

OPTIMAL CONTROL PROBLEMS WITH ELASTIC COLLISIONS

J. M. MURRAY¹

(Received 8 March 1988; revised 3 July 1988)

Abstract

In this paper we consider optimal control problems with linear state constraints where the states can be discontinuous at the boundary. In fact the state vector models the position and velocity of a particle where the collisions with the boundary that cause the discontinuities are elastic. Necessary conditions are derived by looking at limits of approximate problems that are unconstrained.

1. Introduction

There are many instances when problems fit into the setting of optimal control theory except that they allow discontinuities in the state variables. One can find such examples in areas of application from fisheries [2] to rockets [7]. In this paper we study constrained physical problems where discontinuities occur in the velocity component of the state when the mass in question collides with the boundary. An added condition that must be satisfied is that the collisions are elastic so that the proper reflection law must be obeyed. An example of this type of problem can be found in [6].

We derive optimality conditions for the problem by using a limiting procedure involving unconstrained problems with a forcing term. These unconstrained problems fit into the usual optimal control framework and their optimality conditions are already known. The limit of these optimality conditions, as the forcing term becomes larger and larger when the original constraints are violated, leads to optimality conditions for the original problem.

¹School of Mathematics, University of N.S.W., P.O. Box 1, Kensington, N.S.W., 2033, Australia.
© Copyright Australian Mathematical Society 1989, Serial-fee code 0334-2700/89

2. The problem and its approximations

In this paper we shall investigate the behaviour of optimal solutions of the following problem, where an object can undergo elastic collisions with boundaries:

$$\begin{aligned}
 & \min f_0(x(t_1), y(t_1)) \\
 & \text{subject to } \dot{x}(t) = y(t) \\
 & \quad \dot{y}(t) = A(t)x(t) + B(t)y(t) + C(t)u(t) + a(t) \\
 & \quad x(t_0) = x_0, \quad y(t_0) = y_0, \quad x, y \in R^n \\
 & \quad u(t) \in U(t) \quad u \in R^m \\
 & \quad x(t) \in X = \{x: D^T x - d \leq 0\} \quad D \in R^{n \times k}, d \in R^k
 \end{aligned} \tag{P_0}$$

and if $x(t)$ is on the boundary of X with

$$\begin{aligned}
 D_i^T x(t) - d_i &= 0 & D_i \text{ is the } i\text{th column of } D \\
 D_j^T x(t) - d_j &< 0 & j \neq i
 \end{aligned}$$

then $|y(t^-)| = |y(t^+)|$ and $y(t^+) = y(t^-) - \lambda D_i$ for some $\lambda \geq 0$, t_0 and t_1 fixed.

The above minimisation is over all feasible controls, in other words over all functions u that satisfy the constraints and where u is measurable.

This investigation will be carried out by using a penalty approach to convert (P_0) into approximate problems for which optimality results are available. The approximate problems, denoted (P_ε) with $\varepsilon > 0$ being a small parameter, are as follows.

$$\begin{aligned}
 & \min f_0(x(t_1), y(t_1)) \\
 & \text{subject to } \dot{x}(t) = y(t) \\
 & \quad \dot{y}(t) = A(t)x(t) + B(t)y(t) + C(t)u(t) + a(t) \\
 & \quad - \begin{cases} 0 & \text{if } x(t) \in X \\ (\alpha_i/\varepsilon)D_i(D_i^T x - d_i) & \text{if } D_i^T x - d_i > 0 \end{cases}
 \end{aligned} \tag{P_\varepsilon}$$

where $x(t_0) = x_0$, $y(t_0) = y_0$, $u(t) \in U(t)$, and $\alpha_i > 0$, $i = 1, \dots, k$, t_0 and t_1 fixed.

The following will be assumed to hold throughout this paper.

General assumptions. (i) The functions f_0 , $\partial f_0/\partial x$ and $\partial f_0/\partial y$ are continuous.

(ii) The $n \times n$ matrix function A , the $n \times m$ matrix function C , and the n vector function a are measurable. The $n \times n$ matrix function B is differentiable and the components of A , B , C and a are bounded on $[t_0, t_1]$.

(iii) $U(t)$ is compact and convex for all $t \in [t_0, t_1]$ and bounded on $[t_0, t_1]$.

(iv) The graph of U is $\mathcal{L} \times \mathcal{B}$ measurable (see 4.1 of [3])

Since the final value of y may be used in the calculation of the objective value for (P_0) and it is feasible for y to be discontinuous, we must concern ourselves with the case where it is possible for y to be discontinuous at t_1 . What value do we then use for $y(t_1)$: do we assign it the value $y(t_1^-)$ or $y(t_1^+)$? Either case, specially when dealing with approximate solutions, has its advantages and disadvantages, so we take the easy way out and assume that the optimal solutions for (P_0) and (P_ϵ) are such that this will not be a worry.

Assumption 1. Either f_0 is independent of y or there exists a $\delta > 0$ and an $\bar{\epsilon} > 0$ such that if (x_ϵ, y_ϵ) is an optimal solution for (P_ϵ) and (x^*, y^*) is an optimal solution for (P_0) , then $x_\epsilon(t_1) \in X$ and $x_\epsilon(t_1)$ is at least δ distance away from any boundary, for all ϵ such that $0 < \epsilon < \bar{\epsilon}$, and likewise for $x^*(t_1)$.

After all, if the optimal solution (x^*, y^*) did not satisfy this assumption, we could change the final time t_1 to either just before t_1 or just after t_1 , depending on whether we wanted $y^*(t_1^-)$ or $y^*(t_1^+)$, so that the assumption was satisfied.

Assumption 2. To simplify matters we assume that the object on its optimal trajectory only hits one boundary at a time and at a finite number of isolated instants of time in (P_0) .

This also guarantees, at least for ϵ small enough, that we only need consider the case where only one constraint at a time is infeasible for (P_ϵ) ; that is, at any time t , there is only one i such that $D_i^T x(t) - d_i > 0$.

