AMO - Advanced Modeling and Optimization, Volume 7, Number 1, 2005

OPTIMAL NODAL CONTROL OF NETWORKED HYPERBOLIC SYSTEMS: EVALUATION OF DERIVATIVES [1]

MARTIN GUGAT

Lehrstuhl II fuer Angewandte Mathematik, Martensstrasse 3, 91058 Erlangen, Germany (gugat@am.uni-erlangen.de).

Abstract. We consider a networked system defined on a graph where each edge corresponds to a quasilinear hyperbolic system with space dimension one. At the nodes, the system is governed by algebraic node conditions. The system is controlled at the nodes of the graph. Optimal control problems for systems of this type arise in the operation of channel networks, for example in hydraulic flood routing. For the solution of such problems, the evaluation of the derivatives of functions that depend on the state of the system is necessary. For the case of continously differentiable states, we present an adjoint sensitivity calculus that allows to compute directional derivatives in seceral directions by solving only one backward equation. The result is used to numerically solve by a gradient–type method a problem of optimal control for the St. Venant Equations.
Keywords. optimal control, hyperbolic partial differential equation, network, node conditions, adjoint equations, St. Venant Equations
Mathematics Subject Classification. 93B35, 49K40, 35L65.

1 Introduction

In many applications, systems appear that can be modelled as a networked system defined on a graph where the dynamics on each edge are modelled by a quasilinear hyperbolic partial differential equation and at the interior nodes of the graph, these systems are coupled by algebraic node conditions. As examples, consider networks of water channels/ gas pipelines/ roads with traffic flow (see [15, 19, 28]). There are many excellent studies of hyperbolic conservation laws, see for example [6, 10, 16, 18] and the references therein. Classical solutions are studied in [22].

We consider problems of optimal control for networked systems that are controlled at the nodes of the graph. We prove the differentiability with respect to the control function of objective functions that are defined as integrals that depend on the state of the system. We present an adjoint sensitivity calculus that is useful for the evaluation of the derivative of the objective function and allows the statement of optimality conditions.

[1] This work was supported by DFG–research cluster: Real-time optimization of complex systems; grant number Le595/13-1.

In order to compute the derivatives of the objective function, as an intermediate step we need the sensitivity of the system state with respect to changes in the control function. Let $v(u)$ denote the state of our system corresponding to the control u. We want to compute the directional derivatives in direction d,

$$\lim_{h \to 0+} [\, v(u + hd) - v(u)\,] / h = H(u, d). \tag{1.1}$$

To do this, we first have to make sure that the directional derivatives exist. We work with the linearized forward equation to prove the existence of the directional derivatives. We show that H is a continuous function on each edge of our graph if the control function u generates a continuously differentiable system state $v(u)$ and the direction d satisfies certain compatibility conditions with respect to the initial conditions. To evaluate the derivatives, a backwards equation is used. If this approach is used, only one backwards equation has to be solved to compute the directional derivatives in several directions. This is cheaper and more accurate than the approximation of directional derivatives with finite differences, where for each direction, the foward system equation has to be solved again.

Related studies can be for example found in [5] where sensitivity equations are derived formally, in [3, 4] and in the work of S. Ulbrich (e.g. [24, 25]), who focusses in particular on the question how shocks in the state depend on the controls. For the St. Venant equations, a sensitivity analysis has been presented by Sanders and Katapodes [21] in engineering style. In this paper, we consider continuously differentiable solutions. The philosophy behind this approach is that we aim to choose the controls in such a way that they do not generate shocks in the state. That this is indeed possible is a question of controllability which has been studied in [9, 12, 13, 19, 23].

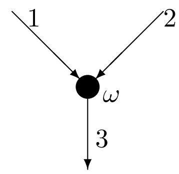
2 Example: Control of water flow to a standstill

Consider a Y–shaped network of three rectangular frictionless horizontal channels with equal length L.

The flow is governed by the De St. Venant (shallow water) equations. The node conditions are chosen to guarantee the conservation of mass and the continuity of the water height. Initially, the flow is stationary with positive velocity. The objective function is the sum of the weighted L^2–norms over $[0, L] \times [T_1, T]$ of the velocities in the three channels (where $T_1 \in (0, T)$ is given) so the aim is to steer the water to rest. The water flow is controlled at the three boundary nodes of the channel network, in the junction there is no control. Arguments similar to those used in the proof of the controllability result in [12] imply that if T_1 is sufficiently large, the optimal value is zero, that is there are control functions that steer the water flow to a standstill in finite time. These control functions are not uniquely determined.

