Rendiconti del Seminario Matematico della Università di Padova

Alberto Bressan

Andrea Marson

A maximum principle for optimally controlled systems of conservation laws

Rendiconti del Seminario Matematico della Università di Padova, tome 94 (1995), p. 79-94

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

REND. SEM. MAT. UNIV. PADOVA, Vol. 94 (1995)

A Maximum Principle for Optimally Controlled Systems of Conservation Laws.

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:

$$u_t + [F(u)]_x = h(t, x, u, z), \qquad u(0, x) = \bar{u}(x). \tag{1.1}$$

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 $V: \mathbb{R} \times \mathbb{R}^m \mapsto \mathbb{R}$, we consider the optimization problem

$$\max_{z \in \mathcal{Z}} J(u(z)), \tag{1.2}$$

where $\mathcal{Z}$ 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:

$$J(u) \doteq \int_{-\infty}^{\infty} V(x, u(T, x))\, dx. \tag{1.3}$$

Necessary conditions will be derived, in order that a control function $\hat{z} = \hat{z}(t, x)$ 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^ε 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 $\hat{z}$, 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^ε remain piecewise Lipschitz continuous, with the same number of jumps as $\hat{u}$. Therefore, when a family $\{z^\varepsilon\}$ of control variations is constructed, it is essential to check that the corresponding solutions $u^\varepsilon = u(z^\varepsilon)$ do not develop a gradient catastrophe before the terminal time T. For this reason, strong regularity assumptions on the optimal control $\hat{z}$ 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 $\mathbb{R}^m$, respectively. We first consider the unperturbed system of conservation laws

$$u_t + [F(u)]_x = 0\,, \tag{2.1}$$

$$u(0, x) = \overline{u}(x)\,, \tag{2.2}$$

under the basic hypotheses

(H1) *The set $\Omega \subseteq \mathbb{R}^m$ is open and convex, $F: \Omega \mapsto \mathbb{R}^m$ is a $\mathcal{C}^1$ 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

$$|r_i(u)| \equiv 1, \qquad \langle l_i(u),\ r_j(u)\rangle \equiv \delta_{ij}\,,$$

where δ_{ij} is the Kronecker symbol. For $u, u' \in \Omega$, define the averaged matrix

$$A(u, u') = \int_0^1 A(\theta u + (1-\theta)u')\,d\theta\,. \tag{2.3}$$

Clearly $A(u, u') = A(u', u)$ and $A(u, u) = A(u)$. For $i = 1, \dots, m$, 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 λ_i do not overlap, i.e. that there exist disjoint intervals $[\lambda_i^-, \lambda_i^+]$, such that

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

Because of the regularity of A, it is possible to choose r_i, l_i to be $\mathcal{C}^1$ functions of u, u', normalized according to

$$|r_i(u, u')| \equiv 1, \qquad \langle l_i(u, u'),\ r_j(u, u')\rangle \equiv \delta_{ij}\,.$$

If ϕ is any function defined on Ω, its directional derivative along r_i at u is denoted by

$$r_i \bullet \phi(u) \doteq [\nabla\phi(u)]\,r_i(u) = \lim_{\varepsilon\to 0} \frac{\phi(u + \varepsilon r_i(u)) - \phi(u)}{\varepsilon}\,.$$

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

$$D\lambda_i(u^+, u^-)\cdot(v^+, v^-) \doteq \lim_{\varepsilon\to 0} \frac{\lambda_i(u^+ + \varepsilon v^+, u^- + \varepsilon v^-) - \lambda_i(u^+, u^-)}{\varepsilon}\,.$$

A similar notation is used for the differentials of the right and left eigenvectors of A.

