METHODS AND APPLICATIONS OF ANALYSIS.
Vol. 12, No. 2, pp. 103–134, June 2005
© 2005 International Press
001

DATA ASSIMILATION FOR CONSERVATION LAWS*

CLAUDE BARDOS† AND OLIVIER PIRONNEAU‡

Abstract. Data Assimilation is important in meteorology and oceanography, because it is a way to improve the models with newly measured data, statically or dynamically. It is a type of inverse problem for which the most popular solution method is least square with regularization and optimal control algorithms. As control theory assumes differentiability, there are mathematical difficulties when viscosity is neglected and the modeling uses a conservation law like the shallow water or Euler equations. In this paper we study the differentiated equations of some systems of conservation laws and show that Calculus of Variation can be applied in a formal and rigorous manner provided that principal values are defined at shocks and equations written in the sense of distribution theory. Numerical illustrations are given for the control of shocks for Burgers' equation and for the shallow water equations in one space dimension.

Thanks. We really appreciate the opportunity of dedicating this article to Joel Smoller as a token of long lasting friendship. In particular we are happy to show that a field were Joel brought important contributions is still very active at the interface between theory and application.

Key words. Inverse problems, data assimilation, optimal control, shocks

AMS subject classifications. Primarly 65M99; Secondary 93B40

1. Introduction.

Data Assimilation is important in meteorology, oceanography and climatology because it is a way to improve the models with newly measured data, statically or dynamically. The problem is well known in control and system theory and a large number of methods applied to environmental sciences are issued from Kalman filters, and system control. Yet data assimilation by least squares and optimal control is the most popular method, known in meteorology as 4d-var methods [26], [17], [28], [7] and [9]. Optimal control could be greatly simplified and the number of parameters could be greatly reduced if there was an explicit analysis of the meteorological fronts as unknowns (controls); but control theory assumes differentiability with respect to the control and there are new mathematical difficulties here in the absence of viscosity: if gradient methods are to be used computationally, one needs to establish that derivatives exists in some sense and that descent methods can use them.

To illustrate our point consider a system of conservation laws in $\mathbb{R}^n$ like

$$\partial_t u + \nabla \cdot F(u) = 0 \quad u(0) = u_0. \tag{1}$$

Data assimilation consists in finding u_0 in a set $\mathcal{U}$ so as to fit as best as possible observations u_d made at time T in a subset $D \subset \mathbb{R}^n$. The problem can be cast into the minimization of a functional

$$\min_{u_0 \in \mathcal{U}} E(u_0) = \frac{1}{2} \int_D |u(T) - u_d|^2 \quad : \quad \text{subject to (1).} \tag{2}$$

*Received November 11, 2004; accepted for publication June 23, 2005.
†LJLL (Laboratoire Jacques-Louis Lions), Université Paris VII, Boite courrier 187, 75252 Paris Cedex 05, France (bardos@math.jussieu.fr).
‡LJLL, Université Paris VI - IUF, 175 rue du Chevaleret 75013, France (Olivier.Pironneau@upmc.fr).

As for all optimal control problems one would like to use the Calculus of Variations, and write

$$\delta E = \int_{\mathbb{R}^n} (u(T) - u_d)\delta u(T)\,, \tag{3}$$

$$\partial_t \delta u + \nabla \cdot (F'_u \delta u) = o(|\delta u|) \tag{4}$$

and then introduce an adjoint state u^* solution of

$$\partial_t u^* + {F'_u}^T \nabla \cdot u^* = 0, \quad u^*(T) = I_D(u - u_d) \tag{5}$$

(where I_D is the indicator function of D) and show by an integration by parts that

$$\delta E = \int_{\mathbb{R}^n} u^*(0) \cdot \delta u(0) + o(\|\delta u(0)\|). \tag{6}$$

The validity of this calculus is usually taken for granted in engineering but it so happens that in the presence of shocks it is not correct.
Indeed take the simple case where $\mathcal{U}$ is the set of $x \to u_0(x, a)$ which is an indicator function of a ball of center $x(a)$ function of a parameter $a \in \mathbb{R}$. Then $\delta u = u'_a \delta a$ but u being discontinuous u' has a Dirac singularity if the shock position depends on a (see below). So right from the beginning the calculus above is wrong because $(u(T) - u_d)\delta u(T)$, for instance, is the product of a discontinuous function by a Dirac measure.

It was already noticed in [1] that (4) made sense even at the shock when understood as a distribution, namely

$$\int_{\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}} (\partial_t u' + \nabla \cdot (\tilde{F}'_u u'))w(x, t; a)\mathrm{d}x\mathrm{d}t\mathrm{d}a = 0 \quad \forall w \in \mathcal{D}(\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}).$$

An integration by parts in x, t, a puts all the derivatives on w and there is no problem for the definition of the integrals; however, because of the discontinuities, line integrals appear at the shocks involving jumps of F'_u so the pointwise value of F'_u at the shock must be defined, hence the notation $\tilde{F}'$ above. When F is quadratic, F' is linear and if $\tilde{F}'$ is the principal value of F', the identity above is true.

In this paper we wish to extend the formalism to systems and to apply it to optimal control and show that the hand-waving calculus of variation above is in fact correct for a class of flux functions F. This obviously required the proof of the differentiability of the solution in some suitable sense, then the construction of the linearized system, the construction of its adjoint and the proof of some duality formula. Formal differentiation of these conservation laws involves only differentiation in the sense of distributions and uses of the Rankine Hugoniot relation. On the other hand it is at the level of the linearized operator that the entropy condition plays a role,in particular in relation to the stability of the linearized system.

Perturbation of non linear hyperbolic problems raises several mathematical issues well explored in the major contributions of Majda [24] and [25].

First a small perturbation of a solution with one shock may in some cases generate solutions with several shocks. To avoid this phenomena it is assumed that the initial perturbation remains in the class of data $u + \delta u$ for which one has at time $t = 0$ the Rankine Hugoniot condition or more precisely, the existence of a scalar function σ defined on the shock curve such that (68) holds.

Second since the problem is non linear global in time estimates in most case are not available.

Keeping in mind earlier interesting results concerning the stability of a shock near the solution of a Riemann problem obtained by Li Da-quian and Yu Wen-ci [22] and [23] for local in time solutions of one dimensional systems of conservation laws, to the best of our knowledge, the result the most adapted to our purpose remains [24] Theorem 1 page 8 which provides the existence of multidimensional shock fronts for a finite time ($0 < t < T$).

Global in time stability results do exist but they are confined either to scalar equations (cf [6]) were no differentiability property is proven or to system in one space variable (cf. for instance [4] and [5]) where a much weaker notion of derivative than in [24] is proved.

The mathematical study of "adjoints" for this type of problem seems to appear for the first time in [5]. In the some specific cases (one dimensional conservation laws, 2×2 systems in one space variables etc...) some complementary results are available for "large time" (see [2], [3] and [19, 20, 8]).

In the present contribution we try to provide an intrinsic definition of this adjoint (avoiding the explicit use of Riemann invariants). We also try to underline the difference between what is a consequence of the notion of weak solution (differentiation of the Rankine Hugoniot relations for instance) and what comes from the entropy relation. This is the reason why we treat explicitly some examples, being aware that they are contained in the general theorems of [24].

The paper is organized as follows. We begin by a systematic analysis of the differentiation of piecewise continuous functions and conservation laws, recalling the formalism developed in [1]. Proofs of the validity of the variational calculus rely on Majda's results [24] [25] as explained in Section 3.5. To obtain more explicit formulas and to underline the role of the entropy condition we describe the treatment of two elementary but basic examples in the section 3.6. This illustration continues with the analytic and numerical computation for the Burgers equation. The following sections are devoted to the compressible Euler equation (one of the goal of the present contribution). There we show both how to use the formal calculus and how to introduce for numerical purpose an Hamilton-Jacobi formulation. We apply the results to a control problem for the shallow water equations (also Euler equations of gas dynamics with $\gamma = 1$ and adiabatic flows [31]) solved numerically with a descent method. Finally in view of forecoming multi-dimensional control applications, we introduce a multi-dimensional Hamilon Jacobi formulation, to general vector conservation laws and give some numerical solutions in simple cases.

2. Preliminaries.

2.1. Derivative of piecewise smooth functions. In the context of the present contribution solutions or functions are assumed to be piecewise smooth therefore as observed in [1] a systematic calculus of derivative can be used even on the shock.

More precisely let Q be an open bounded set of R^d and A an interval of R. Consider a vector valued function from Q to R^d function of a parameter $a \in A$:

$$(z, a) \in \mathcal{Q}_A := Q \times A \to v(z, a) \in R^d. \tag{7}$$

Let S be a smooth surface which cuts Q into two parts Q^+, Q^-. Assume that the restrictions of v, $v^\pm$ to $Q^\pm$ are continuously differentiable. We denote by C^1_S the space of such functions.

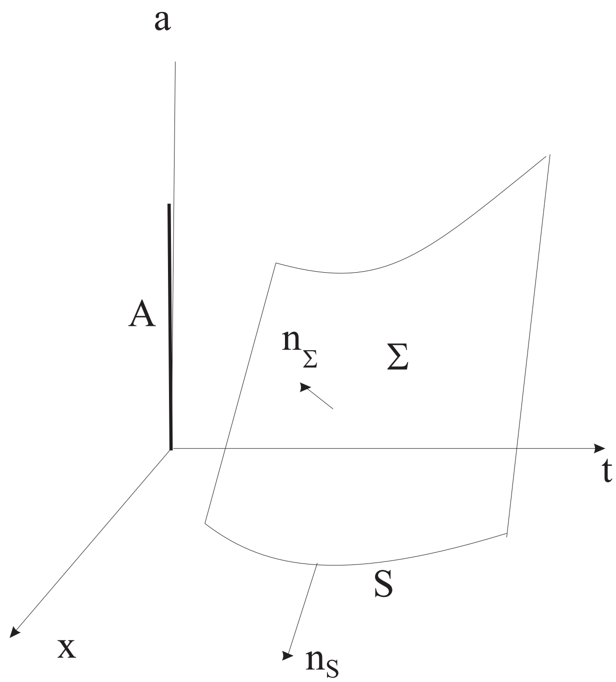

FIG. 1. *A shock curve S in R^2. In the case of Burgers equation for instance, the shock is in the space time frame, so $x_1 = x$, $x_2 = t$. When the parameter a varies the shock describes a surface Σ. The normals to S and Σ are also shown*

Assume also that v is differentiable with respect to a everywhere except on S and denote by v'_a this pointwise derivative. We introduce the following notations (see figure 1):

- $z = z(s, a)$, $s \in I_S \subset R^{d-1}$, a parametrization of S, $z' = \partial z/\partial a$.
- $\Sigma = \{(z, a) \ : \ z \in S, \quad a \in A\}$, $\mathcal{Q}_A{}^{\pm} = \{(z, a) \ : \ z \in Q^{\pm}, \quad a \in A\}$
- n_S, n_Σ: the normals to S and Σ pointing inside Q^+ and $\mathcal{Q}_A{}^+$ respectively.
- $[v] := v^+ - v^-$ is the jump of any function $v \in C^1_S$ across S.

At some point a more specific local description of the shock will be needed with the following parametrization:

$$\tilde{x} = (x_1, x_2, \ldots, x_{n-1}), \qquad x_n = \phi(\tilde{x}, t, a) \qquad (8)$$
$$\Sigma = \{(\tilde{x}, \phi(\tilde{x}, t, a), t, a)\}.$$

PROPOSITION 1. *Let v be a function in C^1_S. The derivative of v with respect to a, in the sense of distributions, v', is*

$$v' = v'_a - [v] z' \cdot n_S \ \delta_S \qquad (9)$$

where v'_a is the pointwise derivative of v with respect to a and δ_S is the Dirac function on S defined by $\int_Q \delta_S w = \int_S w \ \forall w \in \mathcal{D}(Q)$; $z' \cdot n_S$ is the normal component of the a-derivative of the position of the shock.

