

## OPTIMAL CONTROL WITH A COST OF SWITCHING CONTROL

JOHN M. BLATT

(Received 28 January 1976)

(Revised 16 July 1976)

### Abstract

The Pontryagin theory of optimal control is modified by assuming a positive cost associated with switching control from one discrete value to another. The resulting new theory permits a general existence theorem. Pontryagin's maximum principle is replaced by an "indifference principle".

### 1. Introduction

In many practical servo-mechanisms, the control mechanism has only two settings, "off" and "on", but is used to control a process described by a function  $x(t)$  which can take a continuous range of real values.

Consider the following optimal control problem:

Minimize

$$x_0(T) = \int_0^T [(x(t) - \xi)^2 + cv(t)] dt \quad (1.1)$$

subject to

$$\frac{dx}{dt} = v - x \quad (1.2)$$

$$x(0) = 0 \quad (1.3)$$

$$v = 0, 1 \text{ only,} \quad (1.4)$$

$$0 < \xi < 1, 0 < c \text{ and } T > 0 \text{ are given constants} \quad (1.5)$$

Standard optimal control theory [1] applied to this problem gives no result at all for a large class of reasonable cases. To be precise: if

$$\hat{x} = \xi - \frac{1}{2}c \quad (1.6)$$

is positive, and if the planning horizon  $T$  is such that

$$T > \log \left[ \frac{\xi + (c(\xi - \frac{1}{4}c))^{\frac{1}{2}}}{\hat{x}(1 - \hat{x})} \right] \quad (1.7)$$

then *no piecewise continuous optimal control function  $v(t)$  exists*. The “optimal policy”, over a non-zero fraction of the time  $T$ , consists in switching the control “on” and “off” infinitely rapidly, so as to simulate a continuously-variable servo-mechanism kept at the constant intermediate setting  $v = \hat{x}$ .

One way out (the way taken by many mathematical control theorists) is to admit any *measurable* function  $v(t)$ . In our opinion, this amounts to removing the mathematical model from all contact with reality.

A more acceptable way out is provided by the following observation. In practice, there is always some *cost associated with a switch of control*. The cost may be as trivial as wear and tear on the switching mechanism, or it may be as complex as the cost of “loss of confidence in a stop-go national economy”. But some cost exists in all cases.

We therefore associate a positive cost  $\gamma > 0$  with each switch of control. We admit only piecewise continuous controls  $v(t)$ , which (by (1.4)) means *piecewise constant functions  $v(t)$* . A *control policy  $P$*  is defined completely by the set

$$P = \{v(0); k; t_1, t_2, \dots, t_k\} \quad (1.8)$$

where  $v(0) = 0$  or  $1$  is the control setting at time  $t = 0$ ; the non-negative integer  $k$  is the number of times control alters during the planning horizon  $T$ ; and the  $t_i$  are the times at which control alters, satisfying

$$t_0 \equiv 0 < t_1 < t_2 < \dots < t_k < T \equiv t_{k+1} \quad (1.9)$$

Given the policy  $P$ , the control function  $v(t)$  is:

$$v(t) = \frac{1}{2} + (-1)^m [v(0) - \frac{1}{2}] \quad \text{for } t_m < t \leq t_{m+1} \quad (1.10)$$

and  $m = 0, 1, 2, \dots, k$

In order to keep to essentials, we shall phrase the discussion in terms of a system with a scalar state variable  $x(t)$ , and with a two-valued scalar control variable  $v(t) = 0$  or  $1$ , only. Neither restriction is essential, and the generalizations are discussed in section 4.

Let us now state our control problem. Let  $\gamma > 0$ ,  $T > 0$ ,  $x(0)$ ,  $\hat{\alpha}$ ,  $\hat{\beta}$ ,  $\hat{\gamma}$ ,  $\hat{\delta}$ , and  $M > 0$  be given constants, and let  $f(x, v, t)$  and  $f_0(x, v, t)$  be real-valued functions defined on the set  $R$ :

$$R: \hat{\alpha} \leq x \leq \hat{\delta}; \quad v = 0, 1 \text{ only}; \quad 0 \leq t \leq T \quad (1.11a)$$

with the properties

$$f, f_0, \frac{\partial f}{\partial x}, \frac{\partial f_0}{\partial x}, \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f_0}{\partial x^2} \text{ exist everywhere and are} \\ \text{continuous on } R \quad (1.11b)$$

$$\frac{\partial f}{\partial t}, \frac{\partial f_0}{\partial t} \text{ exist everywhere and are continuous on } R \quad (1.11c)$$

$$|f(x, v, t)| \leq M \text{ everywhere inside } R \quad (1.11d)$$

$$\hat{\alpha} < \hat{\beta} \leq x(0) \leq \hat{\gamma} < \hat{\delta} \quad (1.11e)$$

$$f(x, v, t) > 0 \text{ for } \hat{\alpha} \leq x \leq \hat{\beta}; \quad v = 0, 1; \quad \text{and } 0 \leq t \leq T \quad (1.11f)$$

$$f(x, v, t) < 0 \text{ for } \hat{\gamma} \leq x \leq \hat{\delta}; \quad v = 0, 1; \quad \text{and } 0 \leq t \leq T \quad (1.11g)$$

We are to minimize

$$J = x_0(T) + k\gamma = \int_0^T f_0[x(t), v(t), t] dt + k\gamma \quad (1.12)$$

subject to the given value  $x(0)$  at time  $t = 0$ , and subject to

$$\frac{dx}{dt} = f[x(t), v(t), t] \quad (1.13)$$

$$v(t) = \text{piecewise continuous, as given by (1.10)} \quad (1.14)$$

$$f_0(x, v, t) \geq 0 \text{ everywhere inside } R \quad (1.15)$$

This concludes the statement of the problem.

