Semesterproject
Optimal Control for Nonlinear Conservation Laws in the Presence of Shocks

supervised by: Prof. Dr. S. Mishra

Franziska Weber

October 30, 2011

1 Introduction

We consider the following control problem: Let $T > 0$ be finite, $F : \mathbb{R}^n \to \mathbb{R}^n$ be a strictly hyperbolic C^2-flux, and let $u : \mathbb{R} \times (0,T) \to \mathbb{R}^n$, the state variable, solve the system of conservation laws in one space dimension

$$\begin{aligned} &u_t + F(u)_x = 0, \qquad &(x,t) \in \mathbb{R} \times (0,T), \\ &u(x,0) = u^0(x), \qquad &x \in \mathbb{R}. \end{aligned} \tag{1}$$

where the control $u^0 : \mathbb{R} \to \mathbb{R}^n$ varies inside an admissible set $\mathcal{U}_{\mathrm{ad}} \subset L^1(\mathbb{R}, \mathbb{R}^n)$. Given a target function $u^d \in L^2(\mathbb{R}, \mathbb{R}^n)$ we consider the cost functional $J : L^1(\mathbb{R}, \mathbb{R}^n) \to \mathbb{R}$ to be minimized, defined by

$$J(u^0) = \frac{1}{2} \int_{\mathbb{R}} |u(x,T) - u^d(x)|^2 \, dx, \tag{2}$$

where $u(x,T)$ is a weak solution of (1).
We want to solve then the optimization problem: Find $u^{0,\min} \in \mathcal{U}_{\mathrm{ad}}$ such that

$$J(u^{0,\min}) = \min_{u^0 \in \mathcal{U}_{\mathrm{ad}}} J(u^0). \tag{3}$$

In order to perform numerical computations the above continuous problem has to be replaced by a discrete approximation. This means that instead of (1) and the cost functional J a discrete version are considered. We denote the discrete approximation of the solution u of (1) by u_Δ obtained by a discretization of (1) with mesh-sizes Δx and

Δt and by J_Δ and $\mathcal{U}_{\text{ad},\Delta}$ a discrete version of J and $\mathcal{U}_{\text{ad}}$ respectively. We consider the approximate discrete minimization problem

$$J_\Delta(u_\Delta^{0,\min}) = \min_{u_\Delta^0 \in \mathcal{U}_{\text{ad},\Delta}} J_\Delta(u_\Delta^0), \tag{4}$$

where we take J_Δ to be

$$J_\Delta(u_\Delta^0) = \frac{\Delta x}{2} \sum_{j=-\infty}^{\infty} (u_j^{N+1} - u_j^d)^2. \tag{5}$$

u_j^n denotes an approximation of $u(x_j, t^n)$ obtained as a solution of a discrete version of (1) computed on a mesh in $\mathbb{R} \times [0, T]$ given by $(x_j, t^n) = (j\Delta x, n\Delta t)$, $j = -\infty, \ldots, \infty$; $n = 0, \ldots, N+1$ so that $(N+1)\Delta t = T$, $u_\Delta^0 = \{u_j^0\}$ and $u_\Delta^d = \{u_j^d\}$ are discrete versions of the initial data and the target u^d respectively.

In [5] they propose the alternating descent method to approximate the optimization problem (3) in the case of one scalar conservation law with Burgers flux $f(u) = u^2$. We describe this method and compare its performance on the above control problem to the performace on the flux identification problem for scalar conservation laws [6]. We also discuss possible extensions of the method to problems with different cost functionals and to systems of conservation laws in one space dimension.

2 The alternating descent method for scalar conservation laws

We consider (1) for $n = 1$ and $f : \mathbb{R} \to \mathbb{R}$ a strictly convex C^2-flux function. We define the set of admissible initial data $\mathcal{U}_{\text{ad}}$

$$\mathcal{U}_{\text{ad}} = \{f \in L^\infty(\mathbb{R}), supp(f) \subset K, ||f||_\infty \leq C\}, \tag{6}$$

where $K \subset \mathbb{R}$ is a bounded interval and $C > 0$ a constant.

2.1 Existence of minimizers

Theorem 2.1. *[5] Assume $u^d \in L^2(\mathbb{R})$ and $\mathcal{U}_{ad}$ as in (6). Then the minimization problem,*

$$\min_{u^0 \in \mathcal{U}_{ad}} J(u^0), \tag{7}$$

has at least one minimizer $u^{0,min} \in \mathcal{U}_{ad}$. Moreover, uniqueness is in general false for this optimization problem.

The proof uses Oleinik's one-sided Lipschitz condition (OSLC) in order to get a uniform bound on the BV-norm of the minimizing sequence $u_n(\cdot, T)$ locally in space.

Remark 2.1. According to [5] the theorem can also be proved for admissible sets of the form $\mathcal{U}_{\text{ad}} = \{f \in L^1(\mathbb{R}), supp(f) \subset K, ||f||_1 \le C\}$ and for other optimization problems with different functionals J.