We will derive optimality conditions for (P_0) by looking at the limit of the problem (P_ϵ) .

3. The limit of optimal solutions of (P_ϵ)

DEFINITION. If g is a vector function of bounded variation then there exists a one dimensional, nonnegative regular Borel measure θ such that

$$dg(t) = \dot{g}(t) dt + \gamma(t) d\theta(t)$$

for some Borel measurable function γ . We can then define the following norm

$$\|g\|_v = \int_{t_0}^{t_1} |\dot{g}(t)| dt + \int_{t_0}^{t_1} |\gamma(t)| d\theta(t)$$

whose value is independent of the choice of θ and γ .

Suppose $\{z_\epsilon\}$ is a generalised sequence in the space of functions of bounded variation on $[t_0, t_1]$, and where $z_\epsilon(t) \in R^n$ and $z_\epsilon(t) = z_0$ for all ϵ . If z^* is the

weak* limit of $\{z_\varepsilon\}$ with $z^*(t_0) = z_0$ then this means that

$$\lim_{\varepsilon \rightarrow 0} \int_{t_0}^{t_1} p(t) \cdot dz_\varepsilon(t) = \int_{t_0}^{t_1} p(t) \cdot dz^*(t)$$

for all $p \in C_n[t_0, t_1]$, the space of continuous n -vector functions on $[t_0, t_1]$.

Assumption 3. Each (P_ε) has an optimal solution $(x_\varepsilon, y_\varepsilon)$ and $\exists M > 0$ independent of ε such that

$$\|(x_\varepsilon, y_\varepsilon)\|_v \leq M.$$

The set $V = \{(x, y) : \|(x, y)\|_v \leq M\}$ is metrisable in the weak* topology. So if we are interested in the limit points of the generalised sequence $\{(x_\varepsilon, y_\varepsilon)\}$ it suffices to look at sequences contained within the generalised sequence. By Theorem 33 of [4] we have that a subsequence of this sequence converges to a pair $(x^*, y^*) \in V$. Using Theorems 32 and 34 of [4], and by extracting a further subsequence (which we shall still call $\{(x_\varepsilon, y_\varepsilon)\}$) if necessary, we have that

$$(x_\varepsilon(t), y_\varepsilon(t)) \rightarrow (x^*(t), y^*(t)) \text{ a.e.}$$

As implied by the assumption stated earlier, we only consider the case when x^* hits no corners of the boundary. Our first task is to show that (x^*, y^*) is feasible for (P_0) .

PROPOSITION 1. $x^*(t)$ belongs to X for all $t \in [t_0, t_1]$

PROOF. Since $\|(x_\varepsilon, y_\varepsilon)\|_v \leq M$ we have that $|\dot{x}_\varepsilon(t)| \leq M$ and hence $|x_\varepsilon(t)| \leq x_0 + \int_0^t |\dot{x}_\varepsilon(\tau)| d\tau \leq x_0 + M$ for all t and ε . Therefore $\{x_\varepsilon\}$ is bounded and equicontinuous, so there is a subsequence once again called $\{x_\varepsilon\}$ converging uniformly to a limit x . It also converges weak* to x^* . Let us show that these are the same.

We also have a subsequence of $\{\dot{x}_\varepsilon\}$ converging weakly to a limit v . That is, $\lim_\varepsilon \int_{t_0}^{t_1} s(t) \dot{x}_\varepsilon(t) dt = \int_{t_0}^{t_1} s(t) v(t) dt \forall s \in L_n^1[t_0, t_1]$. Since $x_\varepsilon(t) = x_0 + \int_{t_0}^t \dot{x}_\varepsilon(\tau) d\tau \rightarrow x_0 + \int_{t_0}^t v(\tau) d\tau$, we see that $x(t) = x_0 + \int_{t_0}^t v(\tau) d\tau$ and

$$\dot{x}(t) = v(t) \text{ a.e.}$$

We must also have then $\int_{t_0}^{t_1} p(t)(\dot{x}(t) - \dot{x}^*(t)) dt = 0 \forall p \in C_n[t_0, t_1]$. Hence $x = x^*$.

If $x^*(t)$ does not belong to X for all t , there is an interval (a, b) , a number $\delta > 0$ and some i such that $D_i^T x^*(t) - d_i > \delta \forall t \in (a, b)$. We have that

$$\begin{aligned} |\dot{y}_\varepsilon(t)| &\geq (\alpha_i/\varepsilon) |D_i(D_i^T x_\varepsilon(t) - d_i)| - |A(t)x_\varepsilon(t) + B(t)y_\varepsilon(t) + C(t)u_\varepsilon(t) + a(t)| \\ &\geq (\alpha_i/\varepsilon) |D_i(D_i^T x_\varepsilon(t) - d_i)| - N \end{aligned}$$

for some $N \geq 0$ independent of ε (because $\{x_\varepsilon\}$, $\{y_\varepsilon\}$ and $\{u_\varepsilon\}$ are bounded). Since $x_\varepsilon(t)$ converges to $x^*(t)$ for all t , then we must have for ε sufficiently small

that $|\dot{y}_\varepsilon(t)| > (\delta\alpha_i/2\varepsilon)|D_i| - N$ and $\|y_\varepsilon\|_v > (\delta\alpha_i/2\varepsilon)|D_i|(b-a) - N(b-a) > M$, which would be in violation of Assumption 3. Therefore $x^*(t) \in X$ for all $t \in [t_0, t_1]$.

PROPOSITION 2. (x^*, y^*) satisfies the differential equation $\dot{x}(t) = y(t)$.