Although conditions (1.11) appear very restrictive, they are satisfied in most practical applications. For example, consider equations (1.1)–(1.5). Take  $\hat{\alpha} = -2$ ,  $\hat{\beta} = -1$ ,  $\hat{\gamma} = 2$ ,  $\hat{\delta} = 3$ . Then  $x(0) = 0$  satisfies (1.11e). Furthermore,  $f(x, v, t) = v - x$  clearly satisfies (1.11f) and (1.11g), and if we pick  $M = 3$ , say, then (1.11d) is satisfied also. (1.11b) and (1.11c) are satisfied trivially for  $f = v - x$  and  $f_0 = (x - \xi)^2 + cv$ , and so is (1.15). The linear nature of  $f$  and quadratic nature of  $f_0$  are *not* essential.

The main purpose of conditions (1.11) is to ensure that a solution of (1.13) exists and the solution trajectory  $(x(t), v(t), t)$  for  $0 \leq t \leq T$  remains entirely within the region  $R$ . As we shall see shortly, this is always true, no matter how big  $T$  gets.

Although conditions (1.11f)–(1.11g) achieve the purpose stated in the preceding paragraph, the theory also goes through with other less restrictive conditions—the essential thing is that solutions of (1.13) must exist for all controls (1.8)–(1.10), and must remain within some compact region  $R$ .

If we are prepared to just *postulate* existence of a solution (as Pontryagin [1] does), then the conditions can be relaxed even more. But we prefer to give a general and constructive proof of the existence of an optimal solution of our control problem.

Nothing whatever is gained by admitting all measurable control functions  $v(t)$ , rather than just the functions (1.10). We shall show that optimal solutions exist within the piecewise continuous functions (1.10). There is no point in using a sledgehammer to kill a fly, or to use Lebesgue theory for functions which are piecewise constant, with a finite number of points of discontinuity!

The problem (1.10)–(1.15) reduces to a standard Pontryagin problem [1] for  $\gamma = 0$ , a limiting case which we *exclude* explicitly by the condition  $\gamma > 0$ . However we shall need much of Pontryagin's *notation*. We recall that the Pontryagin discussion uses the *Hamiltonian*

$$H(\lambda, x, v, t) = \lambda(t)f(x, v, t) - f_0(x, v, t) \quad (1.16)$$

and the *co-state equation*

$$\frac{d\lambda}{dt} = -\frac{\partial H}{\partial x} = \frac{\partial f_0}{\partial x} - \lambda \frac{\partial f}{\partial x} \quad (1.17)$$

with the *end-point condition*

$$\lambda(T) = 0 \quad (1.18)$$

The Pontryagin optimal policy (if it exists!) consists of maximizing  $H$  with respect to  $v$ , i.e., choose  $v^*(t)$  such that at each time  $t$ , and for both permissible values  $v = 0, 1$ , it is true that

$$H(\lambda, x, v^*, t) \geq H(\lambda, x, v, t) \quad (1.19)$$

It is a consequence of (1.19) and the assumed continuity of  $f$  and  $f_0$  that the Pontryagin optimal control  $v^*(t)$  can switch between  $v^* = 0$  and  $v^* = 1$  only at times  $t'$  such that

$$H(\lambda, x, 1, t') = H(\lambda, x, 0, t') \quad (1.20)$$

Unless this condition holds, one or the other of  $v = 0$  and  $v = 1$  is definitely preferred and must be chosen.

We know that this Pontryagin problem may fail to have admissible

(piecewise constant control) solutions. We shall show that the new problem (1.10)–(1.15) with  $\gamma > 0$  *always* possesses admissible optimal solutions, and we shall characterize these solutions.

In section 2 we prove existence. Section 3 is devoted to theorems characterizing the nature of the optimal solution. Finally, section 4 is devoted to a brief discussion of extensions and generalizations of this theory. A short example is given in Section 5.

## 2. Existence of an optimal control

LEMMA 1. *Given any control policy  $P$  and corresponding control function  $v(t)$ , (1.10), the differential equation (1.13) has a solution  $x(t)$  with  $x(0)$  as specified; furthermore, for all  $t$  in  $0 \leq t \leq T$ , the points  $(x(t), v(t), t)$  lie inside the region  $R$ , (1.11a).*

PROOF. Let  $v(t)$  be given by (1.10); define  $g(x, t)$  by

$$g(x, t) = f(x, v(t), t) \quad (2.1)$$

The differential equation becomes

$$\frac{dx}{dt} = g(x, t) \quad (2.2)$$

Let

$$\Delta t = \text{Min} \left( \frac{\hat{\beta} - \hat{\alpha}}{M}, \frac{\hat{\delta} - \hat{\gamma}}{M} \right) \quad (2.3)$$

Then  $g(x, t)$  is bounded by (1.11d) within the region

$$\Delta R_1: \quad 0 \leq t \leq \Delta t; \quad x(0) - Mt \leq x \leq x(0) + Mt \quad (2.4)$$

We note that (2.3) and (1.11e) ensure that  $\hat{\alpha} \leq x \leq \hat{\delta}$  everywhere in  $\Delta R$ . Furthermore  $g(x, t)$  satisfies a Lipschitz condition

$$|g(x, t) - g(x', t)| \leq A |x - x'| \quad \text{within } \Delta R_1, \quad (2.5)$$

where

$$A = \max_{(R)} \left( \frac{\partial f}{\partial x} \right) \quad (2.6)$$

(We take the maximum over the bigger region  $R$ , for later convenience.) Hence (see Agnew [3], for example), the differential equation (2.2) has a solution which remains inside  $\Delta R_1$ . We shall now show that this solution can be extended from  $\Delta R_1$  to  $R$ .

