JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 32, No. 1, SEPTEMBER 1980

Sensititivity Analysis in Dynamic Optimization[1]

A. H. EVERS[2]

Communicated by G. Leitmann

Abstract. To find the optimal control of chemical processes, Pontryagin's minimum principle can be used. In practice, however, one is not only interested in the optimal solution, which satisfies the restrictions on the control, the initial and terminal conditions, and the process parameters. It is also important to known how the optimal control and the minimum value of the objective function change, due to small variations in all the restrictions and the parameters. It is shown how to determine the effect of these variations directly from the optimal solution. This saves computer time, compared with the more traditional sensitivity analysis based on computing the optimal control for every single variation considered. The theory is applied to a chemical process.

Key Words. Sensitivity analysis, dynamic optimization, optimal control, near-optimal control, Pontryagin's minimum principle.

1. Introduction

In many physical and chemical processes, the dynamics at equations play an important role. The dynamical behavior of the processes depends on the control. If the terminal conditions can be reached by various controls (as a function of time), we want to know by which control a previously defined objective function is minimized. To compute this optimal control, Pontryagin's minimum principle can be used. In practice, it is also important to know how the optimal control and the minimum value of the objective function change for small variations in the restrictions and the parameters. It is shown how to determine the effect of these variations directly and without much additional work from the optimal solution of the problem under nominal restrictions, conditions, and parameters (Section 2). In Section 3,

[1] This paper is based on the author's doctoral study performed at Akzo Research Laboratories, Department of Applied Mathematics, Arnhem, Holland.

[2] Research Assistant, Department of Applied Mathematics, Twente University of Technology, Enschede, Holland.

0022-3239/80/0900-0017$03.00/0

Pontryagin's minimum principle is applied to a chemical process. In a postoptimal sensitivity analysis, the results of Section 2 have been implemented.

2. Sensitivity Analysis

Pontryagin's Minimum Principle. Let us consider the problem of finding an optimal control $u(t)$ that minimizes the integral

$$J=\int_{T_0}^{T_1} F(x, u, q, t)\, dt,$$

where $x(t)$ satisfies a set of differential equations

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

with initial conditions

$$x(T_0)=X_0,$$

and terminal conditions

$$\Omega_j(X_1, T_1)=0, \qquad j=1, \ldots, r;$$

here, t is the independent variable, with initial value T_0 and final value T_1, $x(t)$ is an n-dimensional state variable, X_1 is the final state $x(T_1)$, f is an n-dimensional vector function, $u(t)$ is an m-dimensional control variable, and q is an s-dimensional parameter vector, possibly depending on t.

The independent variable will be referred to as time, but may have another meaning (for example, the spatial coordinate in a chemical tubular reactor).

Assumption 2.1. The functions F and f are once differentiable with respect to x, u, q; their derivatives are continuous with respect to x, u, q.

Assumption 2.2. The functions Ω_j are differentiable with respect to X_1 and T_1 for every j; their derivatives are continuous with respect to X_1 and T_1.

The problem is to find conditions for a piecewise continuous optimal control $u(t)$ that belongs to an admissible set U defined by inequality restrictions

$$\phi_j(u, t)\leq 0, \qquad j=1, \ldots, \mu.$$

A control u is at the boundary of the admissible set U at time t, if at least one of the functions $\phi_j(u, t)$ satisfies $\phi_j(u, t) = 0$. Restrictions on the state $x(t)$ are not considered.

Pontryagin's principle now postulates that the optimal control minimizes the Hamiltonian

$$H(x, u, p, q, t) = F(x, u, q, t) + \sum_{i=1}^{n} p_i(t) f_i(x, u, q, t),$$

with respect to u. The adjoint variable $p(t)$ satisfies by definition

$$\dot{p}_i = -\partial H/\partial x_i, \qquad i = 1, \ldots, n.$$

At the terminal point (X_1, T_1) of the optimal trajectory [the curve $x(t)$ in the (x, t)-space, defined by the optimal control $u(t)$], the following transversality conditions must be satisfied:

$$-p_i(T_1) + \sum_{j=1}^{r} \omega_j(\partial\Omega_j/\partial X_{1i}) = 0, \qquad j = 1, \ldots, n,$$

$$H(T_1) + \sum_{j=1}^{r} \omega_j(\partial\Omega_j/\partial T_1) = 0,$$

where $\omega_1, \omega_2, \ldots, \omega_r$ are constant multipliers.

In the following sections, small variations of the control and state are considered. They are defined as follows.

Definition 2.1. A variation $u(t) + \delta u(t)$ of the optimal control $u(t)$ is called *small* if

$$\int_{T_0}^{T_1} \|\delta u(t)\|\, dt = O(\epsilon),$$

where ϵ is a given small positive number and

$$\|\delta u(t)\| = \left[\sum_{j=1}^{m} \{\delta u_j(t)\}^2\right]^{\frac{1}{2}}.$$

Definition 2.2. A variation $x(t) + \delta x(t)$ of the optimal trajectory $x(t)$ is called *small* if

$$\|\delta x(t)\| = O(\epsilon), \qquad \text{for all } t \in (T_0, T_1),$$

where ϵ is a given small positive number and

$$\|\delta x(t)\| = \left[\sum_{i=1}^{n} \{\delta x_i(t)\}^2\right]^{\frac{1}{2}}.$$

We consider optimal control problems which satisfy the following assumption.

Assumption 2.3. The optimal control depends continuously on the data; the variation in the optimal control due to small variations in the conditions or the parameters is also small; that is, when

$$\|\delta x_0\| < \epsilon, \qquad |\delta q| < \epsilon, \qquad \text{and/or } |\delta \Omega_j| < \epsilon,$$

the new optimal control $v(t)$ satisfies the relations

$$\int_{T_0}^{T_1} \|v(t) - u(t)\| \, dt = O(\epsilon),$$

$$\delta T_1 = O(\epsilon),$$

$$v(t) - u(T_1) = O(\epsilon), \qquad \text{for } t \in (T_1, T_1 + \delta T_1).$$

Formulation of the Problem. In many situations, one is interested in the effect of small changes in the control, the initial condition, the terminal conditions, or the parameters. If an optimal control is difficult to realize, then one might approximate it by a neighboring control that is easier to realize. Sometimes, it is possible to change the initial conditions (for instance, the initial concentrations of the reactions in a chemical process) or the terminal conditions (such as the reaction time or the requirements to be met by the purity of a product of a distillation process).

