Optimal control for conservation laws

C. Castro (UPM), F. Palacios (IMDEA Mathematics) and E. Zuazua (UAM)

Benasque, August 27-September 6, 2007

Motivation: An optimal design problem

Let $\Omega \subset \mathbb{R}^2$ be an exterior domain with smooth boundary S.

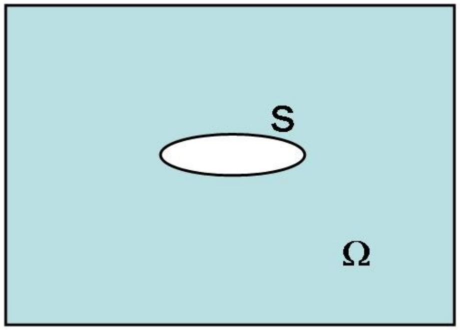


Figure 1: Exterior domain with boundary S

We consider the two-dimensional Euler system

$$\partial_t U + \nabla \cdot F = \partial_t U + \partial_x F_x + \partial_y F_y = 0, \text{ en } \Omega \tag{1}$$

with U, F defined by

$$U = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{pmatrix}, \qquad F_x = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uH \end{pmatrix}, \quad F_y = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho vH \end{pmatrix} \tag{2}$$

where

$$p = (\gamma - 1)\rho \left(E - \frac{1}{2}(u^2 + v^2) \right), \quad H = E + \frac{p}{\rho}. \tag{3}$$

The velocity vector field is denoted by $\mathbf{v}$,

$$\mathbf{v} = \begin{pmatrix} u \\ v \end{pmatrix}.$$

We consider the following boundary condition on S:

$$\mathbf{v} \cdot \mathbf{n}_S = 0, \quad \text{on } S, \tag{4}$$

where $\mathbf{n}_S$ is the exterior unitary normal vector.

System (1) must be completed with some initial conditions

$$U(x, 0) = U_0(x). \tag{5}$$

Finally we introduce the functional

$$J(S) = \int_0^T \int_S g(P) ds\, dt,$$

where $g(s)$ is a given smooth function and the set of admissible designs U_{ad}.

The optimal design problem is then: Find $S^{\min} \in U_{ad}$ such that

$$J(S^{\min}) = \min_{S \in U_{ad}} J(S).$$

In presence of a shock Σ, i.e. a discontinuity of the variables U in the (t, x) variables, the Euler system above must be completed with the Rankine-Hugoniot conditions on the shock:

$$[U]_\Sigma n_\Sigma^t + [(F_x, F_y)]_\Sigma \cdot \mathbf{n}_\Sigma^x = 0, \quad \text{on } \Sigma, \tag{6}$$

where $(n_\Sigma^t, \mathbf{n}_\Sigma^x) = n_\Sigma \in R \times R^2$ is a normal vector to Σ, and $[U]_\Sigma$ is the jump across the shock defined as

$$[U]_\Sigma = \lim_{\varepsilon \to 0} U((x,t) + \varepsilon n_\Sigma) - \lim_{\varepsilon \to 0} U((x,t) - \varepsilon n_\Sigma).$$

Remark Note that shock may possibly meet the boundary S

This problem has been the object of intensive study in aerodynamics since the pioneering works of A. Jameson (80) and O. Pironneu.

Some difficulties

- Huge optimization problems in aerodynamic applications
- Mesh generation in iterative optimization processes
- Mesh sensitivities. How the functional depends on the mesh
- Presence of shocks
- ...

An optimal control problem for Burgers equation

We consider the inviscid Burgers equation:

$$\begin{cases} \partial_t u + \partial_x(\frac{u^2}{2}) = 0, & \text{in } \mathbb{R}\times(0,T), \\ u(x,0) = u^0(x), & x\in\mathbb{R} \end{cases} \tag{7}$$

Given a target $u^d \in L^2(\mathbb{R}))$ we consider the cost functional to be minimized $J : L^1(\mathbb{R}) \to \mathbb{R}$, defined by

$$J(u^0) = \int_{\mathbb{R}} |u(x,T) - u^d(x)|^2 \, dx, \tag{8}$$

where $u(x,t)$ is the unique entropy solution.

We also introduce the set of admissible initial data $\mathcal{U}_{ad} \subset L^1(\mathbb{R})$.

We consider the inverse design problem: Find $u^{0,\min} \in \mathcal{U}_{ad}$ such that

$$J(u^{0,\min}) = \min_{u^0\in\mathcal{U}_{ad}} J(u^0). \tag{9}$$

Main questions

1. Existence of minimizers. We include conditions on the admissible set to guarantee compactness of minimizing sequences. We can consider

$$\mathcal{U}_{ad} = \{f \in L^\infty,\ supp(f) \subset K,\ \|f\|_{L^\infty} \leq C\}\,.$$

2. Uniqueness. A unique minimizer does not exists in general for such problems. Moreover we can have many local minima.

3. Numerical approximation.

   (a) Introduce a suitable discretization for the functional J, J_Δ, the equations, etc.

   (b) Solve the discrete optimization problem: Find $u_\Delta^{0,\min}$ s.t.

$$J_\Delta(u_\Delta^{0,\min}) = \min_{u_\Delta^0 \in \mathcal{U}_\Delta} J_\Delta(u^0),$$

4. Convergence of discrete minimizers when $\Delta \to 0$ (conservative monotone schemes satisfying the discrete one-side Lipschitz condition OSLC).