PROOF. We have that

$$\int_{t_0}^{t_1} p(t)\dot{y}_\varepsilon(t) dt \rightarrow \int_{t_0}^{t_1} p(t) dy^*(t) \quad \forall p \in C_n[t_0, t_1].$$

Hence $y(t_0) + \int_{t_0}^{t_1} \dot{y}_\varepsilon(t) dt \rightarrow y(t_0) + \int_{t_0}^{t_1} dy^*(t)$. In other words, $y_\varepsilon(t_1) \rightarrow y^*(t_1)$. By choosing the continuous functions p appropriately, we also obtain

$$\begin{aligned} p(t_1)y_\varepsilon(t_1) - p(t_0)y_\varepsilon(t_0) - \int_{t_0}^{t_1} \dot{p}(t)y_\varepsilon(t) dt \\ \rightarrow p(t_1)y^*(t_1) - p(t_0)y^*(t_0) - \int_{t_0}^{t_1} \dot{p}(t)y^*(t) dt. \end{aligned}$$

But $y_\varepsilon(t_0) = y^*(t_0) = y_0$ and $y_\varepsilon(t_1) \rightarrow y^*(t_1)$; so

$$\int_{t_0}^{t_1} \dot{p}(t)y_\varepsilon(t) dt \rightarrow \int_{t_0}^{t_1} \dot{p}(t)y^*(t) dt$$

for all differentiable functions p .

Now from before we have that $y_\varepsilon = \dot{x}_\varepsilon \xrightarrow{\text{weakly}} \dot{x}^*$ hence $\int_{t_0}^{t_1} p(t)\dot{x}_\varepsilon(t) dt \rightarrow \int_{t_0}^{t_1} p(t)\dot{x}^*(t) dt$ and from above, $\int_{t_0}^{t_1} p(t)y_\varepsilon(t) dt \rightarrow \int_{t_0}^{t_1} p(t)y^*(t) dt \forall p \in C_n[t_0, t_1]$. But the left hand sides are equivalent for each ε , and so must be the right hand sides. Hence $\dot{x}^*(t) = y^*(t)$ a.e.

We now show that y^* is feasible. Since $y_\varepsilon \rightarrow y^*$ in the weak* topology we also have that $y_\varepsilon(t) \rightarrow y^*(t)$ a.e. So the only way that y^* could be discontinuous at a point \bar{t} is if \dot{y}_ε becomes large inside some interval around \bar{t} . In other words, only if $x_\varepsilon(t)$ lies outside X for some t near \bar{t} . Let us look at such a situation.

The penalty term in the differential equation only has an effect normal to the constraint. By choosing $z(t) = (D_i^T x(t) - d_i)$ and $\zeta(t) = D_i^T y(t)$, the differential equations for (P_ε) become

$$\begin{aligned} \dot{z}(t) &= \zeta(t) \\ \dot{\zeta}(t) &= D_i^T [A(t)x(t) + B(t)y(t) + \Psi(t)u(t) + a(t)] - (\beta_i/\varepsilon)z(t) \end{aligned} \tag{D_\varepsilon}$$

where $\beta_i = \alpha_i D_i^T D_i$ and we only consider a time interval where $x(t)$ lies outside X . To that effect let $t = \tau_{0\varepsilon}$ be such that $z(\tau_{0\varepsilon}) = 0$ and $\zeta(\tau_{0\varepsilon}) = \zeta_{0\varepsilon}$. Let us look at another system of differential equations with the same initial conditions

$$\dot{z}(t) = \zeta(t), \quad \dot{\zeta}(t) = (\beta_i/\varepsilon)z(t). \tag{\bar{D}_\varepsilon}$$

The solution to (\bar{D}_ϵ) is

$$\begin{aligned} \bar{z}_\epsilon(t) &= \zeta_{0\epsilon} \sqrt{\epsilon/\beta_i} \sin \sqrt{\beta_i/\epsilon}(t - \tau_{0\epsilon}) \\ \bar{\zeta}_\epsilon(t) &= \zeta_{0\epsilon} \cos \sqrt{\beta_i/\epsilon}(t - \tau_{0\epsilon}). \end{aligned}$$

Now \bar{z}_ϵ becomes zero again when $t = \tau_{0\epsilon} + \pi\sqrt{\epsilon/\beta_i} = \tau_{1\epsilon}$ and at that time $\bar{\zeta}_\epsilon(\tau_{1\epsilon}) = -\zeta_{0\epsilon}$. As $\epsilon \rightarrow 0$, the interval $[\tau_{0\epsilon}, \tau_{1\epsilon}] \rightarrow \{\bar{\tau}_0\}$, $\bar{z}_\epsilon(t) \rightarrow 0$ and $\bar{\zeta}(\bar{\tau}^+) = -\bar{\zeta}(\bar{\tau}^-)$. How close is the solution to (D_ϵ) for any ϵ ?

Introduce new variables $w = \sqrt{\epsilon/\beta_i}\zeta$ and $\tau = \sqrt{\beta_i/\epsilon}(t - \tau_{0\epsilon})$. Then (D_ϵ) and (\bar{D}_ϵ) can now be written as

$$z' = w, \quad w' = (\epsilon/\beta_i)D_i^T[A(\tau)x(\tau) + B(\tau)y(\tau) + C(\tau)u(\tau) + a(\tau)] - z \tag{D_\epsilon}$$

that is, $(z', w') = F(\tau, z, w)$, and

$$z' = w, \quad w' = -z \tag{\bar{D}_\epsilon}$$

that is, $(z', w') = G(\tau, z, w)$.