If a parameter has been estimated, it is important to know the effect of the error in the estimate in order to decide whether it is necessary to obtain a better estimate or not. It is also possible that the value of a parameter changes with time (catalyst, heat transfer coefficient). In all of these cases, we are interested in the effect of small variations on the objective function J.

Instead of computing the optimal control and the objective function for every single variation, we want to compute sensitivities directly from the data belonging to the normal optimal control. This problem has been studied by Tuel, Lee, and DeRusso (Ref. 1), Courtin and Rootenberg (Ref. 2), Kreindler (Ref. 3), Rootenberg and Courtin (Ref. 4), and Peterson (Ref. 5). But these authors all deal with special cases (such as unrestricted control, differential equations linear in u, comparison between open and closed-loop controllers).

We first define *near-optimal controls*. Then, we apply variational methods to determine the effect of small variations in the control, initial and terminal conditions, and parameters.

Near-Optimal Controls. In order to consider the effect of small variations in the optimal control $u(t)$, we define near-optimal controls.

Definition 2.3. An admissible small variation $u(t)+\delta u(t)$ of the optimal control $u(t)$ is near optimal if

$$\int_{T_0}^{T_1} \{H(x, u+\delta u, q, t)-H(x, u, q, t)\}\, dt = O(\epsilon^2).$$

If the time interval for the varied control is extended, then the variation is near optimal if, in addition,

$$\int_{T_1}^{T_1+\delta T_1} \{H(x, u+\delta u, q, t)-H(x, u, q, t)|_{T_1}\}\, dt = O(\epsilon^2).$$

Theorem 2.1. The optimal control of the problem with slightly varied conditions or parameters is a near-optimal variation of the optimal control of the nominal problem.

Proof. (i) Suppose that v is the optimal control of the varied problem. From Assumption 2.3, it follows that, for a small variation ϵ in the conditions or parameters, the following relations hold:

$$\int_{T_0}^{T_1} \|v(t)-u(t)\|\, dt = O(\epsilon),$$

$$\delta T_1 = O(\epsilon),$$

$$v(t)-u(T_1) = O(\epsilon), \text{ for } t \in (T_1, T_1+\delta T_1).$$

For this varied problem, the optimal tranjectory is $x(t)+\delta x(t)$, and the adjoint variable $p(t)+\delta p(t)$.

(ii) The variation in the state, due to variations

$$\|\delta x_0\| < \epsilon, \qquad |\delta q| < \epsilon, \qquad \int_{T_0}^{T_1} \|\delta u\|\, dt < \epsilon,$$

is

$$\delta x(t) = \delta x_0 + \int_{T_0}^{t} \{f(x+\delta x, u+\delta u, q+\delta q, \tau)-f(x, u, q, \tau)\}\, d\tau.$$

From Assumption 2.1, it follows that

$$\|f(x+\delta x, u+\delta u, q+\delta q, t)-f(x, u, q, t)\| \le L_1(\|\delta x\|+\|\delta u\|+|\delta q|),$$

and so

$$\begin{aligned}\|\delta x(t)\| &\leq \|\delta X_0\| + \int_{T_0}^{t} L_1(\|\delta x\| + \|\delta u\| + |\delta q|)\, d\tau \\ &\leq \epsilon + L_1\{M(t-T_0) + \epsilon + \epsilon(t-T_0)\},\end{aligned} \tag{1}$$

where L_1 is a Lipschitz constant and

$$M = \max_{t\in(T_0,T_1)} \|\delta x(t)\|.$$

Repeated substitution (N times) of the estimation of $\|\delta x(t)\|$ into (1) yields

$$\|\delta x(t)\| \leq (M+\epsilon)[L_1^{N+1}(t-T_0)^{N+1}/(N+1)!] + \epsilon L \sum_{k=0}^{N} [L_1^k (t-T_0)^k/k!],$$

where

$$(M+\epsilon)[L_1^{N+1}(t-T_0)^{N+1}/(N+1)!] \to 0, \qquad \text{if } N \to \infty.$$

Therefore,

$$\|\delta x(t)\| \leq \epsilon L_1 \exp[L_1(t-T_0)],$$

or

$$\delta x(t) = O(\epsilon).$$

With a similar derivation, it can be proved that $\delta p(t)$ is $O(\epsilon)$.

(iii) As $u(t)$ is the optimal control, it follows from the minimum principle that

$$\int_{T_0}^{T_1} \{H(x, v, q, p, t) - H(x, u, q, p, t)\}\, dt \geq 0. \tag{2}$$

But $v(t)$ is the optimal control of the varied problem. Therefore,

$$\int_{T_0}^{T_1} \{H(x+\delta x, u, q+\delta q, p+\delta p, t) - H(x+\delta x, v, q+\delta q, p+\delta p, t)\}\, dt \geq 0,$$

or

$$\begin{aligned}&\int_{T_0}^{T_1} \{H(x,u,q,p,t) - H(x,v,q,p,t)\}\, dt + \int_{T_0}^{T_1} \left\{ \sum_{i=1}^{n} \left[\frac{\partial H}{\partial x_i}(u) - \frac{\partial H}{\partial x_i}(v)\right] \delta x_i \right. \\ &\qquad \left. + \left[\frac{\partial H}{\partial q}(u) - \frac{\partial H}{\partial q}(v)\right] \delta q + \sum_{i=1}^{n} \left[\frac{\partial H}{\partial p_i}(u) - \frac{\partial H}{\partial p_i}(v)\right] \delta p_i \right\} dt \geq 0.\end{aligned}$$

From Assumption 2.1, it follows that

$$\left|\frac{\partial H}{\partial x_i}(u)-\frac{\partial H}{\partial x_i}(v)\right| \leq L_2\|u-v\|;$$

therefore,

$$\left|\int_{T_0}^{T_1} \sum_{i=1}^{n} \left[\frac{\partial H}{\partial x_i}(u)-\frac{\partial H}{\partial x_i}(v)\right] \delta x_i \, dt\right| \leq L_2 n \max_{t\in(T_0,T_1)} \|\delta x(t)\| \int_{T_0}^{T_1} \|u-v\| \, dt = O(\epsilon^2).$$