Discrete problem

Assume that we discretize the Burgers equation using one of the convergent conservative numerical scheme (Lax-Friedrichs, upwind, etc.) and we take

$$J_\Delta(u^0_\Delta) = \frac{\Delta x}{2} \sum_{j=-\infty}^{\infty} (u_j^{N+1} - u_j^d)^2, \tag{10}$$

where $u^0_{\Delta x} = \{u_j^0\}$ and $u^d_\Delta = \{u_j^d\}$ are numerical approximations of $u^0(x)$ and $u^d(x)$ at the nodes x_j, respectively. For example, we can take

$$u_j^0 = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x)dx,$$

where $x_{j\pm 1/2} = x_j \pm \Delta x$.

Let us introduce an approximation of the space $\mathcal{U}_{ad}$, $\mathcal{U}_{ad}^\Delta$ constituted by sequences $u_\Delta = \{v_j\}_{j\in\mathbb{Z}}$ for which the function obtained by piecewise constant interpolation u_Δ, defined by

$$u_\Delta(x) = u_j, \quad x_{j-1/2} < x < x_{j+1/2},$$

satisfies $u_\Delta \in \mathcal{U}_{ad}$.

Problem: Find $u_\Delta^{0,\min}$ such that

$$J_\Delta(u_\Delta^{0,\min}) = \min_{u^0_\Delta \in \mathcal{U}_{ad}^\Delta} J_\Delta(u^0_\Delta). \tag{11}$$

Methods to approximate the gradient

- The discrete approach: differentiable schemes.
- The discrete approach: non-differentiable schemes.
- The continuous approach: Internal boundary conditions on the shock.
- The continuous approach: The alternating descent method.

The discrete approach: Differentiable numerical schemes

Assume that the Burgers equation is approximated by a differentiable conservative numerical scheme

$$u_j^{n+1} = u_j^n - \lambda(g_{j+1/2}^n - g_{j-1/2}^n), \quad j \in \mathbb{Z}, \quad n = 0, ..., N.$$

$$u_j^0 = u_{j,0}, \qquad \lambda = \Delta t / \Delta x$$

where

$$g_{j+1/2}^n = g(u_j^n, u_{j+1}^n)$$

and the numerical flux $g(u, v)$ is differentiable. For example,

$$g^{LF}(u,v) = \frac{u^2+v^2}{4} - \frac{v-u}{2\lambda}, \qquad \text{or} \quad g^{EO}(u,v) = u\frac{u+|u|}{4} + v\frac{v-|v|}{4}$$

The linearized scheme is well-defined as

$$\delta u_j^{n+1} = \delta u_j^n - \lambda \left(\partial_1 g_{j+1/2}^n \delta u_j^n + \partial_2 g_{j+1/2}^n \delta u_{j+1}^n - \partial_1 g_{j-1/2}^n \delta u_{j-1}^n - \partial_2 g_{j-1/2}^n \delta u_j^n \right) = 0,$$
$$j \in \mathbb{Z}, \quad n = 0, ..., N.$$

The derivative of the cost functional

$$\delta J_\Delta = \frac{\Delta x}{2} \sum_{j=-\infty}^{\infty} (u_j^{N+1} - u_j^d)^2, \tag{12}$$

is given by

$$\delta J_\Delta = \Delta x \sum_{j=-\infty}^{\infty} (u_j^{N+1} - u_j^d) \delta u_j^{N+1}, \tag{13}$$

where δu_j^n solves the above linearized system. If we introduce the following adjoint system

$$\begin{aligned} p_j^n &= p_j^{n+1} + \lambda \left(\partial_1 g_{j+1/2}^n (p_{j+1}^{n+1} - p_j^{n+1}) + \partial_2 g_{j-1/2}^n (p_j^{n+1} - p_{j-1}^{n+1}) \right), \\ p_j^{N+1} &= (u_j^{N+1} - u_j^d), \quad j \in \mathbb{Z}, \quad n = 0, ..., N. \end{aligned}$$

it is easy to check that

$$\delta J_\Delta = \Delta x \sum_{j \in \mathbb{Z}} (u_j^{N+1} - u_j^d) \delta u_j^{N+1} = \Delta x \sum_{j \in \mathbb{Z}} p_j^0 \delta u_j^0.$$

Thus, the gradient of J_Δ is given by p_j^0.

Lax-Friedrichs

$$\begin{cases} \frac{u_j^{n+1} - \frac{u_{j-1}^n + u_{j+1}^n}{2}}{\Delta t} + \frac{f(u_{j+1}^n) - f(u_{j-1}^n)}{2\Delta x} = 0, \qquad n = 0, ..., N, \\ u_j^0 = u_{0,j}, \qquad j \in \mathbb{Z}, \end{cases} \tag{14}$$

Adjoint:

$$\begin{cases} \frac{p_j^n - \frac{p_{j+1}^{n+1} + p_{j-1}^{n+1}}{2}}{\Delta t} + u_j^n \frac{p_{j-1}^{n+1} - p_{j+1}^{n+1}}{2\Delta x} = 0, \qquad n = 0, ..., N \\ p_j^{N+1} = p_j^T, \quad j \in \mathbb{Z}, \end{cases} \tag{15}$$

with $p_j^T = (u_j^{N+1} - u_j^d)$.