For the discrete version of the minimization problem (3), we introduce a 3-point numerical approximation scheme for (1):

$$u_j^{n+1} = u_j^n - \lambda(g_{j+1/2}^n - g_{j-1/2}^n), \quad \lambda = \frac{\Delta t}{\Delta x}, \quad j \in \mathbb{Z}, n = 0, \dots, N, \tag{8}$$

where

$$g_{j+1/2}^n = g(u_j^n, u_{j+1}^n), \tag{9}$$

is a consistent numerical flux. For each fixed $(\Delta x, \Delta t)$ (satisfying a suitable CFL condition) the discrete analogue of Theorem 2.1 holds. But to pass to the limit as $\Delta t, \Delta x \to 0$, we need a discrete version of Oleinik's entropy condition:

$$\frac{u_{j+1}^n - u_j^n}{\Delta x} \le \frac{1}{\alpha n \Delta t}. \tag{10}$$

Theorem 2.2. *[5] Assume that u_Δ^n is obtained by a conservative monotone numerical scheme consistent with (1) and satisfying the discrete OSLC (10). Then*

1. *For all $\Delta x, \Delta t > 0$, the discrete minimization problem (4) has at least one solution $u_\Delta^{0,min} \in \mathcal{U}_{ad,\Delta}$.*

2. *Any accumulation point of $u_\Delta^{0,min}$ with respect to the weak-* topology in L^∞, as $\Delta x, \Delta t \to 0$ (with $\Delta t/\Delta x = \lambda$ fixed and under a suitable CFL condition), is a minimizer of the continuous problem (3).*

Remark 2.2. Numerical schemes of the form (8) satisfying the discrete OSLC are Lax-Friedrichs, Engquist-Osher or Godunov.

2.1.1 Extension arbitrary flux functions f and to higher order numerical schemes

According to [2] the OSLC has been proved only for scalar conservation laws with strictly convex flux functions. In the proofs of the theorems above the OSLC is used to obtain a uniform bound on the BV-norm of the sequence, which yields a convergent subsequence by the compactness of the embedding $BV(I) \subset L^2(I)$ for a bounded interval I. If we can obtain the uniform bound in a different way, we do not need the OSLC, for example by restricting to initial data with BV-norm bounded by some constant $C > 0$. Thus, we obtain existence of a minimizer for an arbitrary C^2-flux function f.

Similarly, for the discrete approximations u^n_Δ, if we can show that they are locally uniformly bounded in the discrete BV-norm, this is enough to guarantee convergence of a subsequence to a discrete minimizer. Hence if we restrict to initial data that is BV-bounded, we can use any scheme with the TVB property for the approximation of the solution u.

2.2 Sensitivity analysis (as in [5])

2.2.1 Sensitivity without shocks

Let $u^0 \in C^1_0(\mathbb{R})$ be an initial datum for which there exists a classical solution $u(x,t)$ of (1) in $(x,t) \in \mathbb{R} \times [0,T]$ and which can be extended to a classical solution in $t \in [0, T+\tau]$ for some $\tau > 0$. Then for $\epsilon > 0$ small enough, the solution $u^\epsilon(x,t)$ corresponding to the initial datum

$$u^{\epsilon,0}(x) = u^0(x) + \epsilon\, v^0(x),$$

where $v^0 \in C^1_0(\mathbb{R})$, is also a classical solution in $\mathbb{R} \times (0,T)$ and $u^\epsilon \in C^1(\mathbb{R} \times [0,T])$ can be written as

$$u^\epsilon = u + \epsilon\, v + \mathcal{O}(\epsilon^2) \quad \text{with respect to } C^1\text{-topology},$$

where v is the solution of the linearized equation

$$\begin{aligned} v_t + (f'(u)v)_x &= 0, && (x,t) \in \mathbb{R} \times (0,T), \\ v(x,0) &= v^0(x), && x \in \mathbb{R}. \end{aligned} \tag{11}$$

Assuming u, v, u^d smooth, the Gateaux derivative δJ of J at u^0 in the direction v^0 is given by

$$\delta J = \int_{\mathbb{R}} (u(x,T) - u^d(x)) v(x,T)\, dx,$$

where v solves (11). We introduce the adjoint system

$$\begin{aligned} -p_t - f'(u)p_x &= 0, && (x,t) \in \mathbb{R} \times (0,T), \\ p(x,T) = p^T(x) = u(x,T) - u^d(x), & && x \in \mathbb{R}. \end{aligned} \tag{12}$$

Multiplying (11) by p and integrating by parts, we obtain

$$\int_{\mathbb{R}} p^T(x) v(x,T)\, dx = \int_{\mathbb{R}} p(x,0) v^0(x)\, dx.$$

Hence

$$\delta J = \int_{\mathbb{R}} p(x,0) v^0(x)\, dx$$

and we can choose $v^0 = -p(x,0)$ as descent direction. Note also, that u and p have the same characteristics.

2.2.2 Sensitivity in the presence of shocks

Now we assume that $u(x,t)$ is a weak solution of (1) with a discontinuity along a regular curve $\Sigma = \{(t, \phi(t)), t \in [0,T]\}$ which is Lipschitz continuous outside Σ. Σ divides $\mathbb{R} \times (0,T)$ in two parts: Q^- and Q^+ to the 'left' and the 'right' of Σ respectively. Then the pair (u, ϕ), where ϕ denotes the discontinuity location, satisfies the system:

$$\begin{aligned} u_t + f(u)_x &= 0, && (x,t) \in Q^- \cup Q^+, \\ \phi'(t)[u]_{\phi(t)} &= [f(u)]_{\phi(t)}, && t \in (0,T), \\ \phi(0) &= \phi^0, && \\ u(x,0) &= u^0(x), && x \in \{x < \phi^0\} \cup \{x > \phi^0\}. \end{aligned} \tag{13}$$