We have

$$\begin{aligned} |F(\tau, z, w) - G(\tau, z, w)| &= (\epsilon/\beta_i)D_i^T|A(\tau)x + B(\tau)y + C(\tau)u + a(\tau)| \\ &< (\alpha\epsilon/\beta_i)N \end{aligned}$$

for some number N independent of ϵ , and

$$|G(\tau, z_1, w_1) - G(\tau, z_2, w_2)| = |(z_1, w_1) - (z_2, w_2)|$$

Using Theorem 3.3.1 from [5] we find that any solution (z_ϵ, w_ϵ) to (D_ϵ) is such that $|(z_\epsilon(\tau), w_\epsilon(\tau)) - (\bar{z}_\epsilon(\tau), \bar{w}_\epsilon(\tau))| \leq (\epsilon N/\beta_i)[e^\tau - 1]$ where $(\bar{z}_\epsilon, \bar{w}_\epsilon)$ is the solution to (\bar{D}_ϵ) . This implies

$$|z_\epsilon(\tau_{0\epsilon} + \sqrt{\epsilon/\beta_i}\pi)| \leq N\epsilon/\beta_i[e^\pi - 1], \quad |\zeta_\epsilon(\tau_{0\epsilon} + \sqrt{\epsilon/\beta_i}\pi)| \leq \sqrt{\epsilon/\beta_i}N[e^\pi - 1].$$

So if we have a sequence of solutions (x_ϵ, y_ϵ) to (P_ϵ) where $x_\epsilon(t)$ lies outside X on the interval $(\tau_{0\epsilon}, \tau_{1\epsilon})$ with $D_i^T T_i x_\epsilon(\tau_{0\epsilon}) - d_i = D_i^T x_\epsilon(\tau_{1\epsilon}) - d_i = 0$ and $y_\epsilon(\tau_{0\epsilon}) = y_{0\epsilon}$, then the limit of these solutions (x^*, y^*) is such that $D_i^T x^*(\bar{\tau}_0) - d_i = 0$ and $D_i^T y^*(\bar{\tau}_0^+) = -D_i^T y^*(\bar{\tau}_0^-)$.

This last equality implies $y^*(\bar{\tau}_0^+) + y^*(\bar{\tau}_0^-)$ is orthogonal to D_i . On the other hand, if v is a vector orthogonal to D_i , then we have $v^T y^*(\bar{\tau}_0^+) = v^T y^*(\bar{\tau}_0^-)$ so that $y^*(\bar{\tau}_0^+) - y^*(\bar{\tau}_0^-)$ is orthogonal to v . Combining this information, we see that $dy^*(\bar{\tau}_0)$ is feasible for (P_0) .

It only remains to show that the absolutely continuous part of dy^* is feasible.

PROPOSITION 3. *The limiting solution (x^*, y^*) satisfies the differential equation*

$$\dot{y}(t) = A(t)x(t) + B(t)y(t) + C(t)u(t) + a(t)$$

for some feasible u^* .

PROOF. Let $x^*(t) \in \text{int } X$ on the interval (τ_0, τ_1) . Then on any subinterval $[\gamma_0, \gamma_1]$ we have $\bar{\varepsilon} > 0$ sufficiently small that $x_\varepsilon(t)$ lies in $\text{int } X$ when $\varepsilon < \bar{\varepsilon}$ and so $|\dot{y}_\varepsilon(t)| \leq N \forall t \in [\gamma_0, \gamma_1], \forall \varepsilon < \bar{\varepsilon}$. By similar reasoning to that found in Proposition 1 we have that a subsequence of $\{y_\varepsilon\}$ converges uniformly to y^* on $[\gamma_0, \gamma_1]$ and that $\{y_\varepsilon\}$ converges weakly to y^* there. Thus by Theorem 3.2 of [1] there exists a convex combination of the tail of the sequence $\{y_\varepsilon\}$ that converges almost everywhere on $[\gamma_0, \gamma_1]$ to y^* . Therefore

$$\dot{y}^*(t) = A(t)x^*(t) + B(t)y^*(t) + C(t)u^*(t) + a(t) \text{ a.e. on } [\gamma_0, \gamma_1]$$

where $u^*(t)$ is the limit of the same convex combination of the tail of the sequence $\{u_\varepsilon\}$. This is feasible since $U(t)$ is compact and convex for each t . Since $\dot{y}^*(\cdot) - A(\cdot)x^*(\cdot) - B(\cdot)y^*(\cdot) - a(\cdot)$ is a measurable function on $[\gamma_0, \gamma_1]$, then so is $C(\cdot)u^*(\cdot)$. Hence y^* is feasible on $[\gamma_0, \gamma_1]$. Since this is an arbitrary closed interval within (τ_0, τ_1) it must also be true on (τ_0, τ_1) .

4. Approximation of feasible points in (P_0)

In the last section we looked at the limit (x^*, y^*) of a sequence $\{(x_\varepsilon, y_\varepsilon)\}$ and showed that the limit was feasible. Is the reverse true? That is, is any feasible point in (P_0) a limit of feasible points of (P_ε) as ε goes to zero?

Let (\bar{x}, \bar{y}) be feasible for (P_0) and let \bar{y} have discontinuities at say $\tau_1, \tau_2, \dots, \tau_r$. Since we are really only concerned with optimal solutions, we can assume $\tau_r \neq t_1$ in line with Assumption 1. Choose an $\varepsilon > 0$. Using the same control \bar{u} that generates (\bar{x}, \bar{y}) , the solution to the differential equation for (P_ε) , which we shall call $(\bar{x}_\varepsilon, \bar{y}_\varepsilon)$, is exactly the same as (\bar{x}, \bar{y}) until \bar{x}_ε hits the boundary at τ_1 . Then no matter what control $\bar{u}_\varepsilon(t) \in U(t)$ is used while $\bar{x}_\varepsilon(t)$ lies outside X , we have

$$|\bar{x}_\varepsilon(\tau_1 + \pi\sqrt{\varepsilon/\beta_i})| \leq N\varepsilon/\beta_i[e^\pi - 1]$$

$$|\bar{y}_\varepsilon(\tau_1 + \pi\sqrt{\varepsilon/\beta_i}) - \bar{y}(\tau_1^+)| \leq \sqrt{\varepsilon/\beta_i}N[e^\pi - 1]$$