Further, it can be proved that

$$\int_{T_0}^{T_1} \left[\frac{\partial H}{\partial q}(u)-\frac{\partial H}{\partial q}(v)\right] \delta q \, dt = O(\epsilon^2),$$

and also, using $\partial H/\partial p_i = f_i$, that

$$\int_{T_0}^{T_1} \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}(u)-\frac{\partial H}{\partial p_i}(v)\Big] \delta p_i \, dt = O(\epsilon^2).$$

Therefore,

$$\int_{T_0}^{T_1} \{H(x, u, q, p, t)-H(x, v, q, p, t)\} \, dt \geq 0, \qquad (3)$$

up to terms of order ϵ^2. Combining (2) and (3), we find that

$$\int_{T_0}^{T_1} \{H(x, v, q, p, t)-H(x, u, q, p, t)\} \, dt = O(\epsilon^2).$$

For the extension of the time interval, we have

$$\left|\int_{T_1}^{T_1+\delta T_1} \{H(x, v, q, t)-H(x, u, q, t)|_{T_1}\} \, dt\right| \leq \int_{T_1}^{T_1+\delta T_1} L_3\|v(t)-u(T_1)\| \, dt = O(\epsilon^2).$$

In conclusion, $v(t)$ is a near-optimal variation of $u(t)$.

Effect of Variations in Control, Initial Conditions, Terminal Conditions, and Parameters on *J*. In this section, we first prove a theorem with relation to the neighborhood of the terminal point. Next, we give a general derivation of the effect on the performance criterion of small variations in the control, initial and terminal conditions, and parameters.

First, we study the variation of the terminal point (Fig. 1).

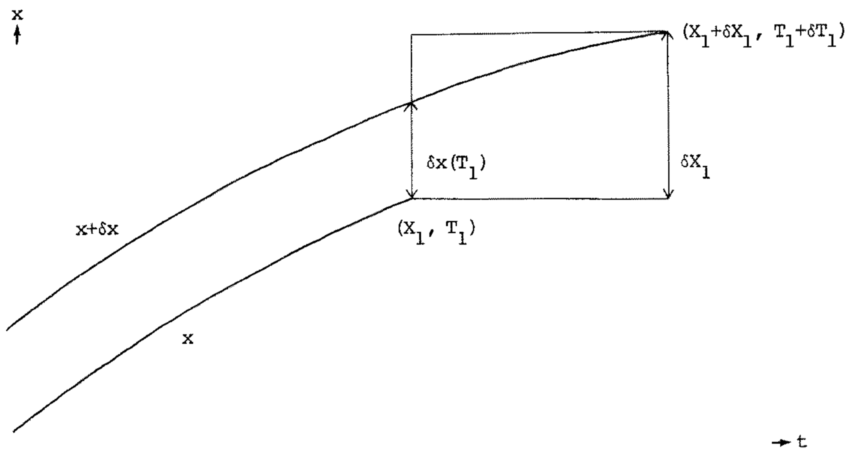

Fig. 1. Variation of the terminal point.

Theorem 2.2. If the terminal point is subject to the end conditions

$$\Omega_j(X_1, T_1) = 0, \qquad j = 1, \ldots, r,$$

then, for a reachable terminal point $(X_1+\delta X_1, T_1+\delta T_1)$ in the neighborhood of the terminal point (X_1, T_1) of the optimal trajectory, the relation

$$-\sum_{i=1}^{n} p_i(T_1)\,\delta X_{1i} + H(T_1)\,\delta T_1 + \sum_{j=1}^{r} \omega_j\,\delta\Omega_j = 0$$

is satisfied up to terms of order ϵ^2.

Proof. At the terminal point (X_1, T_1), the transversality conditions hold:

$$-p_i(T_1) + \sum_{j=1}^{r} \omega_j(\partial\Omega_j/\partial X_{1i}) = 0, \qquad i = 1, \ldots, n,$$

$$H(T_1) + \sum_{j=1}^{r} \omega_j(\partial\Omega_j/\partial T_1) = 0.$$

Multiplying the upper equations by δX_{1i}, summing over i, and adding to the lower equation, multiplied by δT_1, we find that

$$\begin{aligned}
-\sum_{i=1}^{n} p_i(T_1)\,\delta X_{1i} &+ \sum_{i=1}^{n}\sum_{j=1}^{r} \omega_j(\partial\Omega_j/\partial X_{1i})\,\delta X_{1i} + H(T_1)\,\delta T_1 \\
&+ \sum_{j=1}^{r} \omega_j(\partial\Omega_j/\partial T_1)\,\delta T_1 \\
&= -\sum_{i=1}^{n} p_i(T_1)\,\delta X_{1i} + H(T_1)\,\delta T_1 \\
&+ \sum_{j=1}^{r} \omega_j\Big[\sum_{i=1}^{n} (\partial\Omega_j/\partial X_{1i})\,\delta X_{1i} + (\partial\Omega_j/\partial T_1)\,\delta T_1\Big] \\
&= -\sum_{i=1}^{n} p_i(T_1)\,\delta X_{1i} + H(T_1)\,\delta T_1 + \sum_{j=1}^{r} \omega_j\,\delta\Omega_j = 0, \qquad (4)
\end{aligned}$$

in which

$$\delta\Omega_j = \sum_{i=1}^{n} (\partial\Omega_j/\partial X_{1i})\,\delta X_{1i} + (\partial\Omega_j/\partial T_1)\,\delta T_1$$

is the first variation of Ω_j. Equation (4) holds at every reachable point $(X_1+\delta X_1, T_1+\delta T_1)$ in the neighborhood of the terminal point (X_1, T_1) of the optimal trajectory.

Remark 2.1. In the case $\delta T_1 = 0$, the term $H(T_1)\,\delta T_1$ vanishes, and we may replace δX_{1i} by $\delta x_i(T_1)$; for, up to terms of second order in ϵ,

$$\delta X_{1i} = \delta x_i(T_1) + \dot{x}_i(T_1)\,\delta T_1 = \delta x_i(T_1).$$

Next, we present a general derivation of the variation in J as a result of small variations in the control, initial and terminal conditions, and parameters. A change in the control, initial conditions, final time, and/or parameters causes a change in the state. Thus, variations δu, δX_0, δT_1, δq result in variations $\delta x(t)$, δX_1.