In order to analyze the sensitivity of (u, ϕ) with respect to perturbations $(v^0, \delta\phi^0)$ of the initial datum, where $v^0 \in L^1(\mathbb{R})$ a perturbation of the initial profile u^0 and $\delta\phi^0$ a perturbation of the initial discontinuity location ϕ^0, we use generalized tangent vectors as introduced in [4]. In that paper the concept of generalized tangent vectors is introduced for strictly hyperbolic systems of conservation laws with piecewise Lipschitz continuous solutions. We consider now a generalized tangent vector $(v^0, \delta\phi^0) \in L^1(\mathbb{R}) \times \mathbb{R}$ of (u^0, ϕ^0). Let $u^{0,\epsilon} \in \Sigma_{u^0}$ be a path generating $(v^0, \delta\phi^0)$. For ϵ sufficiently small, the solution $u^\epsilon(\cdot, t)$ of (13) with initial datum $u^{0,\epsilon}, \phi^{0,\epsilon}$ is Lipschitz continuous with a single discontinuity at $x = \phi^\epsilon(t)$, for all $t \in [0,T]$. Therefore, $u^\epsilon(\cdot,t)$ generates a generalized tangent vector $(v(\cdot,t), \delta\phi(t)) \in L^1(\mathbb{R}) \times \mathbb{R}$. It satisfies the linearized system [3]

$$\begin{aligned} v_t + (f'(u)v)_x &= 0, && (x,t) \in Q^- \cup Q^+, \\ \delta\phi'(t)[u]_{\phi(t)} + \delta\phi(t)(\phi'(t)[u_x]_{\phi(t)} - [f'(u)u_x]_{\phi(t)}) & && \\ +\phi'(t)[v]_{\phi(t)} - [f'(u)v]_{\phi(t)} &= 0, && t \in (0,T), \\ \delta\phi(0) &= \delta\phi^0, && \\ v(x,0) &= v^0(x), && x \in \{x < \phi^0\} \cup \{x > \phi^0\}. \end{aligned} \tag{14}$$

In [8, 4] the corresponding system of equations is derived for systems of conservation laws in one-space dimension, where the solution u is assumed to be Lipschitz continuous outside Σ.

Remark 2.3. For a linear transport equation $u_t + a\, u_x = 0$ we have that the speed of a discontinuity tisfies (due to the Rankine-Hugoniot condition) $\phi(t) = a$. Hence (14) becomes

$$\begin{aligned} v_t + a\, v_x &= 0, && (x,t) \in Q^- \cup Q^+, \\ \delta\phi'(t) &= 0, && t \in (0,T), \\ \delta\phi(0) = \delta\phi^0 &= \delta\phi(T), && \\ v(x,0) &= v^0(x), && x \in \{x < \phi^0\} \cup \{x > \phi^0\}. \end{aligned}$$

For ϵ small enough, we can write

$$(u_\epsilon, \phi_\epsilon) = (u, \phi) + \epsilon(v, \delta\phi) + \mathcal{O}(\epsilon^2)$$

Definition 2.1. [5] Let $J : L^1(\mathbb{R}) \to \mathbb{R}$ be a functional and $u^0 \in L^1(\mathbb{R})$ be Lipschitz continuous with a single discontinuity at $x = \phi^0$ and such that the solution u of (1) is Lipschitz continuous outside a C^1-curve Σ. J is *Gateaux differentiable in a generalized sense* if for any generalized tangent vector $(v^0, \delta\phi^0)$ and any family $u^{0,\epsilon} \in \Sigma_{u^0}$ associated to $(v^0, \delta\phi^0)$ the following limit exists,

$$\delta J := \lim_{\epsilon \to 0} \frac{J(u^{0,\epsilon}) - J(u^0)}{\epsilon} \tag{15}$$

and depends only on (u^0, ϕ^0) and $(v^0, \delta\phi^0)$. We call δJ *generalized Gateaux derivative of J* in the direction $(v^0, \delta\phi^0)$.

Proposition 2.3. *[5] The generalized Gateaux derivative of J can be written as*

$$\delta J = \int_{\{x<\phi^0\}\cup\{x>\phi^0\}} p(x,0) v^0(x)\, dx + q(0)[u^0]_{\phi^0} \delta\phi^0, \tag{16}$$

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

$$\begin{aligned} -p_t - f'(u) p_x &= 0, && (x,t) \in Q^- \cup Q^+, \\ [p]_\Sigma &= 0, && \\ q(t) &= p(\phi(t), t), && t \in (0,T) \\ q'(t) &= 0, && t \in (0,T) \\ p(x,T) &= u(x,T) - u^d(x), && x \in \{x < \phi(T)\} \cup \{x > \phi(T)\} \\ q(T) = \frac{[\frac{1}{2}(u(\cdot,T) - u^d)^2]_{\phi(T)}}{[u]_{\phi(T)}}&. && \end{aligned} \tag{17}$$

Formula (16) provides a way to compute a first descent direction of J at u^0, namely we can take

$$(v^0, \delta\phi^0) = (-p(x,0), -q(0)[u]_{\phi^0}). \tag{18}$$

The value of $\delta\phi^0$ should be interpreted as the optimal infinitesimal displacement of the discontinuity of u^0. But this $(v^0, \delta\phi^0)$ is not a generalized tangent vector since $p(x,0)$ is not continuous away from $x \neq \phi^0$. $p(x,t)$ takes the same constant value $q(T)$ in the whole region occupied by the characteristics of (1) which meet the discontinuity Σ. Hence p has in general two discontinuities at the boundaries of this region and therefore $p(x,0)$ too. If we take (18) as descent direction, the the new initial datum $u^{0,\text{new}}$ should be obtained from u^0 following a path associated to this descent direction and therefore have the form

$$u^{0,\text{new}} = \begin{cases} u^0 + \epsilon\, v^0 + [u^0]_{\phi^0} \chi_{[\phi^0+\epsilon\delta\phi^0, \phi^0]}, & \text{if } \delta\phi^0 < 0, \\ u^0 + \epsilon\, v^0 - [u^0]_{\phi^0} \chi_{[\phi^0, \phi^0+\epsilon\delta\phi^0]}, & \text{if } \delta\phi^0 > 0, \end{cases} \tag{19}$$

for some ϵ small enough and correctly chosen. But $u^{0,\text{new}}$ has now three discontinuities whereas u^0 has only one. Hence in an iterative process the number of discontinuities the initial datum has will increase in each step and therefore increase the complexity of the solution.