For each $k \in \{1, \dots, m\}$, 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 $r_k \bullet \lambda_k(u) \equiv 0$ 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 $\bar{k} \in \{1, \dots, m\}$, the couples of states u^+, u^- which are connected by a shock of the $\bar{k}$-th characteristic family can be determined by the system of $m-1$ equations

$$\langle l_i(u^+, u^-),\ u^+ - u^- \rangle = 0 \qquad i \neq \bar{k}. \tag{2.4}$$

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

$$\Phi_i(u^-, u^+, w^-, w^+) = 0 \qquad i \neq \bar{k}, \tag{2.5}$$

where

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

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

$$\mathfrak{I} \doteq \{i^+; i \leq \bar{k}\} \cup \{i^-; i \geq \bar{k}\}, \tag{2.6}$$

$$\mathcal{O} \doteq \{j^-; j < \bar{k}\} \cup \{j^+; j > \bar{k}\}, \tag{2.7}$$

if the $\bar{k}$-th characteristic field is genuinely nonlinear, while

$$\mathfrak{I} \doteq \{i^+; i < \bar{k}\} \cup \{i^-; i > \bar{k}\}, \tag{2.8}$$

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

$$\det\left(\frac{\partial \Phi_i(u^-, u^+, w^-, w^+)}{\partial w_j^\pm}\right) \neq 0 \qquad (i \neq \bar{k},\ j^\pm \in \mathcal{O})\,. \tag{2.9}$$

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 $w_j^\pm$, $j^\pm \in \mathcal{O}$:

$$w_j^\pm = W_j(u^-, u^+)(w^{\mathfrak{I}}) \qquad j \neq \bar{k}\,. \tag{2.10}$$

Here $w^{\mathfrak{I}}$ denotes the set of $n+1$ incoming variables $\{w_i^\pm\,;\ i^\pm \in \mathfrak{I}\}$. We remark that, in the case where the $\bar{k}$-th characteristic field is linearly degenerate, one has

$$\frac{\partial \Phi_i}{\partial w_{\bar{k}}^\pm} \equiv 0\,, \tag{2.11}$$

hence all functions $W_{j^\pm}$ do not depend on $w_{\bar{k}}^+$, $w_{\bar{k}}^-$. This is consistent with our definition (2.8) of incoming waves.

Next, consider the perturbed system

$$u_t + [F(u)]_x = h(t, x, u)\,, \tag{2.12}$$

where h is a continuously differentiable function of its arguments. We say that $u = u(t, x)$ is a *piecewise* $\mathcal{C}^1$ solution of (2.12) if there exists finitely many $\mathcal{C}^1$ curves

$$\gamma_\alpha \doteq \{(t, x);\ x = x_\alpha(t),\ t \in [t'_\alpha, t''_\alpha]\}$$

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 γ_α.

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

$$\begin{cases} u(t, x_\alpha \pm) = \lim_{x \to x_\alpha(t) \pm} u(t, x), \\ u_x(t, x_\alpha \pm) = \lim_{x \to x_\alpha(t) \pm} u_x(t, x), \\ t \in]t'_\alpha, t''_\alpha[, \end{cases}$$

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_α if u_x is discontinuous but the function u itself is continuous at each point $(t, x_\alpha(t))$. In the case $u(t, x_\alpha +) \neq u(t, x_\alpha -)$, we say that u has a strong discontinuity, or a jump, at x_α.

3 - Generalized tangent vectors.

Let $u: [a, b] \mapsto \mathbb{R}^m$ 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 Σ_u of all continuous paths $\gamma: [0, \varepsilon_0] \mapsto L^1$ with $\gamma(0) = u$ (with $\varepsilon_0 > 0$ possibly depending on γ), consider the equivalence relation $\sim$ defined by

$$\gamma \sim \gamma' \Leftrightarrow \lim_{\varepsilon \to 0} \frac{\|\gamma(\varepsilon) - \gamma'(\varepsilon)\|_{L^1}}{\varepsilon} = 0. \tag{3.1}$$

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

$$\gamma_{(v, \xi; u)}(\varepsilon) = u + \varepsilon v + \sum_{\xi_\alpha < 0} (u(x_\alpha^+) - u(x_\alpha^-)) \chi_{[x_\alpha + \varepsilon\xi_\alpha, x_\alpha]} - \\ - \sum_{\xi_\alpha > 0} (u(x_\alpha^+) - u(x_\alpha^-)) \chi_{[x_\alpha, x_\alpha + \varepsilon\xi_\alpha]}. \tag{3.2}$$

Up to higher order terms, $\gamma(\varepsilon)$ is thus obtained from u by adding εv and shifting the points x_α, where the discontinuities of u occur, by $\varepsilon \xi_\alpha$. 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 $u:\mathbb{R}\mapsto\mathbb{R}^n$ 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

$$|u(x) - u(x')| \leq L|x - x'| \tag{3.3}$$

whenever the interval $[x, x']$ does not contain any point x_α.

