MATHEMATICAL CONTROL THEORY
BANACH CENTER PUBLICATIONS, VOLUME 14
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

A PENALTY METHOD FOR DERIVING NECESSARY CONDITIONS IN PROBLEMS OF OPTIMAL CONTROL

LEONARD D. BERKOVITZ

Department of Mathematics, Purdue University, West Lafayette, Ind., U.S.A.

1. Introduction

Currently, the standard proofs of necessary conditions in problems of optimal control and mathematical programming are based upon a linearization of the problem about the optimal solution and upon the separation of convex sets. In the case of optimal control problems, the proofs involve detailed studies of the effects of perturbations of the controls and initial data on an optimal trajectory and utilize the Brower fixed point theorem, or its equivalent, to justify the neglect of the higher order terms in the linearized problem. In this paper we shall show how to use a penalty function technique to obtain the maximum principle for relaxed optimal control problems, including the relaxed hereditary control problem. In our opinion, the penalty function proofs are simpler than the standard proofs, both conceptually and technically. Except for the finite-dimensional case, we shall only present the outlines of the proofs here. For full details the reader is referred to [2], [4], [7], [8].

The use of penalty methods was suggested by Courant [5], but was not fully exploited by him. A penalty method was successfully employed in optimal control theory by A. V. Balakrishnan [1], who used it to obtain the maximum principle along particular optimal trajectory-control pairs that are obtained as limits of solutions to penalized problems. The approach described here uses some of Balakrishnan's ideas, and was suggested by E. J. McShane's penalty function proof of the multiplier rule for finite-dimensional problems [9].

2. The finite-dimensional problem

The basic ideas of the penalty function technique are very clear in the finite-dimensional case, where technical difficulties do not obscure the procedure. We therefore first derive necessary conditions for the finite-dimensional programming problem. Our presentation follows McShane [9].

Let $\mathfrak{X}_0$ be an open set in R^n. Let f be a C^1 mapping from $\mathfrak{X}_0$ to R^1, let $g = (g^1, \ldots, g^m)$ be a C^1 mapping from $\mathfrak{X}_0$ to R^m, and let $h = (h^1, \ldots, h^k)$ be a C^1 mapping from $\mathfrak{X}_0$ to R^k. Let

$$\mathfrak{X} = \{x\colon\ x \in \mathfrak{X}_0,\ g(x) \leqslant 0,\ h(x) = 0\},$$

where $y = (y^1, \ldots, y^m) \geqslant 0$ means that each component y^i is nonnegative.

The problem to be considered in this section is the following.

Problem 1. Minimize f over $\mathfrak{X}$.

We shall use the following notation. The inner product of two vectors x and y will be written simply as xy. If A is an $m \times n$ matrix and x is an m-vector, then xA will denote the product of the row vector x with the matrix A. If y is an n-vector, then Ay will denote the product of A with the column vector y. For the real valued function f, the symbol $\nabla f(x)$ will denote, as usual, the gradient of f evaluated at x. For the vector valued function g, the symbol $\nabla g(x)$ will denote the Jacobian matrix of g evaluated at x; i.e., the matrix whose entries are $(\partial g^i/\partial x^j)(x)$. The ordinary euclidean norm will be denoted by $\|\ \|$. Thus $\|x\| = (x\,x)^{1/2}$.

We shall use a penalty technique to prove the following multiplier rule.

Theorem 2.1. *Let $\bar{x}$ be a solution of Problem 1. Then there exists a real number $\lambda^0 \geqslant 0$, a vector $\lambda \in R^m$, $\lambda \geqslant 0$, and a vector μ in R^k such that:*

(i) $\lambda g(\bar{x}) = 0$;

(ii) $(\lambda^0, \lambda, \mu) \neq 0$; *and*

(iii) $\lambda^0\, \nabla f(\bar{x}) + \lambda \nabla g(\bar{x}) + \mu \nabla h(\bar{x}) = 0$.

Proof. Without loss of generality we may assume that $\bar{x} = 0$ and that $f(\bar{x}) = 0$. Let $I = \{i\colon\ g^i(\bar{x}) = 0\}$ and let $J = \{i\colon\ g^i(\bar{x}) < 0\}$. Suppose that the indices are such that $I = \{1, \ldots, r\}$ and $J = \{r+1, \ldots, m\}$, where $0 \leqslant r \leqslant m$, and $J = \emptyset$ if $r = m$. Let $g_I = (g^1, \ldots, g^r)$ and let $g_J = (g^{r+1}, \ldots, g^m)$.

For $\varepsilon > 0$, let $B(\varepsilon)$ denote the closed ball with center at the origin and radius ε; i.e., $B(\varepsilon) = \{x\colon\ \|x\| \leqslant \varepsilon\}$. Since $\mathfrak{X}_0$ is open and g is continuous, there exists an $\varepsilon_0 > 0$ such that $B(\varepsilon_0) \subset \mathfrak{X}_0$ and $g_J(x) < 0$ for $x \in B(\varepsilon_0)$.

The first step in the proof is to define a penalty function on the cartesian product of $B(\varepsilon_0)$ and the positive integers as follows. Let ω be any real valued C^1 function defined on $(-\infty, \infty)$ such that $\omega(u) = 0$ for $u \leqslant 0$, $\omega(u) > 0$ for $u > 0$, and ω is increasing on $[0, \infty)$. The penalty function F is defined as follows:

$$F(x, N) = f(x) + \|x\|^2 + N\Big\{\sum_{i=1}^{r} \omega\big(g^i(x)\big) + \sum_{i=1}^{k} \big(h^i(x)\big)^2\Big\}. \tag{2.1}$$

The second step is to establish the following result.

LEMMA 2.1. *For every $0 < \varepsilon < \varepsilon_0$ there exists an integer $N(\varepsilon)$ such that the minimum of $F(\cdot, N(\varepsilon))$ on the closed ball $B(\varepsilon)$ is attained at a point x_ε in the interior of $B(\varepsilon)$.*

To prove Lemma 2.1 let us suppose for the moment that we have proved the following fact.

LEMMA 2.2. *For every $0 < \varepsilon < \varepsilon_0$ there exists an integer $N(\varepsilon)$ such that $F(x, N(\varepsilon)) > 0$ for all x such that $\|x\| = \varepsilon$.*

Since $F(\cdot, N(\varepsilon))$ is continuous on the compact set $B(\varepsilon)$, it attains its minimum at some point x_ε in $B(\varepsilon)$. From the definition of F we see that $F(0, N(\varepsilon)) = 0$ and from Lemma 2.2 we have that $F(x, N(\varepsilon)) > 0$ for $\|x\| = \varepsilon$. Hence we must have $\|x_\varepsilon\| < \varepsilon$.

Thus to complete the proof of Lemma 2.1 we need to prove Lemma 2.2. To this end, suppose that Lemma 2.2 were false. Then there would exist a sequence of integers $\{N_p\}$ with $N_p \to \infty$ and a sequence of points $\{x_p\}$ with $\|x_p\| = \varepsilon$ such that $F(x_p, N_p) \leqslant 0$. Hence

$$f(x_p) + \|x_p\|^2 \leqslant -N_p \left\{ \sum_{i=1}^{r} \omega(g^i(x_p)) + \sum_{i=1}^{k} (h^i(x_p))^2 \right\}. \tag{2.2}$$

Since for each p, $\|x_p\| = \varepsilon$, there exists a subsequence, which we again label as $\{x_p\}$, and a point x^* with $\|x^*\| = \varepsilon$ such that $x_p \to x^*$. If in (2.2) we divide through by $-N_p$ and let $p \to \infty$ we get, since f, g, h and ω are continuous, that

$$0 \geqslant \sum_{i=1}^{r} \omega(g^i(x^*)) + \sum_{i=1}^{k} (h^i(x^*))^2.$$

Hence for each $i = 1, \ldots, r$, $g^i(x^*) \leqslant 0$ and for each $i = 1, \ldots, k$, $h^i(x^*) = 0$. Since $\|x^*\| < \varepsilon < \varepsilon_0$, $g_J(x^*) < 0$. Thus $x^* \in \mathfrak{X}$. Hence

$$0 = f(0) \leqslant f(x^*). \tag{2.3}$$

On the other hand, from (2.2) we have that $f(x_p) \leqslant -\varepsilon^2$. Since $f(x_p) \to$ $\to f(x^*)$, we have that $f(x^*) \leqslant -\varepsilon^2$, which contradicts (2.3) and Lemma 2.2 is established.