Remark 2.4. In general the system of adjoint equations is not uniquely solvable. However, system (17) is uniquely solvable, because of the second, third and forth equation in (17). They are needed to define the value of p on Σ. The solution of this system is a so-called reversible solution, introduced in [1], which satisfies certain support, monotonicity and total variation properties and an entropy inequality (see [1]). In the class of reversible solutions, the adjoint equation

$$\begin{aligned} -p_t - f'(u)p_x &= 0, \qquad && (x,t) \in \mathbb{R} \times (0,T) \\ p(x,T) &= p^T(x), \qquad && x \in \mathbb{R}, \end{aligned} \tag{20}$$

where p^T is locally Lipschitz, are uniquely solvable. The solution can be constructed by the method of characteristics backwards in time (see [1]).

Remark 2.5. These adjoint equations do not coincide with those in [3]. If one computes the adjoint equations of [3] for the case of a scalar conservation law, one obtains

$$-p_t - f'(u)p_x = 0 \tag{21}$$

outside the line where u is discontinuous, together with

$$\begin{aligned} q'(t) &= \frac{q(t)}{[u]_{\phi(t)}} \cdot \left\{\phi'(t)[u_x]_{\phi(t)} - [f'(u)u_x]_{\phi(t)}\right\} \\ p^+(t) &= \frac{q(t)}{[u]_{\phi(t)}} \cdot \operatorname{sign}(f'(u^+) - \phi'(t)) \\ p^-(t) &= \frac{-q(t)}{[u]_{\phi(t)}} \cdot \operatorname{sign}(f'(u^-) - \phi'(t)) \end{aligned} \tag{22}$$

where u is discontinuous. With Lax' entropy condition, the last two identities become

$$p^+(t) = -\frac{q(t)}{[u]_{\phi(t)}},$$
$$p^-(t) = -\frac{q(t)}{[u]_{\phi(t)}}.$$

So the solution of the adjoint equation does not have a discontinuity in the place where u does have one (at least in the scalar case). Nevertheless it could have discontinuities in other places (especially if $p^T(x)$ is discontinuous). Bressan and Marson choose these adjoint equations because they have the property that the duality product defined by

$$\langle (v(\cdot,t), \delta\phi(t)), (p(\cdot,t), q(t))\rangle := \int_{\mathbb{R}} v(x,t)p(x,t)\,dx + \delta\phi(t)\cdot q(t) \tag{23}$$

remains constant in time, in particular

$$\int_{\mathbb{R}} v^0(x)p(x,0)\,dx + \delta\phi^0 \cdot q(0) = \int_{\mathbb{R}} v(x,T)p^T(x)\,dx + \delta\phi(T)\cdot q(T). \tag{24}$$

If we then choose as terminal condition

$$\begin{aligned} p(x,T) &= u(x,T) - u^d(x), \quad x \in \{x < \phi(T)\} \cup \{x > \phi(T)\}, \\ q(T) &= \left[\frac{1}{2}(u(\cdot,T) - u^d)^2\right]_{\phi(T)} \end{aligned} \tag{25}$$

we get

$$\delta J = \langle (v(\cdot,t), \delta\phi(t)), (p(\cdot,t), q(t))\rangle$$

for any $t \in [0,T]$ and thus a steepest descent direction is

$$(v^0, \delta\phi^0) = (-p(x,0), -q(0)). \tag{26}$$

In contrast to the solution of (17) the solution of (21), (22), (25) might not be constant in the region occupied by the characteristics of u which meet Σ. This is only the case if $p^+(t) = p^-(t) = const.$

2.3 Alternating descent directions

Notation. We define

$$x^- = \phi(T) - f'(u^-(\phi(T)))T \quad x^+ = \phi(T) - f'(u^+(\phi(T)))T$$

and the subsets

$$\widehat{Q}^- = \{(x,t) \in \mathbb{R}\times(0,T) \text{ such that } x < x^- + f'(u^-(\phi(T)))t\}$$
$$\widehat{Q}^+ = \{(x,t) \in \mathbb{R}\times(0,T) \text{ such that } x < x^+ + f'(u^+(\phi(T)))t\}.$$

Proposition 2.4. *[5] Assume that we restrict to the set of paths in Σ_{u^0} for which the associated generalized tangent vectors $(v^0, \delta\phi^0) \in T_{u^0}$ satisfy*

$$\delta\phi^0 = \frac{\int_{x^-}^{\phi^0} \delta u^0 + \int_{\phi^0}^{x^+} \delta u^0}{[u]_{\phi^0}}. \tag{27}$$

Then the solution $(v, \delta\phi)$ of system (14) satisfies $\delta\phi(T) = 0$ and the generalized Gateaux derivative of J in the direction $(v^0, \delta\phi^0)$ can be written as

$$\delta J = \int_{\{x<x^-\}\cup\{x>x^+\}} p(x,0)v^0(x)\,dx \tag{28}$$

where p satisfies the system

$$\begin{aligned} -p_t - f'(u)p_x = 0, \qquad & (x,t) \in \widehat{Q}^- \cup \widehat{Q}^+, \\ p(x,T) = u(x,T) - u^d(x), \qquad & x \in \{x < \phi(T)\} \cup \{x > \phi(T)\}. \end{aligned} \tag{29}$$