Proof. For clarity we do the proof for $d = 3$ only. By definition of derivatives in the sense of distributions,

$$\forall w \in \mathcal{D}(\mathcal{Q}_A) \ : \quad \int_{\mathcal{Q}_A} v'w = -\int_{\mathcal{Q}_A} vw' \tag{10}$$

Since $v^{\pm} := v|_{Q^{\pm}}$ are smooth, one has:

$$\begin{aligned}\int_{\mathcal{Q}_A} vw' &= \int_{\mathcal{Q}_A{}^-} v^- w' + \int_{\mathcal{Q}_A{}^+} v^+ w' \\ &= -\int_{\mathcal{Q}_A{}^-} v^{-\prime} w - \int_{\mathcal{Q}_A{}^+} v^{+\prime} w + \int_{\Sigma} (v^+ - v^-) n_{\Sigma d+1} w.\end{aligned} \tag{11}$$

where $n_{\Sigma d+1}$ is the last component of the normal to Σ. It is easy to see (see [1] for details) that

$$n_\Sigma = \frac{(n_S, -z' \cdot n_S)^T}{\sqrt{1+\dot{z}^2}}, \qquad \mathrm{d}\Sigma = \sqrt{1+\dot{z}^2}\mathrm{d}S\mathrm{d}a. \tag{12}$$

Proposition 2. *Let $w \in (C^0(S))^d$. Denote by $\nabla_S \cdot w_S$ the surface divergence on S:*

$$\nabla_S \cdot w_S = \sum_{i=1}^{d-1} \partial_{s^i}(w \cdot s^i)$$

where $\{s_i\}_1^d$ is an ortho-normal system of tangent vectors at the point where the derivatives are considered. Assume that $w \cdot n_S = 0$. Then

$$\nabla \cdot (w\delta_S) = \delta_S \nabla_S \cdot w_S. \tag{13}$$

Proof.

$$\begin{aligned}&\int_Q w\delta_S \cdot \nabla\hat{w} := \int_S w \cdot \nabla\hat{w} = \sum_{i=1}^{d-1} \int_S \frac{1}{\sqrt{g}} w \cdot s^i \partial_{s^i}\hat{w}\sqrt{g}\mathrm{d}s_1 ... \mathrm{d}s_d \\ &= -\sum_{i=1}^{d-1} \int_S \partial_{s^i}(w \cdot s^i)\hat{w}\mathrm{d}s_1 ... \mathrm{d}s_d\end{aligned} \tag{14}$$

where g is the determinant of the Riemanian metric induced on the surface ($g = 1$ whenever $\{s_i\}_1^d$ are ortho-normal).

Consequently from (9) we have the following property.

Corollary 1. *Consider $v \in (C_S^1)^d$. If $[v \cdot n_S] = 0$ and the pointwise a-derivative, v'_a, has a trace left and right of S, then*

$$\nabla \cdot v' = \nabla \cdot v'_a + ([v'_a] \cdot n - \nabla_S \cdot ([v]z' \cdot n_S)\delta_S \tag{15}$$

where $\nabla \cdot v'_a$ is the pointwise value.

Proof. Using linearity and starting from the formula:

$$v' = v'_a - [v]z' \cdot n_S \ \delta_S \tag{16}$$

one has

$$\nabla \cdot v' = \nabla_x \cdot v'_a - \nabla \cdot ([v] z' \cdot n_S \ \delta_S) \tag{17}$$

and with the hypothesis $[v \cdot n_S] = 0$ the last term is computed according to Proposition 2.

Finally recall that for $w \in C^1_S$,

$$\nabla \cdot w = \nabla_x \cdot w + [w] \cdot n_S \delta_S \tag{18}$$

where $\nabla_x \cdot w$ is the pointwise value; This applied to $w = v'_a$ gives

$$\nabla v'_a = \nabla_x \cdot v'_a + [v'_a] \cdot n_S \delta_S$$

which, in turn, proves the result.

2.2. Rules for derivatives. Consider a smooth mapping $u \mapsto F(u)$ defined on an open set D of $\mathbb{R}^p$ with values in $\mathbb{R}^p$ and $u \in (C_S)^d$ a piecewise smooth function with value in D. Then for any $1 \le i \le d$ the derivative $\partial_{x_i} F(u)$ is the sum of the smooth function $F'_u \partial_{x_i} u$ away from the shock and of the density $n_i [F(u)] \delta_S$. With the introduction of

$$G(u) = \int_0^1 F'_u(\sigma u^+ + (1-\sigma) u^-) d\sigma \tag{19}$$

which coincides with F'_u away from the shock and which is a continuous density on the shock one obtains the formula:

$$\partial_{x_i} F(u) = G(u) \partial_{x_i} u \tag{20}$$

because ∂_{x_i} is its pointwise value plus $-[u] n_i \delta_S$ and

$$G(u)[u] = \int_0^1 F'_u(\sigma u^+ + (1-\sigma) u^-)[u] \mathrm{d}\sigma = \int_{u^-}^{u^+} F'_u \mathrm{d}u = [F(u)]. \tag{21}$$

Definition 1. *For any $v \in C^1_S$ one denotes by $\overline{v}$ the function which is equal to v in $Q^\pm$ and which is equal to $(v^+ + v^-)/2$ on S*

one has the following

Proposition 3. *If $f \in C^1_S(\mathbb{R}^n, \mathbb{R}^{p \times q})$ and and $g \in C^1_S(\mathbb{R}^n, \mathbb{R}^{q \times r})$ are two matrices valued functions one has:*

$$(f \cdot g)' = f' \cdot \overline{g} + \overline{f} \cdot g'. \tag{22}$$

Proof. The formula is true pointwise. At the shock the weight on the Dirac mass of $(f \cdot g)'$ is $-[fg] x' \cdot n_S$ and the right hand side has a Dirac mass of weight $-([f]\bar{g} + [g]\bar{f}) x' \cdot n_S$. Since

$$[fg] = [f] \frac{g^+ + g^-}{2} + [g] \frac{f^+ + f^-}{2} = [f]\bar{g} + [g]\bar{f} \tag{23}$$

equation (23) holds in the sense of distributions.

Other extensions of this type of calculus are useful and can be obtained from (23). For instance to compute $(u^3)'$ one writes $u^3 = u \cdot u^2$ and $u^2 = u \cdot u$; hence

$$(u^3)' = (u \cdot u^2)' = \overline{u^2}u' + (u^2)'\overline{u} = \overline{u^2}u' + (u'\overline{u} + \bar{u}u')\overline{u} = (\overline{u^2} + 2\overline{u}^2)u'. \tag{24}$$

To compute $m' = (\frac{1}{u})'$ one writes

$$0 = (um - 1)' = \overline{m}u' + \overline{u}m' \text{ or } m' = -\frac{1}{\overline{u}}\overline{(\frac{1}{u})}u' \tag{25}$$

and so on. Essentially any rational fraction of implicit functions of powers of u can be differentiated.

In the scalar case one should observe that $G(u)$ is equal to the Volpert ratio : $[F(u)]/[u]$ on the shock.

The interest of the above calculus for our purpose comes not only from the fact that piecewise smooth solutions are natural objects for hyperbolic systems but also from the fact that state laws of fluid mechanic which are Galilean invariant turn out to be rational in term of the conserved quantities (cf. [10]).

2.3. Derivative of quadratic Functionals. For Calculus of Variation, a natural quantity is the least square distance to a desired state u_d:

$$E(u)(a) = \frac{1}{2}\int_{\mathbb{R}^n} |u(x,a) - u_d(x)|^2 \mathrm{d}x \tag{26}$$

and to minimize such a function the natural strategy is the computation of its gradient.

In fact it turns out with that Corollary 1 provides a direct way of computing functionals and this is the object of the

PROPOSITION 4. *For each $a \in A \subset \mathbb{R}$ assume that there is a shock S, i.e. a set*

$$\{x(s,a) : s \in I_s(a) \subset \mathbb{R}^{n-1}\} \subset \mathbb{R}^n; \qquad \Sigma = \{x(s,a) : s \in I_s(a), a \in A\}$$

Assume that Σ is a C^1 oriented manifold in the variable (x,a). Denote by $x'(s,a)$ the $a-$derivative of $x(s,a)$ and by n_S the oriented normal to S.

Let $v \in C^1_S$ and assume (just for simplicity) the existence of a finite number R independent of a such that $|x| > R \Rightarrow v(x,a) = 0$.
Then with

$$F(a) = \int_{\mathbb{R}^n} v(x,a)dx$$

one has:

$$F'(a) = \frac{d}{da}\left(\int_{\mathbb{R}^n} v(x,a)dx\right) = \int_{\mathbb{R}^n} v'_a dx - \int_S [v]z' \cdot n_S dS \tag{27}$$

and the Taylor Formula:

$$|F(a+h) - F(a) - hF'(a)| = o(h). \tag{28}$$

Proof. Introduce a smooth test function $\phi \in \mathcal{D}(\mathbb{R}^n)$ which is equal to 1 on the support of v and with Proposition (1) write:

$$\partial_a \int_{\mathbb{R}^n} v(x,a)\mathrm{d}x = \partial_a \langle v, \phi\rangle = \langle v', \phi\rangle = \langle v'_a - [v]z' \cdot n_S \delta_S, \phi\rangle$$
$$= \int_{\mathbb{R}^n} v'_a \mathrm{d}x - \int_S [v]x' \cdot n_S \mathrm{d}S. \tag{29}$$

This proves (27). The Taylor formula follows from the fact that the function:

$$a \mapsto \int_{\mathbb{R}^n} v'_a \mathrm{d}x - \int_S [v]x' \cdot n_S \mathrm{d}S \tag{30}$$

is continuous with respect to a.

Corollary 2. *For each $a \in A$ let $u \in C^1_S$, defined as above. Let $u_d(x) \in C^0$ independent of a. Assume the existence of R independent of a such that $u(x,a)$ and $u_d(x)$ coincide for $|x| > R$ then for*

$$E(a) = \frac{1}{2}\int |u(x,a) - u_d(x)|^2 dx$$

one has

$$E'(a) = \int_{\mathbb{R}^n} u'_a(u - u_d)dx - \int_S [u](\overline{u} - \overline{u_d})x' \cdot n_S dS \tag{31}$$

and the Taylor formula:

$$|E(a+h) - E(a) - hE'(a)| = o(h). \tag{32}$$

Proof. Apply to the function

$$v(x,a) = |u(x,a) - u_d(x)|^2$$

the previous proposition and the rule of computation of derivatives of Proposition 3.

Remark 1. *Observe that*

$$\mu = u'_a(u - u_d) - [u](\overline{u} - \overline{u_d})x' \cdot n_S \delta S = (\overline{u} - \overline{u_d})u'$$

is a bounded measure and that the formula (31) can be written

$$E'(a) = \int \mu = \int_{\mathbb{R}^n} (\overline{u} - \overline{u_d})u' dx. \tag{33}$$

Remark 2. *When u_d is discontinuous, say $u_d \in C^1_{\tilde{S}}$ and $u \in C^1_S$ then result is still valid so long as $S \neq \tilde{S}$. Otherwise the result is false because $u - u_d \notin C^1_S$ at $a = 0$ if $\lim_{a\to 0} S = \tilde{S}$. In such cases Corollary 2 should be applied at $a \neq 0$ and then let $a \to 0$ in the result. This is seen in the following example.*

Example. Let $H(x)$ be the Heaviside function ($H(x) = x^+/x$). Let $u = 1 - H(x-a)$ and $u_d = 1 - H(x)$. Then

$$E(a) = \frac{1}{2}\int_{\mathbb{R}} (u - u_d)^2 \mathrm{d}x = \frac{|a|}{2} \quad \Rightarrow \quad E'(a) = \frac{1}{2}\mathrm{sign}(a).$$

On the other hand

$$\int_{\mathbb{R}} u'_a(u-u_d)\mathrm{d}x - [u-u_d](\bar{u}-\bar{u}_d)x' = 0 \tag{34}$$

because $u'_a = 0,\ x' = 1,\ [u] = [u_d] = -1,\ \bar{u} = \bar{u}_d = \frac{1}{2}$ at $a = 0$. But if (34) is applied at $a > 0$ for instance then $[u_d] = 0,\ \bar{u}_d = 0$ and the following is correct:

$$E' = \int_{\mathbb{R}} u'_a(u-u_d)\mathrm{d}x - [u](\bar{u}-\bar{u}_d)x' = \frac{1}{2}. \tag{35}$$

At $a < 0$, $[u_d] = 0,\ \bar{u}_d = 1$ and the above formula gives $E' = -\frac{1}{2}$ which is also correct.