where $\beta_i = \alpha_i D_i^T D_i$ and we are assuming that the i th constraint is violated. So let us choose a feasible control \bar{u}_ε and suppose \bar{x}_ε lies outside X on $(\tau_1, \hat{\tau}_{1\varepsilon})$. Then we have that $\bar{x}_\varepsilon(\hat{\tau}_{1\varepsilon})$ and $\bar{y}_\varepsilon(\hat{\tau}_{1\varepsilon})$ satisfy the conditions as above but with the exponent π replaced by $\pi + g(\varepsilon)$ where $g(\varepsilon) \rightarrow 0$. Since U is bounded we also have that $(\bar{x}_\varepsilon(\hat{\tau}_{1\varepsilon}), \bar{y}_\varepsilon(\hat{\tau}_{1\varepsilon}))$ and $(\bar{x}(\hat{\tau}_{1\varepsilon}), \bar{y}(\hat{\tau}_{1\varepsilon}))$ are close. If we now use the control $u_\varepsilon(t) = \bar{u}(t)$ again, we have that $(\bar{x}_\varepsilon(t), \bar{y}_\varepsilon(t))$ stays close to $(\bar{x}(t), \bar{y}(t))$.

Now we merely repeat the above procedure when \bar{x}_ϵ strikes the next boundary and so on. Since there are only a finite number of boundary points, we have that as ϵ goes to zero $(\bar{x}_\epsilon(t), \bar{y}_\epsilon(t))$ converges to $(\bar{x}(t), \bar{y}(t))$ away from $t \in \{\tau_1, \dots, \tau_r\}$, and in particular $(x_\epsilon(t_1), y_\epsilon(t_1))$ converges to $(\bar{x}(t_1), \bar{y}(t_1))$.

Actually the proof of Theorem 3.3.1 of [5] shows that

$$\|\bar{x}_\epsilon(\tau) - \hat{x}_\epsilon(\tau)\|_v \leq N\epsilon/\beta_i[e^\pi - 1], \quad \|\bar{y}_\epsilon(\tau) - \hat{y}_\epsilon(\tau)\|_v \leq \sqrt{\epsilon/\beta_i}N[e^\pi - 1]$$

where $(\hat{x}_\epsilon, \hat{y}_\epsilon)$ solves the system $x'(\tau) = y(\tau), y'(\tau) = -\beta_i/\epsilon x(\tau)$ for $\tau \in [\tau_j, \tau_j + \pi\sqrt{\epsilon/\beta_i}]$. These bounds, in conjunction with the ones we can obtain through similar results on the variation of initial conditions, allow us to bound the variation of (x_ϵ, y_ϵ) to be close to that of (\bar{x}, \bar{y}) ; so given any $M > \|(\bar{x}, \bar{y})\|_v$ there exists an $\bar{\epsilon} > 0$ such that for any $\epsilon < \bar{\epsilon}$ the sequence $\{(x_\epsilon, y_\epsilon)\}$ we have constructed by the aforementioned process is such that $\|(x_\epsilon, y_\epsilon)\|_v < M$. This says in part that $\|(x_\epsilon, y_\epsilon)\|_v \rightarrow \|(\bar{x}, \bar{y})\|_v$, so that the sequence $\{(x_\epsilon, y_\epsilon)\}$ converges in the strong topology, and so it also must converge in the weak* topology. Hence given any feasible (\bar{x}, \bar{y}) from (P_0) , we can construct a sequence $\{(x_\epsilon, y_\epsilon)\}$ where each (x_ϵ, y_ϵ) is feasible for (P_ϵ) and where the sequence converges to (\bar{x}, \bar{y}) in the weak* topology.

Thus every feasible point has an approximating sequence converging to it. Since the objective function f_0 is a continuous function of the endpoints we must have that $\lim_{\epsilon \rightarrow 0} f_0(\bar{x}_\epsilon(t_1), \bar{y}_\epsilon(t_1)) = f_0(\bar{x}(t_1), \bar{y}(t_1))$.

PROPOSITION 4. (x^*, y^*) is optimal for (P_0) .

PROOF. Suppose this were not the case. Then there is another feasible point (\hat{x}, \hat{y}) with a lower objective value. But it too has an approximating sequence $\{(\hat{x}_\epsilon, \hat{y}_\epsilon)\}$ where $\lim_{\epsilon \rightarrow 0} f_0(\hat{x}_\epsilon(t_1), \hat{y}_\epsilon(t_1)) = f_0(\hat{x}(t_1), \hat{y}(t_1))$. However, for each ϵ we have that (x_ϵ, y_ϵ) is optimal for (P_ϵ) so that

$$f_0(\hat{x}_\epsilon(t_1), \hat{y}_\epsilon(t_1)) \geq f_0(x_\epsilon(t_1), y_\epsilon(t_1)),$$

which implies

$$f_0(\hat{x}(t_1), \hat{y}(t_1)) \geq f_0(x^*(t_1), y^*(t_1))$$

contradicting (\hat{x}, \hat{y}) having a lower objective value. So (x^*, y^*) is optimal.

Now let us consider a related question. If (\bar{x}, \bar{y}) is optimal for (P_0) does there exist a sequence $\{(\bar{x}_\epsilon, \bar{y}_\epsilon)\}$ converging to (\bar{x}, \bar{y}) where $(\bar{x}_\epsilon, \bar{y}_\epsilon)$ is optimal for (P_ϵ) ? This is not necessarily true; however there is something just as good. If we change the word "optimal" in the question to an appropriately defined "locally optimal" then the answer is yes. The following definitions will lead to this.

DEFINITIONS. Let (\bar{x}, \bar{y}) be feasible for (P_0) . The completion $\Gamma_{(\bar{x}, \bar{y})}$ of (\bar{x}, \bar{y}) at a point $t \in [t_0, t_1]$ is given by

$$\Gamma_{(\bar{x}, \bar{y})}(t) = \{(x, y) : (x, y) = (\bar{x}(t), \lambda \bar{y}(t^-) + (1 - \lambda)\bar{y}(t^+)), \lambda \in [0, 1]\}.$$

If \bar{y} is continuous (which is the case for the completion of a feasible pair (\hat{x}, \hat{y}) from (P_ϵ)) then $\Gamma_{(\bar{x}, \bar{y})}$ is just the graph of (\bar{x}, \bar{y}) .