It is clear from the above proof that if $\{\varepsilon_n\}$ is a sequence tending to zero, then we may suppose that the integers $N(\varepsilon_n) \to \infty$.

The third step is to derive necessary conditions that are satisfied at a point x_ε at which the unconstrained function $F(\cdot, N(\varepsilon))$ attains its minimum on the set $B(\varepsilon)$.

Since x_ε is an interior point of $B(\varepsilon)$,

$$\frac{\partial F}{\partial x^j}(x_\varepsilon, N(\varepsilon)) = 0, \quad j = 1, \ldots, n,$$

or

(2.4)
$$\frac{\partial f}{\partial x^j}(x_\varepsilon)+2x_\varepsilon^j+\sum_{i=1}^{r} N\omega'\big(g^i(x_\varepsilon)\big)\frac{\partial g^i}{\partial x^j}(x_\varepsilon)+\sum_{i=1}^{k} 2Nh^i(x_\varepsilon)\frac{\partial h^i}{\partial x^j}(x_\varepsilon)=0,$$
$$j=1,\ldots,n.$$

Define

$$L(\varepsilon)=\Big\{1+\sum_{i=1}^{r}\big[N\omega'\big(g^i(x_\varepsilon)\big)\big]^2+\sum_{i=1}^{k}[2Nh^i(x_\varepsilon)]^2\Big\}^{1/2},$$
$$\lambda^0(\varepsilon)=1/L(\varepsilon),$$
$$\lambda^i(\varepsilon)=N\omega'\big(g^i(x_\varepsilon)\big)/L(\varepsilon),\qquad i=1,\ldots,r,$$
$$\lambda^i(\varepsilon)=0,\qquad i=r+1,\ldots,m,$$
$$\mu^i(\varepsilon)=2Nh^i(x_\varepsilon)/L(\varepsilon),\qquad i=1,\ldots,k.$$

Note that $\big(\lambda^0(\varepsilon),\lambda(\varepsilon),\mu(\varepsilon)\big)$ is a unit vector and that $\lambda^0(\varepsilon)>0$, $\lambda(\varepsilon)\geqslant 0$. If we divide through by $L(\varepsilon)$ in (2.4), we obtain:

(2.5)
$$\lambda^0(\varepsilon)\frac{\partial f}{\partial x^j}(x_\varepsilon)+\frac{2x_\varepsilon^j}{L(\varepsilon)}+\sum_{i=1}^{r}\lambda^i(\varepsilon)\frac{\partial g^i(x_\varepsilon)}{\partial x^j}+\sum_{i=1}^{k}\mu^i(\varepsilon)\frac{\partial h^i(x_\varepsilon)}{\partial x^j}=0,$$
$$j=1,\ldots,n.$$

This is our desired necessary condition, which we call the *ε-multiplier rule.*

The final step is to let $\varepsilon\to 0$ and let the ε-multiplier rule tend to the multiplier rule of Theorem 2.1. Thus, let $\{\varepsilon_q\}$ be a decreasing sequence converging to zero. Then the corresponding minimum points $x_{\varepsilon_q}\to 0$. Since the vectors $\big(\lambda^0(\varepsilon_q),\lambda(\varepsilon_q),\mu(\varepsilon_q)\big)$ are unit vectors, there exists a subsequence, which we again label as $\{\varepsilon_q\}$, and a unit vector (λ^0,λ,μ) such that

$$\big(\lambda^0(\varepsilon_q),\lambda(\varepsilon_q),\mu(\varepsilon_q)\big)\to(\lambda^0,\lambda,\mu).$$

Clearly,

$$\lambda^0\geqslant 0,\qquad \lambda_I=(\lambda^1,\ldots,\lambda^r)\geqslant 0,\qquad \lambda_J\stackrel{\text{d}}{=}(\lambda^{r+1},\ldots,\lambda^m)=0,$$

and $\lambda g(0)=0$. Also note that $L(\varepsilon_q)\geqslant 1$. Hence, if we let $\varepsilon_q\to 0$ in (2.5), we obtain the conclusion of Theorem 2.1.

It is easy to see that $\lambda^0>0$, and hence may be assumed to equal one, whenever the following constraint qualification holds [6]. The matrix $\nabla h(\bar{x})$ has rank k and the system $\nabla g_I(\bar{x})z>0$, $\nabla h(\bar{x})z=0$ has a solution z in R^n.

3. The optimal control problem

The first control problem to be considered is the following. Minimize

$$J(\varphi, u) = \int_0^1 f^0(t, \varphi(t), u(t))\,dt \tag{3.1}$$

over all pairs of functions (φ, u), defined on $[0,1]$ such that φ is absolutely continuous and u is measurable, and such that the following hold:

$$\varphi'(t) = f(t, \varphi(t), u(t)) \text{ a.e.,} \tag{3.2}$$

$$u(t) \in \Omega(t) \text{ a.e.,} \tag{3.3}$$

$$(\varphi(0), \varphi(1)) \in \mathfrak{B}. \tag{3.4}$$

Such pairs (φ, u) will be called *admissible pairs*. The function φ will be called an *admissible trajectory* and the function u an *admissible control.*

The following assumptions are made about the data of the problem. The function f^0: $(t, x, z) \to f^0(t, x, z)$ is a mapping from $I_0 \times R^n \times R^m$ to R^1, where I_0 is a fixed open interval containing $[0,1]$. The function f: $(t, x, z) \to f(t, x, z)$ is a mapping from $I_0 \times R^n \times R^m$ to R^n. The functions f^0 and f are (i) differentiable on R^n for fixed (t, z); (ii) continuous on $R^n \times R^m$ for fixed t in I_0; and are (iii) measurable on I_0 for fixed (x, z). Let $f_1(t, x, z)$ denote the matrix of partial derivatives of f with respect to x evaluated at (t, x, z) and let $f_1^0(t, x, z)$ have similar meaning. Let $I = [0,1]$. For every compact subset Γ of $R^n \times R^m$ there is a function μ in $L_2[I]$ such that, for every (t, x, z) in $I \times \Gamma$,

$$\begin{aligned} \|\bar{f}(t, x, z)\| &\leqslant \mu(t), \\ \|\bar{f}_1(t, x, z)\| &\leqslant \mu(t), \end{aligned} \tag{3.5}$$

where $\bar{f} = (f^0, f) = (f^0, f^1, \ldots, f^n)$. For each t in I the set $\Omega(t)$ is a subset of R^m. The set $\mathfrak{B}$ is a fixed set in R^{2n}.

We further assume that $\mathfrak{B}$ is compact and that the sets $\Omega(t)$ are independent of t; i.e., $\Omega(t) = \Omega$, a fixed set in R^m.

There is no loss of generality in assuming that the initial time and terminal time are fixed, since this can always be achieved by a suitable transformation (see [3], pp. 27–28). The condition that $\Omega(t)$ is constant can be replaced by the condition that the union of the sets $\Omega(t)$ as t ranges over I is contained in a compact set Ω_1.

If one were to apply the penalty function procedure of Section 2 to this control problem, one would be tempted to proceed as follows. The role of x is now played by a pair (φ, u) with φ in $W^{1,2}(I)$ and u in $L_\infty[I]$ with values $u(t) \in \Omega$ a.e. The role of the equality constraint $h(x) = 0$ is played by the differential equation (3.2). Thus if $(\bar{\varphi}, \bar{u})$ were a solution of the optimal control problem, one would consider a penalty function of the form

$$F(\varphi, u, N) = \int_0^1 f^0(t, \varphi(t), u(t))\,dt + \|(\varphi, u) - (\bar{\varphi}, \bar{u})\|_{\Delta}^2 + \\ + \left\|\varphi'(t) - f(t, \varphi(t), u(t))\right\|_2^2,$$

where $\|\ \|_2$ denotes the L_2-norm and $\|\ \|_{\Delta}$ denotes "an appropriate norm" on pairs (φ, u).

The first step in the penalty procedure would require that under some norm on (φ, u) the closed ball $B(\varepsilon)$ be compact in some topology, and that for fixed N, the function $F(\cdot, \cdot, N)$ be continuous — or at least lower semicontinuous — on $B(\varepsilon)$ with respect to the chosen topology on $B(\varepsilon)$. If one attempts to proceed in this direction, it soon becomes apparent that none of the usual norms or notions of weak or weak-star convergence will suffice. It turns out that the appropriate procedure is to consider the relaxed control problem in place of the original one. For the relaxed control problem we shall have the requisite compactness and lower semicontinuity. We devote the next section to the formulation of the relaxed control problem.