Let us consider the expression

$$\delta J = \int_{T_0}^{T_1+\delta T_1} F(x+\delta x, u+\delta u, q+\delta q, t)\,dt - \int_{T_0}^{T_1} F(x, u, q, t)\,dt.$$

In the Appendix, it is shown that this expression equals

$$\begin{aligned}
\Big[-\sum_{i=1}^{n} p_i\delta x_i\Big]_{T_0}^{T_1} &+ \int_{T_0}^{T_1} \{H(x+\delta x, u+\delta u, q+\delta q, t) - H(x+\delta x, u, q+\delta q, t) \\
&+ \sum_{l=1}^{s} \frac{\partial H}{\partial q_l}(x, u, q, t)\,\delta q_l\}\,dt + [F(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1}\,\delta T_1 + O(\epsilon^2),
\end{aligned} \qquad (5)$$

where $\delta x_i(T_1)$ follows from (see Fig. 1)

$$\delta X_{1i} = \delta x_i(T_1) + [f_i(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1}\, \delta T_1 + O(\epsilon^2).$$

Therefore,

$$\begin{aligned}
\delta J = \left[\sum_{i=1}^{n} p_i \delta x_i\right]_{T_0} &- \sum_{i=1}^{n} p_i(T_1)\{\delta X_{1i} - [f_i(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1}\, \delta T_1\} \\
&+ \int_{T_0}^{T_1} \left\{H(x, u+\delta u, q, t) - H(x, u, q, t) + \sum_{l=1}^{s} \frac{\partial H}{\partial q_l}(x, u, q, t)\, \delta q_l\right\} dt \\
&+ \int_{T_0}^{T_1} \sum_{i=1}^{n} \left\{\frac{\partial H}{\partial x_i}(x, u+\delta u, q+\delta q, t) - \frac{\partial H}{\partial x_i}(x, u, q+\delta q, t)\right\} \delta x_i\, dt \\
&+ \int_{T_0}^{T_1} \sum_{l=1}^{s} \left\{\frac{\partial H}{\partial q_l}(x, u+\delta u, q, t) - \frac{\partial H}{\partial q_l}(x, u, q, t)\right\} \delta q_l\, dt \\
&+ [F(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1}\, \delta T_1 + O(\epsilon^2).
\end{aligned}$$

It is assumed that

$$\int_{T_0}^{T_1+\delta T_1} \|\delta u\|\, dt < \epsilon, \qquad |\delta T_1| < \epsilon.$$

From Assumption 2.1, it follows that $\partial H/\partial x_i$ and $\partial H/\partial q_l$ satisfy the Lipschitz conditions:

$$\left|\frac{\partial H}{\partial x_i}(x, u+\delta u, q+\delta q, t) - \frac{\partial H}{\partial x_i}(x, u, q+\delta q, t)\right| \le C_1 \|\delta u\|, \tag{6-1}$$

$$\left|\frac{\partial H}{\partial q_l}(x, u+\delta u, q, t) - \frac{\partial H}{\partial q_l}(x, u, q, t)\right| \le C_2 \|\delta u\|, \tag{6-2}$$

in which C_1 and C_2 are constant. Then,

$$\begin{aligned}
&\left|\int_{T_0}^{T_1} \sum_{i=1}^{n} \left\{\frac{\partial H}{\partial x_i}(x, u+\delta u, q+\delta q, t) - \frac{\partial H}{\partial x_i}(x, u, q+\delta q, t)\right\} \delta x_i\, dt\right| \\
&\quad \le \int_{T_0}^{T_1} \sum_{i=1}^{n} \left\{\left|\frac{\partial H}{\partial x_i}(x, u+\delta u, q+\delta q, t) - \frac{\partial H}{\partial x_i}(x, u, q+\delta q, t)\right|\right\} |\delta x_i|\, dt \\
&\quad \le \int_{T_0}^{T_1} \sum_{i=1}^{n} C_1 \|\delta u\| \cdot |\delta x_i|\, dt \le C_1 \left\{\sum_{i=1}^{n} \max |\delta x_i|\right\} \int_{T_0}^{T_1} \|\delta u\|\, dt = O(\epsilon^2),
\end{aligned}$$

as

$$\sum_{i=1}^{n} \max |\delta x_i|, \qquad \int_{T_0}^{T_1} \|\delta u\|\, dt$$

are of order ϵ, and

$$\left|\int_{T_0}^{T_1} \sum_{l=1}^{s} \left\{\frac{\partial H}{\partial q_l}(x, u+\delta u, q, t) - \frac{\partial H}{\partial q_l}(x, u, q, t)\right\} \delta q_l \, dt\right|$$
$$\leq \int_{T_0}^{T_1} \sum_{l=1}^{s} \left\{\left|\frac{\partial H}{\partial q_l}(x, u+\delta u, q, t) - \frac{\partial H}{\partial q_l}(x, u, q, t)\right|\right\} |\delta q_l| \, dt$$
$$\leq \int_{T_0}^{T_1} \sum_{l=1}^{s} C_2 \|\delta u\| \cdot |\delta q_l| \, dt \leq C_2 \left\{\sum_{l=1}^{s} \max|\delta q_l|\right\} \int_{T_0}^{T_1} \|\delta u\| \, dt = O(\epsilon^2).$$

as

$$\sum_{l=1}^{s} \max|\delta q_l|, \qquad \int_{T_0}^{T_1} \|\delta u\| \, dt$$

are of order ϵ.

Assuming $u+\delta u$ to be near optimal on $[T_1, T_1+\delta T_1]$, we have from Theorem 2.2 that

$$-\sum_{i=1}^{n} p_i(T_i)\delta X_{1i} + \sum_{i=1}^{n} p_i(T_1)[f_i(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1} \delta T_1$$
$$+[F(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1} \delta T_1$$
$$= -\sum_{i=1}^{n} p_i(T_1)\,\delta X_{1i} + [H(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1} \delta T_1$$
$$= -\sum_{i=1}^{n} p_i(T_1)\,\delta X_{1i} + [H(x, u+\delta u, q, t)]_{T_1} \delta T_1$$
$$+ \sum_{i=1}^{n} \left[\frac{\partial H}{\partial x_i}(x, u+\delta u, q+\delta q, t)\right]_{T_1} \delta x_i \, \delta T_1$$
$$+ \sum_{l=1}^{s} \left[\frac{\partial H}{\partial q_l}(x, u+\delta u, q, t)\right]_{T_1} \delta q_l \, \delta T_1 + O(\epsilon^2)$$
$$= -\sum_{i=1}^{n} p_i(R_i)\,\delta X_{1i} + \{[H(x, u, q, t)]_{T_1} + O(\epsilon)\} \delta T_1 + O(\epsilon^2)$$
$$= -\sum_{j=1}^{r} \omega_j \, \delta\Omega_j + O(\epsilon^2),$$

as δx_i, δT_1, δq_l are of order ϵ. We therefore find the following result for the variation of J:

$$\delta J=\sum_{i=1}^{n} p_i(T_0)\,\delta x_i(t_0)+\int_{T_0}^{T_1}\{H(x,u+\delta u,q,t)-H(x,u,q,t)\}\,dt$$
$$+\int_{T_0}^{T_1}\sum_{l=1}^{s}\frac{\partial H}{\partial q_l}(x,u,q,t)\,\delta q_l\,dt-\sum_{j=1}^{r}\omega_j\,\delta\Omega_j+O(\epsilon^2). \quad (7)$$