First, we prove that the solution  $x(t)$  satisfies

$$\hat{\beta} \leq x(t) \leq \hat{\gamma} \quad \text{for } 0 \leq t \leq \Delta t \tag{2.7}$$

Suppose (2.7) is false, and let  $t_1$  be such that  $x(t_1) > \hat{\gamma}$ , say. Since  $x(0) < \hat{\gamma}$  by (1.11e) and  $x(t)$  is differentiable by (2.2), there exists a  $t_2$ ,  $0 < t_2 < t_1$ , such that  $x(t_2) = \hat{\gamma}$  and  $x(t) > \hat{\gamma}$  for  $t_2 < t < t_1$ . Now use the mean-value theorem of the differential calculus: there exists  $t_3$ ,  $t_2 < t_3 < t_1$ , such that

$$\left(\frac{dx}{dt}\right)_{t_3} = f(x(t_3), v(t_3), t_3) = \frac{x(t_1) - x(t_2)}{t_1 - t_2} > 0 \tag{2.8}$$

However, the point  $(x(t_3), v(t_3), t_3)$  is inside  $\Delta R$  and hence *a fortiori* inside  $R$ . Furthermore  $\hat{\gamma} < x(t_3) < \hat{\delta}$ . But then (2.8) contradicts our assumption (1.11g)!

Similarly, the assumption  $x(t_1) < \hat{\beta}$  leads to a contradiction with (1.11f). Thus (2.7) is established.

We can now prove the lemma itself. Consider the differential equation (2.2) for the  $n$ th interval,  $(n - 1)\Delta t \leq t \leq n\Delta t$ , where  $n = 1, 2, \dots$  up to the integral part of  $(1 + T/\Delta t)$ . We define  $t_n \equiv n\Delta t$  and the region  $\Delta R_n$ :

$$\Delta R_n: t_{n-1} \leq t \leq t_n; \quad x(t_{n-1}) - M(t - t_{n-1}) \leq x \leq x(t_{n-1}) + M(t - t_{n-1}) \tag{2.9}$$

If  $x(t_{n-1})$  satisfies  $\hat{\beta} \leq x(t_{n-1}) \leq \hat{\gamma}$ , then (2.3) and (2.9) ensure that  $\hat{\alpha} \leq x \leq \hat{\delta}$  everywhere inside  $\Delta R_n$ . Furthermore,  $g(x, t)$  satisfies the Lipschitz condition (2.5) within  $\Delta R_n$ , with the same  $A$ , (2.6). Thus, everything goes through by induction on  $n$ . The same argument which led to (2.7) shows that

$$\hat{\beta} \leq x(t_{n-1}) \leq \hat{\gamma} \quad \text{implies} \quad \hat{\beta} \leq x(t_n) \leq \hat{\gamma}$$

and the leading step of the induction ( $n = 1$ ) has already been established. Q.E.D.

**THEOREM 1.** *There exists an upper bound  $K$  on the number of switches of control ( $k$ ) in any optimal policy.*

**PROOF.** Consider the policy  $P_0: \{v(0) = 0, k = 0\}$ ; i.e., we set  $v(t) = 0$  for all  $0 \leq t \leq T$ . Solve the differential equation (1.13) with initial condition  $x(0)$  given; by lemma 1, this can be done.

Substitute the solution  $x(t)$  for policy  $P_0$  into (1.12). The integral exists [(1.11b) and (1.11c) ensure that  $f_0$  is bounded and continuous] and we get

$$J(P_0) = \int_0^T f_0[x(t), 0, t] dt + 0 \tag{2.10}$$

Now let  $P$  be any policy (1.8). The solution  $x(t)$  of (1.13) exists, by lemma 1. Furthermore, the inequality (1.15) leads to

$$J(P) = \int_0^T f_0[x(t), v(t), t] dt + k\gamma \cong k\gamma \tag{2.11}$$

But unless  $J(P) < J(P_0)$ , there is no point in even considering policy  $P$  as a candidate for optimality. Thus serious candidate policies  $P$  have  $k\gamma \cong J(P) \cong J(P_0)$  which implies

$$k \leq \frac{J(P_0)}{\gamma} \equiv K \tag{2.12}$$

This explicit upper bound on  $k$  establishes the theorem. Note that policies with higher values of  $k$  are *feasible*. But they are never *optimal*.

**THEOREM 2.** *An admissible optimal policy exists.*

**PROOF.** The candidate policies  $P$ , (1.8), form a set. We can give this set a topology as follows: Two policies  $P$  and  $P'$  are  $\epsilon$ -neighbours of each other if

- i)  $k = k'$
- ii)  $|t_1 - t'_1| + |t_2 - t'_2| + \dots + |t_k - t'_k| < \epsilon$

With this definition, the policies  $P$  belong to a topological space  $\pi$ . Furthermore, by theorem 1,  $k$  is bounded, and all the  $t_i$  are bounded ( $0 \leq t_i \leq T$ ). Therefore  $\pi$  is a *compact* space (Heine–Borel theorem).

The cost  $J(P)$  is a *continuous* function from  $\pi$  to the real numbers, and is *bounded below* by  $J(P) \geq 0$ , all  $P$ .

Therefore  $J$  has a minimum and that minimum is attained for at least one policy  $P$  in the policy space. Q.E.D.

### 3. Characterization of optimal controls

Let  $k(\gamma)$  be the value of  $k$  for an optimal policy, for a given switching cost  $\gamma$ . Let  $J_\gamma(P)$  be the cost (1.12) of a policy  $P$ , for this switching cost, and  $J^*(\gamma)$  be the optimal cost.

