Alberto Bressan
Andrea Marson

A maximum principle for optimally controlled systems of conservation laws


<http://www.numdam.org/item?id=RSMUP_1995__94__79_0>

ALBERTO BRESSAN - ANDREA MARSON (*)

ABSTRACT - We study a class of optimization problems of Mayer form, for the strictly hyperbolic nonlinear controlled system of conservation laws
\[ u_t + [F(u)]_x = h(t, x, u, z), \]
where \( z = z(t, x) \) is the control variable. Introducing a family of «generalized cotangent vectors», we derive necessary conditions for a solution \( \hat{u} \) to be optimal, stated in the form of a Maximum Principle.

1. Introduction.

This paper is concerned with a class of optimization problems for a strictly hyperbolic system of conservation laws with distributed control, in one space dimension:

\[ u_t + [F(u)]_x = h(t, x, u, z), \quad u(0, x) = \overline{u}(x). \]

Here \((t, x) \in [0, T] \times \mathbb{R}\), while \( u \in \mathbb{R}^m \) is the state variable, and the control \( z = z(t, x) \) varies inside an admissible set \( Z \subset \mathbb{R}^p \). Given a smooth function \( V: \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R} \), we consider the optimization problem

\[ \max_{z \in Z} J(u(z)), \]

where \( Z \) is the family of all measurable control functions taking values inside \( Z \), \( u(z) \) is the solution of (1.1) corresponding to the control \( z \), and

(*) Indirizzo degli AA.: S.I.S.S.A., Via Beirut 4, Trieste 34014, Italy.
J is a functional which depends on the terminal values of $u$:

$$J(u) = \int_{-\infty}^{\infty} V(x, u(T, x)) \, dx.$$  

Necessary conditions will be derived, in order that a control function $\tilde{z} = \tilde{z}(t, x)$ be optimal for the problem (1.1)-(1.2). The key for obtaining such conditions is to understand how the values $u(t, x)$ of the solution of (1.1) are affected, if the control $z$ is varied in the neighborhood of any given point $(t_0, x_0)$.

Assuming that the optimal solution $\hat{u}$ is piecewise Lipschitz with finitely many lines of discontinuity, the behavior of a slightly perturbed solution $u^\epsilon$ can be described using the calculus for first order generalized tangent vectors developed in [2]. In this paper, we introduce a class of «generalized cotangent vectors» and derive an adjoint system of linear equations and boundary conditions, determining how these covectors are transported backward in time along $\hat{u}$. We then prove a necessary condition for the optimality of a sufficiently regular control $\tilde{z}$, stated in the form of a Maximum Principle.

The main technical problem arising in the proof is the fact that the transport equations for tangent vectors can be justified only under the a-priori assumption that all perturbed solutions $u^\epsilon$ remain piecewise Lipschitz continuous, with the same number of jumps as $\hat{u}$. Therefore, when a family $\{z^\epsilon\}$ of control variations is constructed, it is essential to check that the corresponding solutions $u^\epsilon = u(z^\epsilon)$ do not develop a gradient catastrophe before the terminal time $T$. For this reason, strong regularity assumptions on the optimal control $\tilde{z}$ and on the optimal solution $\hat{u}$ will be used. We conjecture that these requirements could be considerably relaxed.

Our main theorem, stated in § 5 in the form of a Maximum Principle, covers the case of an optimal solution with finitely many, non-intersecting lines of discontinuity. In the light of the analysis in [2], it is expected that similar results should be valid also in the case of interacting shocks.

2. Basic assumptions and notations.

In the following, $|\cdot|$ and $\langle \cdot, \cdot \rangle$ denote the Euclidean norm and inner product on $\mathbb{R}^n$, respectively. We first consider the unperturbed system of conservation laws

$$u_t + [F(u)]_x = 0,$$

$$u(0, x) = \overline{u}(x),$$
under the basic hypotheses

\((H1)\) The set \(\Omega \subset \mathbb{R}^m\) is open and convex, \(F: \Omega \to \mathbb{R}^m\) is a \(C^1\) vector field. The system is strictly hyperbolic, and each characteristic field is either linearly degenerate or genuinely nonlinear.

For the basic theory of discontinuous solutions of conservative systems, we refer to \([5, 6, 7, 8]\).

We denote by \(\lambda_i(u), r_i(u), l_i(u)\) respectively the \(i\)-th eigenvalue and \(i\)-th right and left eigenvector of the Jacobian matrix \(A(u) = DF(u)\), normalized so that

\[ |r_i(u)| \equiv 1, \quad \langle l_i(u), r_j(u) \rangle \equiv \delta_{ij}, \]

where \(\delta_{ij}\) is the Kronecker symbol. For \(u, u' \in \Omega\), define the averaged matrix

\[(2.3)\]

\[ A(u, u') = \int_0^1 A(\theta u + (1 - \theta)u') d\theta. \]

Clearly \(A(u, u') = A(u', u)\) and \(A(u, u) = A(u)\). For \(i = 1, \ldots, m\), the \(i\)-th eigenvalue and eigenvectors of \(A(u, u')\) will be denoted by \(\lambda_i(u, u'), r_i(u, u'), l_i(u, u')\). We assume that the ranges of the eigenvalues \(\lambda_i\) do not overlap, i.e. that there exist disjoint intervals \([\lambda_i^-, \lambda_i^+]\), such that

\[ \lambda_i(u, u') \in [\lambda_i^-, \lambda_i^+], \quad \forall u, u' \in \Omega, \quad i \in \{1, \ldots, m\}. \]