The discrete approach: Non-differentiable numerical schemes

Assume now that the Burgers equation is approximated by a non-differentiable conservative numerical scheme

$$u_j^{n+1} = u_j^n - \lambda(g_{j+1/2}^n - g_{j-1/2}^n), \quad j \in \mathbb{Z}, \quad n = 0, ..., N.$$

$$u_j^0 = u_{j,0}$$

where

$$g_{j+1/2}^n = g(u_j^n, u_{j+1}^n)$$

and the numerical flux $g(u,v)$ is non-differentiable. For example,

$$g^{Up}(u,v) = \frac{1}{4}(u^2 + v^2 - |u+v|(v-u))$$

In this case non-smooth optimization techniques are necessary.
A proposed linearization (Godlewski-Raviart, 1995),

$$\delta g(u,v) = \frac{1}{4}((2u+2v)(\delta u + \delta v) - |u+v|(\delta v - \delta w))$$

The continuous approach in presence of a single shock

Assume that $u(x,t)$ is a weak entropy solution of Burgers equation with a discontinuity along a regular curve $\Sigma = \{(t, \varphi(t)), t \in [0,T]\}$, which is Lipschitz continuous outside Σ. In particular, it satisfies the Rankine-Hugoniot condition on Σ

$$\varphi'(t)[u]_{\varphi(t)} = \left[u^2/2\right]_{\varphi(t)} . \tag{16}$$

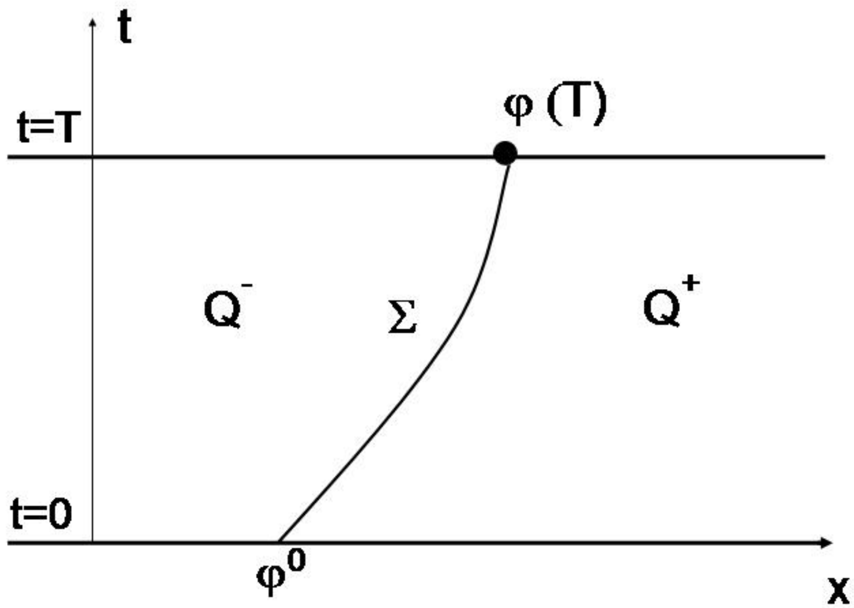


Figure 2: Subdomains Q^- and Q^+.

Then the pair (u, φ) satisfies the system

$$
\begin{cases}
\partial_t u + \partial_x(\frac{u^2}{2}) = 0, & \text{in } Q^- \cup Q^+, \\
\varphi'(t)[u]_{\varphi(t)} = \left[u^2/2\right]_{\varphi(t)}, & t \in (0, T), \\
\varphi(0) = \varphi^0, & \\
u(x, 0) = u^0(x), & \text{in } \{x < \varphi^0\} \cup \{x > \varphi^0\}.
\end{cases} \tag{17}
$$

The **generalized tangent vector** $(\delta u, \delta\varphi)$ satisfies the following linearized system (Bressan and Marson, Ulbrich, Godlewski and Raviart, etc.):

$$
\begin{cases}
\partial_t \delta u + \partial_x(u\delta u) = 0, \quad \text{in } Q^- \cup Q^+, \\
\delta\varphi'(t)[u]_{\varphi(t)} + \delta\varphi(t)\left(\varphi'(t)[u_x]_{\varphi(t)} - [u_x u]_{\varphi(t)}\right) \\
\quad +\varphi'(t)[\delta u]_{\varphi(t)} - [u\delta u]_{\varphi(t)} = 0, \quad \text{in } (0, T), \\
\delta u(x, 0) = \delta u^0, \quad \text{in } \{x < \varphi^0\} \cup \{x > \varphi^0\}, \\
\delta\varphi(0) = \delta\varphi^0,
\end{cases} \tag{18}
$$

with the initial data $(\delta u^0, \delta\varphi^0)$.

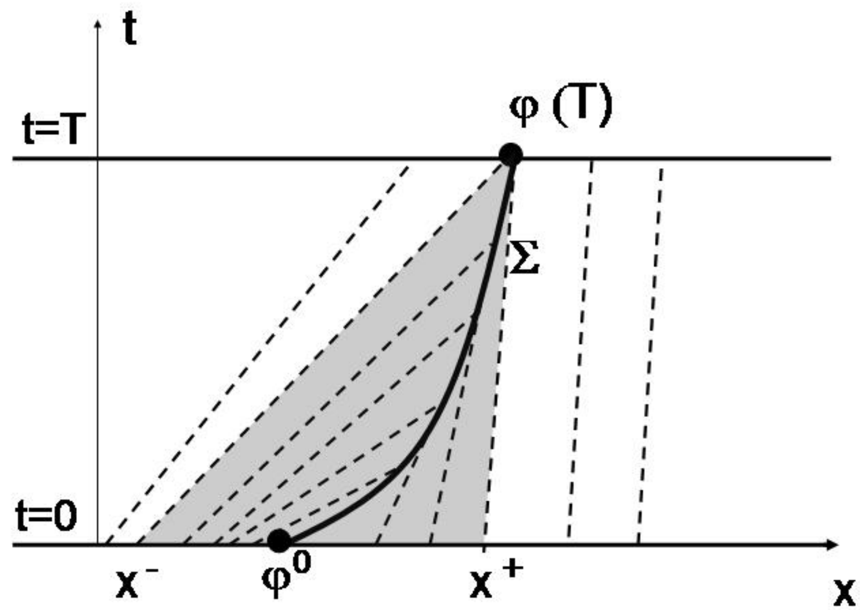


Figure 3: Characteristic lines entering on a shock

Variation of the functional J:

$$J(u^0) = \int_{\mathbb{R}} |u(x,T) - u^d|^2 dx$$

$$\delta J = \int_{\{x<\varphi(T)\}\cup\{x>\varphi(T)\}} (u(x,T) - u^d(x))\delta u(x,T) - \left[\frac{(u(x,T) - u^d(x))^2}{2}\right]_{\varphi(T)} \delta\varphi(T).$$

Lemma The Gateaux derivative of J can be written as

$$\delta J = \int_{\{x<\varphi^0\}\cup\{x>\varphi^0\}} p(x,0)\delta u^0(x)\, dx + q(0)[u^0]_{\varphi^0}\delta\varphi^0, \tag{19}$$

where the adjoint state pair (p, q) satisfies the system

$$\begin{cases} -\partial_t p - u\partial_x p = 0, \quad \text{in } Q^- \cup Q^+, \\ [p]_\Sigma = 0, \\ q(t) = p(\varphi(t), t), \text{ in } t \in (0,T) \\ q'(t) = 0, \text{ in } t \in (0,T) \\ p(x,T) = u(x,T) - u^d, \quad \text{in } \{x < \varphi(T)\} \cup \{x > \varphi(T)\} \\ q(T) = \frac{\frac{1}{2}\left[(u(x,T) - u^d)^2\right]_{\varphi(T)}}{[u]_{\varphi(T)}}. \end{cases} \tag{20}$$

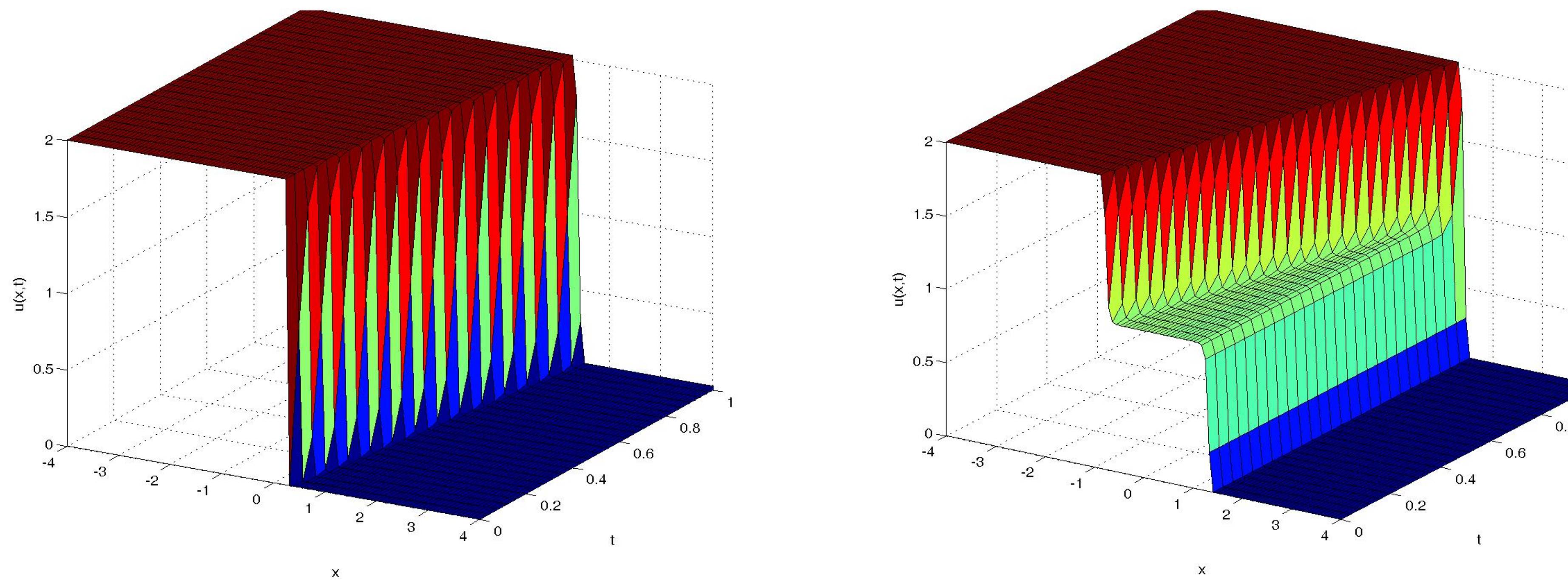


Figure 4: Solution $u(x,t)$ of the Burgers equation with an initial datum having a discontinuity (left) and adjoint solution which takes a constant value in the region occupied by the characteristics that meet the shock (right).