The relaxed problem should not be thought of as an artificial device introduced merely to permit the use of a penalty method to derive necessary conditions. A case can be made for the assertion that the relaxed problem is the proper context in which to consider the optimal control problem since under reasonable assumptions on the data the relaxed problem has a solution. The ordinary control problem requires convexity conditions to guarantee the existence of a solution. This requirement severely limits the class of problems for which a solution can be guaranteed. If relaxed controls are introduced, the existence of a solution is guaranteed for a much broader class of problems.

4. The relaxed problem

A *relaxed control* is a mapping

$$v\colon t \to v(t) = \mu(t, \cdot)$$

from I to the probability measures on Ω such that for every polynomial p the function P defined by

$$P(t) = \int_{\Omega} p(z)\,d\mu(t, z)$$

is Lebesgue measurable on I. An arbitrary ordinary control u can be identified with the relaxed control v_u that assigns to t the atomic measure concentrated at $u(t)$.

We shall use the letter v for relaxed controls and the letter u for ordinary controls.

A *relaxed admissible pair* (φ, v) is an absolutely continuous function φ defined on I and satisfying (3.4) and a relaxed control v such that for a.e. t in I

$$\varphi'(t) = \int_{\Omega} f\big(t, \varphi(t), z\big)\, d\mu(t, z).$$

To simplify notation we define

$$f\big(t, \varphi(t), v(t)\big) = \int_{\Omega} f\big(t, \varphi(t), z\big)\, d\mu(t, z),$$

and similarly for $f^0\big(t, \varphi(t), v(t)\big)$. Thus, an admissible relaxed pair (φ, v) satisfies

$$\varphi'(t) = f\big(t, \varphi(t), v(t)\big).$$

A relaxed pair (φ_0, v_0) is a *solution of the relaxed problem*, or a *relaxed optimal pair*, if

$$\int_0^1 f^0\big(t, \varphi_0(t), v_0(t)\big)\, dt \leqslant \int_0^1 f^0\big(t, \varphi(t), v(t)\big)\, dt$$

for all relaxed admissible pairs (φ, v).

We now summarize some important facts concerning relaxed controls; for details see Warga ([10], pp. 264–273).

Let $\mathfrak{L}$ denote the set of real-valued functions ψ defined on $I \times \Omega$ such that $\psi(\cdot, z)$ is measurable on I for each z in Ω and $\psi(t, \cdot)$ is continuous on Ω for each t in I and such that the function $t \to \max\{|\psi(t, z)|: z \in \Omega\}$ is in $L_1[I]$. The set $\mathfrak{L}$ is a Banach space with norm

$$\|\psi\| = \int_I \big(\max_z |\psi(t, z)|\big)\, dt.$$

Let $\tilde{v}$ be a mapping from I to the set of finite Radon measures on Ω. Thus: $\tilde{v}$: $t \to \nu(t; \cdot)$. Let $\mathfrak{N}$ denote the set of such mappings with the following two properties: (i) For every polynomial p the function $t \to \int_{\Omega} p(z)\, d\nu(t, z)$ is measurable; (ii) If $|\nu(t)|$ denotes the total variation measure of $\nu(t)$, then ess $\sup\{|\nu(t)|(\Omega): t \in I\}$ is finite. Then $\mathfrak{N}$ can be identified with the dual space $\mathfrak{L}^*$ of $\mathfrak{L}$, where each ν in $\mathfrak{N}$ is identified with the functional which we also denote by ν, as follows:

$$\nu(\psi) = \int_I \left(\int_\Omega \psi(t, z) d\nu(t, z) \right) dt, \qquad \psi \in \mathfrak{L}.$$

The norm of an element of $\mathfrak{N}$, viewed as an element of $\mathfrak{L}^*$, will be denoted by $\|\nu\|_L$ and is defined by

$$\|\nu\|_L = \sup\{|\nu(\psi)|: \|\psi\| = 1\}. \tag{4.1}$$

It is also given by

$$\|\nu\|_L = \operatorname{ess\,sup}\{|\nu(t)|(\Omega): t \in I\}.$$

Let $\mathfrak{R}^*$ denote the set of relaxed controls. Then $\mathfrak{R}^*$ is a convex subset of $\mathfrak{L}^*$ and is compact and sequentially compact in the weak star topology of $\mathfrak{L}^*$. ([10], Theorem IV, 2.1 p. 272.)

Let $\varepsilon > 0$, let v_0 be a relaxed control and let

$$B(v_0, \varepsilon) = \{\nu \in \mathfrak{N}: \|\nu - v_0\|_L \leqslant \varepsilon\}.$$

Then $B(v_0, \varepsilon)$ is weak star compact. Moreover, since $\mathfrak{L}$ is separable, $B(v_0, \varepsilon)$ is sequentially weak star compact.

Since the set of relaxed controls is weak star compact, it is weak star closed. Thus the following statement holds.

LEMMA 4.1. *For every $\varepsilon > 0$, the set of relaxed controls v such that $v \in B(v_0, \varepsilon)$ is weak star compact and weak star sequentially compact.*

The next lemma will permit us to carry out passages to the limit at several points in our penalty arguments.

LEMMA 4.2. *Let g be a mapping from $I \times R^n \times R^m$ to R^r, $r \geqslant 1$, having the properties of the functions f^0 and f listed in Section 3. Let $\{\varphi_n\}$ be a sequence of continuous functions converging uniformly on I to a function φ_0. Let $\{v_n\}$ be a sequence of relaxed controls converging weak-star to v_0. Then, for any function X in $L_2[I]$,*

$$\lim_{n\to\infty} \int_0^1 X(t) g\big(t, \varphi_n(t), v_n(t)\big) dt \to \int_0^1 X(t) g\big(t, \varphi(t), v(t)\big) dt.$$

The conclusion of the lemma can be restated as follows. Let $\bar{g}_n(t) = g\big(t, \varphi_n(t), v_n(t)\big)$ and let $\bar{g}_0(t) = g\big(t, \varphi_0(t), v_0(t)\big)$. Then $\bar{g}_n \to \bar{g}_0$ weakly in $L_2[I]$. For a proof of Lemma 4.2 the reader is referred to [4].

5. A penalty function proof of the maximum principle for the relaxed control problem

In this section we sketch the use of the penalty function technique to derive the maximum principle for the relaxed problem with fixed end points. For full details and for the problem with variable end points see [4].

The problem that we consider is to minimize

$$J(\varphi, v) = \int_0^1 f^0\big(t, \varphi(t), v(t)\big)\,dt \tag{5.1}$$

over elements (φ, v) in $W^{1,2}(I)\times\mathfrak{R}^*$ subject to

$$\varphi'(t) = f\big(t, \varphi(t), v(t)\big), \tag{5.2}$$

$$\varphi(0) = x_0, \quad \varphi(1) = x_1, \tag{5.3}$$

where x_0 and x_1 are fixed points in R^n and the notation is as established in the preceding section. The functions f^0 and f are assumed to satisfy the hypotheses set out in Section 3.

Let (φ_0, v_0) be a solution of the problem. Without loss of generality we may assume that $J(\varphi_0, v_0) = 0$.

The first step in the proof is to define a penalty function F on the cartesian product of $W^{1,2}(I)\times\mathfrak{R}^*\times Z$, where Z denotes the positive integers, as follows:

$$F(\varphi, v, N) = J(\varphi, v) + \|\varphi' - \varphi_0'\|_2^2 + \|\varphi(0) - x_0\|^2 + \varepsilon\|v - v_0\|_L + N\big\|\varphi' - f\big(t, \varphi(t), v(t)\big)\big\|_2^2, \tag{5.4}$$

where $\|\ \|_2$ denotes the L_2 norm and $\|\ \|_L$ is the norm defined in (4.1). The pair (φ, v) corresponds to x in (2.1), the functional J corresponds to f in (2.1), the sum $\|\varphi' - \varphi_0'\|_2^2 + \|\varphi(0) - x_0\|^2 + \varepsilon\|v - v_0\|_L$ corresponds to $\|x\|^2 = \|x - 0\|^2$ in (2.1), and the term $N\big\|\varphi' - f\big(t, \varphi(t), v(t)\big)\big\|_2^2$ corresponds to $N\sum_{i=1}^{k}\big(h^i(x)\big)^2$ in (2.1).