(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

$$\langle l_i(u^+, u^-), u^+ - u^-\rangle = 0 \qquad \forall i \neq k_\alpha\,, \tag{3.4}$$

$$u^+ \neq u^-, \qquad \lambda_{k_\alpha}(u^+) \leq \lambda_{k_\alpha}(u^+, u^-) \leq \lambda_{k_\alpha}(u^-), \tag{3.5}$$

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 $\varepsilon \in [0, \varepsilon_0]$, all functions $u^\varepsilon \doteq \gamma(\varepsilon)$ are in PLSD, with jumps at points $x_1^\varepsilon < \ldots < x_N^\varepsilon$ depending continuously on ε. They all satisfy Definition 1 with a Lipschitz constant L independent of ε.

For each $\varepsilon \in [0, \varepsilon_0]$, let $u^\varepsilon = u^\varepsilon(t, x)$ be a piecewise $\mathcal{C}^1$ solution of the system (2.12), with jumps at $x_1^\varepsilon(t) < \ldots < x_n^\varepsilon(t)$. Assume that, at some initial time $\bar{t}$, the family $u^\varepsilon(\bar{t}, \cdot)$ is a R.V. of $u^0(\bar{t}, \cdot)$, generating the tangent vector $(\bar{v}, \bar{\xi})$. Then, as long as the discontinuities in u^ε do not interact and the Lipschitz constants of the u^ε (outside the jumps) remain uniformly bounded, for $t > \bar{t}$ the family $u^\varepsilon(t, \cdot)$ 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 $(v, \xi)(\bar{t}) = (\bar{v}, \bar{\xi})$ of the linear system

$$v_t + A(u)v_x + [DA(u)\cdot v]u_x = h_u(t, x, u)v \tag{3.6}$$

outside the discontinuities of u, coupled with the boundary conditions

$$\begin{aligned}\langle Dl_i(u^+, u^-)\cdot(\xi_\alpha u_x^+ + v^+, \xi_\alpha u_x^- + v^-), (u^+ - u^-)\rangle + \\ + \langle l_i(u^+, u^-), \xi_\alpha u_x^+ + v^+ - \xi_\alpha u_x^- v^-\rangle = 0, \qquad \forall i \neq k_\alpha\,,\end{aligned} \tag{3.7}$$

$$\dot{\xi}_\alpha = D\lambda_{k_\alpha}(u^+, u^-)\cdot(\xi_\alpha u_x^+ + v^+, \xi_\alpha u_x^- + v^-), \tag{3.8}$$

along each line $x = x_\alpha(t)$ where u suffers a discontinuity in the k_α-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 $u_x^i = \langle l_i(u), u_x\rangle$, $v_i = \langle l_i(u), v\rangle$. Differentiating w.r.t. ε 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

$$(3.9) \qquad [DA(u)\cdot v]\, u_x =$$

$$= \sum_{i,j} (r_j \bullet \lambda_i)\, u_x^i v_j r_i + \sum_{i,j} \lambda_i \langle r_j \bullet l_i, u_x\rangle v_j r_i + \sum_{i,j} \lambda_i u_x^i (r_j \bullet r_i)\, v_j .$$

Using (3.9) together with the relations

$$l_{i,t} = \sum_j (r_j \bullet l_i)(-\lambda_j u_x^j + \langle l_j, h\rangle),$$

$$l_{i,x} = \sum_j (r_j \bullet l_i)\, u_x^j, \qquad \lambda_{i,x} = \sum_j (r_j \bullet \lambda_i)\, u_x^j,$$

$$\langle r_j \bullet l_i, r_k\rangle + \langle l_i, r_j \bullet r_k\rangle = r_j \bullet \langle l_i, r_k\rangle \equiv 0,$$

multiplying (3.6) on the left by l_i we find

$$(3.10) \qquad (v_i)_t + (\lambda_i v_i)_x + \sum_{k \neq i} (r_k \bullet \lambda_i)\{u_x^i v_k - u_x^k v_i\} +$$

$$+ \sum_{j \neq k} \langle l_i, [r_j, r_k]\rangle(\lambda_i - \lambda_j)\, u_x^j v_k =$$

$$= -\sum_{j,k} \langle l_i, r_j \bullet r_k\rangle \cdot \langle l_j, h\rangle v_k + \sum_k \langle l_i, r_k \bullet h\rangle v_k \qquad (i = 1, \dots, m).$$

Here $[r_j, r_k] \doteq r_j \bullet r_k - r_k \bullet r_j$ denotes the Lie bracket of the vector fields r_j, r_k.