**THEOREM 3.**  *$J^*(\gamma)$  is a non-decreasing function of  $\gamma$ ;  $k(\gamma)$  is a non-increasing function of  $\gamma$ , and becomes 0 for large enough  $\gamma$ .*

**PROOF.** Let policy  $P$ , with  $k$  switches of control, be optimal for switching cost  $\gamma$ ; and let policy  $P'$ , with  $k'$  switches of control, be optimal for switching cost  $\gamma'$ . Let  $\gamma > \gamma'$ .

Consider policy  $P$  for switching cost  $\gamma'$ . Its cost is, by (1.12)

$$J_{\gamma'}(P) = J_\gamma(P) + (\gamma' - \gamma)k \geq J_{\gamma'}(P') \tag{3.1}$$

which we rewrite in the form

$$k \leq [J_\gamma(P) - J_{\gamma'}(P')]/[\gamma - \gamma'] \tag{3.2}$$

Since  $\gamma > \gamma'$  by assumption, and  $k \geq 0$ , it follows that  $J_\gamma(P) = J^*(\gamma)$  is no smaller than  $J_{\gamma'}(P') = J^*(\gamma')$ . This proves the first assertion of the theorem.

Next, consider policy  $P'$  for switching cost  $\gamma$ . Its cost must exceed  $J_\gamma(P)$ , hence we get

$$J_\gamma(P') = J_{\gamma'}(P') + (\gamma - \gamma')k' \geq J_\gamma(P)$$

which we rewrite as

$$k' \geq [J_\gamma(P) - J_{\gamma'}(P')]/[\gamma - \gamma'] \quad (3.3)$$

Combination of (3.2) and (3.3) proves  $k \leq k'$ , which is the second assertion of the theorem.

Finally, we recall the bound (2.12) on  $k$  established in theorem 1. Clearly, when  $\gamma$  exceeds  $J(P_0)$ , (2.12) implies  $k = 0$  as the only possible value. Q.E.D.

All these results are obvious intuitively. If the switching cost increases, the optimal policy will certainly not use *more* switches of control than before. And if it costs *very* much to switch from  $v = 0$  to  $v = 1$ , the optimal policy consists of not switching at all during the entire time  $T$ . We then choose between  $v(t) = 0$ , all  $t$ , and  $v(t) = 1$ , all  $t$ . One or the other (or perhaps both) are optimal, and that is that.

At this point we also note that the condition  $\gamma > 0$  is *essential*. When  $\gamma = 0$  we have the Pontryagin problem for which there is no existence proof (and for which an explicit counter-example can be constructed). For  $\gamma < 0$ , there is *never* a sensible solution: If someone pays you money to throw a switch, just stand there and throw it forward and back as fast as you can — the faster the better!

**THEOREM 4.** *If the  $\gamma = 0$  (Pontryagin) problem has an admissible optimal solution with  $k_0$  switches of control, then for any  $\gamma > 0$  it is true that  $k(\gamma) \leq k_0$ . Furthermore, if  $k(\gamma) = k_0$ , then the optimal policy for  $\gamma > 0$  equals the Pontryagin policy.*

**PROOF.** The first assertion follows from theorem 3. The second assertion is a consequence of the trivial observation that the switching cost  $k\gamma$  cancels out in a comparison of two policies with the same value of  $k$ .

**THEOREM 5.** *(The indifference principle.) Let  $P$ , (1.8), be an optimal policy. Let  $H$  and  $\lambda(t)$  be defined by (1.16), (1.17), and (1.18). Then at each switching time  $t_i$  of  $P$  ( $i = 1, 2, \dots, k$ ) the Hamiltonian  $H$  is indifferent to the choice of control, i.e.,*

$$H(\lambda(t_i), x(t_i), 1, t_i) = H(\lambda(t_i), x(t_i), 0, t_i) \quad (3.4)$$

PROOF. First compare policy  $P$ , (1.8), with another policy  $P'$  which has  $v'(0) = v(0)$ ,  $k' = k$ ,  $t'_i = t_i$  for  $i \neq j$ , but differs in the  $j$ th switching time:

$$t'_j = t_j - \varepsilon \quad 0 < \varepsilon < t_j - t_{j-1} \quad (3.5)$$

For the sake of definiteness, assume that  $v(t)$ , as determined by (1.10) from the optimal policy, is  $v(t) = 0$  for  $t_{i-1} < t \leq t_i$  and  $v(t) = 1$  for  $t_i < t \leq t_{i+1}$ . (The opposite case, switch of control from  $v = 1$  to  $v = 0$ , goes through with obvious alterations in appropriate places).

In reference [2], it is shown that the varied state function  $x'(t)$  satisfies, for small  $\varepsilon$ ,

$$x'(t) = x(t) + \varepsilon \zeta(t) + o(\varepsilon) \quad (3.6)$$

$$\zeta(t) = 0 \quad \text{for } 0 \leq t \leq t_j - \varepsilon \quad (3.7)$$

$$\zeta(t_j) = f[x(t_j), 1, t_j] - f[x(t_j), 0, t_j] \quad (3.8)$$

$$\frac{d\zeta}{dt} = \left( \frac{\partial f}{\partial x} \right) \zeta \quad \text{for } t_j < t \leq T \quad (3.9)$$

where  $\partial f / \partial x$  in the last equation is evaluated along the optimal trajectory  $(x(t), v(t), t)$ . We define

$$x_0(t) = \int_0^t f_0[x(t'), v(t'), t'] dt' \quad (3.10)$$

so that

$$\frac{dx_0}{dt} = f_0(x, v, t) \quad (3.11)$$

$$x_0(0) = 0 \quad (3.12)$$

By comparison of policies  $P$  and  $P'$  we obtain [2]

$$x'_0(t) = x_0(t) + \varepsilon \zeta_0(t) + o(\varepsilon) \quad (3.13)$$

$$\zeta_0(t) = 0 \quad \text{for } 0 \leq t \leq t_j - \varepsilon \quad (3.14)$$

$$\zeta_0(t_j) = f_0[x(t_j), 1, t_j] - f_0[x(t_j), 0, t_j] \quad (3.15)$$

$$\frac{d\zeta_0}{dt} = \left( \frac{\partial f_0}{\partial x} \right) \zeta_0 \quad \text{for } t_j < t \leq T \quad (3.16)$$

where  $\partial f_0 / \partial x$  is evaluated along the optimal trajectory.