3. Application to Conservation Laws..

3.1. Intrinsic computations. Consider a conservation law characterized by a flux function $F = (F_1, F_2, \dots, F_n) : \mathbb{R}^p \to \mathbb{R}^{n\times p}$:

$$\partial_t u + \nabla \cdot F(u) = 0 \text{ in } Q \quad u(x,0) = u^0(x,a) \ \ \forall x \in \mathbb{R}^n. \tag{36}$$

Let S be the shock surface for a given parameter a. From Proposition 1, assuming that the family of solutions $(x,t) \to u(x,t,a)$ are in $(C^1_S)^p$ for all a, then at each x, t, a the derivative with respect to a, V' of the vector $V = (u, F(u))$ is

$$V' = (u', F(u)') = (u'_a - [u]x' \cdot n_S\delta_S, F'_u u'_a - [F(u)]x' \cdot n_S\delta_S) \tag{37}$$

or with the notation (19)

$$V' = (u', G(u)u') = (u'_a, F'_u) - ([u], G(u)[u])x' \cdot n_S\delta_S. \tag{38}$$

Equation (36) is also the divergence with respect to t, x of V, therefore the derivative (with respect to the parameter a and in the sense of distributions) of (36) is

$$\nabla_{t,x} V' = 0 \tag{39}$$

i.e.

$$\partial_t u' + \nabla \cdot (G(u)u') = 0. \tag{40}$$

We know now that the Rankine Hugoniot condition

$$n_t[u] + n_x \cdot [F(u)] = 0 \tag{41}$$

differentiated are contained in (40). Corollary 1 can be used for (40) giving the relation:

$$\partial_t u'_a + \nabla \cdot (G(u)u'_a) + \left(n_t[u'_a] + n_x \cdot [F'_u(u)u'_a] - \nabla_S \cdot ([(u, F(u))]x' \cdot n_S)\right)\delta_S = 0 \tag{42}$$

which is equivalent to:

$$\begin{aligned}&\text{away from the shock: } \partial_t u'_a + \nabla \cdot F'(u)u'_a = 0\\ &\text{on the shock: } n_t[u'_a] + n_x \cdot [F'_u(u)u'_a] - \nabla_S \cdot ([(u, F(u))]x' \cdot n_S) = 0.\end{aligned} \tag{43}$$

3.2. Adjoint Equation. So far we have established (33) and (40). In order to express E' in terms of u'_0 only we need an adjoint equation. To obtain an adjoint equation, one introduces smooth test functions w^* with compact support in $\mathbb{R}^n \times]0,T[$ and then observe that the relation (40) is by definition:

$$0 = \langle \partial_t u' + \nabla\cdot(G(u)u'), w^*\rangle = -\langle u', \partial_t w^* + G(u)^T \nabla w^*\rangle \ \ \forall w^* \in \mathcal{D}(\mathbb{R}^n \times (0,T)). \tag{44}$$

This leads to the following adjoint problem

$$\partial_t u^* + G^T \nabla u^* = 0. \tag{45}$$

The solutions of the adjoint problem have a finite speed of propagation therefore one can assume that their spatial support is bounded. So one starts with a "smooth" solution of with bounded spatial support and then applies Formula (44) to the function $w^*(x,t) = \theta(t)u^*(x,t),\ \theta \in \mathcal{D}(0,T)$ to obtain:

$$\begin{aligned} 0 &= \langle \partial_t u' + \nabla\cdot(G(u)u'), w^*\rangle \\ &= -\int_0^T (\theta\langle u(.,t), (\partial_t + G^T\nabla)u^*(.,t)\rangle + \theta' < u\ u^* > \mathrm{d}t \\ &= -\int_0^T \theta' < u\ u^* > \mathrm{d}t. \end{aligned} \tag{46}$$

Equation (40) implies that $t \mapsto u(.,t)$ is a continuous function from $]0,T[$ with value in $\mathcal{D}'(\mathbb{R}^n)$, therefore with $\theta(t)$ converging to

$$I(t_1,t_2) = 1 \text{ for } 0 < t_1 < t_2 < T \quad \text{ and } 0 \quad \text{ elsewhere}$$

one deduces from (46) the relation:

$$\langle u(.,t_1)\ u^*(.,t_1)\rangle = \langle u(.,t_2)\ u^*(.,t_2)\rangle. \tag{47}$$

3.3. Differentiation of Flux with Shocks. As said in Section 2.2 the introduction of the notation $\overline{u}$ (see Definition 1) leads to convenient rules of computation for the derivative of the flux functions with respect to an extra parameter a.

For the p-system

$$\begin{aligned} &\partial_t v - \partial_x u = 0 \\ &\partial_t u + \partial_x \frac{1}{v} = 0 \end{aligned}$$

one has:

$$G = \begin{pmatrix} 0 & -1 \\ -\frac{1}{\overline{v}}\overline{(\frac{1}{v})} & 0 \end{pmatrix}. \tag{48}$$

For the derivative of solutions (ρ,u) the isentropic Euler equation (the Eulerian version of the p system and also the shallow water equations [31])

$$\partial_t u + \partial_x(\frac{u^2}{2} + \rho)\,, \qquad \partial_t \rho + \partial_x(\rho u) = 0 \tag{49}$$

one has

$$\partial_t u' + \partial_x(\bar{u}u' + \rho')\,, \qquad \partial_t \rho' + \partial_x(\bar{\rho}u' + \bar{u}\rho') = 0. \tag{50}$$

Finally for the "full" Euler system:

$$\begin{aligned}
&\partial_t\rho + \partial_x(\rho u) = 0\\
&\partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0\\
&\partial_t e + \partial_x(u(\gamma e - (\gamma-1)\frac{\rho u^2}{2})) = 0\\
&p = (\gamma-1)(e - \frac{\rho u^2}{2})
\end{aligned} \tag{51}$$

we use the variables u , $w := u^2$ and follow the rules of Section 2.2 to decompose all products into binary products

$$v = \rho u \;\Rightarrow\; v' = u'\bar\rho + \bar u\rho' \;\Rightarrow\; u' = \frac{v'}{\bar\rho} - \frac{\bar u}{\bar\rho}\rho' \;\Rightarrow\; w' = 2\bar u u' = 2\frac{\bar u}{\bar\rho}v' - 2\frac{\bar u^2}{\bar\rho}\rho' \tag{52}$$

and obtain

$$\begin{aligned}
&\partial_t\rho' + \partial_x v' = 0\,,\\
&\partial_t v' + \partial_x(-\frac{3-\gamma}{2}\frac{\bar u\bar v}{\bar\rho}\rho' + \frac{3-\gamma}{2}(\bar u + \frac{\bar v}{\bar\rho})v' + (\gamma-1)e') = 0\,,\\
&\partial_t e' + \partial_x((-\gamma\frac{\bar u\bar e}{\bar\rho} + (\gamma-1)\frac{\bar u^2\bar v}{\bar\rho})\rho' + (\gamma\frac{\bar e}{\bar\rho} - \frac{\gamma-1}{2}(\overline{u^2} + 2\frac{\bar v\bar u}{\bar\rho}))v' + \gamma\bar u e') = 0.
\end{aligned} \tag{53}$$

Once again on can observe that this complex system reduces to the classical form outside the shock and contains the differentiated Rankine-Hugoniot conditions at the shock.

3.4. Explicit Computation of the "Jump Conditions" in the Linearized and Adjoint Problems. For explicit computation one uses the local parametrization (8)

$$x_n = \phi(\tilde x, t, a). \tag{54}$$

The normal to S is therefore the vector

$$\frac{1}{\sqrt{|\nabla\phi|^2+1}}(-\partial_t\phi, -\nabla_{\tilde x}\phi, 1).$$

The Rankine Hugoniot condition on u for (36) is

$$0 = [V]\cdot n_S := n_t[u] + n_x\cdot[F(u)] = [u]\partial_t\phi + \sum_{1\le i\le n-1}\partial_{x_i}\phi[F_i(u)] - [F_n(u)] \tag{55}$$

and the relation

$$\left((n_t[u'_a] + n_x\cdot[F'_u(u)u'_a]) - \nabla_S\cdot([(u,F(u))]x'\cdot n_S)\right)\delta_S = 0 \tag{56}$$

is

$$\begin{aligned}
&\partial_t(\phi'_a[u]) + \sum_{1\le i\le n-1}\partial_{x_i}([F_i(u)]\phi'_a)\\
&= -(\partial_t\phi[u'_a] + \sum_{1\le i\le n-1}[\partial_u F_i(u)u'_a]\partial_{x_i}\phi - [\partial_u F_n(u)u'_a]).
\end{aligned} \tag{57}$$

Observe that (57) can also be obtained by differentiating (55) with respect to a. Using the fact that u^* is smooth (in particular near the shock) and multiplying

$$\partial_t u^* + G^T \nabla u^* = 0$$

by $[u]$ one obtains on S:

$$[u]\partial_t u^* + [F(u)]\nabla u^* = 0. \tag{58}$$

From the Rankine Hugoniot condition we know that the vector field $([u], [F(u)])$ is tangent to S and therefore (58) is a first order pde on this surface. With the parametrization (8) it is

$$[u]\partial_t u^*(\tilde{x}, \phi(\tilde{x}, a), t) + \sum_{1\le i\le n} [F_i(u)]\partial_{x_i} u^*(\tilde{x}, \phi(\tilde{x}, a), t) = 0. \tag{59}$$

In particular in one space with $\xi(t, a) = x'(t, a)$ this formula is

$$\frac{d}{dt}\left(\xi(t, a)[u]\right) = -(\dot{x}(t, a)[u'_a] - [F'(u)u'_a]) \tag{60}$$

and since $[F(u)] = \dot{x}(t)[u]$, it is

$$[u]\frac{\mathrm{d}u^*}{\mathrm{d}t}(x(t), t) = 0. \tag{61}$$

Summarizing, with the parametrization (8) the linearized system and its adjoint are characterized by the equations:

$$\text{away from the shock: } \partial_t u'_a + \nabla\cdot(\partial_u F(u)u'_a) = 0\,, \tag{62}$$

$$\text{on the shock: } \partial_t(\phi'_a[u]) + \sum_{1\le i\le n-1} \partial_{x_i}([F_i(u)]\phi'_a) \tag{63}$$

$$= -(\partial_t\phi[u'_a] + \sum_{1\le i\le n-1} [\partial_u F_i(u)u'_a]\partial_{x_i}\phi - [\partial_u F_n(u)u'_a]) \tag{64}$$

and

$$\text{away from the shock: } \partial_t u^* + (\partial_u F(u)^T \nabla u^* = 0\,, \tag{65}$$

$$\text{on the shock: } [u]\partial_t u^*(\tilde{x}, \phi(\tilde{x}, a), t) + \sum_{1\le i\le n} [F_i(u)]\partial_{x_i} u^*(\tilde{x}, \phi(\tilde{x}, a), t) = 0 \tag{66}$$

Finally under suitable hypotheses ($u \in C^1_S$, u^* smooth on the shock) one has for any $0 < t_1 < t_2 < T$ the relation (cf. 47) :

$$\begin{aligned}
&\int_{\mathbb{R}^n} u'_a(x, t_1)u^*(x, t)\mathrm{d}x - \int_{R^{n-1}} (\phi'_a[u]\cdot u^*)(\tilde{x}, t_1, a)d\tilde{x} \\
&= \int_{\mathbb{R}^n} u'_a(x, t_2)u^*(x, t)\mathrm{d}x - \int_{R^{n-1}} (\phi'_a[u]\cdot u^*)(\tilde{x}, t_2, a)d\tilde{x}.
\end{aligned} \tag{67}$$

REMARK 3. *This formula corresponds to Formula (1.10) in Majda ([24]). To the best of our knowledge the formula (66) has been only written in [5] for the* 1*d case* *and in terms of Riemann invariants. In the same reference one can find (always in term of Riemann invariants) a* 1d *version of the formula (66).*

REMARK 4. *To use equations (65, 66) to evaluate $\phi'_a(\tilde{x},T)$ (which describes the infinitesimal variation of the shock) at time T) the quantity*

$$\xi^*(\tilde{x},T) = ([u].u^*)(\tilde{x},T)$$

has to be given with no relation with the value of $u^*(x,T)$ on the shock surface (in fact $u^*(x,T)$ may be chosen not continuous across this surface) provided it is continuous for $0 < t < T$. In this situation Equation (66) introduces a new constraint on the solution of the adjoint problem. However for such solutions the expression

$$\langle u'(.,T), u^*(.,T)\rangle$$

makes no sense because $u'(.,T)$ contains a δ-function at points where $u^*(.,T)$ may be discontinuous and this observation appears again in the next section.

3.5. Towards a Proof of the Taylor and Duality formulas. The validity of the Taylor formula (32) is reduced to the differentiability with respect to a of the solution u away from the shock and in the direction n_S, the normal to the shock surface in the (x,a) variables. To fully justify this calculus we use the results contained in [24] and [25].
The first remark is that a small perturbation of a solution with one shock may in some cases generate solutions with other shocks (may be of small amplitude); to avoid this phenomena it is assumed that the initial perturbation remains in the class of data $u+\delta u$ for which one has at time $t=0$ the existence of a scalar function σ defined on the shock curve such that

$$-\sigma[u] + n_S \cdot [F(u)] = 0 \tag{68}$$

The second remark is that the problem being non linear global in time estimates are not available in general.

Strangely enough, up to now the existence of a convex entropy and the decay of this entropy has not been used. Yet it turns out to be essential in the following steps which is the object of Theorem 1 page 8 of [25]; this theorem provides the existence of multidimensional shock fronts for a finite time ($0 < t < T$). With minor changes one can formulate it as follows.

THEOREM 1. *Assume that the system is symmetrizable or in practice that it admits a convex entropy (cf. [11]) Assume that a reference solution $u(x,t)$ exhibits a unique shock S which is not a characteristic (cf Majda [25] hypothesis 4, p10) and is smooth enough away from the shock (in $H^s(\mathbb{R}^n\times]0,T[\backslash S), s \geq 2\mathrm{E}(\frac{n}{2})+7)$ with a stable entropy condition (in the sense of (1.14) in [24], where* E *denotes the integer part).*

For the initial perturbation of $u(x,0,a)$ and its discontinuity point $x(0,a)$ assume that they are small enough ($u(x,0,a) - u(x,0)$ small in H^s away from the shock and satisfies on the shock the compatibility condition (68)).
Then there exists a finite time T (depending on the norm of the perturbation in H^s) such that with the initial data $u(a,x,0)$ and for $0 < t < T$ the system

$$\partial_t u + \nabla\cdot F(u) = 0 \tag{69}$$

has a unique entropic solution with a shock surface $x(t,a)$ close to $x(t,0)$. Further more both $u(x,t)$ (away from the shock) and S are C^1 functions of a.

