An Elementary Derivation of the Maximum Principle

MORTON M. DENN and RUTHERFORD ARIS

University of Minnesota, Minneapolis, Minnesota

The maximum principle of L. S. Pontryagin for the solution of optimization problems in continuous processes has been derived recently by many workers (*2, 4, 6, 8*). The rigorous derivations of the Russian school are difficult to read and comprehend, while those based on the principle of optimality are lacking in rigor for even the simplest cases (*7*). The purpose of this note is to present a simple, yet rigorous, derivation of a slightly weakened result which is computationally equivalent and which has the advantage of easy extension to systems of complicated topology of the sort common in the chemical industry.

The authors consider the state of the system to be defined by the N-dimensional vector differential equation

$$\dot{\mathbf{p}} = \mathbf{f}(\mathbf{p}, \mathbf{q}), \quad O \leq t \leq \theta \tag{1}$$

where $\mathbf{q}(t)$ is the decision vector and θ may or may not be specified. If the system is not autonomous, it is made so by defining an additional state variable

$$\dot{p}^{N+1} = 1, \quad p^{N+1}(O) = 0 \tag{2}$$

and removing the explicit time dependence. If θ is fixed, then one has essentially introduced a condition of the type (2) with the additional requirement that $p^{N+1}(\theta) = \theta$. The goal is to maximize a function

$$P[\mathbf{p}(\theta)] \tag{3}$$

of the components of the state at time θ. This objective is completely general, and other functions may be reduced to this case by defining new state variables.

The authors assume that $\mathbf{f}$ and $\partial \mathbf{f}/\partial p^i$ are continuous in their arguments, that $\mathbf{q}$ is piece-by-piece continuous, and that $\partial \mathbf{f}/\partial q^k$ exists and is continuous everywhere except at a finite number of points, where one-sided limits exist.

If two solutions which remain finite for $\theta < \infty$ are taken, namely $\bar{\mathbf{p}}(t)$ satisfying the equations with

$$\bar{\mathbf{q}}(t), \quad \bar{\mathbf{p}}(O) = \bar{\mathbf{p}}_o \tag{4a}$$

and $\mathbf{p}(t)$ corresponding to the decision vector and initial conditions

$$\mathbf{q}(t) = \bar{\mathbf{q}}(t) + \delta\mathbf{q}(t), \quad \mathbf{p}(O) = \bar{\mathbf{p}}_o + \delta\mathbf{p}_o \tag{4b}$$

where

$$|\delta\mathbf{q}| \text{ and } |\delta\mathbf{p}_o| < \epsilon \tag{5}$$

then

$$|\delta\mathbf{p}(\theta)| = |\mathbf{p}(\theta) - \bar{\mathbf{p}}(\theta)| < K\epsilon \tag{6}$$

and K depends on θ but not on the variations.

The variational equations of (1) are*

$$\delta\dot{p}^i = \frac{\partial f^i}{\partial p^j}\delta p^j + \frac{\partial f^i}{\partial q^j}\delta q^j \tag{7}$$

where the partial derivatives are evaluated along $\bar{\mathbf{p}}(t)$, and, as a consequence of (6), neglected terms are of order ϵ^2. This is a linear, nonhomogeneous, first-order system, whose solution may be written in terms of Green's functions by employing Green's identity (*3, 5*):

$$\tilde{\omega}_i \,\delta p^i \,|_\theta = \tilde{\omega}_i \,\delta p^i|_o + \int_o^\theta \tilde{\omega}_i \frac{\partial f^i}{\partial q^j}\delta q^j \,dt \tag{8}$$

where the Green's vector $\tilde{\omega}$ is defined along the trajectory $\bar{\mathbf{p}}(t)$ by the adjoint differential equation

$$\dot{\tilde{\omega}}_i = -\frac{\partial f^j}{\partial p^i}\tilde{\omega}_j \tag{9}$$