Now define  $\lambda(t)$  by (1.17), (1.18) and let

$$Z(t) = \lambda(t)\zeta(t) - \zeta_0(t) \quad (3.17)$$

At  $t = T$ ; (1.18) gives

$$Z(T) = \lambda(T)\zeta(T) - \zeta_0(T) = -\zeta_0(T) \tag{3.18}$$

At this point, the argument differs slightly from the one for the Pontryagin theory (see reference [2]) since  $x_0(T)$  is *not* equal to the cost  $J$  which we wish to minimize. The relation is  $J = x_0(T) + k\gamma$ . However, the two policies  $P$  and  $P'$  have the *same* value of  $k$ . Thus the difference in cost satisfies

$$J(P') - J(P) = x'_0(T) - x_0(T) \tag{3.19}$$

and this difference must be non-negative since  $P$  is optimal by assumption. Using (3.13), we conclude  $\zeta_0(T) \geq 0$ , and hence (3.18) gives  $Z(T) \leq 0$ .

We now combine (1.17), (3.9), (3.16), and (3.17) to deduce

$$\frac{dZ}{dt} = 0 \quad \text{for } t, < t \leq T \tag{3.20}$$

so that  $Z(t_j) = Z(T) \leq 0$ . We use (1.16), (3.8), (3.15) and (3.17) to express this inequality in the form:

$$H[\lambda(t_j), x(t_j), 0, t_j] \geq H[\lambda(t_j), x(t_j), 1, t_j] \tag{3.21}$$

The next step in the argument employs a different (though related) variation of the control function  $v(t)$ . The varied policy  $P''$  has  $v''(0) = v(0)$ ,  $k'' = k$ ,  $t''_i = t_i$  for  $i \neq j$ , but now the  $j$ th switching time  $t''_j$  is larger than its optimal value:

$$t''_j = t_j + \varepsilon \quad 0 < \varepsilon < t_{j+1} - t_j \tag{3.22}$$

As before, assume the optimal  $v(t)$  is 0 for  $t \leq t_j$ , and 1 for  $t > t_j$ . We get, in place of (3.6)–(3.8):

$$x''(t) = x(t) + \varepsilon \hat{\zeta}(t) + o(\varepsilon) \tag{3.23}$$

$$\hat{\zeta}(t) = 0 \quad \text{for } 0 \leq t \leq t_j \tag{3.24}$$

$$\hat{\zeta}(t_j + \varepsilon) = f[x(t_j + \varepsilon), 0, t_j + \varepsilon] - f[x(t_j + \varepsilon), 1, t_j + \varepsilon] \tag{3.25}$$

Equation (3.9) now holds for  $t > t_j + \varepsilon$ . (3.10)–(3.12) are unchanged, (3.13) to (3.15) become

$$x''_0(t) = x_0(t) + \varepsilon \hat{\zeta}_0(t) + o(\varepsilon) \tag{3.26}$$

$$\hat{\zeta}_0(t) = 0 \quad \text{for } 0 \leq t \leq t_j \tag{3.27}$$

$$\hat{\zeta}_0(t_j + \varepsilon) = f_0[x(t_j + \varepsilon), 0, t_j + \varepsilon] - f_0[x(t_j + \varepsilon), 1, t_j + \varepsilon] \tag{3.28}$$

Equation (3.16) now holds for  $t \geq t_j + \varepsilon$ ; equations (3.17)–(3.21) are unaltered, but instead of (3.21) we get:

$$H[\lambda(t + \varepsilon), x(t + \varepsilon), 1, t + \varepsilon] \geq H[\lambda(t + \varepsilon), x(t + \varepsilon), 0, t + \varepsilon] \quad (3.29)$$

At this stage, we use a continuity argument on (3.29) to allow us to combine it with (3.21).  $\lambda(t)$  and  $x(t)$  satisfy first-order differential equations, and are therefore continuous functions of time. By assumptions (1.11b) and (1.11c) the functions  $f(x, v, t)$  and  $f_0(x, v, t)$  are differentiable functions of  $t$  and of  $x$ . These properties imply, for constant  $v$ :

$$\lim_{\varepsilon \rightarrow 0} H[\lambda(t + \varepsilon), x(t + \varepsilon), v, t + \varepsilon] = H[\lambda(t), x(t), v, t] \quad (3.30)$$

We take this limit on both sides of (3.29) to get

$$H[\lambda(t), x(t), 1, t] \geq H[\lambda(t), x(t), 0, t] \quad (3.31)$$

This inequality can be combined directly with (3.21) to yield the *indifference principle* (3.4). This completes the proof of theorem 5.

We note that the indifference principle is a weaker condition than Pontryagin's maximum principle. The maximum principle implies the indifference principle (see equations (1.19) and (1.20)), but not vice versa. The point is that the maximum principle of the Pontryagin theory *forces* us to switch control when the phase space orbit crosses the indifference curve (3.4), whereas the indifference principle merely *allows* us to switch control at such a point. We are permitted to continue on with the same control as before, into a region of phase space where, according to Pontryagin, this control is no longer optimal.

However, once we have made this decision, we are not allowed to depart from it (i.e., not allowed to alter the current value of  $v$ ) until such time as we again reach the indifference curve (3.4).

Although the new theory requires exploration of many more candidate optimal paths than the Pontryagin theory, this does not make it unworkable. And, unlike the Pontryagin theory, our theorem 2 ensures that an admissible optimal policy always exists.