Because of the regularity of \(A\), it is possible to choose \(r_i, l_i\) to be \(C^1\) functions of \(u, u'\), normalized according to

\[ |r_i(u, u')| \equiv 1, \quad \langle l_i(u, u'), r_j(u, u') \rangle \equiv \delta_{ij}. \]

If \(\phi\) is any function defined on \(\Omega\), its directional derivative along \(r_i\) at \(u\) is denoted by

\[ r_i \cdot \phi(u) \equiv [\nabla \phi(u)] r_i(u) = \lim_{\varepsilon \to 0} \frac{\phi(u + \varepsilon r_i(u)) - \phi(u)}{\varepsilon}. \]

For the differential of the \(i\)-th eigenvalue of the matrix \(A\) in (2.3) we write

\[ D\lambda_i(u^+, u^-) \cdot (v^+, v^-) \equiv \lim_{\varepsilon \to 0} \frac{\lambda_i(u^+ + \varepsilon v^+, u^- + \varepsilon v^-) - \lambda_i(u^+, u^-)}{\varepsilon}. \]
A similar notation is used for the differentials of the right and left eigenvectors of $A$.

For each $k \in \{1, \ldots, m\}$, we assume that either the $k$-th characteristic field is genuinely nonlinear and

$$
\lambda_k(u^+) + \varepsilon_1 |u^+ - u^-| < \lambda_k(u^+, u^-) < \lambda_k(u^-) - \varepsilon_1 |u^+ - u^-|
$$

for some $\varepsilon_1 > 0$ and all $u^+, u^- \in \Omega$ connected by an admissible shock of the $k$-th family, or else that the $k$-th characteristic field is linearly degenerate, so that $r_k \cdot \lambda_k(u) \equiv 0$ and

$$
\lambda_k(u^+) = \lambda_k(u^+, u^-) = \lambda_k(u^-)
$$

whenever $u^+$ and $u^-$ are connected by a contact discontinuity of the $k$-th family.

For every fixed $k \in \{1, \ldots, m\}$, the couples of states $u^+, u^-$ which are connected by a shock of the $k$-th characteristic family can be determined by the system of $m - 1$ equations

\begin{equation}
\langle l_i(u^+, u^-), u^+ - u^- \rangle = 0 \quad i \neq k.
\end{equation}

Differentiating (2.4) w.r.t. $u^+, u^-$, one obtains the system

\begin{equation}
\Phi_i(u^-, u^+, w^-, w^+) = 0 \quad i \neq k,
\end{equation}

where

$$
\Phi_i(u^-, u^+, w^-, w^+) \equiv \sum_{j=1}^{m} \langle Dl_i(u^+, u^-) \cdot (w_j^+ r_j^+, w_j^- r_j^-), u^+ - u^- \rangle + \sum_{j=1}^{m} \langle l_i(u^+, u^-), w_j^+ r_j^+ - w_j^- r_j^- \rangle.
$$

To express the general solution of (2.5), define the sets $\mathcal{I}$ and $\mathcal{O}$ (incoming and outgoing) of signed indices

\begin{equation}
\mathcal{I} \equiv \{i^+; i \leq k\} \cup \{i^-; i \geq k\},
\end{equation}

\begin{equation}
\mathcal{O} \equiv \{j^-; j < k\} \cup \{j^+; j > k\},
\end{equation}

if the $k$-th characteristic field is genuinely nonlinear, while

\begin{equation}
\mathcal{I} \equiv \{i^+; i < k\} \cup \{i^-; i > k\},
\end{equation}

in the linearly degenerate case. Observe that the system of $n - 1$ scalar equations (2.7) is linear homogeneous w.r.t. $w^-, w^+$, with coefficients
which depend continuously on $u^-, u^+$. When $u^- = u^+$ one has
\[
\frac{\partial \Phi_i}{\partial w_j^\pm} = \pm \delta_{ij}.
\]
Therefore, if $u^-$ and $u^+$ are sufficiently close to each other, one has
\[
(2.9) \quad \det \left( \frac{\partial \Phi_i(u^-, u^+, w^-, w^+)}{\partial w_j^\pm} \right) \neq 0 \quad (i \neq \overline{k}, \ j^\pm \in \mathcal{O}).
\]
In turn, when the $(n - 1) \times (n - 1)$ determinant in (2.9) does not vanish, one can solve (2.5) for the $n - 1$ outgoing variables $w_j^\pm$, $j^\pm \in \mathcal{O}$:
\[
(2.10) \quad w_j^\pm = W_j(u^-, u^+) (w^3) \quad j \neq \overline{k}.
\]
Here $w^3$ denotes the set of $n + 1$ incoming variables $\{w_i^\pm; i^\pm \in \mathcal{S}\}$. We remark that, in the case where the $k$-th characteristic field is linearly degenerate, one has
\[
(2.11) \quad \frac{\partial \Phi_i}{\partial w_k^\pm} \equiv 0,
\]
hence all functions $W_j^\pm$ do not depend on $w_k^+, w_k^-$. This is consistent with our definition (2.8) of incoming waves.