The role of the set $B(\varepsilon)$ in Section 2 is now played by the set $\mathfrak{D}(\varepsilon)$ which is defined to be the set of pairs (φ, v) where φ is in $W^{1,2}(I)$ and v is in $\mathfrak{R}^*$ such that

$$\|v - v_0\|_L \leqslant \varepsilon, \quad \|\varphi(0) - x_0\|^2 \leqslant \varepsilon \quad \|\varphi' - \varphi_0'\|_2 \leqslant \varepsilon. \tag{5.5}$$

Note that the set $\mathfrak{D}(\varepsilon)$ is compact with respect to sequential weak star convergence of relaxed controls v and weak convergence of functions φ in $W^{1,2}(I)$ – or what is the same, weak convergence in $L_2(I)$ of the derivatives φ'.

We also note that as a consequence of Lemma 4.2 and of the lower semicontinuity of a norm in the dual space of a Banach space with respect to weak star convergence, the following statement is true:

LEMMA 5.1. *Let $\{(\varphi_n, v_n)\}$ be a sequence such that $v_n \to v_0$ weak star, $\varphi_n' \to \varphi_0'$ weakly in L_2, and $\varphi_n \to \varphi_0$ uniformly. Then, for fixed N,*

$$\liminf_{n\to\infty} F(\varphi_n, v_n, N) \geqslant F(\varphi_0, v_0, N).$$

We also point out that as a consequence of the relation

$$\varphi(t) = \varphi(0) + \int_0^t \varphi'(s)\,ds, \tag{5.6}$$

if $\varphi_n' \to \varphi_0'$ weakly in L^2, and $\varphi_n(0) \to x_0$, then there is a subsequence of $\{\varphi_n\}$ that converges uniformly to $\varphi_0(t)$, where

$$\varphi_0(t) = x_0 + \int_0^t \varphi_0'(s)\,ds.$$

Keeping in mind the observation of the last paragraph, Lemma 5.1, and the compactness of $\mathfrak{D}(\varepsilon)$, it is not difficult to establish the following analogue of Lemma 2.1.

LEMMA 5.2. *For every $\varepsilon > 0$ there exists an integer $N(\varepsilon)$ such that the minimum of $F(\cdot, \cdot, N(\varepsilon))$ over the set $\mathfrak{D}(\varepsilon)$ defined by* (5.5) *is attained at a point $(\varphi_\varepsilon, v_\varepsilon)$ in $W^{1,2}(I) \times \mathfrak{R}^*$ satisfying*

$$\|v_\varepsilon - v_0\|_L < \varepsilon, \qquad \|\varphi_\varepsilon(0) - x_0\| < \varepsilon, \qquad \|\varphi_\varepsilon' - \varphi_0'\|_2 < \varepsilon.$$

We note that $N(\varepsilon)$ is to be chosen so that $N(\varepsilon) \to \infty$ as $\varepsilon \to 0$. The next step in the procedure is to derive an ε-maximum principle, which is the analogue of the ε-multiplier rule of Section 2.

Let η be any function of class C^1 that vanishes outside a compact interval I_1 contained in the interior of I. Then since $\|\varphi_\varepsilon(0) - x_0\| < \varepsilon$ and $\|\varphi_\varepsilon' - \varphi_0'\|_2 < \varepsilon$, for all real θ with $|\theta|$ sufficiently small, $(\varphi_\varepsilon + \theta\eta, v_\varepsilon)$ is in $\mathfrak{D}(\varepsilon)$ and the function Φ_ε defined by the formula

$$\Phi_\varepsilon(\theta) = F\big(\varphi_\varepsilon + \theta\eta, v_\varepsilon, N(\varepsilon)\big)$$

has a minimum at $\theta = 0$. This function is also differentiable with respect to θ at $\theta = 0$, and therefore $\Phi_\varepsilon'(0) = 0$ for all η.

Let $f_1^0(\varepsilon, t) = f_1^0\big(t, \varphi_\varepsilon(t), v_\varepsilon(t)\big)$ and let $f_1(\varepsilon, t)$, $f^0(\varepsilon, t)$, $f(\varepsilon, t)$ have similar meanings. Let

$$\psi(\varepsilon, t) = 2\left[\varphi_\varepsilon'(t) - \varphi_0'(t) + N(\varepsilon)\big(\varphi_\varepsilon'(t) - f(\varepsilon, t)\big)\right] \tag{5.7}$$

and let

$$L(\varphi_\varepsilon, v_\varepsilon)(\eta) = \int_0^1 \Big[\Big\{f_1^0(\varepsilon, t) + \Big(2\big(\varphi_\varepsilon'(t) - \varphi_0'(t)\big) - \psi(\varepsilon, t)\Big) f_1(\varepsilon, t)\Big\}\eta(t) + \psi(\varepsilon, t)\eta'(t)\Big]\,dt. \tag{5.8}$$

Then a straightforward calculation shows that for all η

$$\Phi_\varepsilon'(0) = L(\varphi_\varepsilon, v_\varepsilon)(\eta).$$

From this relation and from $\Phi'(0) = 0$ it follows that

$$(5.9)\qquad \psi(\varepsilon, t) = \int_0^t f_1^0(\varepsilon, s) + \big(2(\varphi_\varepsilon'(s) - \varphi_0'(s)) - \psi(\varepsilon, s)\big) f_1(\varepsilon, s)\, ds + c(\varepsilon),$$

where $c(\varepsilon)$ depends on ε.

Let $M(\varepsilon) = 1 + \|\psi(\varepsilon, 0)\|$ and let $\lambda(\varepsilon, t) = \psi(\varepsilon, t)/M(\varepsilon)$. Then from (5.9) we conclude the following. There exists an absolutely continuous function $\lambda(\varepsilon, \cdot)$ defined on I with range in R^n such that for a.e. t in I

$$(5.10)\qquad \lambda'(\varepsilon, t) = \frac{f_1^0(\varepsilon, t)}{M(\varepsilon)} - f_1(\varepsilon, t)^T \lambda(\varepsilon, t) + 2 f_1(\varepsilon, t)^T \frac{(\varphi_\varepsilon'(t) - \varphi_0'(t))}{M(\varepsilon)},$$

where $M(\varepsilon)$ is a constant greater than one, $\|\lambda(\varepsilon, 0)\| \leqslant 1$, and T denotes transpose.

Let v be an arbitrary relaxed control, and for $-\infty < \theta < \infty$ define

$$v(\theta) = v_\varepsilon + \theta(v - v_\varepsilon).$$

Then for $0 \leqslant \theta \leqslant 1$, $v(\theta)$ is a relaxed control. Since $\|v_\varepsilon - v_0\|_L < \varepsilon$, there exists a $\theta_0 > 0$ such that if $0 \leqslant \theta \leqslant \theta_0$, then $\|v(\theta) - v_0\| \leqslant \varepsilon$. Thus

$$\varrho(\theta) = F\big(\varphi_\varepsilon, v(\theta), N(\varepsilon)\big)$$

is defined for all $0 \leqslant \theta \leqslant \theta_0$ and has a minimum at $\theta = 0$. It is not difficult to show that ϱ has a right-hand derivative $\varrho'(0+)$ at $\theta = 0$, and hence $\varrho'(0+) \geqslant 0$. The only term in the definition of $\varrho(\theta)$ that causes difficulty is the term $\gamma(\varepsilon, \theta)$, defined by the formula

$$\gamma(\varepsilon, \theta) = \|v(\theta) - v_0\|_L = \|(v_\varepsilon - v_0) + \theta(v - v_\varepsilon)\|_L.$$

The function $\gamma(\varepsilon, \cdot)$ is convex in θ and therefore has a right-hand derivative $\gamma'(\varepsilon, 0+)$ at $\theta = 0$. Bounds on $\gamma'(\varepsilon, 0+)$, which show that $\gamma'(\varepsilon, 0+)$ is uniformly bounded for all ε, can be obtained using the chord property of convex functions.