As a consequence the above theorem gives some sufficient conditions (essentially T not too large with respect to the perturbation) which ensures the validity of the formula (32) and (33). In fact the proof uses the linearized problem

$$\partial_t u' + \nabla \cdot (G(u)u') = 0 \tag{70}$$

for which

$$u'(x,t) = u'_a(x,t) - [u]_{x(t),t} x'(t) \cdot n_S \delta_S \tag{71}$$

provides the elements for an " internal boundary condition ".
To use the adjoint method to compute E' one starts from the relation (using for clarity's sake the parametrization (8))

$$E'(a) = \int_{\mathbb{R}^n} u'_a (u - u_d) \mathrm{d}x - \int_{\mathbb{R}^{n-1}} [u](\overline{u} - \overline{u_d})(\tilde{x}, \phi(\tilde{x}, t, a)) \phi'_a d\tilde{x}. \tag{72}$$

Now it has been observed that if u'_a is a solution of the direct problem (62) (64) and u^* a smooth solution of the adjoint (65) (66) one has for $0 < t_1 < t_2 < T$

$$\begin{aligned} &\int_{\mathbb{R}^n} u'_a(x,t_1) u^*(x,t_1) \mathrm{d}x - \int_{R^{n-1}} (\phi'_a [u] \cdot u^*)(\tilde{x}, t_1, a) d\tilde{x} \\ &= \int_{\mathbb{R}^n} u'_a(x,t_2) u^*(x,t_2) \mathrm{d}x - \int_{R^{n-1}} (\phi'_a [u] \cdot u^*)(\tilde{x}, t_2, a) d\tilde{x}. \end{aligned} \tag{73}$$

By density and continuity the above formula would extend to any "strong" solutions (cf. [24]) for the definition of such object) of the adjoint problem. Eventually these solutions have to satisfy:

$$\begin{aligned} &\int_{\mathbb{R}^n} u'_a(x,0) u^*(x,0) \mathrm{d}x - \int_{R^{n-1}} (\phi'_a [u] \cdot u^*)(\tilde{x}, 0, a) d\tilde{x} = \\ &\int_{\mathbb{R}^n} u'_a(x,T)(u(x,T) - u_d(x)) \mathrm{d}x - \int_{R^{n-1}} \phi'_a [u] \cdot (\overline{u} - \overline{u_d})(\tilde{x}, T, a) d\tilde{x}. \end{aligned} \tag{74}$$

Two steps have to be considered.

• Step 1 Prove for $0 < t_1 < t_2$ the formula (73) for solutions (u'_a, ϕ'_a) and u^* of the direct and adjoint problem for which this makes sense. This last statement requires that

$$u'_a \in C_0(]0,T[; L^2(\mathbb{R}^n) \backslash S), \quad \phi'_a, [u]u^*|_S \in C_0(]0,T[; L^2(S)). \tag{75}$$

• Step 2 Prove that formula (74) can be deduced from formula (73) by letting $t_1 \to 0$ and $t_2 \to T$ and assume that the solution of the adjoint problems satisfies:

$$\lim_{t \to T} u^*(x,t) = (u(x,T) - u_d(x)) \text{ in } L^2(\mathbb{R}^n) \tag{76}$$

$$\lim_{t \to T} [u] \cdot u^*(x,t) = [u] \cdot (\overline{u} - \overline{u_d})(\tilde{x}, T, a) \text{ in } L^2(\mathbb{R}^{n-1}_{\tilde{x}}). \tag{77}$$

It has been shown by Majda [24] that in classical physical situations and due to the entropy condition (Theorem 1 page 21 and so on) the direct linearized problem has for any initial data

$$(u(.,0), \phi'_a(.,0)) \in L^2(\mathbb{R}^n) \times L^2(\mathbb{R}^{n-1})$$

a unique strong solution. Furthermore for initial data

$$(u(.,0),\phi'_a(.,0)) \in H^s(\mathbb{R}^n\backslash\mathbb{R}^{n-1}) \times H^s(\mathbb{R}^{n-1})\,, s \geq 1$$

satisfying compatibility condition the solution is in $H^s(\mathbb{R}^n \times [0,T]\backslash\mathbb{R}^{n-1}) \times H^s(S)$. The proof assumes that the original system can be symmetrized (it has a convex entropy) and uses the standard theory of Lax Friedrichs and Kreiss. As remarked in Majda [25] (page 10) similar proof works *mutatis mutandis* for the adjoint system keeping in mind that the behavior of $[u]u^*$ is a non standard extra data determined by ξ^* given at time T. Since both solutions are strong formula (73) is valid. These statements contain the ingredients to prove Step 1 and Step 2.

For sake of clarity and to underline the importance of the entropy condition and to show that explicit formulas can be obtained for direct and adjoint problems, two basic examples are given in the next section : the scalar case and the p system are analyzed in depth.

3.6. Examples of Existence and Uniqueness for the Linearized System and its Adjoint.

3.6.1. The scalar case. One considers the following problem

$$\partial_t u + \partial_x f(u) = 0. \tag{78}$$

In (78) f is assumed to be convex and it is also assumed that the solution u has a unique shock (satisfying the entropy condition) in the domain $\mathbb{R}_x \times]0,T[$ on the curve $S = \{(x(t),t)\}$. As before $Q^\pm$ denotes the two domains left and right of the shock. In this section u'_a is denoted by U. It is solution of the problem:

$$\partial_t U + \partial_x(f'(u)U) = 0 \tag{79}$$

and one has:

$$\frac{d}{dt}([u]_{x(t),t}x'(t)) = [f'(u)U] - \dot{x}(t)[U]. \tag{80}$$

For the scalar problem equations (79) and (80) are uncoupled therefore one has the following easy result:

THEOREM 2. *For any initial data $U_0 \in L^2(\mathbb{R})$ and $\xi_0 \in \mathbb{R}$ there exists a unique solution of (79) and (80). In particular the traces of $U_\pm$ belong to $L^2(S)$ and $\xi(t) = x'(t)$ is a continuous function. Finally for any $(s,t) \in [0,T]^2$ one has*

$$\lim_{s\to t}\int_{\mathbb{R}} U^2(x,s)\,dx = \int_{\mathbb{R}} U^2(x,t)\,dx. \tag{81}$$

Proof. With the entropy condition satisfied, the characteristics enter the shock both from Q^- and Q^+ As a consequence the problems (78,79) are well posed and one has the estimates:

$$\int_{x(T)}^{\infty} U_+^2(x,T)\mathrm{d}x + \int_0^T (\dot{x}(t) - f(u_+))U_+^2(x(t),t)\mathrm{d}t \leq C\int_{x(0)}^{\infty} U_+^2(x,0)\mathrm{d}x\,, \tag{82}$$

$$\int_{-\infty}^{x(T)} U_-^2(x,T)\mathrm{d}x + \int_0^T (f(u_-) - \dot{x}(t))U_-^2(x(t),t)\mathrm{d}t \leq C\int_{-\infty}^{x(0)} U_-^2(x,0)\mathrm{d}x. \tag{83}$$

In the above the essential point is the positivity on the shock curve of the quantities $(\dot{x}(t)-f(u_+))$ and $(f(u_-)-\dot{x}(t))$. The constant C comes from the use of the Gronwall lemma which involves the expressions

$$\int_{x(t)}^{\infty} \frac{1}{2}\frac{\partial f(u)}{\partial x} U_+^2(x,t)\mathrm{d}x. \tag{84}$$

Formula (81) is proven by applying Green's formula both to $Q^\pm \cap (\mathbb{R}_x \times [s,t])$; it implies in particular the strong L^2 continuity of $U(x,t)$.

The adjoint problem is

$$\partial_t U^* + f'(u)\partial_x U^* = 0 \text{ in } Q^- \text{ and } Q^+ \tag{85}$$

with the condition:

$$[u]\frac{dU^*}{dt}(x(t),t) = 0. \tag{86}$$

Since (86) is a scalar equation, it implies the relation:

$$U^*(x(t),t) = \xi^*. \tag{87}$$

Then in $Q^\pm$ the equations are solved as above. In particular they have traces on S in $L^2(S)$ which coincide with the constant ξ^*.
Once again one can observe the following facts. U^* is defined both on Q^+ and Q^- with the same boundary condition (on the shock curve the solution has the same L^2 trace on both side of the shock curve for $0 \le t < T$ and the expression $< u'(.,t)\ U^*(.,t) >$ has a well defined meaning. This remark is valid even if $U^*(x,T)$ is discontinuous at $T, x(T)$. It provides with the relation:

$$< u'(.,0)\ U^*(.,0) >= \int_{\mathbb{R}} U(x,T)U^*(x,T)\mathrm{d}x - [u]x'(T)U^*(T) \tag{88}$$

the relevant information on the variation of the shock in spite of the fact that $< u'_a(.,T)U^*(.,T) >$ is not well defined, in particular for a discontinuous data $U^*(x,T)$.

3.6.2. The p system. The p system:

$$\partial_t v - \partial_x u = 0 \quad \partial_t u + \partial_x \frac{1}{v} = 0 \tag{89}$$

corresponds to the Lagrangian formulation of the isentropic Euler equation and also of the shallow water equations (see. [27]). The linearized equations for $(U,V) := (u'_a, V'_a)$ and the adjoint system are

$$\begin{aligned} \partial_t V - \partial_x U = 0\,, &\qquad \partial_t U - \partial_x \frac{1}{v^2} V = 0 \\ \partial_t V^* - \frac{1}{v^2}\partial_x U^* = 0\,, &\quad \partial_t U^* - \partial_x V^* = 0. \end{aligned} \tag{90}$$

The function

$$\eta(v,u) = -\log v + \frac{u^2}{2}$$

is a convex entropy for this system (with flux $\frac{u}{v}$) and one has

$$\eta"(v,u) = \begin{pmatrix} \frac{1}{v^2} & 0 \\ 0 & 1 \end{pmatrix} \tag{91}$$

which diagonalizes (according to Lax and Friedrichs) the linearized system. This is the reason why the introduction of the new variables ($W := V/v, U$) and ($W^* = vV^*, U^*$) reduce both systems to the following symmetric ones:

$$\begin{aligned} \partial_t W - \frac{1}{v}\partial_x U + (\frac{1}{v}\partial_x u)W = 0\,, &\quad \partial_t U - \frac{1}{v}\partial_x W + (\frac{1}{v^2}\partial_x v)W = 0 \\ \partial_t V^* - \frac{1}{v^2}\partial_x U^* = 0\,, &\quad \partial_t U^* - \partial_x V^* = 0 \end{aligned} \tag{92}$$

preserving for smooth functions (with no shock) the duality relation:

$$\int_{\mathbb{R}} (VV^* + UU^*)(x,t)\mathrm{d}x = \int_{\mathbb{R}} (WW^* + UU^*)(x,t)\mathrm{d}x. \tag{93}$$

As in the scalar case we assume that the underlying solution (u, v) is entropic with only one downwind shock. We assume that the origin of the downwind shock is $x(0) = 0$. Denote by $Q^\pm$ the (x,t)-open sets left and right of the shock; then we have the

Theorem 3. *For square integrable initial data $P_0 = (W_0, U_0) \times \xi_0 \in (L^2(\mathbb{R}))^2 \times \mathbb{R}$, there exists a unique solution $P = (W, U, \xi) \in L^2(Q^- \cup Q^+) \times C(0,T)$ of (92).*

The traces of (W, U) on the right and left side shock curve are well defined in $L^2(S)$ and satisfy

$$[W] + \frac{\sqrt{v_+ v_-}}{\frac{v_+ + v_-}{2}}[U] + \xi(t)\left([\frac{1}{v}\frac{\partial v}{\partial x}] + \frac{\sqrt{v_+ v_-}}{\frac{v_+ + v_-}{2}}[\frac{\partial u}{\partial x}]\right) = 0\,, \tag{94}$$

$$\frac{d\xi}{dt} + \frac{\xi}{\sqrt{v_+ v_-}}\frac{1}{2}\left(\frac{1}{v_+}\frac{\partial v_+}{\partial x} + \frac{1}{v_-}\frac{\partial v_-}{\partial x}\right) = -\frac{1}{\sqrt{v_+ v_-}}\frac{W_+ + W_-}{2}. \tag{95}$$

Remark 5. *Observe that ((94) , (95)) is equivalent to the general condition (60) which is a system of 2 linear equations. To obtain (94) and (95) one combines linearly the two scalar equations of (60) to obtain one with $\frac{d\xi}{dt}$ and one without it and then perform the change of variables.*