Concerning the equations (3.7)-(3.8), for each fixed α 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 $v_i^\pm \doteq \langle l_i(u^\pm), v^\pm\rangle$, so that $v^+ = \sum r_i^+ v_i^+$, $v^- = \sum r_i^- v_i^-$. Comparing (3.7) with (2.5), where $\bar{k} = k_\alpha$, it follows that if (2.9) holds then, for any fixed values $v_i^\pm$ ($i^\pm \in \mathfrak{I}$) of the incoming components, the linear equations (3.7) can be uniquely solved for the $m-1$ outgoing components:

$$v_j^{\pm} = V_\alpha^j(v^{\mathfrak{J}}, \xi_\alpha) \quad j^{\pm} \in \mathcal{O}. \tag{3.11}$$

Observe that the V_α^j are linear homogeneous functions of ξ_α and of the incoming variables $v^{\mathfrak{J}}$. In turn, inserting these values in (3.8), one obtains an expression for the time derivatives

$$\dot{\xi}_\alpha = \Psi_\alpha(v^{\mathfrak{J}}, \xi_\alpha). \tag{3.12}$$

4 - The adjoint equations.

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

Given a piecewise Lipschitz solution $u = u(t, x)$ of (2.12), with jumps along the lines $x = x_\alpha(t)$, $\alpha = 1, \dots, N$, 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

$$\langle (v^*, \xi^*), (v, \xi) \rangle \doteq \int v^*(t, x) \cdot v(t, x)\, dx + \sum_{\alpha=1}^{N} \xi_\alpha^*(t)\, \xi_\alpha(t) \tag{4.1}$$

remains constant in time, for every solution (v, ξ) 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

$$v_t^* + v_x^* A(u) + v^* \widetilde{DA}(u)\, u_x = -v^* \cdot h_u(t, x, u), \tag{4.2}$$

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

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

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 $u_x^i = \langle l_i(u), u_x \rangle$, $v_i^* = \langle v^*, r_i(u) \rangle$. For each fixed α, we shall write $\lambda_i(u^+) \doteq \lambda_i(u(x_\alpha +))$ and $\lambda_i(u^-) \doteq \lambda_i(u(x_\alpha -))$ for the the i-th characteristic speeds to the right and to the left of the α-th discontinuity, respectively. Similarly, we write $v_{i+}^* \doteq v_i^*(x_\alpha +)$, $v_{i-}^* \doteq v_i^*(x_\alpha -)$. In the following, V_α^j, Ψ_α are the linear homogeneous functions introduced at (3.11)-(3.12).

PROPOSITION 1. *Let u be a piecewise $\mathcal{C}^1$ solution of the hyperbolic system* (2.12), *with jumps occurring along the (nonintersecting) lines $x = x_\alpha(t)$. 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*

$$(4.3) \qquad (v_i^*)_t + \lambda_i (v_i^*)_x =$$

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

$$+ \sum_{j,k} \langle l_k, r_j \bullet r_i \rangle \langle l_j, h \rangle v_k^* - \sum_k \langle l_k, r_i \bullet h \rangle v_k^*$$

outside the lines where u is discontinuous, together with the equations

$$(4.4) \qquad \dot{\xi}_\alpha^* = -\xi_\alpha^* \cdot \frac{\partial \Psi_\alpha}{\partial \xi_\alpha} - \sum_{j^\pm \in \mathcal{O}} |\lambda_j(u^\pm) - \dot{x}_\alpha| \, v_{j^\pm}^* \cdot \frac{\partial V_\alpha^j}{\partial \xi_\alpha},$$