There is no relation between the indifference principle and what is called "singular control" in the Pontryagin theory. In singular control, the Pontryagin condition  $H = \text{maximum}$  fails to determine a unique value of the optimal control  $v^*$ , because  $H$  is independent of  $v$  (at least for some range of values of  $v$ ). When a cost of switching exists, and hence the maximum principle must be replaced by the indifference principle, then  $v^*$  is no longer determined uniquely by  $H$ , even when  $H$  has a unique maximum.

Before going on, we note that one of Pontryagin's theorems carries over without change (even though the proof requires slight alterations):

**THEOREM 6.** *If the system is autonomous (i.e.,  $f(x, v, t)$  and  $f_0(x, v, t)$  do not depend upon  $t$  explicitly), then the value of  $H$  does not change with time:  $dH/dt = 0$ .*

**PROOF.** (a) At times  $t$  when the system is *not* on the indifference curve (3.4),  $v(t)$  is constant (either 0 or 1, it does not matter), and straightforward differentiation using (1.12), (1.16), and (1.17) gives

$$\frac{dH}{dt} = 0 \quad \text{for } t \neq t_i \quad i = 1, 2, \dots, k \quad (3.32)$$

(b) At an indifference point, control may or may not switch between  $v = 0$  and  $v = 1$ . If control *does not* switch, then the argument leading to (3.32) is still valid. If control *does* switch, we have  $dH/dt = 0$  for  $t < t_i$  and for  $t > t_i$  (by (3.32)), but we also have

$$\lim_{t \rightarrow t_i^-} H = \lim_{t \rightarrow t_i^+} H \quad (3.33)$$

by the indifference property (3.4) and continuity of  $H$ . Thus  $H$  remains constant as we cross the point  $t = t_i$ . Q.E.D.

**THEOREM 7.** *If the system is “amenable to control” and has only one degree of freedom, then the indifference curve (3.4) in the phase plane contains at most one value of  $\lambda$  for every  $x$  and  $t$ , i.e.,  $\lambda$  is a function of  $x$  and  $t$ .*

**PROOF.** We use (1.16) to write (3.4) as

$$\lambda f(x, 0, t) - f_0(x, 0, t) = \lambda f(x, 1, t) - f_0(x, 1, t) \quad (3.34)$$

We define a state  $x$  at time  $t$  to be “amenable to control” provided that

$$f(x, 0, t) \neq f(x, 1, t) \quad (3.35)$$

If this condition is violated, then the time development of the system (at this time) is not influenced at all by the choice of control variable, i.e.,  $dx/dt$  is the same for  $v = 0$  as for  $v = 1$ .

Provided the state of the system is amenable to control, we can solve (3.34) for  $\lambda$  to get the *indifference curve*:

$$\lambda = \Lambda(x, t) \equiv \frac{f_0(x, 0, t) - f_0(x, 1, t)}{f(x, 0, t) - f(x, 1, t)} \quad (3.36)$$

This proves theorem 7 for points amenable to control. At a point which is not amenable to control,  $\Lambda(x, t)$  becomes infinite. We assume there are at most a finite number of such points in  $0 \leq t \leq T$ . We note that an *autonomous system* (no explicit dependence on  $t$ ) has a *time-independent indifference curve*  $\lambda = \Lambda(x)$  in the phase plane of  $\lambda$  versus  $x$ .

Furthermore, for an autonomous system, theorem 6 allows us to determine the actual orbits in the phase plane. Let the constant value of  $H$  be called  $E$ . Use (1.16) with  $H = E$  and solve for  $\lambda$  to obtain

$$\lambda = \frac{E + f_0(x, v)}{f(x, v)} \quad (3.37)$$

Thus at points for which  $f(x, v) \neq 0$ ,  $\lambda$  is determined as a function of  $x$ ,  $v$ , and the time-independent parameter  $E$ .

We have therefore arrived at two one-parameter families of orbits in the phase plane, one family for  $v = 0$ , the other for  $v = 1$ . At any moment, the system moves along one of these orbits, in a direction determined by the sign of  $f(x, v)$ . This continues until and unless the orbit (with its given  $v$  and  $E$ ) intersects the indifference curve (3.34). At each such intersection point, we have a choice: The optimal path may stay on the same orbital curve (i.e.,  $v$  in (3.37) may remain unaltered), or the optimal path may switch to the orbit (3.37) with the same  $E$  but the opposite value of  $v$ .

These specialized properties (based on theorem 7) are very powerful in simplifying the explicit solution of optimal control problems in this theory. Unfortunately, these special properties do not generalize *usefully* to systems with two or more degrees of freedom (i.e., systems in which  $x$  and  $\lambda$  are vectors with  $m \geq 2$  components). Generalizations of theorems 6 and 7 exist, and are stated in section 4. But these generalizations are less useful; in particular, the indifference principle (3.4) does not provide an explicit formula for the vector  $\lambda$  as a function of the vector  $x$ , and neither does conservation of  $E$  by itself give an explicit formula for the orbits in phase space. In both cases, we get restrictive conditions on  $\lambda$ , not explicit formulas for  $\lambda$ . In all other respects, however, the theory generalizes well; the orbits in state space and/or phase space exist and are determined implicitly (though no longer explicitly) by the theory.

#### 4. Some generalizations

##### (a) DISCRETE CONTROL WITH MORE THAN TWO VALUES.

Suppose the control set includes the values

$$v = 0, 1, 2, \dots, l \quad (4.1)$$

Piecewise constant control policies are then defined by sets:

$$P = \{v_0; k; t_1, v_1, t_2, v_2, \dots, t_k, v_k\} \quad (4.2)$$

with  $t_i$  satisfying (1.9) and the  $v_i$  members of the set (4.1). The control function  $v(t)$  takes the form

$$v(t) = v_i \quad \text{for} \quad t_i < t \leq t_{i+1} \quad (4.3)$$

where  $t_0 \equiv 0$  and  $t_{k+1} \equiv T$ . (4.2) replaces (1.8), (4.3) replaces (1.10).

