remark 4. The problem now is to find more computationally useful formulations of (16) for different control variations. We are also faced with a new type of differential equation in the shape of equation (8) - we shall call this a *quasilinear differential equation*. The properties of its solutions have yet to be investigated.

REFERENCES

1. V.F. Demyanov and A.M. Rubinov. On quasidifferentiable functionals. Soviet Math. Dokl., Vol.21(1), 1980, pp. 14-17.

2. L.N. Polyakova. Necessary conditions for an extremum of quasidifferentiable functions. Vestnik Leningrad Univ. Math., Vol. 13, 1981, pp. 241-247.

3. V.F. Demyanov and A.M. Rubinov. On quasidifferentiable mappings. Math. Operationsforsch u. Statist., Ser. Optimization, Vol. 14(1), 1983, pp. 3-21.

4. V.F. Demyanov (Ed.). Nonsmooth Problems of Optimization Theory and of Control (in Russian). Leningrad University Press, Leningrad, 1982 (see especially chapter lV by V.N. Nikulina and I.R. Shablinskaya).

5. L.S. Pontrjagin, V.G. Baltjanskii, R.V. Gamkrelidse and E.F. Mischenko. The Mathematical Theory of Optimal Processes. Wiley, Chichester, 1962.