The new initial datum is ($\delta\varphi^0 > 0$)

$$u_j^{0,new} = \begin{cases} u_j^0 + \varepsilon\delta u_j^0, & \text{if } j < \varphi^0 \text{ or } j > \varphi^0 + \varepsilon\delta\varphi^0/\Delta x, \\ u_j^0 + \varepsilon\delta u_j^0 + [u_j^0]_{\varphi^0}, & \text{if } \varphi^0 \leq j \leq \varphi^0 + \varepsilon\delta\varphi^0/\Delta x. \end{cases}$$

The main drawbacks of this approach are the following:

1. At any step of the descent algorithm, a numerical approximation of the position of the shock is required.

2. The first component in $(p(x,0), q(0))$ has two discontinuities which are not at the same place at the discontinuity of u^0. Thus, an iterative gradient method based on this gradient generates increasingly complex initial data. Numerical experiments confirm that this actually occurs.

3. A pure displacement of the discontinuity will never be a descent direction computed by this method.

The alternating descent method

Let

$$x^- = \varphi(T) - u^-(\varphi(T))T, \qquad x^+ = \varphi(T) - u^+(\varphi(T))T,$$

and consider the following subsets ,

$$\hat{Q}^- = \{(x,t) \in \mathbb{R} \times (0,T) \text{ such that } x < \varphi(T) - u^-(\varphi(T))t\},$$

$$\hat{Q}^+ = \{(x,t) \in \mathbb{R} \times (0,T) \text{ such that } x > \varphi(T) - u^+(\varphi(T))t\}.$$

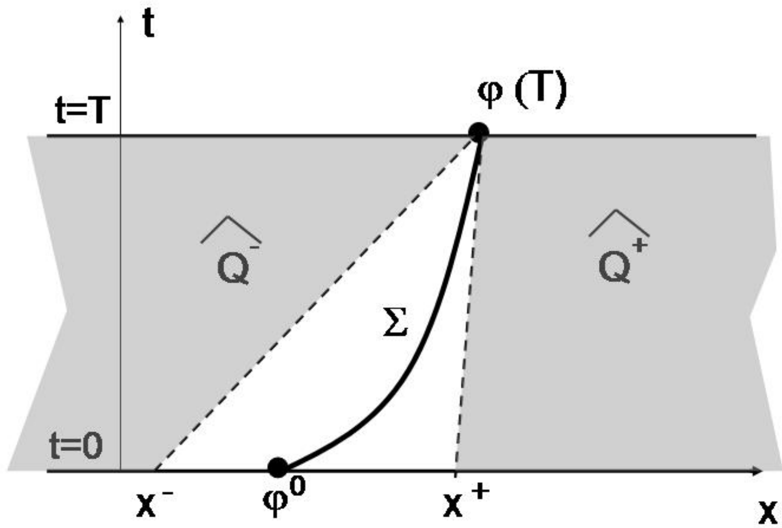


Figure 5: Subdomains $\hat{Q}^-$ and $\hat{Q}^+$

Theorem 1 *Assume that we restrict the generalized tangent vectors* $(\delta u^0, \delta\varphi^0) \in T_{u^0}$ *to those that satisfy,*

$$\delta\varphi^0 = \frac{\int_{x^-}^{\varphi^0} \delta u^0 + \int_{\varphi^0}^{x^+} \delta u^0}{[u]_{\varphi^0}}. \tag{21}$$

Then, the solution $(\delta u, \delta\varphi)$ *of the linearized system satisfies* $\delta\varphi(T) = 0$ *and the generalized Gateaux derivative of* J *in the direction* $(\delta u^0, \delta\varphi^0)$ *can be written as*

$$\delta J = \int_{\{x<x^-\}\cup\{x>x^+\}} p(x,0)\delta u^0(x)\, dx, \tag{22}$$

where p *satisfies the system*

$$\begin{cases} -\partial_t p - u\partial_x p = 0, & \text{in } \hat{Q}^- \cup \hat{Q}^+, \\ p(x,T) = u(x,T) - u^d, & \text{in } \{x < \varphi(T)\} \cup \{x > \varphi(T)\}. \end{cases} \tag{23}$$

Analogously, if we restrict the set of paths in Σ_{u^0} *to those for which the associated generalized tangent vectors* $(\delta u^0, \delta\varphi^0) \in T_{u^0}$ *satisfy* $\delta u^0 = 0$*, then* $\delta u(x,T) = 0$ *and the generalized Gateaux derivative of* J *in the direction* $(\delta u^0, \delta\varphi^0)$ *can be written as*

$$\delta J = -\left[\frac{(u(x,T) - u^d(x))^2}{2}\right]_{\varphi(T)} \frac{[u^0]_{\varphi^0}}{[u(\cdot,T)]_{\varphi(T)}}\delta\varphi^0. \tag{24}$$

Numerical experiments

Experiment 1. We first consider a piecewise constant target profile u^d given by

$$u^d = \begin{cases} 1 \text{ if } x < 0, \\ 0 \text{ if } x \geq 0, \end{cases} \tag{25}$$

and the time $T = 1$. Note that in this case one solution of the optimization problem is obviously given by

$$u^{0,min} = \begin{cases} 1 \text{ if } x < -1/2, \\ 0 \text{ if } x \geq 0. \end{cases} \tag{26}$$

This means that the optimal value $u^{0,min}$ can be attained and the minimum value of J in this case is zero.