A straightforward computation of $\varrho'(0+)$ and the use of the inequality $\varrho'(0+) \geqslant 0$ give:

$$(5.11)\qquad \int_0^1 H\big(\varepsilon; t, \varphi_\varepsilon(t), v_\varepsilon(t), \lambda(\varepsilon, t)\big)\, dt + \varepsilon\gamma'(\varepsilon, 0+)$$
$$\geqslant \int_0^1 H\big(\varepsilon; t, \varphi_\varepsilon(t), v(t), \lambda(\varepsilon, t)\big)\, dt,$$

where

$$(5.12)\qquad H\big(\varepsilon, t, \varphi_\varepsilon(t), v_\varepsilon(t), \lambda(\varepsilon, t)\big) = -f^0\big(t, \varphi_\varepsilon(t), v_\varepsilon(t)\big)/M(\varepsilon) +$$
$$+ \lambda(\varepsilon, t) f\big(t, \varphi_\varepsilon(t), v_\varepsilon(t)\big) + 2\big(\varphi_\varepsilon'(t) - \varphi_0'(t)\big) f\big(t, \varphi_\varepsilon(t), v_\varepsilon(t)\big)/M(\varepsilon)$$

and $H(\varepsilon, t, \varphi_\varepsilon(t), v(t), \lambda(\varepsilon, t))$ is defined similarly with $v(t)$ replacing $v_\varepsilon(t)$.

Relations (5.10)–(5.12) constitute the "ε-maximum principle".

We next let $\varepsilon \to 0$ and pass to the limit in (5.10)–(5.12) to obtain the necessary conditions satisfied by the optimal pair (φ_0, v_0).

Let

$$\lambda^0 = \liminf_{\varepsilon \to 0} (1/M(\varepsilon)).$$

Then since $M(\varepsilon) \geqslant 1$, λ^0 is finite and $\lambda^0 \geqslant 0$. Let $\{\varepsilon_n\}$ be a sequence tending to zero such that $(1/M(\varepsilon_n)) \to \lambda^0$. Since $(\varphi_\varepsilon, v_\varepsilon)$ satisfies (5.5), it follows from this relation and from (5.6) that $\varphi'_{\varepsilon_n} \to \varphi'_0$ strongly in L_2, $v_{\varepsilon_n} \to v_0$ strongly in $\mathfrak{L}^*$ and $\varphi_{\varepsilon_n} \to \varphi_0$ uniformly on I.

We obtain the convergence of a subsequence of the functions $\lambda(\varepsilon, \cdot)$ as follows. The homogeneous equation corresponding to (5.10) is

$$\lambda'(\varepsilon, t) = -f_1(\varepsilon, t)^T \lambda(\varepsilon, t).$$

Let $\Lambda(\varepsilon, t)$ denote the matrix of fundamental solutions of this system with $\Lambda(\varepsilon, 0) = I$, where I is the $n \times n$ identity matrix. Then $\|\Lambda(\varepsilon, t)\|$ is uniformly bounded for all $0 \leqslant \varepsilon \leqslant 1$ and $0 \leqslant t \leqslant 1$. Since $\|\lambda(\varepsilon, 0)\| \leqslant 1$, it then follows that $\|\lambda(\varepsilon, t)\|$ is uniformly bounded. From this we can conclude that the functions $\|\lambda'(\varepsilon, t)\|$ have equiabsolutely continuous integrals. This then enables us to extract a subsequence of $\{\varepsilon_n\}$, which we again label $\{\varepsilon_n\}$, and an absolutely continuous function λ such that $\lambda(\varepsilon_n, t) \to \lambda(t)$ uniformly on $[0, 1]$.

Passage to the limit in (5.10)–(5.12) can now easily be justified and we obtain the maximum principle for the relaxed problem.

THEOREM 5.1 (Maximum Principle). *Let f^0 and f be as in Section 3. Let (φ_0, v_0) be a solution of the relaxed problem. Then there exists a constant $\lambda^0 \geqslant 0$ and an absolutely continuous function λ such that $(\lambda^0, \lambda(t)) \neq 0$ on I and such that*

$$\lambda'(t) = \lambda^0 f_1^0(t) - f_1(t)^T \lambda(t),$$

$$\int_0^1 H(t, \varphi_0(t), v_0(t), -\lambda^0, \lambda(t))\, dt \geqslant \int_0^1 H(t, \varphi_0(t), v(t), -\lambda^0, \lambda(t))\, dt,$$

where

$$f_1^0(t) = f_1^0(t, \varphi_0(t), v_0(t)), \quad f_1(t) = f_1(t, \varphi_0(t), v_0(t)),$$

where

$$H(t, \varphi_0(t), v_0(t), -\lambda^0, \lambda(t)) = -\lambda^0 f^0(t, \varphi_0(t), v_0(t)) + \lambda(t) f(t, \varphi_0(t), v_0(t)),$$

and where $H(t, \varphi_0(t), v(t), -\lambda^0, \lambda(t))$ is defined similarly, except that $v(t)$ replaces $v_0(t)$.

6. The bounded state problem

We consider the following problem. Minimize $J(\varphi, v)$ defined by (5.1) subject to (5.2), (5.3), and the additional state constraint

$$G(\varphi(t)) \leqslant 0 \tag{6.1}$$

for all $t \in I$. Here G is assumed to be a real valued C^2 function defined on R^n. Our treatment will follow Medhin [7], [8], where a more general problem is treated and where full details can be found. As in Section 5, we shall only sketch the method.

As before let (φ_0, v_0) denote a solution of the relaxed problem. We assume that $\nabla G(x) \neq 0$ at all points x whose distance from some point of the set $\{x: x = \varphi_0(t), 0 \leqslant t \leqslant 1\}$ is less than ε_1, where $0 < \varepsilon_1 < 1$. We adopt the notation ∇G for the gradient of G instead of G_1 to emphasize that G_1 is a gradient.

Let Γ denote those functions in $W^{1,2}(I)$ that satisfy (6.1). We again consider the penalty function F defined by (5.4). The arguments used in Section 5 can now be used to establish the following result.

LEMMA 6.1. *For every $0 < \varepsilon \leqslant \varepsilon_1$ there exists an integer $N(\varepsilon)$ such that the minimum of $F(\cdot, \cdot, N(\varepsilon))$ over the set $\mathfrak{D}(\varepsilon) \cap \Gamma$ is attained at a point $(\varphi_\varepsilon, v_\varepsilon)$ satisfying*

$$\|v_\varepsilon - v_0\|_L < \varepsilon, \qquad \|\varphi_\varepsilon(0) - x_0\| < \varepsilon, \qquad \|\varphi_\varepsilon' - \varphi_0'\|_2 < \varepsilon.$$

Since $(\varphi_\varepsilon, v_\varepsilon)$ minimizes $F(\cdot, \cdot, N(\varepsilon))$ over $\mathfrak{D}(\varepsilon) \cap \Gamma$, in order to obtain a necessary condition we cannot vary φ_ε arbitrarily. We must ensure that the comparison function $\varphi_\varepsilon + \theta\eta$ also satisfies (6.1). To this end, let

$$\xi(t) = -\nabla G(\varphi_\varepsilon(t)) / \|\nabla G(\varphi_\varepsilon(t))\|^2 \tag{6.2}$$

and let z be a continuous nonnegative piecewise C^1 scalar function on I such that $z(0) = z(1) = 0$. Let

$$\eta(t) = z(t)\,\xi(t). \tag{6.3}$$

Note that ξ, and hence η, depend on ε. It is easy to see that for η so defined and for sufficiently small $\theta \geqslant 0$, $G(\varphi_\varepsilon(t) + \theta\eta(t)) \leqslant 0$. Also, for such θ, we have $(\varphi_\varepsilon + \theta\eta, v_\varepsilon) \in \mathfrak{D}(\varepsilon)$.

As before, the function Φ_ε defined by the formula $\Phi_\varepsilon(\theta) = F(\varphi_\varepsilon + \theta\eta, v_\varepsilon, N(\varepsilon))$ is differentiable with respect to θ. Now, however, Φ_ε is only known to have a minimum at $\theta = 0$ over the set $0 \leqslant \theta \leqslant \theta_0$ for some $\theta_0 > 0$. Hence, $\Phi_\varepsilon'(0) \geqslant 0$. Thus, we now get

$$\Phi_\varepsilon'(0) = L(\varphi_\varepsilon, v_\varepsilon)(\eta) = L(\varphi_\varepsilon, v_\varepsilon)(z\xi) \geqslant 0 \tag{6.4}$$

for all z. From this relation it is possible to conclude, with the help of

(5.8) and some further manipulation, that

$$L(\varphi_\varepsilon, v_\varepsilon)(z\xi) = \int_0^1 \tilde{\varrho}(\varepsilon, t) z'(t)\,dt \geqslant 0, \tag{6.5}$$

where $\tilde{\varrho}$ is a nonincreasing function defined as follows:

$$\tilde{\varrho}(\varepsilon, t) = \tilde{\psi}(\varepsilon, t)\xi(t) - \int_0^t \{[f_1^0(\varepsilon, s) + 2(\varphi_\varepsilon' - \varphi_0')f_1(\varepsilon, s) - \\ - \tilde{\psi}(\varepsilon, s) f_1(\varepsilon, s)]\xi(s) + \tilde{\psi}(\varepsilon, s)\xi'(s)\}\,ds + d(\varepsilon). \tag{6.6}$$

Here $\tilde{\psi}$ is defined by (5.7). Note that $\tilde{\psi}$ does not satisfy (5.9), even though ψ and $\tilde{\psi}$ are both defined by (5.7).