If one defines the Hamiltonian function by

$$H(t) = \tilde{\omega} \cdot \mathbf{f} = \tilde{\omega}_i f^i \tag{10}$$

then (8) may be written

$$\tilde{\omega}_i \,\delta p^i \,|_\theta = \tilde{\omega}_i \,\delta p^i|_o + \int_o^\theta \frac{\partial H}{\partial q^i}\delta q^i \,dt \tag{11}$$

Now suppose that $\bar{\mathbf{p}}(t)$ is the trajectory which maximizes P. Then it is necessary that

$$\delta P = \frac{\partial P}{\partial p^i}\delta p^i(\theta) \tag{12}$$

be zero for unrestricted variations, or be negative for all allowable variations at constraints. If one now defines a boundary condition for the Green's vector as

$$\tilde{\omega}_i(\theta) = \frac{\partial P}{\partial p^i} \tag{13}$$

then (11) may be rewritten as

$$\delta P = \tilde{\omega}_i \,\delta p^i \,|_o + \int_o^\theta \frac{\partial H}{\partial q^i}\delta q^i \,dt \tag{14}$$

The terms are independent and must each satisfy the necessary conditions. For arbitrary $\delta\mathbf{q}$ it is necessary and sufficient that at all but a finite number of points

$$\frac{\partial H}{\partial q^j} = 0 \tag{15}$$

Sufficiency is obvious; necessity follows by considering the special variation $\delta q^i = \epsilon \dfrac{\partial H}{\partial p^i}$, $\epsilon > 0$. A similar proof requires that at constraints the optimal path satisfy

* The summation convention, where summation is indicated over all values of an index repeated once as a subscript and once as a superscript, is used throughout.

Morton M. Denn is at the University of Delaware, Newark, Delaware, and Rutherford Aris is at Cambridge University, Cambridge, England (1964-65).

$$\frac{\partial H}{\partial q^j} > 0 \qquad (16)$$

where q^j is taken to be increasing at the constraint. Furthermore, it follows at once from (10) that along the optimal path

$$\dot{H}(t) = 0 \qquad (17)$$

or

$$H = \text{const.} = c \qquad (18)$$

The maximum principle states that along an optimal trajectory the decision vector $\bar{q}(t)$ is chosen at each point (with the possible exception of a finite number) to maximize the Hamiltonian (10). The authors have proved the following slightly weaker result: *along an optimal trajectory (with the possible exception of a finite number of points) the Hamiltonian (10) is stationary at interior values of components of the optimal decision vector $\bar{q}(t)$ and a maximum with respect to components of $\bar{q}(t)$ which lie at a constraint.* Since interior maxima are generally found from stationary solutions, the strong and weak principles are computationally equivalent.

Furthermore, if $\mathbf{q}(t)$ is continuous in $\theta - \epsilon \leqq t \leqq \theta$ for any $\epsilon > 0$, then the partal derivatives $\partial p^i(\theta)/\partial p^j(t)$ exist. Since $P[\mathbf{p}(\theta)]$ must be invariant with respect to choice of initial position on the optimal path, then

$$\frac{dP}{dt_o} = \frac{\partial P}{\partial p^i} \frac{\partial p^i(\theta)}{\partial p^j(t_o)} f^j\left(\mathbf{p}(t_o), \mathbf{q}(t_o)\right) = 0, \qquad \theta - \epsilon \leqq t_o \leqq \theta \qquad (19)$$

Taking the limit as $t_o \rightarrow \theta$ one has

$$\frac{\partial P}{\partial p^i} f^i \Big|_\theta = 0 = \tilde{\omega}_i f^i \Big|_\theta = H(\theta) \qquad (20)$$

and hence, for free end conditions

$$c = 0 \qquad (21)$$

If $\mathbf{p}_o$ is specified, then $\delta\mathbf{p}_o$ is identically zero and the first term on the right of (14) vanishes identically. If p_o^k is completely free, then δp_o^k is arbitrary and one obtains the natural boundary condition