$\log(J^\Delta)$	-3	-4	-5	-6	-7
Lax-Friedrichs	14	39	> 1000		
Engquist-Osher	26	85	288	> 1000	
Roe	18	33	54	114	> 1000
Imposing b.c.	5	6	9	21	> 1000
Alternating descent	3	3	3	Not attained	

$\log(J^\Delta)$	-3	-4	-5	-6	-7
Lax-Friedrichs	15	49	> 1000		
Engquist-Osher	115	673	> 1000		
Roe	185	> 1000			
Imposing b.c.	5	6	52	440	> 1000
Alternating descent	3	3	3	3	Not attained

Table 1: Experiment 1. Number of iterations needed for a descent algorithm to obtain the value of $\log(J)$ indicated in the upper row, by the different methods presented above. The upper table corresponds to $\Delta x = 1/20$ and the lower one to $\Delta x = 1/80$. In both cases $\lambda = \Delta t/\Delta x = 1/2$.

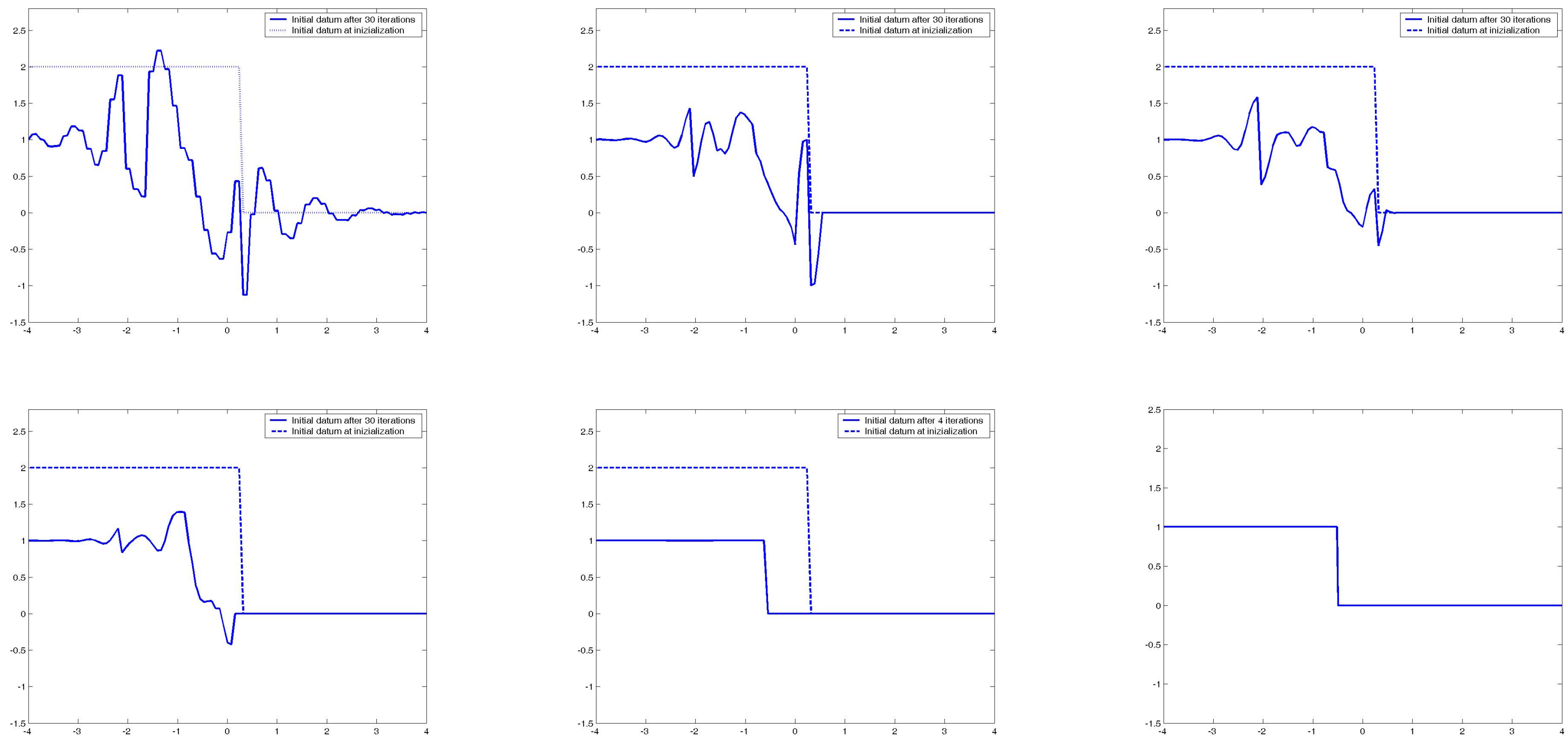
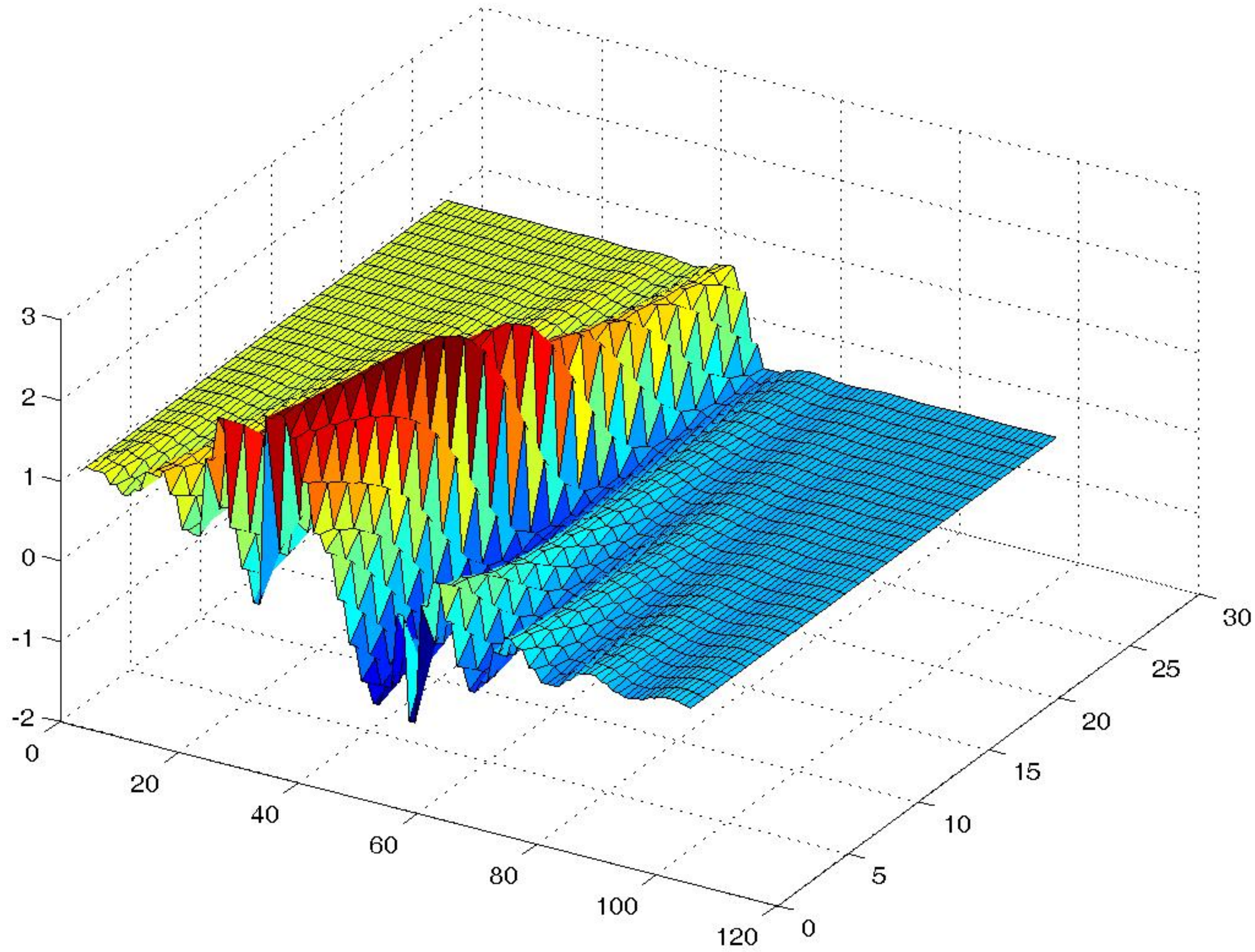