We number the channels in the Y–shaped graph in the following way: channel 1 is the top left–hand side channel, channel 2 is the top right–hand–side

Figure 1: A junction of three channels



channel and channel 3 is the remaining channel. Assume that all channels are directed to the boundary node of channel 3, that is the end L of each channel is closer to this boundary node that the end 0 of this channel (see Figure 1). Let $b_i > 0$ denote the constant width of channel i. Assume that

$$b_1 = b_2 = b_3/2.$$

2.1 De St. Venant Equations

The channel is parametrized lengthwise by $x \in [0, L]$. Let $A_i(x, t)$ denote the wetted cross section at the point x at time t and let $Q_i(x, t)$ denote the corresponding flow rate. Then the conservation of mass yields the equation

$$\frac{d}{dt} A_i(x,t) + \frac{d}{dx} Q_i(x,t) = 0$$

for $i \in \{1, 2, 3\}$. Let $U_i = Q_i/A_i$ denote the average velocity over the cross section of the channel and $h_i = A_i/b_i$ the average water height. The conservation of energy implies the equation

$$\frac{d}{dt} U_i(x,t) + \frac{d}{dx} \left[U_i(x,t)^2/2 + g h_i(x,t) \right] = 0.$$

In terms of the functions U_i and h_i, the quasilinear system can be written as

$$\partial_t \begin{pmatrix} h_i \\ U_i \end{pmatrix} + \begin{pmatrix} U_i & h_i \\ g & U_i \end{pmatrix} \partial_x \begin{pmatrix} h_i \\ U_i \end{pmatrix} = 0.$$

Let $c_i(x,t) = \sqrt{g h_i(x,t)}$ denote the corresponding wave celerity, where $g > 0$ is the gravitational constant. The eigenvalues of the system matrix are $U_i + c_i$ and $U_i - c_i$. In terms of the functions U_i and c_i the system equation is

$$\partial_t \begin{pmatrix} c_i \\ U_i \end{pmatrix} + \begin{pmatrix} U_i & c_i/2 \\ 2c_i & U_i \end{pmatrix} \partial_x \begin{pmatrix} c_i \\ U_i \end{pmatrix} = 0.$$

With the Riemann invariants $R^i_+ = U_i + 2c_i$, $R^i_- = U_i - 2c_i$ and the diagonal matrix

$$D^i(R^i_+, R^i_-) := \begin{pmatrix} \frac{3}{4}R^i_+ + \frac{1}{4}R^i_- & 0 \\ 0 & \frac{1}{4}R^i_+ + \frac{3}{4}R^i_- \end{pmatrix},$$

the de St. Venant equations can be written in the diagonal form

$$\partial_t \begin{pmatrix} R^i_+ \\ R^i_- \end{pmatrix} + D^i(R^i_+, R^i_-)\, \partial_x \begin{pmatrix} R^i_+ \\ R^i_- \end{pmatrix} = 0. \tag{2.1}$$

For the matrix defined later in (7.1) we have

$$M^i(R^i_+, R^i_-) = \begin{pmatrix} \frac{3}{4}\partial_x R^i_+ & \frac{1}{4}\partial_x R^i_+ \\ \frac{1}{4}\partial_x R^i_- & \frac{3}{4}\partial_x R^i_- \end{pmatrix}.$$

The adjoint backwards system is for $i \in \{1, 2, 3\}$

$$\partial_t(\mu^i_+,\, \mu^i_-) + \partial_x\left((\mu^i_+,\, \mu^i_-)D^i(R^i_+,\, R^i_-)\right) \tag{2.2}$$

$$= (\mu^i_+,\, \mu^i_-)M^i(R^i_+, R^i_-) + (\partial_+ f^i(R^i_+, R^i_-),\, \partial_- f^i(R^i_+, R^i_-))$$

with $f^i(R^i_+, R^i_-) = \alpha(t)(R^i_+ + R^i_-)^2/8 = \alpha(t)U_i^2/2$ (the integrand in the objective function), hence

$$\partial_+ f^i(R^i_+, R^i_-) = \partial_- f^i(R^i_+, R^i_-) = \alpha(t)(R^i_+ + R^i_-)/4.$$

The end conditions for μ are for $i \in \{1, 2, 3\}$ (see (7.3))

$$\mu^i_+(x, T) = \mu^i_-(x, T) = 0, \;\; x \in [0, L].$$

2.2 Objective Function

We consider the optimization problem where for all nodes ω, the functions $f^{\omega,e}$ equal zero and for the edges i ($i \in \{1, 2, 3\}$) of the Y–shaped graph we have

$$f^i(R^i_+, R^i_-, x, t) = (1/8)\alpha(t)(R^i_+ + R^i_-)^2 = \alpha(t)U_i^2/2$$

where the weight function $\alpha(t) \geq 0$ has support $\alpha \subset [T_1, T]$, with $0 < T_1 < T$.