$$(4.5) \qquad v_{i^\pm}^* = \frac{1}{|\lambda_i(u^\pm) - \dot{x}_\alpha|} \left\{ \xi_\alpha^* \cdot \frac{\partial \Psi_\alpha}{\partial v_{i^\pm}} + \sum_{j^\pm \in \mathcal{O}} |\lambda_j(u^\pm) - \dot{x}_\alpha| \, v_{j^\pm}^* \cdot \frac{\partial V_\alpha^j}{\partial v_{i^\pm}} \right\}, \qquad i^\pm \in \mathfrak{I},$$

along each line $x = x_\alpha(t)$. Then, for every solution (v, ξ) 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 $\dot{x} = \lambda_i(u)$, the time derivative of (4.1) can be computed as

$$(4.6) \qquad \frac{d}{dt} \left[\int \sum_i v_i^* v_i \, dx + \sum_\alpha \xi_\alpha^* \xi_\alpha \right] =$$

$$= \sum_{\alpha=0}^{N} \sum_i \int_{x_\alpha(t)}^{x_{\alpha+1}(t)} [(v_i^* v_i)_t + (\lambda_i(u) v_i^* v_i)_x] \, dx +$$

$$+ \sum_\alpha \left[\sum_{j^\pm \in \mathcal{O}} |\lambda_j(u^\pm) - \dot{x}_\alpha| \cdot v_{j^\pm}^* v_{j^\pm} - \sum_{i^\pm \in \mathfrak{I}} |\lambda_i(u^\pm) - \dot{x}_\alpha| \cdot v_{i^\pm}^* v_{i^\pm} \right]$$

$$+ \sum_\alpha (\dot{\xi}_\alpha^* \xi_\alpha + \xi_\alpha^* \dot{\xi}_\alpha).$$

From (4.3) and (3.10), a straightforward computation shows that

$$\sum_i (v_i^* v_i)_t + \sum_i (\lambda_i(u) v_i^* v_i)_x = 0 . \tag{4.7}$$

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

Next, we observe that, for each α, the functions V_α^j, Ψ_α in (3.11)-(3.12) are linear homogeneous w.r.t. the independent variables ξ_α, $v_{i^\pm}$, $i^\pm \in \mathfrak{I}$. Therefore, we can write

$$\begin{cases} \dot{\xi}_\alpha = \dfrac{\partial \Psi_\alpha}{\partial \xi_\alpha} \cdot \xi_\alpha + \displaystyle\sum_{i^\pm \in \mathfrak{I}} \dfrac{\partial \Psi_\alpha}{\partial v_{i^\pm}} \cdot v_{i^\pm} , \\ v_{j^\pm} = \dfrac{\partial V_\alpha^j}{\partial \xi_\alpha} \cdot \xi_\alpha + \displaystyle\sum_{i^\pm \in \mathfrak{I}} \dfrac{\partial V_\alpha^j}{\partial v_{i^\pm}} \cdot v_{i^\pm} , \qquad j^\pm \in \mathcal{O} . \end{cases} \tag{4.8}$$

From (4.6), using (4.8) and factoring out the terms ξ_α, $v_{i^\pm}$, we obtain

$$\frac{d}{dt}\left[\int \sum_i v_i^* v_i \, dx + \sum_\alpha \xi_\alpha^* \xi_\alpha \right] = \sum_\alpha \sum_{i^\pm \in \mathfrak{I}} \cdot \tag{4.9}$$

$$\cdot \left[\sum_{j^\pm \in \mathcal{O}} |\lambda_j(u^\pm) - \dot{x}_\alpha| \, v_{j^\pm}^* \cdot \frac{\partial V_\alpha^j}{\partial v_{i^\pm}} - |\lambda_i(u^\pm) - \dot{x}_\alpha| \cdot v_{i^\pm}^* + \xi_\alpha^* \cdot \frac{\partial \Psi_\alpha}{\partial v_{i^\pm}} \right] \cdot v_{i^\pm} +$$

$$+ \sum_\alpha \left[\dot{\xi}_\alpha^* + \xi_\alpha^* \cdot \frac{\partial \Psi_\alpha}{\partial \xi_\alpha} + \sum_{j^\pm \in \mathcal{O}} |\lambda_j(u^\pm) - \dot{x}_\alpha| \, v_{j^\pm}^* \cdot \frac{\partial V_\alpha^j}{\partial \xi_\alpha} \right] \cdot \xi_\alpha = 0 ,$$