We see that the variation of the criterion consists of four terms showing the effect of: a change in the initial conditions (δJ_1), a variation of the control (δJ_2), a variation of the parameter (δJ_3), and a change in the terminal conditions (δJ_4). All terms are expressed for the values belonging to the nominal problem. By definition, δJ_2 is of second order in ϵ if the varied control is near optimal.

For example, if one considers a change in the inital conditions and a near-optimal control such that the terminal conditions are satisfied, then the effect on J is δJ_1. The effect of a change in a boundary value of the control is δJ_2, provided that the terminal conditions are satisfied. If the actual value of a parameter differs from the value used in the model, the error in J is expressed by $\delta J_3+\delta J_4$. If the terminal conditions can be satisfied by a near-optimal control, the effect on J is δJ_3.

3. Application to a Chemical Process

Description of the Model. Consider a chemical reaction in a semi-batch reactor with five chemical reactants, described by

$$\dot{x}_1=\phi-k_1x_1x_5-k_2x_1x_2-k_4x_1,$$
$$\dot{x}_2=k_1x_1x_5-k_2x_1x_2,$$
$$\dot{x}_3=k_2x_1x_2,$$
$$\dot{x}_4=k_4x_1,$$
$$\dot{x}_5=-2k_1x_1x_5-k_2x_1x_2-k_4x_1-\phi,$$

with initial conditions

$$x_i(0)=0, \qquad i=1,2,3,4, \qquad x_5(0)=x_{50},$$

and terminal condition

$$T_1=150\text{ min},$$

where x_i = concentration of chemical reactant i, k_i = reaction constant, with

$$k_i=k_{i0}\exp(-E_i/Ru),$$

k_{i0} = Arrhenius' constant, E_i = activation energy, R = gas constant, u = temperature (control), and ϕ = flow of reactant 1 into the reactor. Only in the first 70 min of the reaction is reactant 1 metered (ϕ = constant); afterward, $\phi = 0$, see Fig. 2.

The control variable u occurs only in the reaction constants k_i and is restricted by

$$u \leq c_1 + c_2 t, \qquad u \leq c_3, \qquad u \leq c_4 - c_5 t,$$

where t is time and $c_1, \ldots, c_5$ are positive constants (see Fig. 2). The

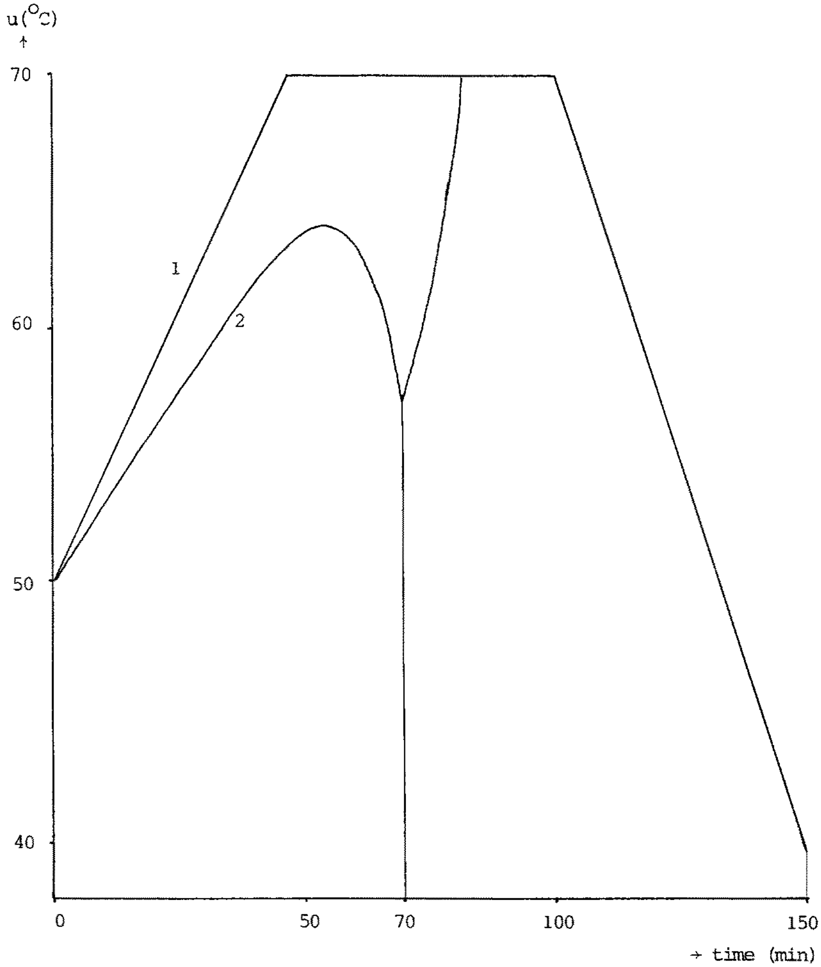

Fig. 2. Practical temperature profile 1 and optimal profile 2. At $t = 70$ min, the metering of reactant 1 is stopped. The practical profile coincides with the upper limit.

performance criterion to be minimized is

$$J = x_1(T_1) - x_2(T_1) = \int_{T_0}^{T_1} (\dot{x}_1 - \dot{x}_2)\, dt.$$

Numerical Solution. For the integration of the differential equations and the adjoint differential equations, a fourth-order Runge–Kutta procedure has been applied. First, a temperature profile was estimated (in this case, the profile used in practice was taken). Then, the differential equations were integrated forward; afterward, the adjoint differential equations were integrated backward with a double step.