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

$$\delta J = -\left[\frac{(u(\cdot,T) - u^d)^2}{2}\right]_{\phi(T)} \frac{[u^0]_{\phi^0}}{[u(\cdot,T)]_{\phi(T)}}\delta\phi^0. \tag{30}$$

Remark 2.6. Formula (28) provides a simplified expression for the generalized Gateaux derivative of J when considering descent directions that do not move the shock position at $t = T$. Note that system (29) does not allow to determine the value of p outside the region $\widehat{Q}^- \cup \widehat{Q}^+$, i.e. the region under the influence of the shock by the characteristic lines emanating from it, but for the evaluation of the generalized Gateaux derivative (28) this is not necessary.
In the same way, formula (30) provides a simplified expression of the generalized Gateaux derivative of J when considering descend directions $(v^0, \delta\phi^0)$ that correspond to a pure translation of the shock at $t = T$ and do not change the profile of the solution.

This suggests a decomposition of the set of generalized tangent vectors

$$T_{u^0} = T^1_{u_0} \oplus T^2_{u^0},$$

where $T^1_{u_0}$ contains the tangent vectors for which (27) holds and $T^2_{u^0}$ those for which $v^0 = 0$. Thus, we have two classes of descent directions for J at u^0. They are not optimal in the sense that they are not the steepest descent directions but they have three other important properties:

1. They are both descent directions.

2. They allow to split the profile and the shock location of the initial data.

3. They are true generalized tangent vectors and therefore keep the structure of the initial data without increasing its complexity.

When considering generalized tangent vectors belonging to $T^1_{u^0}$, we can choose as descent direction,

$$v^0 = \begin{cases} -p(x,0) & \text{if } x < x^-, \\ -\lim_{x\to x^-, x<x^-} p(x,0) & \text{if } x^- < x < \phi^0, \\ -\lim_{x\to x^+, x>x^+} p(x,0) & \text{if } \phi^0 < x < x^+, \\ -p(x,0) & \text{if } x^+ < x, \end{cases} \tag{31}$$

$$\delta\phi^0 = -\frac{\int_{x^-}^{\phi^0} v^0 + \int_{\phi^0}^{x^+} v^0}{[u]_{\phi^0}}, \tag{32}$$

while for $T^2_{u^0}$ a good choice is

$$v^0 = 0, \quad \delta\phi^0 = \left[\frac{(u(\cdot,T) - u^d)^2}{2}\right]_{\phi(T)} \frac{[u(\cdot,T)]_{\phi(T)}}{[u^0]_{\phi^0}}. \tag{33}$$

Remark 2.7. It seems that Zuazua is messing up minus signs a little bit:-).

Remark 2.8. For the adjoint equations (21), (22), (25) as in [3] the first part of Proposition 2.4 holds as well. The expression (30) in the second part changes slightly, namely we have for generalized tangent vectors $v^0, \delta\phi^0)$ with $v^0 = 0$, using (24)

$$\delta J = q(T) \cdot \delta\phi(T) = q(0) \cdot \delta\phi^0 \tag{34}$$

and hence a good choice of a generalized tangent vector in $T^2_{u^0}$ is

$$v^0 = 0, \quad \delta\phi^0 = -q(0). \tag{35}$$

2.3.1 The alternating descent method

The alternating descent algorithm uses the idea of splitting the descent direction into two generalized tangent vectors from $T^1_{u^0}$, $T^2_{u^0}$ respectively. We describe the algorithm for the case where we have a target function u^d which is Lipschitz continuous except for a single discontinuity at $x = x^d$ with negative jump, i.e. $[u^d]_{x^d} < 0$. To initialize the algorithm, we choose an initial datum u^0 in such a way that the solution at time $t = T$ has a profile similar to u^d, i.e. a Lipschitz continuous function with a single discontinuity of negative jump, located on an arbitrary point $x \in \mathbb{R}$. Then we approximate a minimizer of J alternating the following two steps: First we perturb the initial datum u^0 by only moving the discontinuity of the solution u of (1) at time $t = T$, regardless of its value to both sides of the discontinuity. To do so we choose a descent direction of the form (33) and find the optimal step size ϵ by which the initial datum must modified in the direction given by (33) (one-dimensional optimization problem, can be solved by bisection, Armijo's rule etc.). As a second step, we perturb the resulting $u^{0,(1)}$ without changing the position of the discontinuity of $u(x, T)$, i.e. we choose a descent direction of the form (31) to modifiy the value of the solution at time $t = T$ to both sides of the discontinuity. In practice, the deformations of the second step will slightly move the position of the discontinuity because of its nonlinear dependance on the parameter ϵ. For this reason, one has to iterate the two steps to assure a simultaneous better placement of the shock and a better fitting of the value of the solution away from it [5].

The main advantage of this method is, that for an initial datum u^0 which is Lipschitz continuous outside a single discontinuity located at ϕ^0, the new initial datum obtained by the algorithm above will again have a single discontinuity in general. Therefore, the iterative optimization process will not introduce new discontinuities in u^0.

Remark 2.9. In the case of a scalar conservation law in one space dimension $\delta\phi^0$ is a scalar quantity and it is therefore not necessary to compute expression (33)/ (35) explicitely. They are multiplied by the parameter ϵ which is determined by an optimization algorithm as bisection, Armijo's rule etc. as mentioned above. Hence we can just as well take $\delta\phi^0 = 1$ and determine the optimal ϵ for this choice.