We will call *the δ -tube around (\bar{x}, \bar{y})* the set $\Gamma_{(\bar{x}, \bar{y})}^\delta$ given by

$$\Gamma_{(\bar{x}, \bar{y})}^\delta = \{(t, x, y) : \exists(\hat{t}, \hat{x}, \hat{y}) \text{ with } (\hat{x}, \hat{y}) \in \Gamma_{(\bar{x}, \bar{y})}(\hat{t}) \text{ with } |(t, x, y) - (\hat{t}, \hat{x}, \hat{y})| \leq \delta, \hat{t} \in [t_0, t_1]\}$$

A pair (\bar{x}, \bar{y}) will be said to be *locally optimal* for its respective problem (P_0) or (P_ϵ) if there exists a δ such that $f_0(\bar{x}(t_1), \bar{y}(t_1)) \leq f_0(x(t_1), y(t_1))$ for all feasible pairs (x, y) that also satisfy $(x(t), y(t)) \in \Gamma_{(\bar{x}, \bar{y})}^\delta(t) \forall t \in [t_0, t_1]$.

THEOREM 1. *If (\bar{x}, \bar{y}) is locally optimal for (P_0) and \bar{y} is not discontinuous at t_1 then there exists a sequence $\{(\bar{x}_\epsilon, \bar{y}_\epsilon)\}$ converging to it in the weak* topology where $(\bar{x}_\epsilon, \bar{y}_\epsilon)$ is locally optimal for (P_ϵ) .*

PROOF. Define the sets

$$\begin{aligned} \mathcal{A}_{(\bar{x}, \bar{y})}^\delta &= \{(x(t_1), y(t_1)) : (x, y) \text{ is feasible for } (P_0) \text{ and belongs to } \Gamma_{(\bar{x}, \bar{y})}^\delta\} \\ \mathcal{A}_{\epsilon(\bar{x}, \bar{y})}^\delta &= \{(x(t_1), y(t_1)) : (x, y) \text{ is feasible for } (P_\epsilon) \text{ and belongs to } \Gamma_{(\bar{x}, \bar{y})}^\delta\}. \end{aligned}$$

Let us show that for every $\delta > 0$ and $\epsilon > 0$ the set $\mathcal{A}_{\epsilon(\bar{x}, \bar{y})}^\delta$ is compact. Each of the sets is bounded so it is only necessary to show closure. If we take any sequence $\{(x_{\epsilon k}^\delta(t_1), y_{\epsilon k}^\delta(t_1))\}$ from $\mathcal{A}_{\epsilon(\bar{x}, \bar{y})}^\delta$ and look at its limit $\{(\bar{x}_\epsilon^\delta(t_1), \bar{y}_\epsilon^\delta(t_1))\}$ then we can show that it is contained in $\mathcal{A}_{\epsilon(\bar{x}, \bar{y})}^\delta$ by looking at the weak* limit of the sequence $\{(x_{\epsilon k}^\delta, y_{\epsilon k}^\delta)\}$ and showing that the limit exists, is feasible for (P_ϵ) and belongs to $\Gamma_{(\bar{x}, \bar{y})}^\delta$. This is done by essentially mirroring the proof in the first section when we showed that the weak* limit of $\{(x_\epsilon, y_\epsilon)\}$ was feasible for (P_0) . The details of this procedure are omitted.

Now (\bar{x}, \bar{y}) is locally optimal so $\exists \delta_1 > 0$ such that $f_0(\bar{x}(t_1), \bar{y}(t_1)) \leq f_0(x(t_1), y(t_1))$ for all $(x(t_1), y(t_1)) \in \mathcal{A}_{(\bar{x}, \bar{y})}^\delta$, for all $\delta \leq \delta_1$. By the previous section there exists a sequence $\{(x_\epsilon, y_\epsilon)\}$ converging weak* to (\bar{x}, \bar{y}) , where each (x_ϵ, y_ϵ) is feasible for (P_ϵ) and given the way it was constructed, there exists an $\bar{\epsilon} > 0$ such that for all $\epsilon \leq \epsilon_1$ we will also have $(x_\epsilon, y_\epsilon) \in \Gamma_{(\bar{x}, \bar{y})}^{\delta_1}$. By looking at a decreasing sequence $\{\delta_j\}$ converging to zero and a corresponding sequence $\{\epsilon_j\}$ we obtain a nonempty sequence of sets $\{\mathcal{A}_{\epsilon_j(\bar{x}, \bar{y})}^{\delta_j}\}$ each of which is nonempty. They are also compact and since f_0 is continuous there exists an $(x_j(t_1), y_j(t_1))$ belonging to $\{\mathcal{A}_{\epsilon_j(\bar{x}, \bar{y})}^{\delta_j}\}$ for all j , which minimises f_0 on $\mathcal{A}_{\epsilon_j(\bar{x}, \bar{y})}^{\delta_j}$. Since δ_j goes

to zero we have that the sequence $\{(x_j, y_j)\}$ converges weak* to (\bar{x}, \bar{y}) and each (x_j, y_j) is locally optimal for (P_{ϵ_j}) .

5. Necessary conditions for (P_0)

What conditions must an optimal solution for (P_0) satisfy? We have just discovered that any optimal solution for (P_0) will have an approximating sequence $\{(x_\epsilon, y_\epsilon)\}$ of locally optimal solutions for (P_ϵ) . Each of the problems (P_ϵ) fits into the general framework of optimal control problems analysed in the literature so we can easily write down their necessary conditions. From there we shall try to show that the optimal solution (x^*, y^*) for (P_0) must satisfy the limit of these necessary conditions, thereby giving us necessary conditions for (P_0) .