2.3 Node Conditions

At the interior node ω, we require the conservation of mass and the continuity of the water height, that is for each fixed time the water height in the end points of the channels that are adjacent to the node must be equal.

Then the first node condition is

$$A_1(L, t)\, U_1(L, t) + A_2(L, t)\, U_2(L, t) = A_3(0, t)\, U_3(0, t) \text{ for all } t \geq 0. \tag{2.3}$$

Our second node condition is

$$h_1(L, t) = h_2(L, t) = h_3(0, t), \; t \geq 0. \tag{2.4}$$

Our node conditions can be transformed to a system of linear equations for the Riemann invariants. Since the channels are rectangular, (2.3) and (2.4) imply the equation $b_1 U_1 + b_2 U_2 = b_3 U_3$. Since $U_i = (R_+^i + R_-^i)/2$, this implies

$$b_1(R_+^1(L,t) + R_-^1(L,t)) + b_2(R_+^2(L,t) + R_-^2(L,t)) = b_3(R_+^3(0,t). + R_-^3(0,t)) \tag{2.5}$$

Since $\sqrt{gh_i} = c_i = (R_+^i - R_-^i)/4$, (2.4) implies

$$R_+^1(L,t) - R_-^1(L,t) = R_+^2(L,t) - R_-^2(L,t) = R_+^3(0,t) - R_-^3(0,t). \tag{2.6}$$

With the numbering of the edges in the Y–shaped graph given above, the solution of this system of linear equations is given by the equation

$$\begin{pmatrix} R_-^1(L,t) \\ R_-^2(L,t) \\ R_+^3(0,t) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} R_+^1(L,t) \\ R_+^2(L,t) \\ R_-^3(0,t) \end{pmatrix}. \tag{2.7}$$

Let $\lambda_+^i = \frac{3}{4} R_+^i + \frac{1}{4} R_-^i$, $\lambda_-^i = \frac{1}{4} R_+^i + \frac{3}{4} R_-^i$. The adjoint interior node conditions (see (7.4), (7.6)) are

$$\begin{pmatrix} \mu_+^1(L,t), & \mu_+^2(L,t), & \mu_-^3(0,t) \end{pmatrix} = \begin{pmatrix} \mu_-^1(L,t), & \mu_-^2(L,t), & \mu_+^3(0,t) \end{pmatrix} B^\omega(t) \tag{2.8}$$

with the matrix $B^\omega(t)$ defined as

$$-\begin{pmatrix} \lambda_-^1(L,t) & 0 & 0 \\ 0 & \lambda_-^2(L,t) & 0 \\ 0 & 0 & -\lambda_+^3(0,t) \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\lambda_+^1(L,t)} & 0 & 0 \\ 0 & \frac{1}{\lambda_+^2(L,t)} & 0 \\ 0 & 0 & -\frac{1}{\lambda_-^3(0,t)} \end{pmatrix}.$$

In Section 8, a numerical solution of this problem is presented. The optimization problem is solved with a a gradient–type method. For the evaluation of the gradient, a discretized form of the representation of the directional derivatives given in Theorem 3 below is used. In Section 9, we present some remarks about the analytical solution of our example problem.

3 Notation

Consider a finite directed graph $G = (V, E)$ with vertices (nodes) V and edges E. This graph is taken to represent a network of coupled systems. Each edge $e \in E$ of the graph corresponds to an interval $[0, L_e]$, where the system is governed by a quasilinear system in diagonal form

$$v_t^e(x,t) + D^e(v^e(x,t),x,t)\, v_x^e(x,t) + S^e(v^e(x,t),t) = 0, \tag{3.1}$$

$$e \in E,\ x \in [0, L_e],\ t \in [0,T]$$

where the solution has n components, D^e is a continuously differentiable map to the space of $n \times n$ diagonal matrices,

$$D^e(w,x,t) = \operatorname{diag}(d_{ii}^e(w,x,t)) = \left[d_{ij}^e(w,x,t)\right]_{ij}$$

and S^e is continuously differentiable. We consider only controls that generate states where the diagonal entries of D^e are nowhere zero and the sign of such an entry is constant on $[0, L_e] \times [0, T]$. Define

$$G^e(v^e) = v^e_t + D^e(v^e)\ v^e_x + S^e(v^e). \tag{3.2}$$

The initial state of the system is given by an initial condition of the form

$$v^e(x, 0) = v^e_0(x),\ e \in E,\ x \in [0, L_e] \tag{3.3}$$

where for all edges $e \in E$, the function v^e_0 is continuously differentiable.