Next, consider the perturbed system
\[
(2.12) \quad u_t + [F(u)]_x = h(t, x, u),
\]
where $h$ is a continuously differentiable function of its arguments. We say that $u = u(t, x)$ is a piecewise $C^1$ solution of (2.12) if there exists finitely many $C^1$ curves
\[
\gamma_a = \{(t, x); x = x_a(t), t \in [t_a', t_a'']\}
\]
in the $t$-$x$-plane, such that

(i) The function $u$ is a continuously differentiable solution of (2.12) on the complement of the curves $\gamma_a$. 

(ii) Along each curve $x = x_a(t)$, the right and left limits

$$
\begin{align*}
    u(t, x_a^+) &= \lim_{x \to x_a(t)^+} u(t, x), \\
    u_x(t, x_a^+) &= \lim_{x \to x_a(t)^+} u_x(t, x), \\
    t \in [t'_a, t''_a],
\end{align*}
$$

exist and remain uniformly bounded. Moreover, the usual Rankine-Hugoniot and the entropy admissibility conditions hold.

For the uniqueness of solutions of (2.12) within this class of functions, we refer to [3, 4, 10]. We say that $u$ has a weak discontinuity along $x_a$ if $u_x$ is discontinuous but the function $u$ itself is continuous at each point $(t, x_a(t))$. In the case $u(t, x_a^+) \neq u(t, x_a^-)$, we say that $u$ has a strong discontinuity, or a jump, at $x_a$.

3 - Generalized tangent vectors.

Let $u : [a, b] \mapsto \mathbb{R}^n$ be a piecewise Lipschitz continuous function with discontinuities at points $x_1 < \ldots < x_N$. Following [2], we define the space $T_u$ of generalized tangent vectors to $u$ as the Banach space $L^1 \times \mathbb{R}^n$. On the family $\Sigma_u$ of all continuous paths $\gamma : [0, \varepsilon_0] \mapsto L^1$ with $\gamma(0) = u$ (with $\varepsilon_0 > 0$ possibly depending on $\gamma$), consider the equivalence relation $\sim$ defined by

$$
\gamma \sim \gamma' \iff \lim_{\varepsilon \to 0} \frac{\|\gamma(\varepsilon) - \gamma'(\varepsilon)\|_{L^1}}{\varepsilon} = 0.
$$

We say that a continuous path $\gamma \in \Sigma_u$ generates the tangent vector $(v, \xi) \in T_u$ if $\gamma$ is equivalent to the path $\gamma(v, \xi; u)$ defined as

$$
\gamma(v, \xi; u)(\varepsilon) = u + \varepsilon v + \sum_{\xi_a < 0} (u(x_a^+) - u(x_a^-)) \chi_{[x_a + \varepsilon \xi_a, x_a]} -
\sum_{\xi_a > 0} (u(x_a^+) - u(x_a^-)) \chi_{[x_a, x_a + \varepsilon \xi_a]}.
$$

Up to higher order terms, $\gamma(\varepsilon)$ is thus obtained from $u$ by adding $\varepsilon v$ and shifting the points $x_a$, where the discontinuities of $u$ occur, by $\varepsilon \xi_a$. In order to derive an evolution equation satisfied by these tangent vectors, one needs to consider more regular paths $\gamma \in \Sigma_u$, taking values inside the set of all piecewise Lipschitz functions.
DEFINITION 1. In connection with the system (2.12), we say that a function $u : \mathbb{R} \to \mathbb{R}^n$ is in the class PLSD of Piecewise Lipschitz functions with Simple Discontinuities if it satisfies the following conditions.

(i) $u$ has finitely many discontinuities, say at $x_1 < x_2 < \ldots < x_N$, and there exists a constant $L$ such that
\begin{equation}
|u(x) - u(x')| \leq L|x - x'| \tag{3.3}
\end{equation}
whenever the interval $[x, x']$ does not contain any point $x_a$.

(ii) Each jump of $u$ consists of a contact discontinuity or of a single, stable shock. More precisely, for every $\alpha \in \{1, \ldots, N\}$, there exists $k_\alpha \in \{1, \ldots, m\}$ such that
\begin{equation}
\langle l_i (u^+, u^-), u^+ - u^- \rangle = 0 \quad \forall i \neq k_\alpha, \tag{3.4}
\end{equation}
\begin{equation}
u_+ \neq u_-, \quad \lambda_{k_\alpha} (u^+) \leq \lambda_{k_\alpha} (u^+, u^-) \leq \lambda_{k_\alpha} (u^-), \tag{3.5}
\end{equation}
where $u^+, u^-$ denote respectively the right and left limits of $u(x)$ as $x \to x_\alpha$.

DEFINITION 2. Let $u$ be a PLSD function. A path $\gamma \in \Sigma_u$ is a Regular Variation (R.V.) for $u$ if, for $\epsilon \in [0, \epsilon_0]$, all functions $u^\epsilon = u^\epsilon (t, x)$ are in PLSD, with jumps at points $x_1^\epsilon < \ldots < x_N^\epsilon$ depending continuously on $\epsilon$. They all satisfy Definition 1 with a Lipschitz constant $L$ independent of $\epsilon$.