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

REMARK 1. The equations (4.4)-(4.5) determine the incoming variables $v_{i^\pm}^*$, $i^\pm \in \mathfrak{I}$, in terms of the outgoing variables $v_{j^\pm}^*$, $j^\pm \in \mathcal{O}$. Therefore, the Cauchy problem for the adjoint linear system (4.3)-(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_α, the jump of u consists of a contact discontinuity in the k_α-th characteristic family, then the equations (4.5) determine only the $m-1$ incoming components $v_{i^\pm}^*$, $i^\pm \in \mathfrak{I}$, with $\mathfrak{I}$ defined by (2.14). In this case, the equations (4.9) still hold, because the functions Ψ_α, V_α^j do not depend on $v_{k_\alpha^\pm}$.

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 $\hat{z}$, 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^ε 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^ε 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 $\mathcal{C}^1$ on $[0, T] \times \mathbb{R}$, with finitely many, noninteracting jumps, say at

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

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

Otherwise stated, if $x = x_\alpha(t)$ 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 τ such that

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

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 $(T, x_\alpha(T))$ is written

$$\Delta V\big(x_\alpha(T)\big) \doteq \lim_{x \to x_\alpha(T)+} V\big(x, \hat{u}(T, x)\big) - \lim_{x \to x_\alpha(T)-} V\big(x, \hat{u}(T, x)\big).$$

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 $\mathcal{C}^1$ optimal control, and assume that the corresponding optimal solution $\hat{u} = \hat{u}(t, x)$ of* (1.1) *is piecewise $\mathcal{C}^1$ and satisfies the additional regularity assumptions* (H2).

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

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

$$\xi_\alpha^*(T) = \Delta V(x_\alpha(T)) \qquad \alpha = 1, \ldots, N. \tag{5.2}$$

Then the maximality condition

$$v^*(t, x)\cdot h(t, x, \hat{u}(t, x), \hat{z}(t, x)) = \max_{z \in Z} v^*(t, x)\cdot h(t, x, \hat{u}(t, x), z) \tag{5.3}$$

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

PROOF. 1) If the conclusion of the theorem fails, then in the t-x-plane there exists a point (τ, η) where v^*, $\hat{u}$ are continuous, such that

$$v^*(\tau, \eta)\cdot h(\tau, \eta, \hat{u}(\tau, \eta), \hat{z}(\tau, \eta)) < v^*(\tau, \eta)\cdot h(\tau, \eta, \hat{u}(\tau, \eta), z^\natural), \tag{5.4}$$

for some admissible control value $z^\natural \in Z$.

By continuity, and by possibly changing the value of η, we can choose $\delta > 0$ such that $\hat{u}$ is $\mathcal{C}^1$ on a neighborhood of the segment

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

and, in addition,

$$v^*(\tau, x)\cdot h(\tau, x, \hat{u}(\tau, x), \hat{z}(\tau, x)) < v^*(\tau, x)\cdot h(\tau, x, \hat{u}(\tau, x), z^\natural) \tag{5.5}$$

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

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

$$\mathcal{D}_\varepsilon \doteq \left\{(t, x);\ \tau - \varepsilon\varphi\left(\frac{x - \eta}{\delta}\right) < t < \tau\right\}$$

and the control function

$$z^\varepsilon(t, x) \doteq \begin{cases} z^\natural & \text{if } (t, x) \in \mathcal{D}_\varepsilon, \\ \hat{z}(t, x) & \text{if } (t, x) \notin \mathcal{D}_\varepsilon. \end{cases}$$

Call u^ε the corresponding solution of (1.1).