The switching cost now becomes a *matrix*  $\gamma(\alpha, \beta)$  ( $\alpha, \beta = 0, 1, \dots, l$ ), where  $\gamma(\alpha, \beta)$  is the cost of switching from  $v = \alpha$  to  $v = \beta$ . It is not necessary to assume symmetry, i.e.,  $\gamma(\alpha, \beta)$  may differ from  $\gamma(\beta, \alpha)$ . The cost (1.12) is replaced by

$$J = x_0(T) + \sum_{i=1}^k \gamma(v_{i-1}, v_i) \quad (4.4)$$

The essential assumption for the generalization is a lower bound

$$\gamma(\alpha, \beta) \geq \gamma > 0 \quad (4.5)$$

on all switching costs. With such a lower bound, the existence theorem still holds.

Theorems 3 and 4 lose their meanings unless all  $\gamma(\alpha, \beta)$  are equal; in that case the theorems carry through.

Theorem 5, the indifference principle, needs to be restated slightly. Using the notation (4.2) and (4.3), the assertion (3.4) is replaced by

$$H[\lambda(t_i), x(t_i), v_i, t_i] = H[\lambda(t_i), x(t_i), v_{i-1}, t_i] \quad (4.6)$$

That is, the Hamiltonian is indifferent with respect to the change of control (from  $v_{i-1}$  to  $v_i$ ) which actually takes place at  $t = t_i$ ; it need not be indifferent with respect to all conceivable changes of control.

Theorem 6 survives, in theorem 7 we now have not just one indifference curve, but  $l(l+1)/2$  of them. They are given by

$$\Lambda_{\alpha\beta}(x, t) = \frac{f_0(x, \alpha, t) - f_0(x, \beta, t)}{f(x, \alpha, t) - f(x, \beta, t)} \quad (4.7)$$

with  $\alpha = 0, 1, 2, \dots, l$ ,  $\beta = \alpha + 1, \alpha + 2, \dots, l$ .

If the state variable  $x(t)$  is a *scalar* and the system is autonomous, constancy of  $H$  determines the permissible orbits and (4.7) gives a set of fixed curves in the  $\lambda - x$  phase plane. A switch of control from  $v = \alpha$  to  $v = \beta$ , or vice versa, can occur only if the phase point  $(x(t), \lambda(t))$  falls on the curve  $\lambda = \Lambda_{\alpha\beta}(x)$ . Thus, quite a bit of the theory goes through as before.

#### (b) VECTOR STATE VARIABLE.

Now let  $x(t)$  be a vector with components  $(x_1(t), x_2(t), \dots, x_n(t))$ ,  $n \geq 2$ . The differential equation (1.13) is replaced by the system

$$\frac{dx_i}{dt} = f_i[x(t), v(t), t] \quad i = 1, 2, \dots, n \quad (4.8)$$

We note that, as long as the control set is *discrete*, there is no point in making  $v$  a vector.

Conditions for ensuring the existence of a solution to the differential equations can be written down. We shall not do so here.

Thereafter, for the case  $v = 0, 1$ , only, theorems 1, 2, 3, 4, 5 and 6 go through with obvious minor modifications (in particular,  $\lambda(t)$  becomes a vector, also). Theorem 7 does *not* go through as it stands. Rather, the indifference condition yields:

$$\sum_{i=1}^n \lambda_i [f_i(\mathbf{x}, 0, t) - f_i(\mathbf{x}, 1, t)] = f_0(\mathbf{x}, 0, t) - f_0(\mathbf{x}, 1, t) \tag{4.9}$$

This is a linear relation between the components of the vector  $\lambda$ . It is *not* enough to determine  $\lambda$  explicitly (unless  $n = 1$ ). Similarly, conservation of the Hamiltonian (for an autonomous system) yields

$$\sum_{i=1}^n \lambda_i(t) f_i(\mathbf{x}(t), v(t)) = E + f_0(\mathbf{x}(t), v(t)) \tag{4.10}$$

This is also not enough to determine  $\lambda$  explicitly.

(c) DISCONTINUOUS  $f(x, v, t)$ .

The condition (1.11b) implies that  $f(x, v, t)$  and  $f_0(x, v, t)$  are continuous functions of the time variable  $t$ , for all  $0 \leq t \leq T$ . This is often an inconvenient restriction. We would like to be able to work with functions  $f$  and/or  $f_0$  which may jump at a finite number of values of  $t$ , i.e., which are piecewise  $C^1$  rather than  $C^1$  over the whole interval.

In the Pontryagin case ( $\gamma = 0$ ) this is permissible, as proved in reference [2].

Here, however, we must investigate the possibility that a time  $t = \theta$  at which  $f(x, v, t)$ , say, is discontinuous becomes equal to one of the switching times  $t_i$  of the optimal policy. The proof of the indifference principle relied heavily on continuity arguments, see (3.30). At first sight, therefore, it looks hopeless.

However, it *is* possible to relax the requirements somewhat. Let  $\theta$  be a point of discontinuity, and define jump sizes  $\Delta f$  and  $\Delta f_0$  by

$$\Delta f(x, v) = \lim_{\epsilon \rightarrow 0} [f(x, v, \theta + \epsilon) - f_0(x, v, \theta - \epsilon)] \tag{4.11}$$

$$\Delta f_0(x, v) = \lim_{\epsilon \rightarrow 0} [f_0(x, v, \theta + \epsilon) - f_0(x, v, \theta - \epsilon)] \tag{4.12}$$

They need not both be non-zero, but at least one of them is non-zero if  $\theta$  is a jump point.