For each $\epsilon \in [0, \epsilon_0]$, let $u^\epsilon = u^\epsilon (t, x)$ be a piecewise $C^1$ solution of the system (2.12), with jumps at $x_1^\epsilon (t) < \ldots < x_N^\epsilon (t)$. Assume that, at some initial time $t$, the family $u^\epsilon (t, \cdot)$ is a R.V. of $u^0 (t, \cdot)$, generating the tangent vector $(\tilde{v}, \tilde{\xi})$. Then, as long as the discontinuities in $u^\epsilon$ do not interact and the Lipschitz constants of the $u^\epsilon$ (outside the jumps) remain uniformly bounded, for $t > t$ the family $u^\epsilon (t, \cdot)$ is still a R.V. of $u^0 (t, \cdot)$ and generates a tangent vector $(v(t, \cdot), \xi(t))$. According to Theorem 2.2 in [2], this vector can be determined as the unique broad solution with initial condition $(v, \xi)(t) = (\tilde{v}, \tilde{\xi}(t))$ of the linear system
\begin{equation}
v_t + A(u) v_x + [DA(u) \cdot v] u_x = h_u (t, x, u) v \tag{3.6}
\end{equation}
outside the discontinuities of $u$, coupled with the boundary conditions
\begin{equation}
\langle Dl_i (u^+, u^-) \cdot (\xi_a u_x^+ + v^+, \xi_a u_x^- + v^-), (u^+ - u^-) \rangle +
\langle l_i (u^+, u^-), \xi_a u_x^+ + v^+ - \xi_a u_x^- v^- \rangle = 0, \quad \forall i \neq k_\alpha, \tag{3.7}
\end{equation}
\begin{equation}
\dot{\xi}_a = D\lambda_{k_\alpha} (u^+, u^-) \cdot (\xi_a u_x^+ + v^+, \xi_a u_x^- + v^-), \tag{3.8}
\end{equation}

A maximum principle for optimally controlled systems etc. 85
along each line \( x = x_{\alpha}(t) \) where \( u \) suffers a discontinuity in the \( k_{\alpha} \)th characteristic family. We recall that a broad solution of a semilinear hyperbolic system is a locally integrable function whose components satisfy the appropriate integral equations along almost all characteristics. See [1, 6] for details.

For future applications, it is convenient to derive a version of (3.6)-(3.8) involving the components \( u^+_x = \langle l_i(u), u_x \rangle, v_i = \langle l_i(u), v \rangle \). Differentiating w.r.t. \( \varepsilon \) the equation

\[
A(u + \varepsilon v)u_x = \sum_{i=1}^{n} \lambda_i(u + \varepsilon v)\langle l_i(u + \varepsilon v), u_x \rangle r_i(u + \varepsilon v),
\]

one obtains

\[
(3.9) \quad [DA(u) \cdot v] u_x = \sum_{i,j} (r_j \cdot \lambda_i) u_x^j v_j r_i + \sum_{i,j} \lambda_i (r_j \cdot l_i, u_x) v_j r_i + \sum_{i,j} \lambda_i u_x^j (r_j \cdot l_i) v_j.
\]

Using (3.9) together with the relations

\[
l_{i,t} = \sum_j (r_j \cdot l_i) \left( -\lambda_j u^j_x + \langle l_j, h \rangle \right),
\]

\[
l_{i,x} = \sum_j (r_j \cdot l_i) u^j_x, \quad \lambda_{i,x} = \sum_j (r_j \cdot \lambda_i) u^j_x,
\]

\[
\langle r_j \cdot l_i, r_k \rangle + \langle l_i, r_j \cdot r_k \rangle = r_j \cdot \langle l_i, r_k \rangle \equiv 0,
\]

multiplying (3.6) on the left by \( l_i \) we find

\[
(3.10) \quad (v_i)_t + (\lambda_i v_i)_x + \sum_{k,i} (r_k \cdot \lambda_i) \{ u^i_x v_k - u^k_x v_i \} + \sum_j \langle l_j, [r_j, r_k] \rangle (\alpha_i - \lambda_j) u^j_x v_k =
\]

\[
= -\sum_j \langle l_j, r_j \cdot r_k \rangle \cdot \langle l_j, h \rangle v_k + \sum_k \langle l_i, r_k \cdot h \rangle v_k \quad (i = 1, \ldots, m).
\]

Here \([r_j, r_k] \equiv r_j \cdot r_k - r_k \cdot r_j\) denotes the Lie bracket of the vector fields \( r_j, r_k \).

Concerning the equations (3.7)-(3.8), for each fixed \( \alpha \) call \( u^-, u^+ \) the limits of \( u(t, x) \) as \( x \to x_{\alpha}(t) \) from the left and from the right, respectively. Similarly, define the components \( v_i^\pm = \langle l_i(u^\pm), v^\pm \rangle \), so that \( v^+ = \sum r^+_i v^+_i, v^- = \sum r^-_i v^-_i \). Comparing (3.7) with (2.5), where \( k = k_{\alpha} \), it follows that if (2.9) holds then, for any fixed values \( v_i^\pm (i^\pm \in \mathcal{I}) \) of the incoming components, the linear equations (3.7) can be uniquely solved
for the \( m - 1 \) outgoing components:

\begin{equation}
(3.11) \quad v_j^\pm = V^j_{\alpha}(v^\alpha, \xi_\alpha) \quad j^\pm \in \mathcal{O}.
\end{equation}

Observe that the \( V^j_{\alpha} \) are linear homogeneous functions of \( \xi_\alpha \) and of the incoming variables \( v^\alpha \). In turn, inserting these values in (3.8), one obtains an expression for the time derivatives

\begin{equation}
(3.12) \quad \dot{\xi}_\alpha = \Psi^\alpha_{\alpha}(v^\alpha, \xi_\alpha).
\end{equation}

4 - The adjoint equations.