Figure 6: Experiment 1. Initialization (dashed line) and initial data obtained after 30 iterations (solid line) with Lax-Friedrichs (upper left) , Engquist-Osher (upper right), Roe (middle left), the continuous approach imposing a boundary condition on the shock (middle right) and the generalized tangent vectors decomposition method schemes (lower left). A minimizer u^0 of the continuous functional is given in the lower right figure.



3
2
1
0
-1
-2
0
20
40
60
80
100
120
0
5
10
15
20
25
30

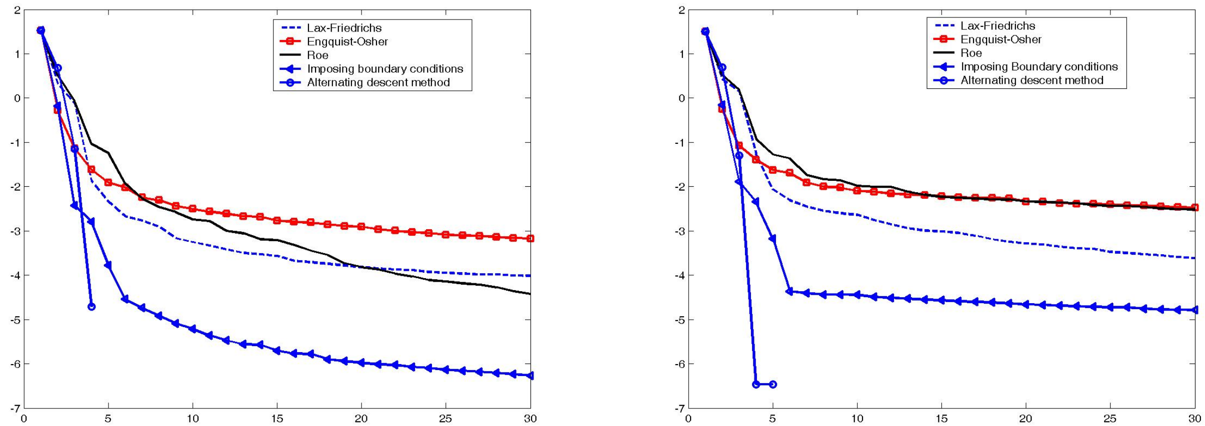


Figure 7: Experiment 1. Log of the value of the functional versus the number of iterations in the descent algorithm for the Lax-Friedrichs, Engquist-Osher and Roe schemes, the continuous approach imposing the internal boundary condition on the shock and the alternating descent method proposed in this article. The upper figure corresponds to $\Delta x = 1/20$ and the lower one to $\Delta x = 1/80$. We see that the last method stabilizes in a few iterations and it is much more efficient when consider small enoughvalues of Δx in order to be able to resolve the shock sufficiently well.