At the points where p and x were calculated, the values of the Hamiltonian and its first and second derivative with respect to u were determined. A better estimation of the temperature profile is given by

$$u + \delta u = u - (\partial H/\partial u)/(\partial^2 H/\partial u^2).$$

The profile thus found must be corrected when the constraints are exceeded. Then, we can integrate again the differential equations, etc. If there was no significant improvement of the objective function J, the procedure was stopped and a sensitivity analysis was made.

Results of the Optimization. The integration of the differential equations was made with a stepsize of 2.5 min (60 steps). The iteration was stopped when

$$|\delta J| < 10^{-6};$$

this is the numerical error in J. The optimal profile and the concentrations of reactants 1 and 2 are given in Figs. 2–4. We can draw three conclusions from this optimization.

(i) The effect on the objective function is negligible.[3] For the practical temperature profile, $J = -0.858830$. For the optimal temperature profile, $J = -0.859075$. Therefore, the improvement is $\Delta J = -0.000245$, namely 0.03% of J.

(ii) The practical temperature profile is too high in the first 80 min (see Fig. 2 for the optimal profile).

(iii) The reaction lasts too long.

With the practical profile, 99% of the final concentration of reactant 2 is reached after 80 min. Thus, we expect that it will not have a serious effect on

[3] The use of six decimals in J is necessary for the sensitivity analysis, because the process is very stable. The digits are numerically significant; but only the first two decimals are of practical importance.

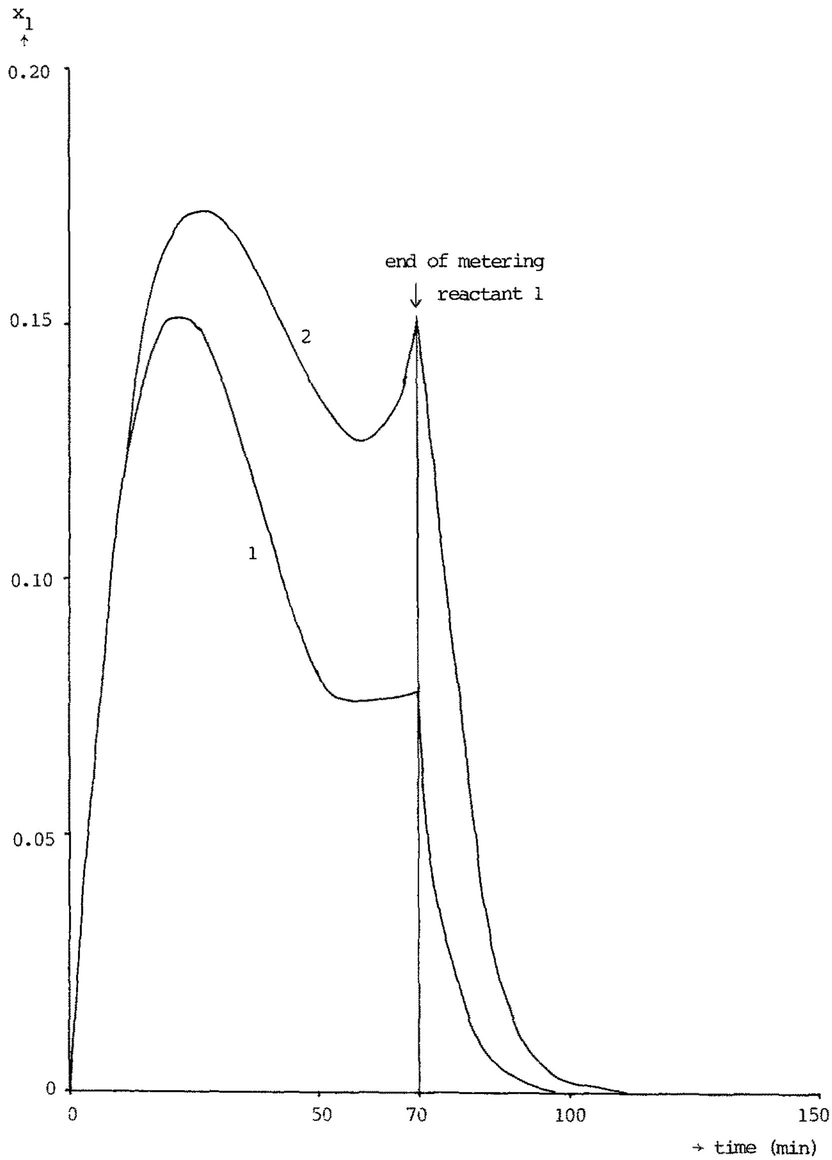

Fig. 3. Fraction of reactant 1 (x_1) for practical profile 1 and optimal profile 2.

the values of the objective function when the total reaction time decreases from 150 to 120 min. Then, cooling starts immediately after the metering of reactant 1 has stopped. Indeed, computation of the optimal temperature profile for a batch time of 120 min shows an increase of the objective