Let the function \( u : \mathbb{R} \to \mathbb{R}^m \) be piecewise Lipschitz continuous with \( N \) points of jump. We then define the space of \textit{generalized cotangent vectors} (or adjoint vectors) to \( u \) as the Banach space \( T^*_u = L^\infty(\mathbb{R}) \times \mathbb{R}^N \). Elements of \( T^*_u \) will be written as \((v^*, \xi^*)\) and regarded as row vectors.

Given a piecewise Lipschitz solution \( u = u(t, x) \) of (2.12), with jumps along the lines \( x = x_\alpha(t), \alpha = 1, \ldots, N \), we shall derive an adjoint system of linear equations on \( T^*_u \) whose solutions \((v^*(t, \cdot), \xi^*(t))\) have the property that the duality product

\begin{equation}
(4.1) \quad \langle (v^*, \xi^*), (v, \xi) \rangle = \int v^*(t, x) \cdot v(t, x) \, dx + \sum_{a=1}^N \xi^*_\alpha(t) \xi_\alpha(t)
\end{equation}

remains constant in time, for every solution \((v, \xi)\) of the linear system (3.6)-(3.8).

Assume that (4.1) holds for every solution \( v \) of (3.6) which vanishes on a neighborhood of all lines \( x = x_\alpha(t) \). Then an integration by parts shows that, away from the discontinuities of \( u \), the function \( v^* \) must satisfy

\begin{equation}
(4.2) \quad v^*_t + v^*_x A(u) + v^* \overline{DA}(u) u_x = -v^* \cdot h_u(t, x, u),
\end{equation}

where, referred to a standard basis \( \{e_1, \ldots, e_m\} \) of \( \mathbb{R}^m \), \( \overline{DA}(u) u_x \) is the \( m \times m \) matrix whose \((j, i)\) entry is

\[ [\overline{DA}(u) u_x]_{ji} = \sum_{k=1}^m \left( \frac{\partial A_{jk}(u)}{\partial u_k} - \frac{\partial A_{j}(u)}{\partial u_i} \right) \partial u_k. \]

In order to formulate also a suitable set of boundary conditions, valid along the lines \( x = x_\alpha(t) \), it is convenient to work with the components \( u^\alpha_i = \langle l_i(u), u_x \rangle, \quad v_i^* = \langle v^*, r_i(u) \rangle \). For each fixed \( \alpha \), we shall write \( \lambda_i(u^+) = \lambda_i(u(x_\alpha +)) \) and \( \lambda_i(u^-) = \lambda_i(u(x_\alpha -)) \) for the the \( i \)-th characteristic speeds to the right and to the left of the \( \alpha \)-th discontinuity, respectively. Similarly, we write \( v_i^* = v_i^*(x_\alpha +), \quad v_i^* = v_i^*(x_\alpha -) \). In the
following, \( V_t^j, \Psi_a \) are the linear homogeneous functions introduced at (3.11)-(3.12).

**Proposition 1.** Let \( u \) be a piecewise \( C^1 \) solution of the hyperbolic system (2.12), with jumps occurring along the (nonintersecting) lines \( x = x_a(t) \). Assume that the map \( t \mapsto (v^*(t, \cdot), \xi^*(t)) \in T_u^* \), with \( v^* = \sum l_i(u) v_i^* \), provides a solution to the linear system

\[
(v_i^*)_t + \lambda_i(v_i^*)_x =
\]

\[
= \sum_{k \neq i} [(r_i \cdot \lambda_k) u_x^k v_i^* - (r_k \cdot \lambda_i) u_x^k v_i^*] + \sum_{j \neq k} \langle l_k, [r_j, r_i] \rangle (\lambda_k - \lambda_j) u_x^j v_k^* + \sum_{j, k} \langle l_k, r_j \cdot r_i \rangle \langle l_j, h \rangle v_k^* - \sum_{k} \langle l_k, r_i \cdot h \rangle v_k^*
\]

outside the lines where \( u \) is discontinuous, together with the equations

\[
\dot{\xi}_i^* = -\xi_i^* \cdot \frac{\partial \Psi_a}{\partial \xi_a} - \sum_{j \in \mathcal{O}} |\lambda_j(u^\pm) - \dot{x}_a| |v_j^*| \cdot \frac{\partial V_i^j}{\partial v_i^j},
\]

\[
v_i^* = \frac{1}{|\lambda_i(u^\pm) - \dot{x}_a|} \left\{ \xi_i^* \cdot \frac{\partial \Psi_a}{\partial v_i^j} + \sum_{j \in \mathcal{O}} |\lambda_j(u^\pm) - \dot{x}_a| |v_j^*| \cdot \frac{\partial V_i^j}{\partial v_i^j} \right\},
\]

along each line \( x = x_a(t) \). Then, for every solution \((v, \xi)\) of (3.6)-(3.8), the product (4.1) remains constant in time.