Proof of the Theorem. It is done in two steps.
First the problem is considered with $\xi(t)$ given and then a fixed point argument is applied. One can introduce the function

$$\tilde{P} = e^{-\lambda t}P(t) = e^{-\lambda t}(W(.,t), U(.,t), \xi(.,t)) = (\tilde{W}(.,t), \tilde{U}(.,t), \tilde{\xi}(.,t))$$

and the systems becomes (with the omission of the tildes)

$$\lambda W + \partial_t W - \frac{1}{v}\partial_x U + (\frac{1}{v}\partial_x u)W = 0\,, \tag{96}$$

$$\lambda U + \partial_t U - \frac{1}{v}\partial_x W + (\frac{1}{v^2}\partial_x v)W = 0\,, \tag{97}$$

$$\lambda\xi + \frac{d\xi}{dt} + \frac{1}{\sqrt{v_+v_-}}\frac{W_+ + W_-}{2} + \frac{\xi}{\sqrt{v_+v_-}}\frac{\left(\frac{1}{v_+}\frac{\partial v_+}{\partial x} + \frac{1}{v_-}\frac{\partial v_-}{\partial x}\right)}{2} = 0\,, \tag{98}$$

$$[W] + \frac{\sqrt{v_+v_-}}{\frac{v_++v_-}{2}}[U] + \xi(t)\left([\frac{1}{v}\frac{\partial v}{\partial x}] + \frac{\sqrt{v_+v_-}}{\frac{v_++v_-}{2}}[\frac{\partial u}{\partial x}]\right) = 0. \tag{99}$$

Next one observes that in Q^+ the system (96) and (97) is symmetric positive (λ has been taken large enough) and that on S the boundary matrix is definite positive:

$$\begin{bmatrix} W \\ U \end{bmatrix}(n_t I - \frac{n_x}{v_+}\begin{pmatrix} 0, & 1, \\ 1, & 0 \end{pmatrix}\begin{bmatrix} W \\ U \end{bmatrix}) = \left(\frac{W^2+U^2}{\sqrt{v_+v_-}} - \frac{2}{v_+}UW\right)\frac{1}{\sqrt{1+\dot{x}(t)^2}} \geq c_1(W^2+U^2). \tag{100}$$

This is due to the entropy condition which implies that $v_+ > v_-$. This corresponds also to the fact that the linearized Riemann invariants, "leave" the boundary and finally it is a simple example of the condition (1.14) given by Majda [24]. As a consequence any pair of data $(W_0, U_0) \in L^2(\mathbb{R}^+)$ defines a solution of (96) and (97) in Q^+ with a trace $(W_+, W_-) \in L^2(S)$ which (with λ given but large enough) satisfies for any $0 < t \leq T$ the estimate:

$$\int_{x(t)}^{\infty} \frac{1}{2}(W(x,t)^2 + U(x,t)^2)\mathrm{d}x + c_1\int_S \frac{1}{2}(W(x(s),s)^2 + U(x(s),s)^2)\mathrm{d}s \leq \int_0^{\infty}\frac{1}{2}(W(x,0)^2 + U(x,0)^2)\mathrm{d}x. \tag{101}$$

For the proof one should first check that the above estimate is valid for smooth solution and then proceed by regularization as in the standard Lax-Phillips theory. Next observe that on the $-$ side of S the boundary matrix is no more positive definite it corresponds to the quadratic form:

$$\begin{bmatrix} W \\ U \end{bmatrix}(n_t^- I - \frac{n_x^-}{v_-}\begin{pmatrix} 0, & 1, \\ 1, & 0 \end{pmatrix}\begin{bmatrix} W \\ U \end{bmatrix}) = \left(-\frac{W^2+U^2}{\sqrt{v_+v_-}} + \frac{2}{v_-}UW\right)\frac{1}{\sqrt{1+\dot{x}(t)^2}} \tag{102}$$

and a direct computation shows that the relations:

$$W_- + \frac{\sqrt{v_+v_-}}{\frac{v_++v_-}{2}}U_- = 0 \tag{103}$$

defines a maximal positive subspace for this matrix.

More precisely or more explicitly the relation (103) and the entropy condition $v_- < v_+$ imply (with $c > 0$) the inequality:

$$\left(-\frac{W^2+U^2}{\sqrt{v_+v_-}} + \frac{2}{v_-}UW\right) \geq \frac{v_+v_-}{\frac{v_++v_-}{2}}(2\frac{v_+}{v_-} - 1 - \frac{v_+v_-}{v_-\frac{v_++v_-}{2}})U^2 \geq cU^2. \tag{104}$$

Consequently the problem is well posed in Q^- with data on $\mathbb{R}_x^- \times (-\infty, 0)$ and boundary condition (103) on S. With regularization and superposition one obtains the same result with the following boundary condition,

$$W_- + \frac{\sqrt{v_+ v_-}}{\frac{v_+ + v_-}{2}} U_- = g \tag{105}$$

provided $g \in L^2(S)$ and with suitable constants (for all $0 < t < T$) the following estimate is valid:

$$\begin{aligned}
&\int_{-\infty}^{x(t)} \frac{1}{2}(W(x,t)^2 + U(x,t)^2)\mathrm{d}x + c_1 \int_S (W_-(x(s),s)^2 + U_-(x(s),s)^2) dS \\
&+ c_2^-(\lambda) \int_{\infty}^{x(t)} \frac{1}{2}(W(x,s)^2 + U(x,s)^2)\mathrm{d}x\mathrm{d}t \\
&\leq \int_{-\infty}^{0} \frac{1}{2}(W(x,0)^2 + U(x,0)^2)\mathrm{d}x + \int_S |g(x(s),s)|^2 dS.
\end{aligned} \tag{106}$$

In particular to construct a fixed point argument, for $\xi(t) \in L^2(S)$ given, we choose

$$g_\xi = W_+ + \frac{\sqrt{v_+ v_-}}{\frac{v_+ + v_-}{2}} U_+ - \xi(t) \left([\frac{1}{v}\frac{\partial v}{\partial x}] + \frac{\sqrt{v_+ v_-}}{\frac{v_+ + v_-}{2}} [\frac{\partial u}{\partial x}] \right) \tag{107}$$

to enforce the condition (94). Then for the corresponding solution denoted (W_-^ξ, U^ξ) one has:

$$\begin{aligned}
&\int_{-\infty}^{x(T)} \frac{1}{2}(W_-^\xi(x,T)^2 + U_-^\xi(x,T)^2)\mathrm{d}x + c_1 \int_S \frac{1}{2}(W_-^\xi(x(s),s)^2 + U_-^\xi(x(s),s)^2) dS \\
&+ c_2^-(\lambda) \int_{Q^-} \frac{1}{2}(W_-^\xi(x,t)^2 + U_-^\xi(x,t)^2)\mathrm{d}x\mathrm{d}t \leq \int_{-\infty}^{0} \frac{1}{2}(W(x,0)^2 + U(x,0)^2)\mathrm{d}x \\
&+ c_3 \int_S |\xi(s)|^2 dS + c_4 \int_S \frac{1}{2}(W_+(x(s),s)^2 + U_+(x(s),s)^2) dS.
\end{aligned} \tag{108}$$

Then one determines the function $h(\xi)$ by the formula

$$\lambda h(\xi) + \frac{dh(\xi)}{dt} + \frac{W_+^\xi + W_-^p}{2\sqrt{v_+ v_-}} + \frac{h(\xi)}{2\sqrt{v_+ v_-}} \left(\frac{\partial \log v_+}{\partial x} + \frac{\partial \log v_-}{\partial x} \right) = 0\,, \quad h(\xi)(0) = \xi(0).$$

The above estimate show that for λ large enough the mapping $\xi \mapsto h(\xi)$ is a contraction in $L^2(S)$. Therefore there is a unique $\xi \in L^2(S)$ for which one has $\xi = h(\xi)$ and with the standard tools of symmetric systems this completes the proof.

Away from the shock curve the equations for the adjoint problem are:

$$\partial_t V^* - \frac{1}{v^2}\partial_x U^* = 0\,, \quad \partial_t U^* - \partial_x V^* = 0 \tag{109}$$

and the "interior boundary condition" on the shock curve is

$$[v]\frac{dV^*}{dt} + [u]\frac{dU^*}{dt} = 0. \tag{110}$$

THEOREM 4. *For any set of "final " data* $(V_T^*(x), U_T^*(x)) \in (L^2(\mathbb{R}))^2$ *and for* $\xi^*(T) \in \mathbb{R}$ *given there exists for* $0 < t \leq T$ *a unique solution* $(V^*(x,t), U^*(x,t), \xi^*(t))$ *function* $\xi^*(t)$ *and there exists a unique solution of the adjoint problem:*

$$\partial_t V^* - \frac{1}{v^2}\partial_x U^* = 0\,, \partial_t U^* - \partial_x V^* = 0 \tag{111}$$

with "final data" $(V^*(x,T), U^*(x,T)) = (V_T^*(x), U_T^*(x))$ *and with interior boundary data:*

$$[v]V^* + [u]U^* = \xi^*(t) = \xi^*(T) - \int_t^T (\frac{d[v]}{dt}V^* + \frac{d[u]}{dt}U^*)(x(s),s)dS. \tag{112}$$

Proof. As for the linearized problem the proof is done in several steps.
First the unknown will be $e^{\lambda(T-t)}(V^*, U^*)$ (keeping the same notation) and $W^* = vV^*$ is introduced and the problem is transformed in the following one.

$$-\lambda W^* + \partial_t W^* - \frac{1}{v}\partial_x U^* - \frac{\partial v}{\partial x}\frac{W^*}{v} = 0 \tag{113}$$

$$-\lambda U^* + \partial_t U^* - \frac{1}{v}\partial_x W^* - \frac{2}{v^2}\partial_x v W^* = 0 \tag{114}$$

with data at time $t = T$ and "internal boundary condition"

$$[v]V^* + [u]U^* = e^{-\lambda(T-t)}\xi^*(T) - \int_t^T e^{-\lambda(s-t)}(\frac{d[v]}{dt}V^* + \frac{d[u]}{dt}U^*)(x(s),s)ds. \tag{115}$$

Next the problem is considered with the internal boundary condition:

$$[v]V^* + [u]U^* = g \text{ given.} \tag{116}$$

It is first solved in Q^- because on the shock curve one backward characteristic leaves the domain and with the relation (116) the problem is well posed. In fact on the lateral boundary of Q^- (116) is equivalent to:

$$\frac{(v_+ - v_-)}{v_-}W^* + (u_+ - u_-)U^* = g \tag{117}$$

and for this time reverse problem the boundary quadratic form is:

$$\begin{aligned} &n_t((W^*)^2 + (U^*)^2) - \frac{n_x}{v_-}W^*U^* \\ &= \frac{-1}{\sqrt{1+(\dot{x}(t))^2}}(\frac{1}{\sqrt{v_-v_+}}((W^*)^2 + (U^*)^2) + \frac{1}{v_-}W^*U^*). \end{aligned} \tag{118}$$

For this form the relation

$$\frac{(v_+ - v_-)}{v_-}W^* + (u_+ - u_-)U^* = 0 \tag{119}$$

implies the estimates:

$$((W^*)^2 + (U^*)^2) + \frac{1}{v_-}W^*U^*) \geq c((W^*)^2 + (U^*)^2). \tag{120}$$

Consequently (114), has data at T in $(L^2(-\infty, x(T)))^2$ and for boundary condition

$$\frac{(v_+ - v_-)}{v_-} W^* + (u_+ - u_-) U^* = g \in L^2(S)$$

a unique solution in Q^- with trace $(W^*, U^*) \in L^2(S)$. This provides values for $(V_-^* = \frac{W^*}{v_-}, U_-^*)$. Now to ensure the continuity of (V^*, U^*) one considers in Q^+ the problem (114), with data at T in $(L^2(x(T), \infty))^2$ and (W^*, U^*) given on the boundary by the formula:

$$W_+^* = v_+ V_+^* = v_+ V_-^* = \frac{v_+}{v_-} W_-^* , \quad U_+^* = U_-^*. \tag{121}$$

Since the two characteristics go into Q^+ at this point of the boundary the problem is as above well posed.
From the above construction one sees that both (W_-^*, U_-^*) and (W_+^*, U_+^*) have a trace on S in $L^2(S)$ and that the trace of (V^*, U^*) on both side of S coincide.
Therefore as in the differentiated problem one can construct an a continuous maps $g \mapsto h^*(g)$ in $L^2(S)$ as follows.
First solve with g given in $L^2(S)$ the problem

$$-\lambda W_g^* + \partial_t W_g^* - \frac{1}{v} \partial_x U_g^* - \frac{\partial v}{\partial x} \frac{W_g^*}{v} = 0 \tag{122}$$

$$-\lambda U_g^* + \partial_t U_p^* - \frac{1}{v} \partial_x W_g^* - \frac{2}{v^2} \partial_x v W_g^* = 0 \tag{123}$$

with the "internal boundary" condition:

$$[v] V_g^* + [u] U_g^* = g. \tag{124}$$

As observed this is a well posed problem and the trace (V_g^*, U_g^*) is well defined in $L^2(S)$ Then define $h^*(g)$ as the solution of:

$$-\lambda h^*(g)(t) + \frac{dh^*(g)}{dt} + (\frac{d[v]}{dt} V_g^* + \frac{d[u]}{dt} U_g^*)(x(s), s) ds = 0 , \quad h^*(T)(g) = \xi^*(T). \tag{125}$$