3 Generalization of the method

3.1 Different choice of the cost functional J

The method seems to be generalizable to other cost functionals which depend only on $x, u(\cdot, T), u^d$. Nevertheless the generalized Gateaux derivative has to be computed again for each choice of J. This will also yield different terminal conditions for the adjoint equations. As an example we consider the cost functional [3]

$$J(u) = \int_{\mathbb{R}} V(x, u(T, x))\, dx, \tag{36}$$

where V is continuously differentiable (see [3]). In this case the second equation in (12) (terminal condition) becomes

$$p(x, T) = p^T(x) = \nabla_u V(x, u(x, T)), \quad x \in \mathbb{R} \tag{37}$$

and in the case of shocks the last two equations in (17) become

$$\begin{aligned} p(x, T) = p^T(x) &= \nabla_u V(x, u(x, T)), \qquad x \in \{x < \phi(T)\} \cup \{x > \phi(T)\} \\ q(T) &= [V(\phi(T), u(\phi(T), T))]_{\phi(T)}. \end{aligned} \tag{38}$$

Next we consider the modified cost functional

$$J_\rho(u) = J_0(u) + \frac{\rho}{2}|\mu_1(u(\cdot,T)) - \mu_1(u^d)|^2 \tag{39}$$

where

$$J_0(u) = \frac{1}{2}\int_{\mathbb{R}} |u(x,T) - u^d|^2\,dx$$

and $\mu_1(f)$ is the first moment of the function $f : \mathbb{R} \to \mathbb{R}$:

$$\mu_1(f) = \int_{\mathbb{R}} x f(x)\,dx.$$

Let $u^\epsilon \in \Sigma_u$ denote a continuous path generated by the tangent vector $(v, \delta\phi) \in T_u$, i.e. at first order, u^ϵ looks as

$$u^\epsilon = u + \epsilon v - \operatorname{sign}(\delta\phi)[u]_{\phi(t)}\chi_{[\min(\phi,\phi+\epsilon\delta\phi),\max(\phi,\phi+\epsilon\delta\phi)]}.$$

Then we have

$$\begin{aligned}\lim_{\epsilon\to 0}\frac{J_\rho(u^\epsilon) - J_\rho(u)}{\epsilon} &= \int_{\mathbb{R}}(u(x.T) - u^d(x))v(x,T)\,dx - \frac{[(u-u^d)^2]_{\phi(T)}}{2}\delta\phi(T) \\ &\quad + \rho\mu_1(u(\cdot,T) - u^d)(\mu_1(v) - [u]_{\phi(T)}\cdot\delta\phi(T)\cdot\phi(T)).\end{aligned} \tag{40}$$

Therefore the terminal condition for the adjoint system in the case where we control the initial data should be

$$\begin{aligned} p(x,T) &= u(x,T) - u^d(x) + \rho\mu_1(u(\cdot,T) - u^d)\cdot x, \qquad x \in \mathbb{R}\backslash\phi(T) \\ q(T) &= -\left[\frac{(u(\phi(T),T) - u^d(\phi(T))^2}{2}\right]_{\phi(T)} - \rho\mu_1(u(\cdot,T) - u^d)[u]_{\phi(T)}\cdot\phi(T). \end{aligned} \tag{41}$$

3.2 Several shocks

The method can also be applied to target functions having several discontinuities. The generalized tangent vectors will then have the form $(v, \underline{\delta\phi})$, where $\underline{\delta\phi} = (\delta\phi^1, \ldots, \delta\phi^k)$ is a vector with the shock locations. The equations for v and the adjoint p are derived in [3] for the case of noninteracting shocks. Anyway, in the case of scalar conservation laws we do not need to consider the case of interacting shocks, an interaction of two shocks will there always result in one emanating shock.

3.3 Comparison with the performance of the method on the flux identification problem

In [6] the alternating descent method is applied to the flux identification problem for a scalar conservation law, i.e.

$$\begin{aligned} u_t + f(u)_x &= 0, \qquad && (x,t) \in \mathbb{R} \times (0,T) \\ u(x,0) &= u^0(x), \qquad && x \in \mathbb{R}, \end{aligned} \tag{42}$$

where $T > 0$, u^0 given, the flux f is the control parameter and $J : C^1(\mathbb{R}) \to \mathbb{R}$ the cost functional to be minimized

$$J(f) = \frac{1}{2} \int_{\mathbb{R}} |u(x,T) - u^d(x)|^2 \, dx. \tag{43}$$

f is chosen in a set of admissible functions $\mathcal{U}_{\text{ad}} \subset C^1(\mathbb{R})$. For the numerical approximations they choose a finite dimensional subset

$$\mathcal{U}_{\text{ad}}^M = \left\{ f = \sum_{j=1}^{M} \alpha_j f_j, \ \alpha_j \in \mathbb{R}, \ \|f\|_{W^{2,\infty}} \leq C \right\} \tag{44}$$

with $f_1, \ldots, f_M \in W^{2,\infty}([-\|u^0\|_{L^\infty}, \|u^0\|_{L^\infty}])$ and $C > 0$ a priori chosen. They use the same methods as in [5] to approximate the optimization problem, namely the 'discrete approach' (with Lax-Friedrichs and Roe scheme), the 'continuous approach' and the alternating descent method.

When we compare the approximate solutions $u(x,T)$ they compute in the numerical examples in [6] with the different methods we observe that none of the methods yields oscillatory solutions. The alternating descent method Zuazua and Castro propose converges faster in the four numerical experiments they conduct but in most cases the other methods do not perform much worse either. The reason why there are no oscillations might be that each method reduces the minimization problem to a finite dimensional one by looking for an optimal flux f in the span of $f_1, \cdots, f_m$ where $f_i : [-\|u^0\|_{L^\infty}, \|u^0\|_{L^\infty}] \to \mathbb{R}$ are polynomial functions. For any method this yields in each step a flux function of the form $f = \sum_{j=1}^{m} \alpha_j f_j$ where α_j are constants. They use the adjoint equations and possibly differentiated numerical schemes only for the computation of the α_j. The approximate solution $u(x,T)$ is computed from the same initial data u^0 (for every method and in every step) using either the Lax-Friedrichs or the Roe scheme by numerically approximating the flux $f = \sum_{j=1}^{m} \alpha_j f_j$ which varies from step to step and method to method but is in any case a polynomial. From the properties of the Lax-Friedrichs and the Roe scheme we expect that the computed solution is bounded and non-oscillatory.