The relations (6.4) and (6.5) are not entirely satisfactory in that they involve an inequality instead of an equality. The inequality is a consequence of using variations that go in the direction of decreasing G at a point $\varphi(t)$. That is, the variations "point into" to constraint set. A condition involving equality can be obtained by using variations h that "move along" surfaces of constant G. Thus if $\varphi(t)$ is on the boundary of the constraint set, i.e., $G(\varphi(t)) = 0$, then the point $\varphi(t) + h(t)$ will "stay on the boundary". This result is contained in the next lemma.

LEMMA 6.2. *Let the function ζ be in $W^{1,2}(I)$ and let*

$$h(t) = \zeta(t) - \big(\nabla G(\varphi_\varepsilon(t)) \cdot \zeta(t)\big)\nabla G(\varphi_\varepsilon(t))/\|\nabla G(\varphi_\varepsilon(t))\|^2. \tag{6.7}$$

Then

$$L(\varphi_\varepsilon, v_\varepsilon)(h) = \frac{d}{d\theta} F\big(\varphi_\varepsilon + \theta h, v_\varepsilon, N(\varepsilon)\big)\big|_{\theta=0} = 0.$$

The proof is carried out by adding a penalty term to F as follows. Define

$$F^*\big(\varphi, v_\varepsilon, N(\varepsilon), K\big) = K \int_0^1 \omega\Big(G\big(\varphi(t)\big)\Big)dt + F\big(\varphi, v_\varepsilon, N(\varepsilon)\big),$$

where ω is as in Section 2. For fixed K consider the problem of minimizing F^* over all φ in $W^{1,2}(I)$ such that $\|\varphi - x_0\| \leqslant \varepsilon$ and $\|\varphi' - \varphi_0'\|_2 \leqslant \varepsilon$.

If there exists an integer $K(\varepsilon)$ such that for all φ satisfying the preceding inequalities the following holds:

$$F^*\big(\varphi, v_\varepsilon, N(\varepsilon), K(\varepsilon)\big) \geqslant F\big(\varphi_\varepsilon, v_\varepsilon, N(\varepsilon)\big),$$

then since $F\big(\varphi_\varepsilon, v_\varepsilon, N(\varepsilon)\big) = F^*\big(\varphi_\varepsilon, v_\varepsilon, N(\varepsilon), K(\varepsilon)\big)$, the function φ_ε furnishes the required minimum and is an interior point. Straightforward differentiation and the special form of h give the desired result.

If no such integer $K(\varepsilon)$ exists, then the argument is rather tricky and cannot be summarized here. The reader is referred to Medhin [7], [8] for details.

We shall obtain the desired "ε-Euler equation" by using an arbitrary variation ζ in $W^{1,2}(I)$ with $\zeta(0) = \zeta(1) = 0$. We write $\zeta = h - y$, where h is defined as in (6.7) and

$$y(t) = -\big(\nabla G(\varphi_\varepsilon(t))\cdot\zeta(t)\big)\nabla G(\varphi_\varepsilon(t))/\|\nabla G(\varphi_\varepsilon(t))\|^2.$$

The function y is a variation of the form (6.3) with $z(t) = \big(\nabla G(\varphi_\varepsilon(t))\cdot\zeta(t)\big)$. Thus we have written an arbitrary ζ as the sum of two "orthogonal" variations, one "along a surface of constant G" and another directed "into the constraint set".

From Lemma 6.2 and from (5.9) we have

$$L(\varphi_\varepsilon, v_\varepsilon)(h) = L(\varphi_\varepsilon, v_\varepsilon)(\zeta + y) = L(\varphi_\varepsilon, v_\varepsilon)(\zeta) + L(\varphi_\varepsilon, v_\varepsilon)(y) = 0.$$

Since $y(t) = z(t)\xi(t)$ with $z(t) = \nabla G(\varphi_\varepsilon(t))\cdot\zeta(t)$, it follows from the equality in (6.5) that

$$\text{(6.8)}\quad L(\varphi_\varepsilon, v_\varepsilon)(\zeta) + \int_0^1 \tilde{\varrho}(\varepsilon, t)\,\nabla G(\varphi_\varepsilon(t))\cdot\zeta'(t)\,dt + \\ + \int_0^1 \tilde{\varrho}(\varepsilon, t)\varphi_\varepsilon'(t)\cdot\partial_x\nabla G(\varphi_\varepsilon(t))\cdot\zeta(t)\,dt = 0,$$

where $\partial_x(\nabla G(x))$ is the jacobian matrix of $\nabla G(x)$, or the Hessian matrix of G. Recall that equation (6.8) holds for all $\zeta \in W^{1,2}(I)$ satisfying $\zeta(0) = \zeta(1) = 0$.

Now substitute the defining expression for $L(\varphi_\varepsilon, v_\varepsilon)(\zeta)$ given by (5.8) into (6.8) and substitute the defining expression for $\tilde{\varrho}(\varepsilon, t)$ given by (6.6) into (6.8). Next integrate by parts and apply the fundamental lemma of the calculus of variations to get the "ε-Euler equation"

$$\text{(6.9)}\quad \tilde{\psi}(\varepsilon, t) + \tilde{\varrho}(\varepsilon, t)\,\nabla G(\varphi_\varepsilon(t)) = \int_0^t \{f^0(\varepsilon, s) - \tilde{\psi}(\varepsilon, s)f_1(\varepsilon, s) + \\ + 2(\varphi_\varepsilon' - \varphi_0')f_1(\varepsilon, s) + \tilde{\varrho}(\varepsilon, s)\varphi_\varepsilon'(s)\,\partial_x\nabla G(\varphi_\varepsilon(s))\}\,ds + C(\varepsilon).$$

From (6.9) it is possible to conclude that $\tilde{\varrho}(\varepsilon, \cdot)$ is bounded, and is constant on intervals (α, β) on which φ_ε is interior to the constraint set; i.e., intervals on which $G(\varphi_\varepsilon(t)) < 0$. By proper choice of $C(\varepsilon)$ we may take $\tilde{\varrho}(\varepsilon, t) \geqslant 0$. Recall that we have already shown that $\tilde{\varrho}(\varepsilon, \cdot)$ is nonincreasing.

If in (6.9) we choose $C(\varepsilon)$ appropriately, then divide through by an a ppropriate normalizing factor $M(\varepsilon)$ and define

$$\lambda(\varepsilon, t) = [\tilde{\psi}(\varepsilon, t) + \tilde{\varrho}(\varepsilon, t)\nabla G(\varphi_\varepsilon(t))]/M(\varepsilon),$$
$$\varrho(\varepsilon, t) = \tilde{\varrho}(\varepsilon, t)/M(\varepsilon),$$

then we may choose an appropriate sequence $\{\varepsilon_n\}$ such that $\varepsilon_n \to 0$ and such that we may pass to the limit in (6.9). We shall get the following relation in the limit

$$\lambda(t) = \int_0^t \{\lambda^0 f_1^0(s, \varphi_0(s), v_0(s)) - \lambda(s) f_1(s, \varphi_0(s), v_0(s)) + \\ + \varrho(s)\,\partial_x \nabla G(\varphi_0(s)) \cdot f(s, \varphi_0(s), v_0(s))\} ds + \lambda(0). \tag{6.10}$$

Here ϱ is a bounded, nonincreasing, nonnegative function which is obtained as the limit of $\tilde{\varrho}(\varepsilon, \cdot)/M(\varepsilon)$. Moreover, $\lambda^0 \geqslant 0$.

Using arguments similar to those used to establish the integrated form of the maximum principle in Theorem 5.1, we can establish a similar result here. Namely,

$$\int_0^1 -\lambda^0 f^0(t, \varphi_0'(t), v_0(t)) + [\lambda(t) - \varrho(t)\,\nabla G(\varphi_0(t))] \cdot f(t, \varphi_0(t), v_0(t))\, dt \\ \geqslant \int_0^1 -\lambda^0 f^0(t, \varphi_0(t), v(t)) + [\lambda(t) - \varrho(t)\,\nabla G(\varphi_0(t))] \cdot f(t, \varphi_0(t), v(t))\, dt \tag{6.11}$$

for all v in $\mathfrak{R}^*$.