With λ large enough the mapping $h^*(g)$ turns out to be a contraction in $L^2(S)$ and the existence of the solution follows.
The construction shows that both (v_a', u_a') and (V^*, U^*) are strong solutions (in the sense of [18]) therefore by density one has for $0 \leq t_1 < t_2 < T$ the duality formula:

$$\begin{aligned} &\int (v_a' V^* + u_a' U^*)(x, t_2) \mathrm{d}x - ([v] V^* + [u] U^*) x'(x(t_2), t_2) \\ &= \int (v_a' V^* + u_a' U^*)(x, t_1) \mathrm{d}x - ([v] V^* + [u] U^*) x'(x(t_1), t_1) \end{aligned} \tag{126}$$

which with t_1 taken equal to 0 and t_2 converging to T gives:

$$\begin{aligned} &\int (v_a' V^* + u_a' U^*)(x, 0) \mathrm{d}x - ([v] V^* + [u] U^*) x'(x(0), 0) \\ &= \int (v_a' V^* + u_a' U^*)(x, T) \mathrm{d}x - \xi^* x'(x(T), T). \end{aligned} \tag{127}$$

Remark 6. *Observe that if* $[v]$ *and* $[u]$ *are constant on the shock one has*

$$[v] V^* + [u] U^* = \xi^*(t) = \xi^*(T).$$

4. Data Assimilation with Burgers' Equation. Consider solutions of the Burgers' equation with a parameter a in the initial condition

$$\partial_t u + \partial_x \frac{u^2}{2} = 0, \quad \text{in } \mathbb{R} \times (0,T)\,, \quad u(x,0) = u_0(a) \tag{128}$$

and assume that there is only one shock (which satisfies the entropy condition $t \to x(t;a)$) and $u^- \geq u^+ \geq 0$ with the convention that $-$ means upwind (left when $u \geq 0$) and $+$ downwind (right)
Given $b < c$ and $x \to u_d(x)$ piecewise continuous with left and right values at points of discontinuities and bounded, we wish to find a solution of

$$\min_a E(a) := \frac{1}{2} \int_b^c (u(T) - u_d)^2 \mathrm{d}x. \tag{129}$$

According to (33)

$$\begin{aligned} E'(a) &= \int_b^c u'(\bar{u} - \bar{u}_d)\mathrm{d}x \\ &= \int_b^c u'_a(x,T)(u(x,T) - u_d(x))\mathrm{d}x - ([u](\bar{u} - \bar{u}_d))_{x(T),T}\, x'(T). \end{aligned} \tag{130}$$

Then $u' = \frac{\partial u}{\partial a}$ is a distribution which satisfies

$$\partial_t u' + \partial_x(\overline{u}u') = 0, \quad \text{in } \mathbb{R} \times (0,T) \quad u'(x,0) = u'_0(a). \tag{131}$$

To compute the right hand side of (130) one introduces adjoint state solution of

$$\partial_t u^* + \overline{u}\partial_x u^* = 0, \quad \text{in } \mathbb{R} \times (0,T) \quad u^*(x,T) = \bar{u}(x,T) - \bar{u}_d(x). \tag{132}$$

It was shown above that this equation is understood in the strong sense on the shock, hence the presence of $\bar{u}$. Written at $(x(t),t)$ it is an autonomous equation, so (132) is in fact

$$\text{away from the shock} \quad \partial_t u^* + \overline{u}\partial_x u^* = 0 \tag{133}$$

$$\text{on the shock} \qquad \frac{du^*}{dt} = 0. \tag{134}$$

The last one can be integrated because $[u] \neq 0$:

$$u^*(x(t),t) = \bar{u}(x(T),T) - \bar{u}_d(x(T)). \tag{135}$$

The quantity $\langle u'\ u^* \rangle$ is independent of t, i.e.

$$\int u'_a(x,t)u^*(x,t)\mathrm{d}x - x(t)[u]u^*(x(t),t) = \text{constant and one has} \tag{136}$$

$$E' = \int_{\mathbb{R}} u'_a(x,0)u^*(x,0)\mathrm{d}x - x'(0)([u]u^*)_{x(0),0}. \tag{137}$$

4.1. An Analytical Example. We choose $u_0(x,a) = 1 - H(x-a)$ where H is the Heaviside function. The solution of Burgers' equation is $u = 1 - H(x - (a + \frac{t}{2})$ so $u' = \delta(x - (a + \frac{t}{2}))$. Take $u_d = 0,\ b = -1,\ c = +\infty$. Then

$$E = \frac{1}{2}\int_{-1}^{1} u^2 \mathrm{d}x \quad \Rightarrow \quad E' = \frac{1}{2}\frac{d}{da}\int_{-1}^{a+\frac{T}{2}} = \frac{1}{2}. \tag{138}$$

On the other hand u^* solution of the adjoint equation

$$\partial_t u^* + \overline{u}\partial_x u^* = 0 \tag{139}$$

is determined as a classical solution away from the shock and on the shock curve the boundary condition

$$\frac{du^*}{dt}(x(t), t) = 0 \tag{140}$$

with the boundary condition

$$\text{away from the shock } u^*(x,T) \equiv 0 \text{ and on the shock} \lim_{t\to 0} u^*(x(t),t) = \frac{1}{2}. \tag{141}$$

Therefore one checks that E' is indeed given by

$$\langle u'(x,0)\ u^*(x(0,0))\rangle = \frac{1}{2}. \tag{142}$$

4.2. A Computational Example. The computation of Dirac functions with a numerical scheme can be attempted (see [14]) but it is not possible with simple finite difference schemes like the Lax-Wendroff method for (131); indeed it finds the correct position of the Dirac masses but does not give the right weight.

A simple computational trick is to compute a primitive of u' because it has no Dirac singularities, only shocks, and it satisfies

$$\begin{aligned} &U'(x,0) = \int_{-\infty}^{x} u'(x,0)\mathrm{d}x \\ &\partial_t U' + \bar{u}\partial_x U' = 0, \quad \text{in } \mathbb{R}\times(0,T). \end{aligned} \tag{143}$$

The Lax-characteristic scheme

$$U'^{m+1}_i = U'^{\,m}_{j(i)} \text{ with } j(i) \text{ the nearest index to } i - \frac{\delta t}{2\delta x}(u^m_{i+1} + u^m_{i-1}) \tag{144}$$

gives the result shown on figure 2. Notice that the two characteristics left and right of $(x(T),T)$ define a "shaded" region in which there are no characteristics going from t to 0. It is the extra condition (87) (or in this section (140)) which fixes the value of u^* in this region.
To compute the adjoint is quite difficult because unless a mesh point falls exactly on $x(T)$, the shaded region will not be computed correctly; this is a major numerical difficulty which renders the use of adjoint impractical until this problem is solved. Here we propose to "thickened" the discontinuity to spread the value of u^* at $x(T)$ on several grid points. The shock line needs also to be "thickened" because $u^*(T)$ uses

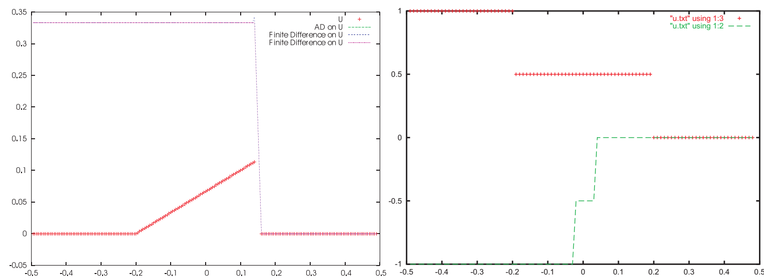

FIG. 2. *On the left side U' is the curve with '+'. The other curve is in fact made of 2 curves almost on top of each other which show u for two close values of a. Substracting these two curves and dividing by δa gives an approximation of u, but it is much less precise than dU'/dx. On the right the adjoint u^* is shown, as computed by (143). It is equal to 1 on the left, 0.5 in the "shaded triangle" and 0 on the right. The other curve is the velocity used in (143), i.e. $-u$.*

values of u at $x(T)$ which will not be correctly computed otherwise. More precisely the adjoint problem solved is (132) with

$$\begin{aligned}&u^*(T) = u^\epsilon(T) \text{ where } u^\epsilon \text{ is given by (131) with}\\ &u(0) = \bar u_0(0) \text{ on } (-\epsilon, +\epsilon), \quad u(0) = u_0 \text{ elsewhere.}\end{aligned} \tag{145}$$

The same Lax-characteristic scheme is applied to (132). The results are shown on figure 2.
Notice that artificial viscosity replaces shocks by smooth curves so it is not a good idea because it will assign to $u^*(x(T),T)$ some value between $u^- - u_d^-$ and $u^+ - u^+$ but not necessarily $\bar u - \bar u_d$. This is not to be confused with a regular calculus of variation on the problem regularized by artificial viscosity, which on the other had works, but nevertheless will be sensitive to the viscosity coefficient (see [12]).

5. Hamilton-Jacobi Formulation. As we proved in the section 3.1 the linearized system for the full Euler equation is

$$\begin{aligned}&\partial_t \rho' + \partial_x v' = 0\\ &\partial_t v' + \partial_x(-\frac{3-\gamma}{2}\frac{\bar u \bar v}{\bar\rho}\rho' + \frac{3-\gamma}{2}(\bar u + \frac{\bar v}{\bar \rho})v' + (\gamma-1)e') = 0\\ &\partial_t e' + \partial_x((-\gamma\frac{\bar u \bar e}{\bar\rho} + (\gamma-1)\frac{\bar u^2 \bar v}{\bar\rho})\rho' + (\gamma\frac{\bar e}{\bar\rho} - \frac{\gamma-1}{2}(\overline{u^2} + 2\frac{\bar v \bar u}{\bar \rho}))v' + \gamma \bar u e') = 0.\end{aligned} \tag{146}$$

For numerical purpose, just as for Burgers' equation, we need to use primitive variables so as to transform the Dirac masses into shocks. Let

$$R'(x,t) = \int_{-\infty}^x \rho'(y,t)\mathrm{d}y, \;\; V'(x,t) = \int_{-\infty}^x v'(y,t)\mathrm{d}y, \;\; P'(x,t) = \int_{-\infty}^x p'(y,t)\mathrm{d}y \tag{147}$$

In matrix form after integration in x the linearized Euler system is

$$\partial_t \begin{pmatrix} R'\\ V'\\ E'\end{pmatrix} = \begin{pmatrix} 0 & 1 & 0\\ -\frac{3-\gamma}{2}\frac{\bar u \bar v}{\bar\rho} & \frac{3-\gamma}{2}(\bar u + \frac{\bar v}{\bar\rho}) & \gamma - 1\\ -\gamma\frac{\bar u \bar e}{\bar \rho} + (\gamma-1)\frac{\bar u^2 \bar v}{\bar\rho} & \gamma\frac{\bar e}{\bar\rho} - \frac{\gamma-1}{2}(\overline{u^2} + 2\frac{\bar v \bar u}{\bar\rho}) & \gamma \bar u\end{pmatrix} \partial_x \begin{pmatrix} R'\\ V'\\ E'\end{pmatrix}.$$

Notice that a standard linearization would have given the matrix:

$$G = \begin{pmatrix} 0 & 1 & 0 \\ -\frac{3-\gamma}{2}u^2 & (3-\gamma)u & \gamma - 1 \\ -\gamma\frac{ue}{\rho} + (\gamma-1)u^3 & \gamma\frac{e}{\rho} - \frac{3}{2}(\gamma-1)u^2 & \gamma u \end{pmatrix} \tag{148}$$

however both matrices are equal outside the shock (where the over-lines can be removed). Matrix (148) has for eigenvalues $\lambda = u, u+c, u-c$ with $c^2 = u^2\gamma(\gamma-1)(\frac{e}{\rho u^2} - \frac{1}{2})$ and eigenvectors, for instance,

$$[1\ ,\quad \lambda\ ,\quad \frac{3-\gamma}{\gamma-1}(\frac{u^2}{2} - (u - \frac{\lambda}{3-\gamma})\lambda)].$$

A numerical scheme is easy to build with this information.