$$\tilde{\omega}_k(O) = 0 \qquad (22)$$

If $\mathbf{p}_o$ is constrained by

$$h(\mathbf{p}_o) = 0 \qquad (23)$$

then

$$\delta h = \frac{\partial h}{\partial p_o^i} \delta p_o^i = 0 \qquad (24)$$

and one obtains the transversality condition

$$\tilde{\omega}_k(O) = \eta \frac{\partial h}{\partial p_o^k} \qquad (25)$$

Similar results obtain for $t = \theta$ and for combinations of conditions at the end points. Thus for the fixed time problem one has $\tilde{\omega}_{N+1}(\theta)$ unspecified, since $p^{N+1}(\theta)$ is fixed. Since $\tilde{\omega}_{N+1}(t)$ is a constant, say $-k$, one has

$$\mathbf{H}(t) = H(t) + k \neq 0 \qquad (26)$$

For the fixed problem one may then define a Hamiltonian $\mathbf{H}$ based on f^i, $i = 1, \ldots, N$, which is a nonzero constant.

The Green's function approach easily treats such complications as constrained trajectories and systems with time delay, producing with simplicity the slightly weakened equivalents of results obtained by Pontryagin and coworkers. Such generalizations are immediate and will not be treated here.

The use of the Green's function approach in the construction of computational techniques is well known (*1, 3, 5*); the simplicity of derivation of necessary conditions suggests the Green's function as providing a unified approach to optimal systems, since the solution of the linear variational equations is necessary to the solution of the optimization problem. A Green's function treatment of discrete systems leading to analogous results for necessary conditions first deduced by Katz (*4*) and to computational techniques will be discussed elsewhere (*1*).

A further advantage of the Green's function approach is the ease with which it treats systems of complicated topology. As a simple example, suppose that t is the holding time in a pipeline chemical reactor, and that a fraction λ of the effluent $\mathbf{p}(\theta)$ is recycled directly to the feed; that is

$$\mathbf{p}_o = \mathbf{p}^* + \lambda\, \mathbf{p}(\theta) \qquad (27)$$

If $\mathbf{p}^*$ is fixed, then

$$\delta\mathbf{p}_o = \lambda \delta\mathbf{p}(\theta) \qquad (28)$$

and Green's identity takes the form

$$[\tilde{\omega}_i(\theta) - \lambda\tilde{\omega}_i(O)]\, \delta p^i(\theta) = \int_o^\theta \frac{\partial H}{\partial q^j} \delta q^j\, dt \qquad (29)$$

Thus it is seen that the maximum principle still obtains but the condition on the Green's functions is now

$$\tilde{\omega}_i(\theta) - \lambda\, \tilde{\omega}_i(O) = \frac{\partial P}{\partial p^i} \qquad (30)$$

Such problems for both continuous and discrete systems, and, in particular, the computational problems involved, will be treated elsewhere (*1*).

NOTATION

c = constant
$\mathbf{H}, H$ = Hamiltonian functions
P = profit function
$\mathbf{p}$ = state vector
$\mathbf{q}$ = decision vector
t = timelike variable

Greek Letters

$\delta\mathbf{p}, \delta\mathbf{q}$ = infinitesimal variations
ϵ = infinitesimal
η = constant in transversality condition
λ = fraction of product recycled
$\tilde{\omega}$ = Green's vector

LITERATURE CITED

1. Denn, M. M., and Rutherford Aris, *Ind. Eng. Chem. Fund.*, to be published.
2. DeSoer, C. A., *J. Franklin Inst.*, **271**, 361 (1961).
3. Goodman, T. R., and G. N. Lance, *Math. Tables and Other Aids to Comp.*, **10**, 82 (1956).
4. Katz, S., *Ind. Eng. Chem. Fund.*, **1**, 226 (1962).
5. Kelley, H. J., "Optimization Techniques," G. Leitmann, ed., Academic Press, New York (1962).
6. Kopp, R., *ibid.*
7. Pontryagin, L. S., V. A. Boltyanskii, R. V. Gamkelidze, and E. F. Mishchenko, "The Mathematical Theory of Optimal Processes," Wiley, New York (1962).
8. Rozonoer, L., *Automation and Remark Control* (trans. *Avt. i Tekmekh.*), **20**, 1288, 1405, 1517 (1959).

Accepted April 13, 1964.