In summary, we have the following result:

THEOREM 6.1. *Let* (φ_0, v_0) *minimize* $J(\varphi, v)$, *defined by* (5.1), *subject to* (5.2), (5.3) *and* (6.1). *Then there exists a constant* $\lambda^0 \geqslant 0$, *an absolutely continuous function* λ *defined on* I, *and a function* ϱ *defined on* I *such that the following hold*:

(i) $\lambda^0 + \varrho(0) > 0$;

(ii) *The function* ϱ *is nonnegative, nonincreasing, continuous from the right on* $(0, 1)$ *and is constant on any subinterval on which* $G(\varphi_0(t)) < 0$.

(iii) *Relations* (6.10) *and* (6.11) *hold.*

In [7] and [8] a more general problem with intermediate and terminal constraints and with additional vector constraints of the form

$$P(t, \varphi(t), v(t)) = 0, \quad R(t, \varphi(t), v(t)) \leqslant 0 \quad \text{and} \quad Q(t, v(t)) = 0$$

is treated. The intermediate and terminal constraints are of the form

$$M_1(\varphi^1(\tau_0), \ldots, \varphi^n(\tau_0), \ldots, \varphi^1(\tau_{\sigma+1}), \ldots, \varphi^n(\tau_{\sigma+1})) = 0,$$
$$M_2(\varphi^1(\tau_0), \ldots, \varphi^n(\tau_0), \ldots, \varphi^1(\tau_{\sigma+1}), \ldots, \varphi^n(\tau_{\sigma+1})) \leqslant 0,$$

where φ^i, $i = 1, \ldots, n$ is the ith coordinate function of φ and $0 = \tau_0 < \tau_1 < \ldots < \tau_{\sigma+1} = 1$ is a partition of $[0, 1]$.

7. Hereditary optimal control problems

In this section we summarize the treatment of the relaxed hereditary optimal control problems given by Bates [2], who applied the penalty technique to this problem. For full details see Bates [2].

Let I denote the unit interval $[0,1]$ and let $\tilde{I}$ denote the interval $[-r,1]$, where r is a given positive number. Let $C(\tilde{I})$ denote the metric space of continuous functions on $\tilde{I}$ with range in R^n and with supremum norm. Let $\mathfrak{U}$ denote the class of all Borel measurable functions defined on $\tilde{I}$ with range in R^m.

Let the following be given:

(i) real valued functions $h^0, h^1, \ldots, h^n$ defined on $I \times C(\tilde{I}) \times R^m$,

(ii) real valued functions $g^0, g^1, \ldots, g^n$ defined on $I \times \tilde{I} \times R^n \times R^m$, and

(iii) real valued functions $\omega^0, \omega^1, \ldots, \omega^n$ defined and measurable on $\tilde{I}$ for every fixed $t \in I$.

For any (t, φ, z) in $I \times C(\tilde{I}) \times R^m$ the notation $h^i(t, \varphi(\cdot), z)$ will indicate that the value of h^i depends on t, z, and some or all of the values $\varphi(s)$, $-r \leqslant s \leqslant t$.

Define functions f^i, $i = 0, 1, \ldots, n$ on $I \times C(\tilde{I}) \times \mathfrak{U}$ as follows:

$$f^i(t, \varphi(\cdot), u(\cdot)) = h^i(t, \varphi(\cdot), u(t)) + \int_{-r}^{t} d_s \omega^i(t, s) g^i(t, s, \varphi(s), u(s)),$$

where the integral is a Lebesgue–Stieltjes integral. The notation again indicates that the value of the function depends on some or all values of $\varphi(s)$ and $u(s)$, where $-r \leqslant s \leqslant t$.

Let $\mathrm{AC}(\tilde{I})$ denote the subset of $C(\tilde{I})$ consisting of all absolutely continuous functions. Let $\mathfrak{S}$ be a given compact subset of $\mathrm{AC}([-r, 0])$. Let x_0 and x_1 be fixed points in R^n and let Ω be a fixed compact subset of R^m. A pair (φ, u) in $C(\tilde{I}) \times \mathfrak{U}$ is called *admissible* if the following conditions are satisfied:

(i) $\varphi'(t) = f(t, \varphi(\cdot), u(\cdot))$ for a.e. t in I;

(ii) $\varphi(t) = \sigma(t)$ for each t in $[-r, 0]$, for some $\sigma \in \mathfrak{S}$;

(iii) $\varphi(0) = x_0$, $\varphi(1) = x_1$;

(iv) $u(t) \in \Omega$ for a.e. t in $\tilde{I}$.

The hereditary optimal control problem is the following: Minimize

$$J(\varphi, u) = \int_0^1 f^0(t, \varphi(\cdot), u(\cdot))\,dt$$

over all admissible (φ, u).

In [2], condition (iii) is replaced by the more general condition, $(\varphi(0), \varphi(1)) \in \mathfrak{B}$, where $\mathfrak{B}$ is some fixed set in R^m.

As in the ordinary control problem and in the bounded state problem, the penalty method requires the introduction of relaxed controls. The definition of a relaxed control is unchanged. The relaxed hereditary control problem is the following one.

Minimize

$$J(\varphi, v) = \int_0^1 f^0\big(t, \varphi(\cdot), v(\cdot)\big)dt \tag{7.1}$$

over elements (φ, v) in $\big(W^{1,2}(I)\cap C(\tilde{I})\big)\times\mathfrak{R}^*$ subject to

$$\varphi'(t) = f\big(t, \varphi(\cdot), v(\cdot)\big) \quad \text{for a.e. } t\in I, \tag{7.2}$$

$$\varphi(t) = \sigma(t), \quad -r\leqslant t\leqslant 0 \quad \text{for some } \sigma\in\mathfrak{S}, \tag{7.3}$$

$$\varphi(0) = x_0, \quad \varphi(1) = x_1, \tag{7.4}$$

where for $i = 0, 1, \ldots, n$,

$$\begin{aligned} f^i\big(t, \varphi(\cdot), v(\cdot)\big) &= h^i\big(t, \varphi(\cdot), v(t)\big) + \int_{-r}^{t} d_s\omega^i(t, s)g^i\big(t, s, \varphi(s), v(s)\big) \\ &= \int_\Omega h^i\big(t, \varphi(\cdot), z\big)d\mu(t, z) + \\ &\quad + \int_{-r}^{t} d\omega^i(t, s)\int_\Omega g^i\big(t, s, \varphi(s), z\big)d\mu(s, z). \end{aligned} \tag{7.5}$$

In [2] it is shown that this problem has a solution under rather mild assumptions on the data of the problem.

We now state the assumptions under which the necessary conditions are derived.

ASSUMPTION 1. The function $\bar{h} = (h^0, h)$ is differentiable with respect to φ for fixed (t, z). Both $\bar{h}$ and its Fréchet derivative with respect to φ, $d\bar{h}$, are Borel measurable in t for fixed (φ, z) in $C(\tilde{I})\times R^m$ and are continuous in (φ, z) for fixed t in I. For every compact subset Γ of R^n there exists a function μ in $L_2[I]$ such that for any ψ in $C(\tilde{I})$ and any (t, φ, z) in $I\times C(\tilde{I}, \Gamma)\times\Omega$,

$$\begin{aligned} &\|\bar{h}\big(t, \varphi(\cdot), z\big)\| \leqslant \mu(t), \\ &\big\|d\bar{h}\big(t, \varphi(\cdot), z; \psi\big)\big\| \leqslant \mu(t)\|\psi\|_\infty, \end{aligned} \tag{7.6}$$

where $C(\tilde{I}, \Gamma)$ denotes the space of continuous functions on $\tilde{I}$ with range in Γ. Condition (7.6) is the analogue of (3.5).

ASSUMPTION 2. The function $\bar{g}$: $(t, s, y, z)\to\bar{g}(t, s, y, z)$ is continuous on $I\times\tilde{I}\times R^n\times R^m$ and is continuously differentiable with respect to y in R^n.