At the nodes $\omega \in V$ of the graph, the values of the functions v^e are coupled through algebraic node conditions. For a node $\omega \in V$, let $E_0(\omega)$ denote the set of adjacent edges. For $e \in E_0(\omega)$, let $x_e(\omega) \in \{0, L_e\}$ denote the end–point of the interval $[0, L_e]$ that corresponds to the vertex.

There are two types of nodes: Boundary nodes, where the set $E_0(\omega)$ has only one element and interior nodes, where $E_0(\omega)$ has at least two elements. Let $V_B \subset V$ denote the set of boundary nodes. For $\omega \in V_B$, let $e(\omega)$ denote the adjacent edge. Let V^0 denote the set of boundary nodes that are at the end zero of the adjacent edge and let V^L denote the set of boundary nodes that are at the end L_e of the adjacent edge e.

3.1 Boundary and interior node conditions

For a boundary node $\omega \in V^0$, we assume that the boundary conditions have the form

$$v^{e(\omega)}_+(0, t) = B_0(t, \omega, v^{e(\omega)}_-(0, t)) + \beta(t, \omega)\ u^{e(\omega)}(t, \omega) \tag{3.4}$$

and for $\omega \in V^L$

$$v^{e(\omega)}_-(L_{e(\omega)}, t) = B_0(t, \omega, v^{e(\omega)}_+(L_{e(\omega)}, t)) + \beta(t, \omega)\ u^{e(\omega)}(t, \omega). \tag{3.5}$$

Here v^e_+ contains the components of v^e that correspond to positive eigenvalues in D^e and v^e_- the components corresponding to negative eigenvalues. $B_0(\cdot, \omega, \cdot)$ is a continuously differentiable function and $u(\cdot, \omega)$ is the corresponding control function. $\beta(\cdot, \omega)$ is a continuously differentiable real–valued function. The control function $u^e(\cdot, \omega)$ is continuously differentiable. The number of components of u depends on the number of positive and negative eigenvalues of D^e. For a discussion of boundary conditions for hyperbolic systems, see [2].

To state the conditions for the interior nodes $\omega \in V \backslash V_B$ of the graph, where several edges are connected, we need the following notation: For $e \in E$, define the index sets

$$\begin{aligned} I^e_+ &= \{i \in \{1, ..., n\} : d^e_{ii}(x, t) > 0,\ (x, t) \in [0, L_e] \times [0, T]\}, \\ I^e_- &= \{i \in \{1, ..., n\} : d^e_{ii}(x, t) < 0,\ (x, t) \in [0, L_e] \times [0, T]\} \end{aligned}$$

where $d^e_{ij}(x, t)$ are the entries of the matrix $D^e(v^e(x, t), x, t)$.

We assume that the interior node conditions can be written in the following form, where $B_0(\cdot, \omega, \cdot)$ is a continuously differentiable function and $\beta(\cdot, \omega)$ is a continuously differentiable real–valued function:

For $e \in E_0(\omega)$ with $x_e(\omega) = 0$ and $i \in I_+^e$:

$$v_i^e(0,t) = B_0(e, i, t, \omega, v_j^f(x_f(\omega), t) : \tag{3.6}$$

$f \in E_0(\omega), x_f(\omega) = 0$ and $j \in I_-^f$ or $x_f(\omega) = L_f$ and $j \in I_+^f) \ + \ \beta(t,\omega)\ u_i^e(t,\omega)$.

For $e \in E_0(\omega)$ with $x_e(\omega) = L_e$ and $i \in I_-^e$:

$$v_i^e(L_e,t) = B_0(e, i, t, \omega, v_j^f(x_f(\omega), t) : \tag{3.7}$$

$f \in E_0(\omega), x_f(\omega) = 0$ and $j \in I_-^f$ or $x_f(\omega) = L_f$ and $j \in I_+^f) \ + \ \beta(t,\omega)\ u_i^e(t,\omega)$.

Hence for the indeces with characteristics that flow out of the node for the forward equation, that is the indeces i for which the values of v_i^e at the node are not determined from the values in the interior of $[0, L_e] \times [0, T]$, the values are determined from the remaining components of the solution that are determined from the characteristics that flow into the node.