Let (x^*, y^*) be optimal for (P_0) and $\{(x_\epsilon, y_\epsilon)\}$ a sequence of locally optimal solutions for (P_ϵ) that converges weak* to (x^*, y^*) . For each $\epsilon > 0$ define the Hamiltonian H_ϵ as

$$H_\epsilon(t, x, y, p, q, u) = p^\top y + q^\top (A(t)x + B(t)y + C(t)u + a(t)) - \begin{cases} 0 & \text{if } x \in X \\ (\alpha_i/\epsilon)q^\top D_i(D_i^\top x - d_i) & \text{if } D_i^\top x - d_i > 0 \end{cases}$$

Using Theorems 5.2.1 of [3], we see that for each (x_ϵ, y_ϵ) , there exist absolutely continuous vector functions p_ϵ and q_ϵ that satisfy

$$\begin{aligned} -\dot{p}(t) &\in A(t)^\top q(t) - \begin{cases} 0 & \text{if } x_\epsilon(t) \in \text{int } X \\ [0, (\alpha_i/\epsilon)D_i D_i^\top q(t)] & \text{if } D_i^\top x_\epsilon(t) - d_i = 0 \\ (\alpha_i/\epsilon)D_i D_i^\top q(t) & \text{if } D_i^\top x_\epsilon(t) - d_i > 0 \end{cases} \\ -\dot{q}(t) &= p(t) + B(t)^\top q(t) \\ p(t_1) &= -(\partial f_0/\partial x)(x_\epsilon(t_1), y_\epsilon(t_1)) \\ q(t_1) &= -(\partial f_0/\partial y)(x_\epsilon(t_1), y_\epsilon(t_1)) \end{aligned}$$

and the control u_ϵ that generates (x_ϵ, y_ϵ) must maximise the Hamiltonian H_ϵ , which is equivalent to $q^\top C(t)u_\epsilon(t) \geq q^\top C(t)u \forall u \in U(t)$.

Now we have a sequence $\{(p_\epsilon, q_\epsilon)\}$.

Assumption 4. There exists a $K > 0$ such that $\|(p_\epsilon, q_\epsilon)\|_v \leq K, \forall \epsilon > 0$.

With this assumption we know that the sequence $\{(p_\epsilon, q_\epsilon)\}$ converges weak* to a vector function (p^*, q^*) of bounded variation. It will suffice to show that (p^*, q^*) satisfies the limit of the previous conditions with (x_ϵ, y_ϵ) replaced by (x^*, y^*) . These will then be necessary conditions for (P_0) .

If $A_i^\top x_\epsilon(t) - d_i > 0$ on some interval of time then (p_ϵ, q_ϵ) must satisfy, for the same times,

$$-\dot{p} = A(t)^\top q(t) - (\alpha_i/\epsilon)D_i D_i^\top q(t), \quad -\dot{q}(t) = p(t) + B(t)^\top q(t).$$

Let $r(t) = p(t) + B(t)^T q(t)$. Then these equations become

$$-\dot{r}(t) = B(t)^T r(t) - \dot{B}(t)^T q(t) + A(t)^T q(t) - (\alpha_i/\varepsilon) D_i D_i^T q, \quad -\dot{q}(t) = r.$$

If we let $w(t) = D_i^T q(t)$, $\xi(t) = -D_i^T r(t)$ and $\beta_i = \alpha_i D_i^T D_i$ then these convert to

$$\begin{aligned} \dot{w}(t) &= \xi(t) \\ \dot{\xi}(t) &= D_i^T [B(t)^T r(t) - \dot{B}(t)^T q(t) + A(t)^T q(t)] - (\beta_i/\varepsilon) w(t) \end{aligned} \tag{F_\varepsilon}$$

which is of the same form as (D_ε) . So its solution is close to the solution of

$$\dot{w}(t) = \xi(t), \quad \dot{\xi}(t) = (-\beta_i/\varepsilon) w(t) \tag{\bar{F}_\varepsilon}$$

If $D_i^T x_\varepsilon(\tau_{0\varepsilon}) - d_i = 0$ and $D_i^T x_\varepsilon(t) - d_i > 0$ for some interval of time beyond $\tau_{0\varepsilon}$, then we know that $D_i^T x_\varepsilon(t) - d_i = 0$ again approximately when $t = \tau_{0\varepsilon} + \pi\sqrt{\varepsilon/\beta_i} = \tau_{1\varepsilon}$, and on the interval $(t_{0\varepsilon}, \tau_{1\varepsilon})$ the solution $(\bar{w}, \bar{\xi})$ to (\bar{F}_ε) is given by

$$\begin{aligned} \bar{w}(t) &= C_1 \cos(\sqrt{\beta_i/\varepsilon}(t - \tau_{0\varepsilon})) + C_2 \sin(\sqrt{\beta_i/\varepsilon}(t - \tau_{0\varepsilon})) \\ \bar{\xi}(t) &= -C_1 \sqrt{\beta_i/\varepsilon} \sin(\sqrt{\beta_i/\varepsilon}(t - \tau_{0\varepsilon})) + C_2 \sqrt{\beta_i/\varepsilon} \cos(\sqrt{\beta_i/\varepsilon}(t - \tau_{0\varepsilon})) \end{aligned}$$

If $\bar{w}(\tau_{0\varepsilon}) = w_{0\varepsilon}$ and $\bar{\xi}(\tau_{0\varepsilon}) = \xi_{0\varepsilon}$ then $C_1 = w_{0\varepsilon}$ and $C_2 = \sqrt{\varepsilon/\beta_i} \xi_{0\varepsilon}$, so

$$\begin{aligned} \bar{w}(t) &= w_{0\varepsilon} \cos(\sqrt{\beta_i/\varepsilon}(t - \tau_{0\varepsilon})) + \sqrt{\varepsilon/\beta_i} \xi_{0\varepsilon} \sin(\sqrt{\beta_i/\varepsilon}(t - \tau_{0\varepsilon})) \\ \bar{\xi}(t) &= -w_{0\varepsilon} \sqrt{\beta_i/\varepsilon} \sin(\sqrt{\beta_i/\varepsilon}(t - \tau_{0\varepsilon})) + \xi_{0\varepsilon} \cos(\sqrt{\beta_i/\varepsilon}(t - \tau_{0\varepsilon})). \end{aligned}$$

However the maximum value of $|\xi(t)|$ on $[\tau_{0\varepsilon}, \tau_{1\varepsilon}]$ is approximately $w_{0\varepsilon} \sqrt{\beta_i/\varepsilon}$ and unless $w_{0\varepsilon}$ goes to zero with ε then we shall violate the last assumption. In terms of the original q variable this says $D_i^T q^*(\tau_0) = 0$.