3) For each $\varepsilon \geqslant 0$ sufficiently small, the curve

$$\gamma^\varepsilon : x \mapsto (\tau - \varepsilon\varphi((x - \eta)/\delta), x) \tag{5.6}$$

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

$$u^\varepsilon \to \widehat{u}, \qquad u_x^\varepsilon \to \widehat{u}_x, \tag{5.7}$$

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

$$\xi_\alpha(\tau) = 0 \quad \forall \alpha, \qquad v(\tau, x) = 0 \quad \text{if } |x-\eta| > \delta, \tag{5.8}$$

$$v(\tau, x) = [h(\tau, x, \widehat{u}(\tau, x), z^\natural) - h(\tau, x, \widehat{u}(\tau, x), \widehat{z}(\tau, x))]\cdot \varphi\left(\frac{x-\eta}{\delta}\right) \tag{5.9}$$

$$\text{if } |x-\eta| \leq \delta.$$

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

Observing that both u^ε and $\widehat{u}$ are $\mathcal{C}^1$ on $\mathcal{D}_\varepsilon$, we can subtract the equations satisfied by u^ε and $\widehat{u}$ one from the other, and obtain

$$(u^\varepsilon - \widehat{u})_t = -A(u^\varepsilon)(u_x^\varepsilon - \widehat{u}_x) - \tag{5.10}$$
$$-[A(u^\varepsilon) - A(\widehat{u})]\widehat{u}_x + h(t, x, u^\varepsilon, z^\varepsilon) - h(t, x, \widehat{u}, \widehat{z}).$$

Because of the uniform limits (5.7) and the fact that $u^\varepsilon = \widehat{u}$ on the lower boundary γ^ε of $\mathcal{D}_\varepsilon$, from (5.10) it follows

$$u^\varepsilon(\tau, x) - \widehat{u}(\tau, x) = \tag{5.11}$$
$$= \int_{\tau-\varepsilon\varphi((x-\eta)/\delta)}^{\tau} \{h(\tau, x, \widehat{u}(\tau, x), z^\natural) - h(\tau, x, \widehat{u}(\tau, x), \widehat{z}(\tau, x)) + \Phi_\varepsilon(t, x)\}\, dt,$$

with

$$\lim_{\varepsilon\to 0} \sup_{(t,x)\in \mathcal{D}_\varepsilon} |\Phi_\varepsilon(t, x)| = 0. \tag{5.12}$$

Together, (5.11) and (5.12) imply

$$(5.13) \qquad \lim_{\varepsilon \to 0} \frac{u^\varepsilon(\tau, x) - \widehat{u}(\tau, x)}{\varepsilon} =$$

$$= \{h(\tau, x, \widehat{u}(\tau, x), z^\natural) - h(\tau, x, \widehat{u}(\tau, x), \widehat{z}(\tau, x))\} \cdot \varphi\left(\frac{x - \eta}{\delta}\right),$$

uniformly for $x \in [\eta - \delta, \ \eta + \delta]$. This establishes (5.9).

5) The regularity assumptions (H2) on the optimal solution $\widehat{u}$ guarantee that the perturbations u^ε all have the same number of lines of discontinuity, and that the derivatives u_x^ε 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 $\widehat{u}(t, \cdot)$ 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

$$(5.14) \qquad \langle (v^*(T), \xi^*(T)), (v(T), \xi(T)) \rangle =$$

$$= \langle (v^*(\tau), \xi^*(\tau)), (v(\tau), \xi(\tau)) \rangle = \int_{\eta - \delta}^{\eta + \delta} v^*(\tau, x) \cdot$$

$$\cdot \| \{h(\tau, x, \widehat{u}(\tau, x), z^\natural) - h(\tau, x, \widehat{u}(\tau, x), \widehat{z}(\tau, x))\} \cdot \varphi\left(\frac{x - \eta}{\delta}\right) dx > 0 .$$

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^ε implies

$$(5.15) \qquad \int [V(x, u^\varepsilon(T, x)) - V(x, \widehat{u}(T, x))] \, dx =$$

$$= \varepsilon \cdot \left\{ \int \nabla_u V(x, \widehat{u}(T, x)) \cdot v(T, x) \, dx + \sum_\alpha \Delta V(x_\alpha(T)) \cdot \xi_\alpha(T) \right\} + o(\varepsilon) =$$

$$= \varepsilon \cdot \langle (v^*(T), \xi^*(T)), \ (v(T), \xi(T)) \rangle + o(\varepsilon),$$

where $o(\varepsilon)$ denotes an infinitesimal of higher order w.r.t. ε.

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