We observe the following:

1. Different numerical approximation and descent methods lead to different solutions.

2. For the first four methods the initial datum u^0 we obtain after the iteration process presents strong oscillations. That is not the case for the alternating descent method.

3. Numerical methods that ignore the presence of the shock (Lax-Friedrichs, Engquist-Osher and Roe) descend more slowly than those that take into account the sensitivity with respect to the shock position (by imposing the boundary condition on the shock or the alternating descent method).

4. For fixed Δx the alternating descent method stabilizes quickly in a few iterations. This is due to the fact that the descent direction is computed for the continuos system and not for the discrete one, and therefore Δx needs to be small for that computation to be valid at the discrete level as well.

5. For smaller values of Δx the only method that remains effective is the alternating descent method. The other methods descent more slowly.

Experiment 2. We consider the same target u^d as in the previous experiment but with different initial data. We see that different initialization functions u^0, with more or less discontinuities, do not alter the efficiency of the alternating descent method. The numerical results are presented in Figure 8.

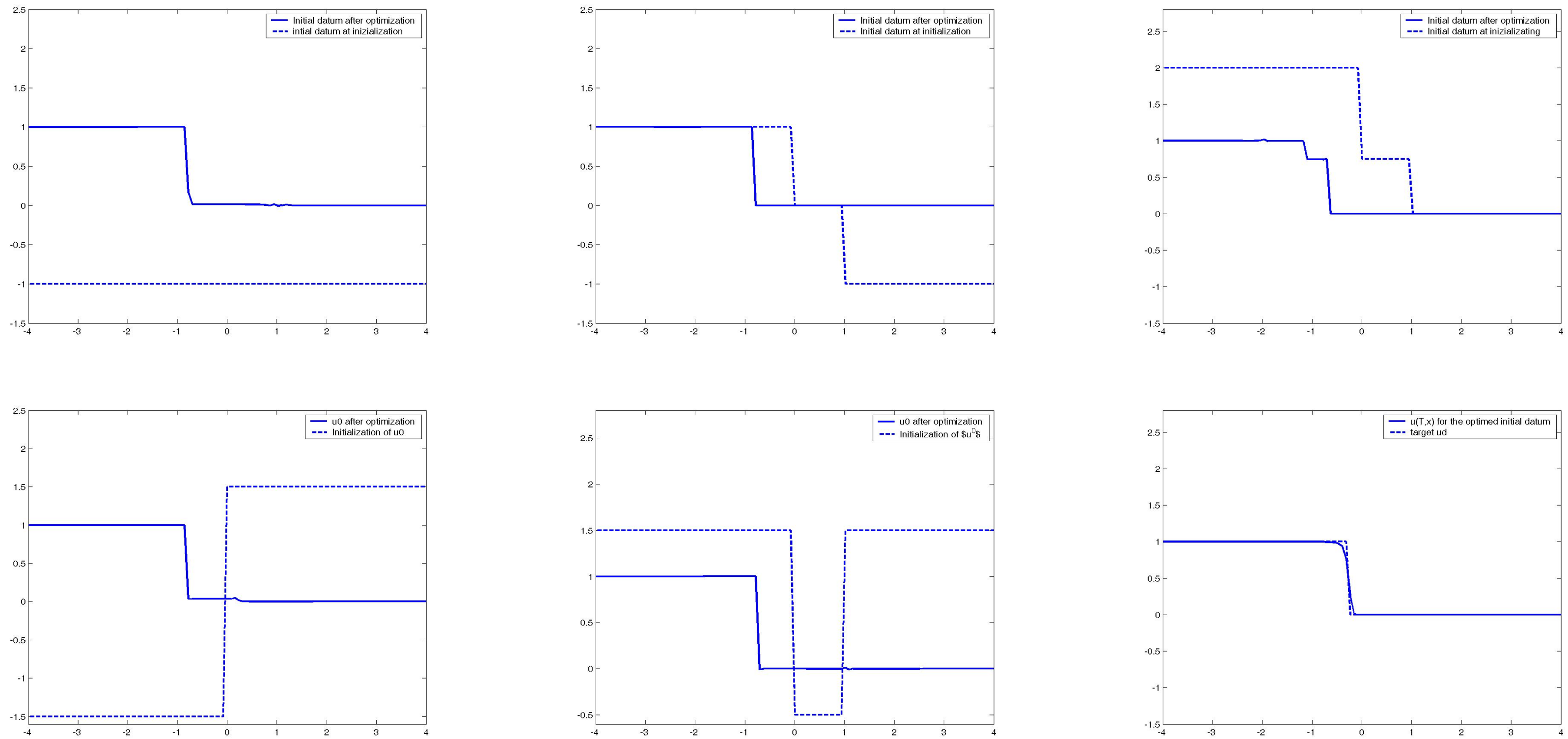


Figure 8: Experiment 2. The four upper figures and the lower left one show the initial data obtained once the descent iteration stops (solid) with different initialization functions u^0 (dashed) with the alternating descent method proposed in this article. In the lower right figure, the target $u^d(x)$ (dashed) and the solution $u(x,T)$ (here $T = 1$) corresponding to the obtained u^0 are drawn for the last initialization. The function $u(x,T)$ one obtains for the other initializations is very similar to this one.