In the first control problem we considered ((1),(2),(3)), where the control is the initial data, in the discrete method the differentiated numerical schemes are used to compute an approximate solution of the adjoint equations which gives the initial data for the next step. The differentiated numerical schemes are not TVB. This might be the reason why the computed optimal initial data is sometimes oscillatory (but the approximate solution $u(x, T)$ computed with the scheme could still be more regular if we use a MCC-scheme to compute it from the initial data). The with the so-called continuous approach computed initial data is probably more oscillatory because in each step a variation which is not a generalized tangent vector is added to the initial data from the previous step and this increases, as Zuazua and Castro explain in the paper, the number of discontinuities. Moreover, their numerical algorithms detect only one discontinuity in the approximation $u(x, T)$ in each step whereas this number might have increased during the iteration. This might be reasons why the alternating descent method performs considerably better than the other methods in the optimization problem with the initial data as the control parameter.

3.4 Generalization to systems of conservation laws in one space dimension

3.4.1 Problems/ Questions

If $n > 1$ in (1) a generalized tangent vector will have the form $(v; \delta\phi^1, \cdots, \delta\phi^k)$ where v is a piecewise Lipschitz function having discontinuities only where u^0 has and $\delta\phi^1, \ldots, \delta\phi^k$ are perturbations of the discontinuity locations $\phi^1, \ldots, \phi^k$ in the initial data u^0. In general, we cannot assume that there is only one discontinuity.

- The shocks might interact. In the scalar case we could choose an initial data such that the shocks do not interact in the time interval $[0, T]$. This might be different now, as we do not have a bound on the total variation of the solution at time T. In the scalar case the interaction of two shocks results in one shock. In the case of systems, the interaction can produce several new shocks. If we now perturb one shock location ϕ^j more than ϕ^{j+1} (or ϕ^{j-1}) for example, this might result in a new shock interaction which we did not have before. Hence perturbing $\phi^1, \ldots, \phi^k$ while setting $v = 0$ might not only move the shock positions at time T but also yield new discontinuities at $t = T$.

- How to find the descent direction that does not move the discontinuities in $u(x, T)$? In the case of a scalar conservation law, Zuazua [5] derives explicit expressions for them by using the method of characteristics and that the reversible solution of the adjoint equations is constant in the region occupied by the characteristics that meet the shock (see Proposition 2.4). For systems of conservation laws we do not have an analogue of the reversible solution. Moreover, the geometry is more complicated which might it make difficult to use the method of characteristics.

3.4.2 Different approach: Viscous approximation of adjoint equations

It seems difficult to generalize the alternating descent method to systems of conservation laws, therefore one could try to tackle the problem by a different approach, namely by using instead of the differentiated Lax-Friedrichs and Roe schemes a more stable numerical scheme for the discretization of the adjoint equations (the discrete approach in Zuazua's paper). The adjoint system of equations is non-conservative and might not admit a unique solution. Instead of (1) we could consider a viscous approximation which has a more regular solution:

$$\begin{aligned} u_t + F(u)_x &= \eta(B(u)u_x)_x, \quad (x,t) \in \mathbb{R} \times (0,T) \\ u(x,0) &= u^0(x), \qquad x \in \mathbb{R}, \end{aligned} \tag{45}$$

where $\eta > 0$ denotes the viscosity parameter, $F : \mathbb{R} \to \mathbb{R}^n$ a C^2-flux and $B : \mathbb{R} \to \mathbb{R}^{n\times n}$ the viscosity matrix (corresponding to the physical viscosity of the equations). The linearized system reads

$$\begin{aligned} v_t + A(u)v_x + (DA(u)v)u_x &= \eta\big(B(u)v_x + (DB(u)v)u_x\big)_x, \quad (x,t) \in \mathbb{R} \times (0,T) \\ v(x,0) &= v^0(x), \qquad x \in \mathbb{R}. \end{aligned} \tag{46}$$

$A(u) := F'(u)$, $DA(u) := DF'(u)$, the Jacobi matrix of $A(u)$. We choose the adjoint equations such that the L^2-inner product of v and the solution of the adjoint equation p is kept constant in time:

$$\frac{d}{dt}\int_{\mathbb{R}} p(x,t)v(x,t)\,dx = 0.$$

This means that p should be the (weak) solution of the following equations:

$$\begin{aligned} p_t + p_x A(u) + p\,\widetilde{DA}(u)u_x &= -\eta\big(p_{xx}B(u) + p_x\widetilde{DB}(u)u_x\big), \quad (x,t) \in \mathbb{R} \times (0,T) \\ p(x,T) &= p^T(x), \qquad x \in \mathbb{R}, \end{aligned} \tag{47}$$

where

$$\big(\widetilde{DM}(u)u_x\big)_{ij} = \sum_{k=1}^{n}\left(\frac{\partial M_{ij}(u)}{\partial u_k} - \frac{\partial M_{ik}(u)}{\partial u_j}\right)\frac{\partial u_k}{\partial x}. \tag{48}$$