**Proof.** For notational convenience, we set \( x_0(t) = -\infty, x_{N+1}(t) = +\infty \). Integrating each component \( v_i^* v_i \) along the corresponding characteristic lines \( x = \lambda_i(u) \), the time derivative of (4.1) can be computed as

\[
\frac{d}{dt} \left[ \int \sum_i v_i^* v_i \, dx + \sum_a \xi_a^* \xi_a \right] =
\]

\[
= \sum_{a=1}^{N} x_{a+1}(t) - x_a(t) \sum_{a} [(v_i^* v_i)_t + (\lambda_i(u) v_i^* v_i)_x] \, dx +
\]

\[
+ \sum_a \left[ \sum_{j \in \mathcal{O}} |\lambda_j(u^\pm) - \dot{x}_a| |v_j^*| v_j^* - \sum_{i \in \mathcal{I}} |\lambda_i(u^\pm) - \dot{x}_a| |v_i^*| v_i^* \right] + \sum_a (\xi_a^* \xi_a + \xi_a^* \xi_a^*).\]
From (4.3) and (3.10), a straightforward computation shows that
\begin{equation}
\sum_i (v_i^* v_i)_t + \sum_i (\lambda_i(u) v_i^* v_i)_x = 0.
\end{equation}

Therefore, all the integrals on the right hand side of (4.6) equal zero.

Next, we observe that, for each \( \alpha \), the functions \( V^i_\alpha, \Psi_\alpha \) in (3.11)-(3.12) are linear homogeneous w.r.t. the independent variables \( \xi_\alpha, v_i^z, i^z \in \mathfrak{I} \). Therefore, we can write
\begin{equation}
\begin{aligned}
\dot{\xi}_\alpha = \frac{\partial \Psi_\alpha}{\partial \xi_\alpha} \cdot \dot{\xi}_\alpha + \sum_{i^z \in \mathfrak{I}} \frac{\partial \Psi_\alpha}{\partial v_i^z} \cdot v_i^z, \\
v_{j^z} = \frac{\partial V^j_\alpha}{\partial \xi_\alpha} \cdot \dot{\xi}_\alpha + \sum_{i^z \in \mathfrak{I}} \frac{\partial V^j_\alpha}{\partial v_i^z} \cdot v_i^z, & \quad j^z \in \mathfrak{O}.
\end{aligned}
\end{equation}

From (4.6), using (4.8) and factoring out the terms \( \xi_\alpha, v_i^z \), we obtain
\begin{equation}
\frac{d}{dt} \left[ \int \sum_i v_i^* v_i \, dx + \sum_a \xi^*_\alpha \xi_\alpha \right] = \sum_a \sum_{i^z \in \mathfrak{I}} \cdot \\
\cdot \left[ \sum_{j^z \in \mathfrak{O}} \left| \lambda_j(u^z) - \dot{x}_\alpha \right| v_{j^z}^* \cdot \frac{\partial V^j_\alpha}{\partial v_i^z} - \left| \lambda_i(u^z) - \dot{x}_\alpha \right| v_i^* + \xi^*_\alpha \frac{\partial \Psi_\alpha}{\partial v_i^z} \right] \cdot v_i^z + \\
+ \sum_a \left[ \xi^*_\alpha + \xi^* \cdot \frac{\partial \Psi_\alpha}{\partial \xi_\alpha} + \sum_{j^z \in \mathfrak{O}} \left| \lambda_j(u^z) - \dot{x}_\alpha \right| v_{j^z}^* \cdot \frac{\partial V^j_\alpha}{\partial \xi_\alpha} \right] \cdot \xi_\alpha = 0,
\end{equation}

because of (4.4), (4.5). This proves Proposition 1.

REMARK 1. The equations (4.4)-(4.5) determine the incoming variables \( v_i^*, i^z \in \mathfrak{I} \), in terms of the outgoing variables \( v_j^*, j^z \in \mathfrak{O} \). Therefore, the Cauchy problem for the adjoint linear system (4.8)-(4.5) is well posed if one assigns the terminal values \( (v^*(T, \cdot), \xi^*(T)) \) and seeks a solution defined backward in time.

REMARK 2. If, at \( x_\alpha \), the jump of \( u \) consists of a contact discontinuity in the \( k_\alpha \)-th characteristic family, then the equations (4.5) determine only the \( m - 1 \) incoming components \( v_i^*, i^z \in \mathfrak{I} \), with \( \mathfrak{I} \) defined by (2.14). In this case, the equations (4.9) still hold, because the functions \( \Psi_\alpha, V^j_\alpha \) do not depend on \( v_{k^z_\alpha} \).
5 - A Maximum Principle.

Consider again the optimization problem (1.2) for the system (1.1). We assume that $F$ satisfies the basic hypotheses (H1) in § 2 and that the functions $h = h(t, x, u, z)$ and $V = V(x, u)$ in (1.1), (1.3) are continuously differentiable. Let $\hat{u}$ be an optimal solution, corresponding to the control $\hat{z}$. In order to derive necessary conditions on $\hat{u}$, we shall construct a family of controls $\{z^\varepsilon; \varepsilon \in [0, \varepsilon_0]\}$, obtained by changing the values of $\hat{z}$ in a neighborhood of a given point $(t_0, x_0)$. We then study how the corresponding solution $u^\varepsilon$ behaves at the terminal time $T$.

By the results in [2], the change in $\hat{u}(T, \cdot)$ can be described up to first order in terms of a generalized tangent vector, provided that all solutions $u^\varepsilon$ remain piecewise Lipschitz continuous, with the same number of discontinuities. To ensure this condition, some stronger regularity assumption on the solution $\hat{u}$ will be used. Namely

(H2) The function $\hat{u} = \hat{u}(t, x)$ is piecewise $C^1$ on $[0, T] \times \mathbb{R}$, with finitely many, noninteracting jumps, say at

$$x_1(t) < \ldots < x_N(t), \quad t \in [0, T].$$

Any two weak discontinuities of $\hat{u}$ can interact with these jumps only at distinct points.