At the exit time $\tau_{1\varepsilon}$ we find that $\bar{w}(\tau_{1\varepsilon}) = -w_{0\varepsilon}$, $\bar{\xi}(\tau_{1\varepsilon}) = -\xi_{0\varepsilon}$. This gives us that $D_i^T r(\tau_0^+) = -D_i^T r(\tau_0^-)$. On the other hand, if we look at vectors \hat{w} and $\hat{\xi}$ that are of the form $v^T q$ and $-v^T r$, where v is a vector orthogonal to D_i , we would find that $\hat{w}(\tau_0^+) = \hat{w}(\tau_0^-)$ and $\hat{\xi}(\tau_0^+) = \hat{\xi}(\tau_0^-)$. Hence, we must have

$$r(\tau_0^+) = r(\tau_0^-) + \mu(\tau_0) D_i \quad \text{for some } \mu(\tau_0) \geq 0.$$

This implies $dp^*(\tau_0) = \mu(\tau_0) D_i$ and $|p^*(\tau_0^*)| = |p^*(\tau_0^-)|$.

For the case when $x^*(t) \in \text{int } X$, the proof that (p^*, q^*) satisfies

$$-\dot{p}(t) = A^T(t)q(t), \quad -\dot{q}(t) = p(t) + B^T(t)q(t)$$

follows in similar fashion to the proof that x^* and y^* satisfied their differential equations.

By taking the convex combination limit and applying it to the inequality

$$q_\varepsilon^T(t)C(t)u_\varepsilon(t) \geq q_\varepsilon^T(t)C(t)u \quad \forall u \in U(t)$$

we obtain the same inequality for the optimal control u^* of (P_0) which we can write in terms of a Hamiltonian H for (P_0) where $H(t, x, y, p, q, u) = p^\top y + q^\top (A(t)x + B(t)y + C(t)u + a(t))$. We then have

$$H(t, x^*(t), y^*(t), p^*(t), q^*(t), u^*(t)) = \max_{u \in U(t)} H(t, x^*(t), y^*(t), p^*(t), q^*(t), u) \text{ a.e.}$$

The fact that $(p_\varepsilon(t), q_\varepsilon(t))$ converges to $(p^*(t), q^*(t))$ almost everywhere, in particular everywhere except around discontinuities of y^* , and when x^* is not on the boundary at the final time t_1 tells us that $(p_\varepsilon(t_1), q_\varepsilon(t_1)) \rightarrow (p^*(t_1), q^*(t_1))$. Similarly $(x_\varepsilon(t_1), y_\varepsilon(t_1)) \rightarrow (x^*(t_1), y^*(t_1))$ and because of the continuity of $\partial f_0/\partial x$ and $\partial f_0/\partial y$ we have

$$p^*(t_1) = -(\partial f_0/\partial x)(x^*(t_1), y^*(t_1)), \quad q^*(t_1) = -(\partial f_0/\partial y)(x^*(t_1), y^*(t_1)).$$

We can now write down the analogous conditions to the necessary conditions for the unconstrained problems.

THEOREM 2 (*Necessary conditions for (P_0)*). *If (x^*, y^*) is optimal for (P_0) and Assumptions 1, 2, 3 and 4 are satisfied then there exists a bounded variation vector function p^* and an absolutely continuous vector function q^* that satisfy*

$$-\dot{p}^*(t) = A(t)^\top q^*(t), \quad -\dot{q}^*(t) = p^*(t) + B(t)^\top q^*(t)$$

and if $y^*(t)$ is discontinuous then $|p^*(t^+)| = |p^*(t^-)|$, $dp^*(t) = \mu(t)D_i$ for some $\mu(t) \geq 0$ and $D_i^\top q^*(t) = 0$. Also $p^*(t_1) = -\partial f_0/\partial x(x^*(t_1), y^*(t_1))$, $q^*(t_1) = -\partial f_0/\partial y(x^*(t_1), y^*(t_1))$ and the control u^* that generates (x^*, y^*) satisfies

$$H(t, x^*(t), y^*(t), p^*(t), q^*(t), u^*(t)) = \max_{u \in U(t)} H(t, x^*(t), y^*(t), p^*(t), q^*(t), u)$$

almost everywhere on $[t_0, t_1]$.

Acknowledgement

The author is grateful to Professor R. T. Rockafellar for his contribution to this paper.

References

- [1] L. D. Berkovitz, *Optimal control theory* (Springer-Verlag, New York, 1974).
- [2] C. W. Clark, F. H. Clarke and G. R. Munro, "The optimal exploitation of renewable resource stocks", *Econometrica* **47** (1979) 25-47.
- [3] F. H. Clarke, *Optimization and nonsmooth analysis* (Wiley, New York, 1983).

- [4] L. M. Graves, *The theory of functions of real variables* (McGraw-Hill, New York, 1946).
- [5] E. Hille, *Lectures on ordinary differential equations* (Addison-Wesley, 1969).
- [6] W. S. Levine and F. E. Zajac, "An example of optimal control of a system with discontinuous state", in *Analysis and optimization of systems* (eds. A. Bensoussan and J. L. Lions), (1984) 534–541.
- [7] L. W. Neustadt, "A general theory of minimum-fuel space trajectories", *SIAM J. Con. Opt.* **3** (1965) 317–356.