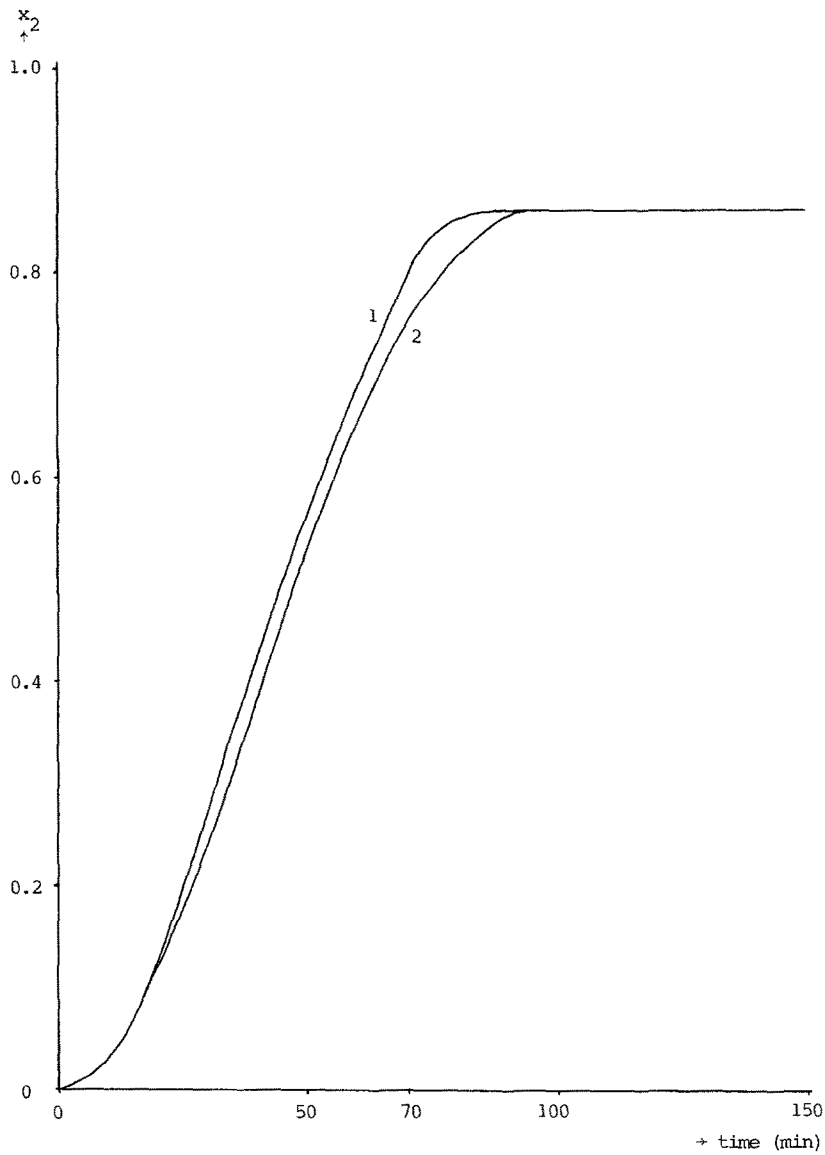

Fig. 4. Fraction of reactant 2 (x_2) for practical profile 1 and optimal profile 2.

function of 0.6% (so, the result is slightly worse). Thus, it is possible to run more batches a day; but this will only be advantageous if the greater amount of reaction product can be further handled in the plant.

Table 1. Results for the sensitivity analysis.

Variable	Nominal value	δV	δJ_1	δJ_2	$\|(\delta J_1 - \delta J_2)/J\|$ ×100%
V1	1.0×10^{10}	+1%	0.001038	0.001033	0.0006%
V2	0.27	+0.01	−0.005448	−0.005246	0.024%
V3	70 min	+2.5 min	−0.000010	−0.000007	0.0003%
V4	70°C	+3°C	−0.000054	−0.000042	0.002%
V5	150 min	+5 min	−0.000095	−0.000095	<0.0001%

V1 = Arrhenius' constant k_{20}, parameter.
V2 = initial concentration of reactant 5, initial condition (x_{50}).
V3 = metering time of reactant 1, total dose kept constant, parameter.
V4 = maximum temperature, restriction to control.
V5 = total reaction time T_1, terminal condition.

Sensitivity Analysis. Partial results of the sensitivity analysis are given in Table 1. They were calculated with the theory of Section 2 (δJ_1). The sensitivity analysis is also made in the traditional way by substituting one by one the nominal variables by slightly varied variables and calculating the new optimal profile and objective function. A comparison with the nominal objective function J gives us the real sensitivity δJ_2.

From the table, we see that an error of 1% in the estimation of k_{20} causes an error of 0.1% in J. An increase of the initial fraction of reactant 5 from 0.27 to 0.28 gives an improvement of 0.6% in J. Furthermore, we cannot expect a better result by changing the metering time of reactant 1, by increasing the upper temperature limit, or by increasing the total reaction time.

Conclusions. From the above results, the following conclusions can be derived.

(i) The effects, calculated with the senstivity analysis of Section 2, of variations in the various restrictions, control, and parameters are sufficiently accurate and give a good indication of the variations that are important. The saving in computer costs is considerable. The total costs including the calculation of seventeen sensitivities was Dfl 60 (ALGOL program, IBM 370/158). If the classical sensitivity analysis had been applied, starting with the nominal optimal temperature profile, the cost would have been Dfl 475.

(ii) For the process considered in the example, we can conclude that the best method to improve the profit is to raise the initial concentration of reactant 5. Also, it is not attractive to use a better temperature profile, because the gain in J is negligible. Finally, the total time for one batch can be lowered by 30 min without losing much of the gain (less than 1%). This makes is possible to handle more batches a day if the rest of the plant can be adapted.

4. Appendix: Derivation of Equation (5)

The derivation described in this section is similar to the derivation of the minimum principle and holds for every control u. Let us consider

$$\begin{aligned}\delta J &= \int_{T_0}^{T_1+\delta T_1} F(x+\delta x, u+\delta u, q+\delta q, t)\,dt - \int_{T_0}^{T_1} F(x, u, q, t)\,dt \\ &= \int_{T_0}^{T_1} \{F(x+\delta x, u+\delta u, q+\delta q, t) - F(x, u, q, t)\}\,dt \\ &\quad + \int_{T_1}^{T_1+\delta T_1} F(x+\delta x, u+\delta u, q+\delta q, t)\,dt \\ &= \int_{T_0}^{T_1} \{F(x+\delta x, u+\delta u, q+\delta q, t) - F(x+\delta x, u, q+\delta q, t) \\ &\quad + F(x+\delta x, u, q+\delta q, t) - F(x, u, q, t)\}\,dt \\ &\quad + \int_{T_1}^{T_1+\delta T_1} F(x+\delta x, u+\delta u, q+\delta q, t)\,dt.\end{aligned}$$

From

$$\begin{aligned}F(x+\delta x, u, q+\delta q, t) - F(x, u, q, t) &= \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(x, u, q, t)\,\delta x_i \\ &\quad + \sum_{l=1}^{s} \frac{\partial F}{\partial q_l}(x, u, q, t)\,\delta q_l,\end{aligned}$$

it follows that

$$\begin{aligned}\delta J = \int_{T_0}^{T_1} \Big\{&F(x+\delta x, u+\delta u, q+\delta q, t) - F(x+\delta x, u, q+\delta q, t) \\ &+ \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(x, u, q, t)\,\delta x_i + \sum_{l=1}^{s} \frac{\partial F}{\partial q_l}(x, u, q, t)\,\delta q_l\Big\}\,dt \\ &+ \int_{T_1}^{T_1+\delta T_1} F(x+\delta x, u+\delta u, q+\delta q, t)\,dt + O(\epsilon^2),\end{aligned}$$

up to terms of order ϵ^2.