The Shallow water equations. To illustrate our purpose we take the simpler case of Saint-Venant's equations (corresponding approximately to $\gamma = 1$ in Euler equations after replacing the energy equation by an entropy equation); flows of rivers of depth ρ and velocity u satisfy these equations:

$$\partial_t u + \partial_x(\frac{u^2}{2} + \rho) = 0\,, \quad \partial_t\rho + \partial_x(\rho u) = 0. \tag{149}$$

We choose for initial conditions

$$\rho(0) = (0.5 + a)I_{x<0} + 0.5, \qquad u(0) = 0.$$

The differentiated system with respect to a is

$$\partial_t u' + \partial_x(\bar{u}u' + \rho') \qquad \partial_t\rho' + \partial_x(\bar{\rho}u' + \bar{u}\rho') = 0 \tag{150}$$

and after integration in x,

$$\partial_t U' + \bar{u}\partial_x U' + \partial_x R' = 0 \qquad \partial_t R' + \bar{u}\partial_x R' + \bar{\rho}\partial_x U' = 0.$$

Numerical Scheme. The system is rewritten as

$$\partial_t W + G\partial_x W = 0 \quad \text{with} \quad W = \begin{pmatrix} R' \\ U' \end{pmatrix} \quad \text{and} \quad G = \begin{pmatrix} \bar{u} & \bar{\rho} \\ 1 & \bar{u} \end{pmatrix}. \tag{151}$$

It can be diagonalized

$$\partial_t S + \Lambda\partial_x S + Q^{-1}(\partial_t Q + G\partial_x Q)S = 0 \;\text{ with } S = Q^{-1}W, \;\; \Lambda = Q^{-1}GQ \tag{152}$$

and with,

$$\Lambda = \begin{pmatrix} \bar{u} + \sqrt{\bar{\rho}} & 0 \\ 0 & \bar{u} - \sqrt{\bar{\rho}} \end{pmatrix} \quad Q = \begin{pmatrix} \sqrt{\bar{\rho}} & -\sqrt{\bar{\rho}} \\ 1 & 1 \end{pmatrix}. \tag{153}$$

Notice also that

$$Q^{-1} = \begin{pmatrix} \frac{1}{2\sqrt{\bar{\rho}}} & \frac{1}{2} \\ -\frac{1}{2\sqrt{\bar{\rho}}} & \frac{1}{2} \end{pmatrix}, \quad W = \begin{pmatrix} \sqrt{\bar{\rho}}(S_1 - S_2) \\ S_1 + S_2 \end{pmatrix}, \quad S = \begin{pmatrix} \frac{W_1}{2\bar{\rho}} + \frac{W_2}{2} \\ -\frac{W_1}{2\bar{\rho}} + \frac{W_2}{2} \end{pmatrix} \tag{154}$$

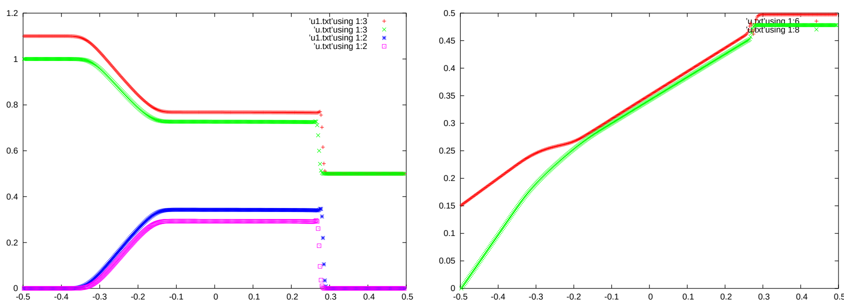

FIG. 3. On the left two results are shown (u in blue and magenta, ρ in red and green) at $T = 0.3$. The difference is due to a change $\delta a = 0.1$ of value in a at $t = 0$, the shock moves by 0.012 with a shock strength of 0.25. On the right is shown R' (in green) and U' (in red). It is seen that $[R'] = 0.03$ giving, with $\delta a = 0.1$, a change in shock position equal 0.012, as expected.

so that

$$Q^{-1}(\partial_t Q + G\partial_x Q)S = \frac{S_1 - S_2}{4}\partial_x \begin{pmatrix} -\bar{u} + 2\sqrt{\bar{\rho}} \\ \bar{u} + 2\sqrt{\bar{\rho}} \end{pmatrix}. \tag{155}$$

Then the upwinding is applied to the system written in S. Results are shown on Figure 3 which displays two computations, one for $a = 0$ and one for $a = 0.1$. The x-primitive R of ρ' of the a-derivative of ρ is shown on the right of the figure and compared with its a-finite difference approximation (only the values around the shock count).

Control of the Shock Position. To test the methodology on a control problem we set the following case:

- Define
$$\rho_d^0 = \frac{1}{2}(I_{x<0} + 1), \quad u_d^0 = 0$$
- Computed $\rho_d(x) = \rho(x, T)$, $u_d(x) := u(x, T)$ by solving (149) with initial conditions equal to ρ_d, u_d.
- Solve by a descent method the problem
$$\min_{a,b,c} E(\rho, u) = \int_{\mathbb{R}} (|\rho(T) - \rho_d|^2 + |u - u_d|^2)\mathrm{d}x \tag{156}$$
where ρ, u are computed from (49) with initial conditions
$$\rho^0 = 1 + b - \frac{1}{2}I_{x>a}, \quad u^0 = c.$$

As before, when the number of unknown is small (3 in our case) it is better not to use the adjoint state because it is numerically too difficult to compute. We have used R', U' directly to compute the derivative of E and a descent method with fixed step size (not the best of course but the easiest to validate the method). The results are shown on figure 4.

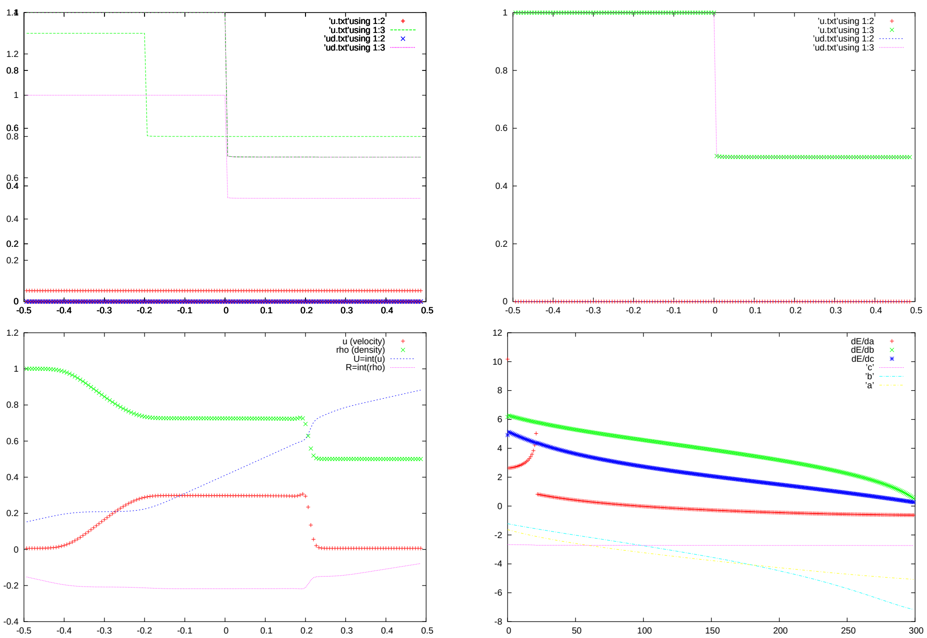

FIG. 4. Top left: Initial (u^0 in red and ρ^0 in green) and target (u_d^0 in blue and ρ_d^0 in magenta) velocities and densities. Top right: Computed (in red and green) and target (in blue and magenta) velocities and densities; this shows that the problem has been solved with very good precision. Bottom left : Solution (u, ρ, U, R) at time T=0.3. Bottom right: Convergence: Gradients (in red, green and blue) and parameters (in magenta, light blue and yellow) with respect to a, b, c versus iteration number (all use log scales).

6. A General Hamilton-Jacobi Formulation. For numerical applications it is difficult to device an efficient scheme for the linearized system because of the Dirac measures. In one dimension of space the trick was to integrate all equations in x. Let us give a similar trick for the general case. We begin with the 2D case.

6.1. A Constructive Argument in 2D. Consider a system of conservation laws for $W(x,t) \in \mathbb{R}^p$:

$$\partial_t W + \partial_x F_1(W) + \partial_y F_2(W) = 0 \quad W(0) = W_0. \tag{157}$$

We assume that the system admits a convex entropy and that there exists one and only solution which satisfies the corresponding entropy condition.

Then obviously it is also

$$\nabla_{t,x,y} \cdot \begin{pmatrix} W^T \\ F_1(W)^T \\ F_2(W)^T \end{pmatrix} = 0 \tag{158}$$

so, there exist $a, b, c \in \mathbb{R}^p$ functions of x, t such that

$$\begin{pmatrix} W & F_1(W) & F_2(W) \end{pmatrix} = \begin{pmatrix} \partial_x c - \partial_y b & \partial_y a - \partial_t c & \partial_t b - \partial_x a \end{pmatrix} \tag{159}$$

The last two equations spell out as

$$\partial_t c + F_1(\partial_x c - \partial_y b) - \partial_y a = 0,$$

$$\partial_t b - F_2(\partial_x c - \partial_y b) - \partial_x a = 0. \tag{160}$$

Now to remove a we notice that the change of variable $c \to c - \partial_y \int^t a,\ b \to b - \partial_x \int^t a$, leaves $\partial_x c - \partial_y b$ unchanged; hence with $V = (c, -b) \in \mathbb{R}^{p\times p}$ and $\mathcal{F} = (F_1, F_2) : \mathbb{R}^p \to \mathbb{R}^{p\times p}$ the system is

$$\partial_t V + \mathcal{F}(\nabla\cdot V) = 0 \quad V(0) = V_0 \ \text{ wíth } \nabla\cdot V_0 = W_0. \tag{161}$$

One way to construct V_0 is to set

$$V_0 = \nabla\Phi \ \text{ wíth } -\Delta\Phi = -W_0. \tag{162}$$

The shallow water equations in $2D$ *.* The shallow water equations

$$\partial_t\rho + \nabla\cdot v = 0\,, \quad \partial_t v + \nabla\cdot(\frac{v\otimes v}{\rho} + \frac{\rho^2}{2}) = 0 \tag{163}$$

lead to the formulas:

$$\nabla\cdot V = W = \begin{pmatrix}\rho\\ v_1\\ v_2\end{pmatrix} \qquad \mathcal{F}(W) = \begin{pmatrix} W_2 & W_3\\ \frac{W_2^2}{W_1} + \frac{W_1^2}{2} & \frac{W_2W_3}{W_1}\\ \frac{W_2W_3}{W_1} & \frac{W_3^2}{W_1} + \frac{W_1^2}{2}\end{pmatrix}. \tag{164}$$

The linearized equation is

$$\partial_t V' + \mathcal{F}'(W)\nabla\cdot V' = 0 \tag{165}$$

with $\nabla\cdot V' = W' = (\rho',\ v_1',\ v_2')$ and

$$\mathcal{F}'(W)W' = \left|\begin{pmatrix} 0 & 1 & 0\\ W_1 - (\frac{W_2}{W_1})^2 & 2\frac{W_2}{W_1} & 0\\ -\frac{W_2W_3}{W_1^2} & \frac{W_3}{W_1} & \frac{W_2}{W_1}\end{pmatrix}\begin{pmatrix}W_1'\\ W_2'\\ W_3'\end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 1\\ -\frac{W_2W_3}{W_1^2} & \frac{W_3}{W_1} & \frac{W_2}{W_1}\\ W_1 - \frac{W_3^2}{W_1^2} & 0 & 2\frac{W_3}{W_1}\end{pmatrix}\begin{pmatrix}W_1'\\ W_2'\\ W_3'\end{pmatrix}\right|. \tag{166}$$

6.2. The General Case. In more than two dimensions of space ($n > 2$) the same trick works but the proof is not constructive. Consider a system of conservation laws in $\mathbb{R}^n$ for $W(x,t) \in \mathbb{R}^p$ with flux $W \mapsto \mathcal{F}(W) \in \mathbb{R}^{n\times p}$:

$$\partial_t W + \nabla\cdot\mathcal{F}(W) = 0 \ \ W(0) = W_0. \tag{167}$$

Let $V \in \mathbb{R}^{p\times n}$ be a solution of

$$\begin{aligned} &\partial_t V + \mathcal{F}(\nabla\cdot V) = 0\\ &V(0) = V_0 \ \text{ chosen such that } \ \nabla\cdot V_0 = W_0\\ &(\text{ for instance: } V_0 = \nabla\Phi \ \text{ wíth } -\Delta\Phi = -W_0). \end{aligned} \tag{168}$$

We have the following facts:

1. $W = \nabla\cdot V$ satisfies (167). Indeed just take the div of (168) to obtain (167).
2. There is at least one solution of (168). Indeed if W is a solution of (167) then the following is a solution

$$V(x,t) = V_0(x) + \int_0^t \mathcal{F}(W(x,\tau),\tau)\mathrm{d}\tau \tag{169}$$

3. If (167) has a unique solution then two distinct solutions V_1, V_2 of (168) have $\nabla \cdot V_1 = \nabla \cdot V_2$.

REMARK 7. *Since the solutions are not continuous they have to satisfy an entropy condition, at least, for uniqueness and stability. For the Hamilton Jacobi system this condition has to be written in term of ∇V . However the numerical computation (see (166)) will involve the matrix $F'(W)$ and it is in the construction of upwind scheme for this matrix that the entropy condition will be enforced. Even with entropy condition uniqueness is an open problem.*

Comparison with Standard Hamiltonian formulation. If the above is applied to a system which has a classical Hamilton-Jacobi equation, what is the connection the vector H-J formulation given above?