Otherwise stated, if $x = x_a(t)$ is the location of a jump in $\hat{u}$ and $y_i(t)$, $y_j(t)$ denote the position of two weak discontinuities (where $\hat{u}$ is continuous but $\hat{u}_x$ jumps), then there exists no time $\tau$ such that

$$x_a(\tau) = y_i(\tau) = y_j(\tau), \quad y_i(t) < y_j(t) \quad \text{for} \quad t < \tau.$$

In the following, $\nabla_u V$ denotes the gradient of $V = V(x, u)$ w.r.t. $u$, while the jump of $V$ at the point $(T, x_a(T))$ is written

$$\Delta V(x_a(T)) = \lim_{x \to x_a(T) (+)} V(x, \hat{u}(T, x)) - \lim_{x \to x_a(T) (-)} V(x, \hat{u}(T, x)).$$

**Theorem 1 (Maximum Principle).** In connection with the optimization problem (1.1)-(1.3), let the functions $h$, $V$ be continuously differentiable and let $F$ satisfy the basic hypotheses (H1). Let $\hat{z} = \hat{z}(t, x)$ be a $C^1$ optimal control, and assume that the corresponding optimal solution $\hat{u} = \hat{u}(t, x)$ of (1.1) is piecewise $C^1$ and satisfies the additional regularity assumptions (H2).

Define the adjoint vector $(v^*, \xi^*)$ as the solution of the linear system (4.3)-(4.5), with terminal conditions:

(5.1) \[ v^*(T, x) = \nabla_u V(x, \hat{u}(T, x)) \]

(5.2) \[ \xi^*_a(T) = \Delta V(x_a(T)) \quad \alpha = 1, \ldots, N. \]
Then the maximality condition

\[(5.3) \quad v^*(t, x) \cdot h(t, x, \tilde{u}(t, x), \tilde{z}(t, x)) = \max_{z \in Z} v^*(t, x) \cdot h(t, x, \tilde{u}(t, x), z)\]

holds at each point \((t, x)\) where both \(v^*\) and \(\tilde{u}\) are continuous.

**Proof.** 1) If the conclusion of the theorem fails, then in the \(t-x\)-plane there exists a point \((\tau, \eta)\) where \(v^*, \tilde{u}\) are continuous, such that

\[(5.4) \quad v^*(\tau, \eta) \cdot h(\tau, \eta, \tilde{u}(\tau, \eta), \tilde{z}(\tau, \eta)) < v^*(\tau, \eta) \cdot h(\tau, \eta, \tilde{u}(\eta, \eta), z^\delta),\]

for some admissible control value \(z^\delta \in Z\).

By continuity, and by possibly changing the value of \(\eta\), we can choose \(\delta > 0\) such that \(\tilde{u}\) is \(C^1\) on a neighborhood of the segment

\[S = \{(t, x); t = \tau, x \in [\eta - \delta, \eta + \delta]\}\]

and, in addition,

\[(5.5) \quad v^*(\tau, x) \cdot h(\tau, x, \tilde{u}(\tau, x), \tilde{z}(\tau, x)) < v^*(\tau, x) \cdot h(\tau, x, \tilde{u}(\tau, x), z^\delta),\]

\[\forall x \in [\eta - \delta, \eta + \delta].\]

2) We now construct a family of piecewise \(C^1\) control variations \(z^\varepsilon\) as follows. Choose a \(C^\infty\) function \(\varphi : \mathbb{R} \rightarrow [0, 1]\) whose support is precisely the interval \([-1, 1]\). For \(\varepsilon > 0\) small, define the open domain

\[\Omega_\varepsilon = \left\{(t, x); t - \varepsilon \varphi\left(\frac{x - \eta}{\delta}\right) < t < \tau\right\}\]

and the control function

\[z^\varepsilon(t, x) = \begin{cases} z^\delta & \text{if } (t, x) \in \Omega_\varepsilon, \\ \tilde{z}(t, x) & \text{if } (t, x) \notin \Omega_\varepsilon.\end{cases}\]

Call \(u^\varepsilon\) the corresponding solution of (1.1).

3) For each \(\varepsilon \geq 0\) sufficiently small, the curve

\[(5.6) \quad \gamma^\varepsilon : x \mapsto (\tau - \varepsilon \varphi((x - \eta)/\delta), x)\]

is space-like, and crosses all characteristics transversally. Hence the solution \(u^\varepsilon\) is well defined and the map \(x \mapsto u^\varepsilon(\tau, x)\) is \(C^1\) on a neighbor-
hood of the interval \([\eta - \delta, \eta + \delta]\). Moreover, as \(\varepsilon \to 0\), one has

\[
(5.7) \quad u^\varepsilon \to \tilde{u}, \quad u^\varepsilon_x \to \tilde{u}_x,
\]

in the space \(L^\infty([0, r] \times \mathbb{R})\). As a consequence, the family \(\{u^\varepsilon(\tau, \cdot); \varepsilon \in [0, \varepsilon_0]\}\) is clearly a Regular Variation of \(\tilde{u}(\tau, \cdot)\). We claim that it generates the tangent vector \((v, \xi) \in \mathbb{L}^1 \times \mathbb{R}^N\), with

\[
(5.8) \quad \xi_a(\tau) = 0 \quad \forall a, \quad v(\tau, x) = 0 \quad \text{if} \quad |x - \eta| > \delta,
\]

\[
(5.9) \quad v(\tau, x) = [h(\tau, x, \tilde{u}(\tau, x), z^A) - h(\tau, x, \tilde{u}(\tau, x), \tilde{z}(\tau, x))] \cdot \varphi \left( \frac{x - \eta}{\delta} \right)
\]

\[
\text{if} \quad |x - \eta| \leq \delta.
\]