The theorems still go through if  $f$  and  $f_0$  are *piecewise  $C^1$*  and if the *jump sizes are independent of the control setting*:

$$\Delta f(x, 0) = \Delta f(x, 1) \quad (4.13)$$

$$\Delta f_0(x, 0) = \Delta f_0(x, 1) \quad (4.14)$$

A detailed look at the proof of theorem 5 shows that this weaker condition suffices.

Other generalizations suggest themselves readily enough, but they are not discussed in this paper.

### 5. A specific example

Consider problem (1.1)–(1.5), with the numerical values  $\xi = 0.6$ ,  $c = 0.1$ ,  $T = 5.0$ . Equation (1.6) yields  $\hat{x} = 0.55$ , and the condition (1.7) becomes  $T > 1.2217$ , which is satisfied. Thus this Pontryagin problem has no optimal solution (one is led to a “chattering control”).

Now change the criterion from (1.1) to (1.12), with a cost  $\gamma = 0.4$  of every change of control.

The candidate optimal solutions corresponding to the policies (1.8) are pieced together from solutions of the differential equations (1.2)' and (1.17), with piecewise constant  $v$ . Let  $v_m$  stand for the value of  $v(t)$  in the interval  $t_m < t \leq t_{m+1}$ , as defined by (1.10). Let  $A_m$  and  $B_m$  be constants,  $m = 0, 1, \dots, k$ . Then in the interval  $t_m < t \leq t_{m+1}$  we have:

$$x(t) = v_m + A_m \exp(t_m - t) \quad (5.1)$$

$$\lambda(t) = B_m \exp(t - t_m) - A_m \exp(t_m - t) - 2v_m + 2\xi \quad (5.2)$$

According to the indifference principle,  $\lambda(t)$  at  $t = t_m$  must be such that  $H$  is independent of  $v$ ; this happens for  $\lambda = c$ . These conditions become

$$\lim_{t \rightarrow t_m^+} \lambda(t) = B_m - A_m - 2v_m + 2\xi = c \quad (5.3)$$

$$\text{for } m = 1, 2, \dots, k - 1, k$$

and, using the notation  $z_m \equiv \exp(t_{m+1} - t_m)$ ,

$$\lambda(t_{m+1}) = \lim_{t \rightarrow t_{m+1}^-} \lambda(t) = B_m z_m - \frac{A_m}{z_m} - 2v_m + 2\xi = c \quad (5.4)$$

$$\text{for } m = 0, 1, 2, \dots, k - 1$$

Furthermore,  $x(t)$  must be continuous at  $t = t_{m+1}$ , giving the conditions:

$$v_m + A_m z_m = v_{m+1} + A_{m+1} \quad \text{for } m = 0, 1, 2, \dots, k - 1 \quad (5.5)$$

So far, we have  $3k$  equations correcting the  $3k + 2$  unknowns  $A_m, B_m$  (with  $m = 0, 1, \dots, k$ ) and  $z_m$  (with  $m = 0, 1, \dots, k - 1$ ). The remaining two equations are obtained from the initial condition  $x(0) = 0$  and the "free final state" condition  $\lambda(T) = 0$ . Thus, there are enough equations to specify at most one candidate optimal solution for each choice of  $v(0)$  and  $k$  in (1.8). Evaluation of (1.12) is straightforward, and we know that the optimal solution can be obtained by finite enumeration, i.e., there is an upper bound on  $k$ .

For the numerical values  $\xi = 0.6$ ,  $c = 0.1$ ,  $T = 5.0$ , and  $\gamma = 0.4$ , the optimal policy is

$$v(0) = 1, \quad k = 0, \quad (5.6)$$

that is,  $v(t) = 1$  for all  $0 \leq t \leq T$ . This is not a Pontryagin policy: near  $t = T$ , the final condition  $\lambda(T) = 0$  together with continuity of  $\lambda(t)$  ensure that  $\lambda(t) < c$ , hence every Pontryagin policy has  $v^*(t) = 0$  near  $t = T$ .

In general, for given  $\xi, c, T$  and for sufficiently high switching cost  $\gamma$  (the minimum value depends on  $\xi, c$ , and  $T$ ) it is clear that any switching of control becomes suboptimal. There then remains a choice between only two candidate policies,  $v(0) = 0$  or  $v(0) = 1$ , both with  $k = 0$ . For  $\hat{x} > 0.5$ ,  $v(0) = 1$  turns out to be optimal, for  $\hat{x} < 0.5$ ,  $v(0) = 0$  is optimal. It may be surprising at first sight that the comparatively small switching cost  $\gamma = 0.4$  suffices to drive us into this "no-switching" regime for a planning time  $T = 5.0$ , which so greatly exceeds the critical time  $T_c = 1.2217$  at which all Pontryagin policies break down into chattering control. We state, without giving the proof, that  $\gamma = 0.4$  is large enough to prevent all switching for  $\xi = 0.6$  and  $c = 0.1$ , no matter how large  $T$  is (i.e., even in the limit  $T \rightarrow \infty$ ).

### Acknowledgements

The author is grateful to Prof. Murray Kemp and to Drs. J. D. Gray and K. L. Teo for valuable discussions.

### References

- [1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gramkrelidze and E. F. Mischenko, *The Mathematical Theory of Optimal Processes*, New York: Interscience Publishers, Inc., (1962).
- [2] John M. Blatt and J. D. Gray, to be published.
- [3] R. P. Agnew, *Differential Equations*, (McGraw-Hill Book Co., New York and London, 1942).

Department of Applied Mathematics,  
University of New South Wales,  
Kensington,  
Australia.