Consider the following example

$$\partial_t u_1 + \frac{1}{2}\partial_x(u_1^2 + u_2^2) = 0\,, \quad \partial_t u_2 + \frac{1}{2}\partial_y(u_1^2 + u_2^2) = 0. \tag{170}$$

This system preserves for any solution the property of being a gradient. Therefore with initial data $(u_1,(x,0),u_2(x,0)) = \nabla\phi$ it has the scalar Hamilton-Jacobi formulation

$$\partial_t\phi + \frac{1}{2}(\partial_x\phi^2 + \partial_y\phi^2) = 0. \tag{171}$$

Written in vector form with $u = \nabla\phi$ it is

$$\partial_t\vec{u} + \nabla\cdot\mathcal{F}(\vec{u}) = 0 \text{ with } \mathcal{F}(\vec{u}) = \frac{1}{2}\begin{pmatrix} |u|^2 & 0 \\ 0 & |u|^2 \end{pmatrix}. \tag{172}$$

Accordingly V satisfies

$$\begin{aligned} &\partial_t V + \mathcal{F}(\nabla\cdot V) = 0, \text{ i.e. with } u = \partial_x V_{11} + \partial_y V_{12}, \quad v = \partial_x V_{21} + \partial_y V_{22} \\ &\partial_t V_{11} + \frac{1}{2}(u^2+v^2), \quad \partial_t V_{12} = 0\,, \\ &\partial_t V_{21} = 0, \quad \partial_t V_{22} + \frac{1}{2}(u^2+v^2) = 0. \end{aligned} \tag{173}$$

So with the initial conditions

$$V_{11} = V_{22} = \phi(0), \quad V_{12} = V_{21} = 0$$

equation (171) is recovered. However notice that it is written twice instead of one time.

6.3. Numerical Examples.

Example 1: linear convection. For given velocity field u in an open set Ω denote by n the exterior normal to $\partial\Omega$, let $\Sigma = \{x,t) \;:\; u(x,t)\cdot n < 0,\ x \in \partial\Omega\}$ and consider

$$\partial_t\rho + \nabla\cdot(\rho u) = 0, \quad \rho(0) = \rho^0, \quad \rho = \rho_\Sigma \text{ on } \Sigma. \tag{174}$$

Then the equations for V are

$$\partial_t\vec{V} + \mathcal{F}(\rho) = 0, \quad \nabla\cdot\vec{V} = \rho, \quad \mathcal{F}(\rho) = \rho\vec{u}. \tag{175}$$

To construct a numerical scheme we rely on the fact that the Jacobian matrix is the same as the one for W. So the same principles for the construction of numerical

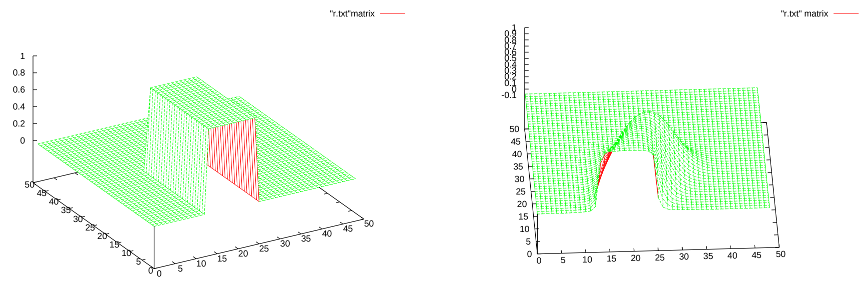

FIG. 5. Convection of a passive scalar entering from the lower right boundary (left) when the convection velocity is aligned with the grid and (right) when it is not. The result is obtained by integrating (175) using (176).

approximations will apply: upwinding via characteristics, numerical flux etc. In this case the problem is linear and an upwind finite difference explicit scheme could be

$$\vec{V}_{i,j}^{m+1} = \vec{V}_{i,j}^{m} - \frac{\delta t}{h}((V_1{}^m{}_{i,j} - V_1{}^m{}_{i-1,j} + V_2{}^m{}_{i,j+1} - V_2{}^m{}_{i,j+1})\vec{u}_{i,j}. \tag{176}$$

Results are shown on figure 5.

Example 2: Shallow Water. Equations (164) are solved with initial conditions corresponding to a stone thrown into a lake at rest. The scheme we have used is similar to (176) but rotated and applied to the diagonalized form of the equations (see (166)). Results are shown on figure 6 for an initial condition which is $\rho = 0$ everywhere except in a disk at the center where $\rho = 1$ and case a) $u = (0, 0)$ everywhere and case b) $u = (1, 0)$ in the disk.

Conclusion. In this paper we have studied the differentiability of entropy weak solutions of conservation laws in the presence of shocks. In general differentiability is not in the usual sense because of Dirac singularities due to changes in the shock positions. We have proposed to use distribution theory to extend the notion of derivative and shown that most of the known results can be reinterpreted with this tool.
We have proposed also a notation which extends the classical calculus of variations. With it, adjoint equations are unambiguously defined when the shock is simple. Integration by part between the adjoint variable and the derivative of the state system is also justified for piecewise smooth solutions with a single shock at all times.
We have shown on simple examples that the results can be used for control; however the adjoint equation is rather difficult to integrate numerically and this is still a limitation of the method. When adjoints are not used (i.e. there are only a few parameters in the optimization problem) then the equation for the derivative can be replaced by its Hamilton-Jacobi primitive which contains shocks instead of Dirac masses. Alternatively Ulbrich proposes in [29, 30] the notion of shift differentiability and it will be compared in a fore coming article.

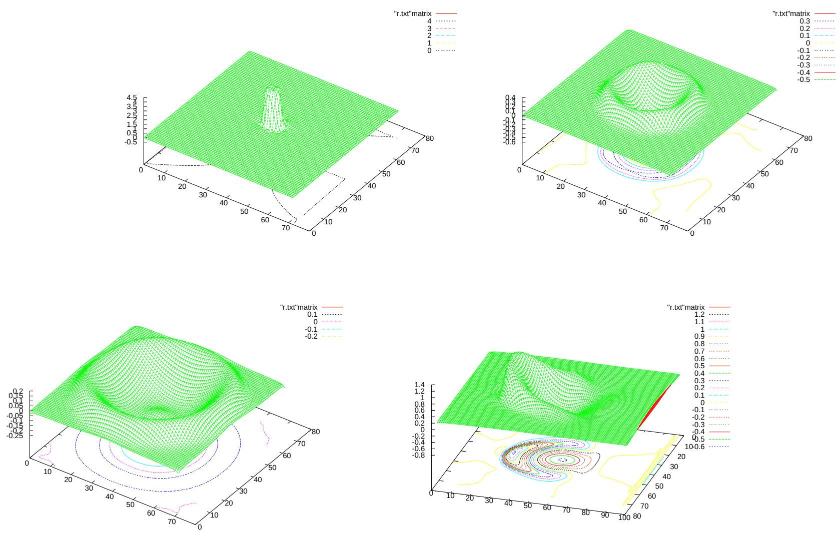

FIG. 6. *T*op and bottom left: Flow at 3 instants of time (including t=0, left) due to a stone falling in a lake and obtained by integrating the Hamilton-Jacobi formulation of the Shallow Water equations. Bottom right: same but the stone has an initial horizontal velocity, thus creating shock propagation.

REFERENCES

[1] BARDOS C. AND PIRONNEAU O., *A formalism for the differentiation of conservation laws*, C.R. Acad. Sci. Paris, Série I, 335:10 (2002), pp. 839–845.

[2] BOUCHUT F. AND JAMES F., *Differentiability with Respect to Initial Data for a Scalar Conservation Law*, International Series of Numerical Mathematics, 129 (1999), pp. 113–118.

[3] BOUCHUT F. AND JAMES F., *One-Dimensional Transport Equations with Discontinuous Coefficients*, Nonlinear Analysis, Theory, Methods & Applications, 32:7 (1998), pp. 891–933.

[4] BRESSAN A. AND MARSON A., *A Maximum Principle for Controlled Systems of Conservation Laws*, Rend. Sem. Mat. Univ. Padova, 94 (1995), pp. 79–94.

[5] BRESSAN A. AND MARSON A., *A Variational Calculus for Discontinuous Solutions of Systems of Conservation Laws*, Comm. in PDE., 20 (9 and 10) (1995), pp. 1491–1552.

[6] CAMPOS-PINTO M., COHEN A., DAHMEN W. AND DEVORE R., *On the Stability of Nonlinear Conservation Laws in the Hausdorff Metric*, IGPM Report, RWTH Aachen, January 2004.

[7] COURTIER P., *Dual formulation of 4D Var Assimilation*, Q. J. Royal Met. Soc., 123 (1997), pp. 2449–2461.

[8] G. CRASTA AND P.G. LEFLOCH, *A class of nonconservative and non strictly hyperbolic systems*, Comm. Pure Appl. Anal., 1 (2002), pp. 513–530.

[9] DAESZCU D.N. AND NAVON I. M., *Adaptative observations in the context of 4D-Var data assimilation*, Meteorlogy and Atmospheric Physics, 85 (2004), pp. 205–226.

[10] DESPRÉS B. AND JAOUEN S. AND RAVIART P.A., *Resolution numerique du problème de Riemann pour un modèle de transition de phase*, Technical report, CEA, 2000.

[11] FRIEDRICHS K. AND LAX P., *Systems of conservation equations with a convex extension*, Proc. Natl. Acad. Sci. USA, 68 (1971), pp. 1686–1688.

[12] GILES M.A. AND N.A. PIERCE, *Analytic adjoint solutions for the quasi-one-dimensional euler equations* Journal of Fluid Mechanics, 426 (2001), pp. 327–345.

[13] GODLEWSKI E. AND OLAZABAL M. AND P.A. RAVIART., *On the linearization of hyperbolic systems of conservation laws*, Application to stability. In *Équations aux dérivées partielles et applications*, pp. 549–570. Gauthier-Villars, Éd. Sci. Méd. Elsevier, Paris, 1998.

[14] GOSSE L. AND JAMES F., *Numerical approximations of one-dimensional linear conservation equations with discontinuous coefficients*, Math. Comp., 69 (2000), pp. 987–1015.

[15] Griewank, A. and Juedes, D. and Utke J., *Algorithm 755: ADOL-C: a package for the automatic differentiation of algorithms written in C/ C++,* j-TOMS, 22:2 (1996), pp. 131–167.

[16] James F. and Sepulveda M., *Convergence Results for the Flux Identification in a Scalar Conservation Law*, SIAM-SICON, 37:3 (1999), pp. 869–891.

[17] Le Dimet, F.-X. and Tallagrand, O., *Variational Algorithms for Analysis and Assimilation of Meteorological Observations: Theoretical Aspects*, Tellus, 38 (1986), pp. 97–110.

[18] Lax P. and Phillips R., *Local Boundary Conditions for Dissipative Symmetric Linear Differential Operators*, Com. in Pure and Appl. Math., Vol. XIII (1993), pp. 425–455, (1960).

[19] Le Floch P., *An existence and uniqueness result for two nonstrictly hyperbolic systems*, IMA Volumes in Math. and its Appl., Springer Verlag, Vol. 27, pp. 126–138, (1990).

[20] Le Floch P. and Xin Z., *Uniqueness via the Adjoint Problems for Systems of Conservation Laws*, Com. in Pure and Appl. Math., Vol. XLVI (1993), pp. 1499–1533.

[21] Lions, J.-L., *Optimal Control of Distributed Systems*, Springer Verlag, (1970).

[22] Li Da-qian and Yu Wen-ci, *Some existence Theorems for Quasilinear Hyperbolic Systems of Partial Differential Equations in Two Independent Variables, II*, Scientia Sinica, 13 (1964), pp. 551–564.

[23] Li Da-qian and Yu Wen-ci, *Boundary Value Problems for Quasilinear Hyperbolic Systems*, Durham: Duke University Math. Serie V (1985).

[24] Majda A., *The stability of multidimensional shock fronts*, Memoirs of the AMS , Volume 41 number 275, January 1983.

[25] Majda A., *The existence of multi-dimensional shock fronts*, Memoirs of the AMS, Volume 43 number 281, May 1983.

[26] Sasaki, Y., *A fundamental Study on the Numerical Prediction Based on the Variational Principle*, J. Met. Soc. Japan, 33 (1955), pp. 262–275.

[27] Smoller J., *Shock Waves and Reaction-Diffusion Equations*, Springer Verlag Band, 258 (1982).

[28] Swanson, K. and Vautard, R. and Pires C., *Four-dimensional Variational Assimilation and Predictability in a Quasi-geostrophic Model*, Tellus, 50A (1997), pp. 369–390.

[29] Ulbrich S., *A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms*, SIAM J. Control Optim., 41:3 (2002), pp. 740–797.

[30] Ulbrich S., *Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws*, Optimization and control of distributed systems., Systems Control Lett., 48:3-4 (2003), pp. 313–328.

[31] Whitham, G.B., *Linear and non linear waves*, Wiley, New-York, 1974.