4) Since the curves (5.6) are space-like, for all \(\varepsilon \geq 0\) one has \(u^\varepsilon(\tau, x) = \tilde{u}(\tau, x)\) whenever \(|x - \eta| \geq \delta\). This clearly implies (5.8).

Observing that both \(u^\varepsilon\) and \(\tilde{u}\) are \(C^1\) on \(\partial_\varepsilon\), we can subtract the equations satisfied by \(u^\varepsilon\) and \(\tilde{u}\) one from the other, and obtain

\[
(5.10) \quad (u^\varepsilon - \tilde{u})_t = -A(u^\varepsilon)(u^\varepsilon_x - \tilde{u}_x) - \]

\[
-[A(u^\varepsilon) - A(\tilde{u})] \tilde{u}_x + h(t, x, u^\varepsilon, z^\varepsilon) - h(t, x, \tilde{u}, \tilde{z}) .
\]

Because of the uniform limits (5.7) and the fact that \(u^\varepsilon = \tilde{u}\) on the lower boundary \(\gamma^\varepsilon\) of \(\partial_\varepsilon\), from (5.10) it follows

\[
(5.11) \quad u^\varepsilon(\tau, x) - \tilde{u}(\tau, x) =
\]

\[
= \int_{\tau - \varepsilon \varphi((x - \eta)/\delta)}^{\tau} \{h(\tau, x, \tilde{u}(\tau, x), z^A) - h(\tau, x, \tilde{u}(\tau, x), \tilde{z}(\tau, x)) + \Phi_\varepsilon(t, x)\} \, dt ,
\]

with

\[
(5.12) \quad \lim_{\varepsilon \to 0} \sup_{(t, x) \in \partial_\varepsilon} |\Phi_\varepsilon(t, x)| = 0 .
\]
Together, (5.11) and (5.12) imply
\begin{align}
\lim_{\varepsilon \to 0} \frac{u^\varepsilon(\tau, x) - \tilde{u}(\tau, x)}{\varepsilon} = \\
= \left\{ h(\tau, x, \tilde{u}(\tau, x), z^h) - h(\tau, x, \tilde{u}(\tau, x), \tilde{z}(\tau, x)) \right\} \cdot \varphi \left( \frac{x - \eta}{\delta} \right),
\end{align}
uniformly for $x \in [\eta - \delta, \eta + \delta]$. This establishes (5.9).

5) The regularity assumptions (H2) on the optimal solution $\tilde{u}$ guarantee that the perturbations $u^\varepsilon$ all have the same number of lines of discontinuity, and that the derivatives $u_x^\varepsilon$ remain uniformly bounded, for $\varepsilon > 0$ suitably small.

By the results in [2], we conclude that, for all $t \in [\tau, T]$ the family $u^\varepsilon(t, \cdot)$ is a R.V. of $\tilde{u}(t, \cdot)$ which generates a tangent vector $(v, \xi(t))$. This vector is determined as the unique broad solution of the corresponding linear system (3.6)-(3.8). Using Proposition 1, together with (5.8)-(5.9) and then (5.5), we now compute
\begin{align}
\langle (v^*(T), \xi^*(T)), (v(T), \xi(T)) \rangle = \\
= \int_{\eta - \delta}^{\eta + \delta} \langle v^*(\tau), \xi^*(\tau) \rangle \cdot v^*(\tau, x) \cdot \\
\cdot \left\{ h(\tau, x, \tilde{u}(\tau, x), z^h) - h(\tau, x, \tilde{u}(\tau, x), \tilde{z}(\tau, x)) \right\} \cdot \varphi \left( \frac{x - \eta}{\delta} \right) \, dx > 0.
\end{align}

6) In order to derive a contradiction, it now suffices to interpret (5.14) at the light of the definitions (3.1), (3.2) and (5.2). Indeed, the regularity of the functions $V$ and $u^\varepsilon$ implies
\begin{align}
\int \left[ V(x, u^\varepsilon(T, x)) - V(x, \tilde{u}(T, x)) \right] \, dx = \\
= \varepsilon \cdot \left\{ \int \nabla V(x, \tilde{u}(T, x)) \cdot v(T, x) \, dx + \sum_a \Delta V(x_a(T)) \cdot \xi_a(T) \right\} + o(\varepsilon) = \\
= \varepsilon \cdot \langle (v^*(T), \xi^*(T)), (v(T), \xi(T)) \rangle + o(\varepsilon),
\end{align}
where $o(\varepsilon)$ denotes an infinitesimal of higher order w.r.t. $\varepsilon$.

By (5.14), for $\varepsilon > 0$ sufficiently small the quantity in (5.15) is strictly positive. This contradicts the optimality of $\tilde{u}$, proving the theorem.
REFERENCES


Manoscritto pervenuto in redazione il 4 marzo 1994.