ASSUMPTION 3. Each function ω^i, $i = 0, 1, \ldots, n$, is measurable on $I \times \tilde{I}$, and for each $t \in I$, $\omega^i(t, \cdot)$ is of bounded variation on $\tilde{I}$. Each $\omega^i(t, \cdot)$ is continuous from the right as a function of s, and vanishes for $s \geqslant t$. For each t in I, let $V\omega^i(t, \cdot)$ denote the total variation of $\omega^i(t, \cdot)$ on $\tilde{I}$. Then there exists a function $\mu_1 \in L_2[I]$ such that $V\omega^i(t, \cdot) \leqslant \mu_1(t)$ for a.e. t in I.

For each $i = 0, 1, \ldots, n$ and $t \in I$, let $\omega^i(t, \cdot) = \omega_1^i(t, \cdot) - \omega_2^i(t, \cdot)$, where ω_1^i and ω_2^i are monotone functions of s. For any Borel set $E \subset \tilde{I}$ define

$$\gamma_j^i(E) = \int_0^1 \int_E d_s \omega_j^i(t, s) dt, \qquad j = 1, 2; \qquad i = 1, \ldots, n.$$

ASSUMPTION 4. For any Borel set $E \subset I$ of Lebesgue measure zero, $|\gamma^i|(E) = 0$, where $|\cdot|$ denotes the total variation measure and $\gamma^i = \gamma_1^i - \gamma_2^i$, $i = 1, \ldots, n$.

For fixed (t, φ, z), $d\bar{h}$ is a linear functional on $C(\tilde{I})$. By virtue of (7.6) $d\bar{h}$ is bounded for a.e. $t \in I$. Hence for a.e. t in I and each (φ, z) in $C(\tilde{I}) \times E^m$, $d\bar{h}\big(t, \varphi(\cdot), z; \cdot\big)$ is an element in the dual of $C(\tilde{I})$. Therefore, by the Riesz representation theorem, there exists a function Γ_1 defined on $\tilde{I}$, of bounded variation and continuous from the right, such that for each ψ in $C(\tilde{I})$,

$$d\bar{h}\big(t, \varphi(\cdot), z; \psi\big) = \int_{-r}^{t} d_s \Gamma_1(t, \varphi, z; s)\psi(s).$$

Here $\bar{h} = (h^0, h^1, \ldots, h^n)$ and $\Gamma_1 = (\Gamma_1^0, \Gamma_1^1 \ldots \Gamma_1^n)$. By Γ_1 being of bounded variation we mean that each component is of bounded variation. Since $\bar{h}$ does not depend on $\varphi(s)$ for $s > t$, we must have $\Gamma_1(t, \varphi, z; s)$ equal to a constant for $s \geqslant t$. If we take the value of the constant to be zero then $\Gamma_1(t, \varphi, z; \cdot)$ is uniquely determined. Hence we may write

$$d\bar{h}\big(t, \varphi(\cdot), z; \psi\big) = \int_{-r}^{t} d_s \Gamma_1(t, \varphi, z; s)\psi(s).$$

It follows from Assumption 1 that Γ_1 is a measurable function of t for each fixed (φ, z, s) and is a continuous function of (φ, z) for each fixed (t, s). Hence, for any relaxed control v, we may now write

$$\Gamma_1(t, \varphi, v; s) = \int_{\Omega} \Gamma_1(t, \varphi, z; s) d\mu(t, z).$$

Let $\varpi = (\omega^0, \omega^1, \ldots, \omega^n) = (\omega^0, \omega)$. By virtue of the continuity and differentiability properties of $\bar{g}$, the mapping

$$(t, \varphi, v) \to \int_{-r}^{t} d_s \varpi(t, s) \bar{g}\big(t, s, \varphi(s), v(s)\big)$$

has a Fréchet derivative with respect to φ whose action on ψ in $C(\tilde{I})$ is given by

$$\int_{-r}^{t} d_s\varpi(t,s)\bar{g}_y\big(t,s,\varphi(s),v(s)\big)\psi(s),$$

for any (t,φ,v) in $I\times C(\tilde{I})\times\mathfrak{R}^*$. Here $\bar{g}_y$ denotes the jacobian matrix of $\bar{g}$ with respect to y.

Define

$$\Gamma_2(t,\varphi,v;s) = -\int_{s}^{t} d_\tau(t,\tau)\bar{g}_y\big(t,\tau,\varphi(\tau),v(\tau)\big)$$

for $-r\leqslant s<t$ and equal to zero for $s\geqslant t$. Define $\Gamma=\Gamma_1+\Gamma_2$. Then the Fréchet derivative of $\bar{f}$ is given by

$$d\bar{f}\big(t,\varphi(\ ),v(\ );\psi\big) = \int_{-r}^{t} d_s\Gamma(t,\varphi,v;s)\psi(s)$$

for any (t,φ,v) in $I\times C(\tilde{I})\times\mathfrak{R}^*$ and any ψ in $C(\tilde{I})$.

The maximum principle for the relaxed hereditary optimal control problem has the following form.

THEOREM 7.1. *Let (φ_0, v_0) be optimal for the relaxed hereditary control problem. Then there exists a vector function $\bar{\lambda}=(\lambda^0,\lambda)=(\lambda^0,\lambda^1,\ldots,\lambda^n)$ of bounded variation on I such that*

(i) *$\lambda^0\leqslant 0$ is constant;*

(ii) *λ is continuous at $t=1$ and $\big(\lambda^0,\lambda(1)\big)\neq 0$;*

(iii) *λ satisfies the integral equation*

$$\lambda(s)+\int_{s}^{1}\bar{\lambda}(t)\Gamma(t,s)dt = \lambda(1).$$

for each s in I, where $\Gamma(t,s)=\Gamma(t,\varphi_0,v_0;s)$;

(iv) *For every relaxed control v,*

$$\int_0^1\bar{\lambda}(t)\bar{f}\big(t,\varphi_0(\cdot),v_0(\cdot)\big)dt\geqslant\int_0^1\bar{\lambda}(t)\bar{f}\big(t,\varphi_0(\cdot),v(\cdot)\big)dt.$$

The general outline of the proof of this theorem should be clear to the reader by now. We shall merely review the major steps and refer the reader to [2] for the details, which require careful attention.

First we define a penalty function on $\big(W^{1,2}(I)\cap C(\tilde{I})\big)\times Z$ as follows:

$$F(\varphi,v,N) = J(\varphi,v)+\|\varphi'-\varphi_0'\|_2^2+\|\varphi(0)-\varphi_0(0)\|^2+\varepsilon\|v-v_0\|_L+ +N\big\|\varphi'-f\big(t,\varphi(\cdot),v(\cdot)\big)\big\|_2^2.$$

We again show that for all $0 < \varepsilon \leqslant \varepsilon_1$, where $\varepsilon_1 > 0$ is fixed, there is an $N(\varepsilon)$ such that $F(\varphi, v, N(\varepsilon))$ attains its minimum at an interior point of the set $\mathfrak{D}(\varepsilon)$ defined by (5.5). Denote the interior minimum by $(\varphi_\varepsilon, v_\varepsilon)$. Since $(\varphi_\varepsilon, v_\varepsilon)$ is an interior point of $\mathfrak{D}(\varepsilon)$, we can obtain a necessary condition satisfied by $(\varphi_\varepsilon, v_\varepsilon)$ by rather elementary calculations. We then let $\varepsilon \to 0$. The pair $(\varphi_\varepsilon, v_\varepsilon)$ will tend to (φ_0, v_0) in an appropriate sense, and the necessary condition obtained for the ε-problem will tend to the necessary condition of Theorem 7.1.

References

[1] A. V. Balakrishnan, *A computational approach to the maximum principle*, J. Comput. System Sci. 5 (1971), 163–191.

[2] G. R. Bates, *Hereditary optimal control problems*, Ph. D. Thesis, Purdue University, 1977.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer-Verlag, New York 1974.

[4] —, *A penalty function proof of the maximum principle*, Appl. Math. Optim. 2 (1976), 291–303.

[5] R. Courant, *Calculus of Variations and Supplementary Notes and Exercises Revised and Amended by J. Moser*, New York University, 1956–57.

[6] O. L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York 1969.

[7] N. G. Medhin, *Necessary conditions for optimal control problems with bounded state by a penalty method*, Ph. D. Thesis, Purdue University, 1980.

[8] —, *Necessary conditions for optimal control problems with bounded state by a penalty method*, SIAM J. Control Optimization, to appear.

[9] E. J. McShane, *The Lagrange multiplier rule*, Amer. Math. Monthly 80 (1973), 922–925.

[10] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York 1972.