The term

$$\sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(x, u, q, t)\, \delta x_i$$

will now be eliminated. For this, we consider functions p_i that satisfy, by definition,

$$\dot{p}_i = -\partial H/\partial x_i = -\partial F/\partial x_i - \sum_{j=1}^{n} p_j(\partial f_j/\partial x_i).$$

From

$$\begin{aligned}(d/dt)\left[\sum_{i=1}^{n} p_i\, \delta x_i\right] &= \sum_{i=1}^{n} [(d/dt)(p_i\, \delta x_i)] = \sum_{i=1}^{n} \dot{p}_i\, \delta x_i + \sum_{i=1}^{n} p_i\, \delta \dot{x}_i \\ &= -\sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(x, u, q, t)\, \delta x_i \\ &\quad - \sum_{i=1}^{n}\sum_{j=1}^{n} p_j \frac{\partial f_j}{\partial x_i}(x, u, q, t)\, \delta x_i + \sum_{i=1}^{n} p_i\, \delta \dot{x}_i,\end{aligned}$$

up to terms of order ϵ^2, and from

$$\begin{aligned}\delta \dot{x}_i &= \delta f_i = f_i(x+\delta x, u+\delta u, q+\delta q, t) - f_i(x, u, q, t) \\ &= f_i(x+\delta x, u+\delta u, q+\delta q, t) - f_i(x+\delta x, u, q+\delta q, t) \\ &\quad + f_i(x+\delta x, u, q+\delta q, t) - f_i(x, u, q, t) \\ &= f_i(x+\delta x, u+\delta u, q+\delta q, t) - f_i(x+\delta x, u, q+\delta q, t) \\ &\quad + \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x, u, q, t)\, \delta x_j + \sum_{l=1}^{s} \frac{\partial f_i}{\partial q_l}(x, u, q, t)\, \delta q_l,\end{aligned}$$

we obtain

$$\begin{aligned}\sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(x, u, q, t)\, \delta x_i = -(d/dt)\Bigg\{ &\sum_{i=1}^{n} p_i\, \delta x_i - \sum_{i=1}^{n}\sum_{j=1}^{n} p_j \frac{\partial f_j}{\partial x_i}(x, u, q, t)\, \delta x_i \\ &+ \sum_{i=1}^{n} p_i f_i(x+\delta x, u+\delta u, q+\delta q, t) \\ &- \sum_{i=1}^{n} p_i f_i(x+\delta x, u, q+\delta q, t) \\ &+ \sum_{i=1}^{n}\sum_{j=1}^{n} p_i \frac{\partial f_i}{\partial x_j}(x, u, q, t)\, \delta x_j \\ &+ \sum_{i=1}^{n}\sum_{l=1}^{s} p_i \frac{\partial f_i}{\partial q_l}(x, u, q, t)\, \delta q_l + O(\epsilon^2).\end{aligned}$$

Therefore,

$$\delta J = \int_{T_0}^{T_1} \Big[F(x+\delta x, u+\delta u, q+\delta q, t) - F(x+\delta x, u, q+\delta q, t)$$

$$-(d/dt)\Big\{ \sum_{i=1}^{n} p_i \, \delta x_i \Big\}$$

$$+ \sum_{i=1}^{n} p_i f_i(x+\delta x, u+\delta u, q+\delta q, t) - \sum_{i=1}^{n} p_i f_i(x+\delta x, u, q+\delta q, t)$$

$$+ \sum_{i=1}^{n} \sum_{l=1}^{s} p_i \frac{\partial f_i}{\partial q_l}(x, u, q, t)\, \delta q_l + \sum_{l=1}^{s} \frac{\partial F}{\partial q_l}(x, u, q, t)\, \delta q_l \Big]\, dt$$

$$+ \int_{T_1}^{T_1+\delta T_1} F(x+\delta x, u+\delta u, q+\delta q, t)\, dt + O(\epsilon^2)$$

$$= -\Big[\sum_{i=1}^{n} p_i \delta x_i \Big]_{T_0}^{T_1} + \int_{T_0}^{T_1} \Big\{ H(x+\delta x, u+\delta u, q+\delta q, t)$$

$$-H(x+\delta x, u, q+\delta q, t) + \sum_{l=1}^{s} \frac{\partial H}{\partial q_l}(x, u, q, t)\, \delta q_l \Big\}\, dt$$

$$+ \int_{T_1}^{T_1+\delta T_1} F(x+\delta x, u+\delta u, q+\delta q, t)\, dt + O(\epsilon^2).$$

From the mean-value theorem of integral calculus, we have that

$$[F(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1+\xi}\, \delta T_1 + O(\epsilon^2)$$
$$= [F(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1}\, \delta T_1 + [dF/dt]_{T_1}\, \xi\, \delta \cdot T_1 + O(\epsilon^2),$$

where $\xi\, \delta T_1$ is $O(\epsilon^2)$ and

$$0 \le \xi \le \delta T_1.$$

Therefore,

$$\delta J = -\Big[\sum_{i=1}^{n} p_i \, \delta x_i \Big]_{T_0}^{T_1} + \int_{T_0}^{T_1} \Big\{ H(x+\delta x, u+\delta u, q+\delta q, t)$$

$$-H(x+\delta x, u, q+\delta q, t)$$

$$+ \sum_{l=1}^{s} \frac{\partial H}{\partial q_l}(x, u, q, t)\, \delta q_l \Big\}\, dt + [F(x+\delta x, u+\delta u, q+\delta q, t)]_{T_1}\, \delta T_1 + O(\epsilon^2).$$

This derivation holds for every control $u(t)$.

References

1. TUEL, W. G., LEE, I., and DERUSSO, P. M., *Synthesis of Optimal Control Systems with Sensitivity Constraints,* Proceedings of the IFAC 3rd Congress, London, England, 1966.
2. COURTIN, P., and ROOTENBERG, J., *Performance Index Sensitivity of Optimal Control Systems,* IEEE Transactions on Automatic Control, Vol. 16, pp. 275–277, 1971.
3. KREINDLER, E., *On Performance Sensitivity of Optimal Control Problems,* International Journal of Control, Vol. 15, pp. 481–486, 1972.
4. ROOTENBERG, J., and COURTIN, P., *Sensitivity of Optimal Control Systems with Bang-Bang Control,* International Journal of Control, Vol. 18, pp. 537–543, 1973.
5. PETERSON, D. W., *On Sensitivity in Optimal Control Problems,* Journal of Optimization Theory and Applications, Vol. 13, pp. 56–73, 1974.