Assuming v, p smooth (can we do so? if u is smooth, does it also hold for the linearized equation and for the adjoint?), we compute formally

$$
\begin{aligned}
0 &= \frac{d}{dt}\int_{\mathbb{R}} p(x,t)v(x,t)\,dx \\
&= \int p_t \cdot v + p \cdot v_t\,dx \\
&= \int p_t \cdot v + p \cdot \Big\{ -A(u)v_x - (DA(u)v)u_x + \eta\big(B(u)v_x + (DB(u)v)u_x\big)_x \Big\}\,dx \\
&= \int p_t \cdot v + p_x A(u)v + p(DA(u)u_x)v - p(DA(u)v)u_x - \eta\, p_x\big(B(u)v_x + (DB(u)v)u_x\big)\,dx \\
&= \int p_t \cdot v + p_x A(u)v + p(\widetilde{DA}(u)u_x)v + \eta\Big\{ p_{xx}B(u)v + p_x\big((DB(u)u_x)v - (DB(u)v)u_x\big)\Big\}\,dx \\
&= \int \Big\{ p_t + p_x A(u) + p(\widetilde{DA}(u)u_x) + \eta\big(p_{xx}B(u) + p_x(\widetilde{DB}(u)u_x)\big)\Big\} \cdot v\,dx.
\end{aligned}
$$

Hence p should be the solution to (47) at least away from possible discontinuities. If we choose B to be the physical viscosity of the system (45), we could maybe use a fancy Ulrik-Sid scheme [7] to discretize (45) and (47).

Remark 3.1. In the scalar case (47) is

$$
\begin{aligned}
p_t + f'(u)\,p_x &= -\eta\, b(u)\, p_{xx}, \qquad && (x,t) \in \mathbb{R}\times(0,T) \\
p(x,T) &= p^T(x), && x \in \mathbb{R}
\end{aligned}
\tag{49}
$$

Remark 3.2. If we discretize (47) with a simple first order finite difference scheme, we obtain

$$
\begin{aligned}
p_j^{n-1} &= p_j^n + \frac{\Delta t}{2\Delta x}\{(p_{j+1}^n - p_{j-1}^n)A(u_j^n) + p_j^n\widetilde{DA}(u_j^n)(u_{j+1}^n - u_{j-1}^n)\} \\
&+ \eta\frac{\Delta t}{\Delta x^2}\Big\{(p_{j+1}^n - 2p_j^n + p_{j-1}^n)B(u_j^n) + \frac{p_{j+1}^n - p_{j-1}^n}{2}\widetilde{DB}(u_j^n)\frac{u_{j+1}^n - u_{j-1}^n}{2}\Big\}, \\
p_j^N &= p^T(x_j)
\end{aligned}
\tag{50}
$$

for $j \in \mathbb{N}$, $n \in 0, \ldots, N$, $(N \cdot \Delta t = T)$, where

$$
(\widetilde{DB}(u_s)(u_{s+1} - u_{s-1}))_{ij} = \sum_{k=1}^{n}\bigg(\frac{\partial B_{ij}(u_s)}{\partial u^k} - \frac{\partial B_{ik}(u_s)}{\partial u^j}\bigg)(u_{s+1}^k - u_{s-1}^k).
$$

We can choose $\eta = c\Delta x$ for a constant c such that the numerical viscosity becomes smaller when the stepsize Δx is decreased.

3.4.3 Questions/Problems

- What regularity can we assume of u and its adjoint p in the viscous equation? This depends on $B(u)$. (If $B(u) = I$, the solution u to system (45) is smooth, but what about the adjoint system of equations?) If p has discontinuities what equations do they satisfy? Under what conditions is the adjoint viscous system of equations uniquely solvable?

- The optimal gradient direction $-p(\cdot, 0)$ computed by the means of the adjoint direction might have a different profile than the current control u^0, i.e. shocks in different locations as u^0. Hence when iterating the method, we might still end up with a very irregular 'optimal' control u^0 as in [5]. Then we could maybe kind of apply the alternating descent method by splitting the perturbation v into profile and shock locations (the question is how algorithmically...) and first use gradient directions that only move the shock positions and as soon as the shock locations are more or less in the right places, we perturb the profile.

References

[1] F. Bouchut and F. James. One-dimensional transport equations with discontinuous coefficients. *Nonlinear Analysis, Theory, Methods and Applications*, 32(7):891–933, 1998.

[2] Y. Brenier and S. Osher. The discrete one-sided Lipschitz condition for convex scalar conservation laws. *SIAM Journal on Numerical Analysis*, 25(1):8–23, 1988.

[3] A. Bressan and A. Marson. A maximum principle for optimally controlled systems of conservation laws. *Rendiconti del Seminario Matematico della Universita di Padova*, 94:79–94, 1995.

[4] A. Bressan and A. Marson. A variational calculus for discontinuous solutions of systems of conservation laws. *Communications in Partial Differential Equations*, 20:1491–1552, 1995.

[5] C. Castro, F. Palacios, and E. Zuazua. An alternating descent method for the optimal control of the inviscid Burgers equation in the presence of shocks. *Mathematical Models and Methods in Applied Sciences (M3AS)*, 18:369–416, 2008.

[6] C. Castro and E. Zuazua. Flux identification for 1-d scalar conservation laws. *Mathematics of Computation, American Mathematical Society*, 80:2025–2070, 2011.

[7] U.S. Fjordholm and S. Mishra. Accurate numerical discretizations of non-conservative hyperbolic systems.

[8] E. Godlewski and P.-A. Raviart. The linearized stability of solutions of nonlinear

hyperbolic systems of conservation laws, A general numerical approach. *Mathematics and Computers in Simulation*, 50:77–95, 1999.