Obviously, the sign of the corresponding eigenvalue of the system matrix and the end–point (0 or L_e) determine wether a characteristic curve goes out of the node or into the node. Again the control functions u are assumed to be continuously differentiable.

4 The Linearized Problem

In a problem of optimal control, we can regard the system equation as an equality constraint of the form $G^e(v^e) = 0$ for all $e \in E$. On each edge $e \in E$, the linearized equation is

$$H_t^e + [D^e(v^e)]\ H_x^e + [\nabla_v D^e(v^e) H^e]\ v_x^e + [(S^e)'(v^e)]\ H^e = 0. \tag{4.1}$$

Here for $v \in R^n$ the matrix $(S^e)'(v)$ is defined by the equation

$$(S^e)'(v,t) = \left(\frac{\partial S_i^e}{\partial v_j}(v,t) \right)_{ij} \tag{4.2}$$

and denotes the derivative of S^e and

$$[\nabla_v D^e(v^e) H^e] = \operatorname{diag} \left(\sum_{j=1}^{n} \partial_{v_j^e} d_{ii}(v^e) H_j^e \right)$$

is a diagonal matrix.

With the appropriate initial and boundary/interior node conditions this yields the initial boundary value problem for H.

Since the given initial state v_0 does not depend on the boundary control function, the initial condition for H is

$$H^e(x,0) = 0,\ e \in E,\ x \in [0, L_e]. \tag{4.3}$$

It is important that in general the partial derivative v_x^e that appears in the linearized problem cannot be expected to be continuously differentiable if the function v^e is continously differentiable. Therefore in general we cannot expect that the linearized problem has a continuously differentiable solution. We can only expect that the solution H^e is continuous.

4.1 Boundary and interior node conditions for the linearized problem

For a boundary node $\omega \in V^0$, the linearized boundary conditions have the form

$$H_+^{e(\omega)}(0,t) = B(t,\omega)H_-^{e(\omega)}(0,t) + \beta(t,\omega)\ d^{e(\omega)}(t,\omega) \tag{4.4}$$

and for $\omega \in V^L$

$$H_-^{e(\omega)}(L_{e(\omega)},t) = B(t,\omega)H_+^{e(\omega)}(L_{e(\omega)},t) + \beta(t,\omega)\ d^{e(\omega)}(t,\omega) \tag{4.5}$$

where the matrix $B(t,\omega)$ contains the corresponding partial derivatives of B_0. The functions $d(\cdot,\omega)$ are continuously differentiable with $d(0,\omega) = 0$.

Let b_{ik}^{eg} denote the partial derivative

$$b_{ik}^{eg}(t,\omega,v_j^f(x_f(\omega),t) : f \in E_0(\omega), x_f(\omega) = 0 \text{ and } j \in I_-^f \text{ or } x_f(\omega) = L_f \text{ and } j \in I_+^f)$$

$$= \frac{\partial}{\partial v_k^g} B_0(e,i,t,\omega,v_j^f(x_f(\omega),t) :$$

$$f \in E_0(\omega), x_f(\omega) = 0 \text{ and } j \in I_-^f \text{ or } x_f(\omega) = L_f \text{ and } j \in I_+^f).$$

The interior node conditions are as follows:

For $e \in E_0(\omega)$ with $x_e(\omega) = 0$ and $i \in I_+^e$:

$$H_i^e(0,t) = \sum_{\substack{f \in E_0(\omega):\\ x_f(\omega)=0,\\ j \in I_-^f}} b_{ij}^{ef}(t,\omega)H_j^f(0,t) + \sum_{\substack{f \in E_0(\omega):\\ x_f(\omega)=L_f,\\ j \in I_+^f}} b_{ij}^{ef}(t,\omega)H_j^f(L_e,t) + \beta(t,\omega)\ d_i^e(t,\omega). \tag{4.6}$$

Again the functions b_{ij}^{ef} are given by the corresponding partial derivatives of B_0. The sum over the partial derivatives appears in (4.6) compared with (3.6) on account of the linearization.

For $e \in E_0(\omega)$ with $x_e(\omega) = L_e$ and $i \in I_-^e$:

$$H_i^e(L_e,t) = \sum_{\substack{f \in E_0(\omega): \\ x_f(\omega)=0, \\ j \in I_-^f}} b_{ij}^{ef}(t,\omega) H_j^f(0,t) + \sum_{\substack{f \in E_0(\omega): \\ x_f(\omega)=L_f, \\ j \in I_+^f}} b_{ij}^{ef}(t,\omega) H_j^f(L_e,t) + \beta(t,\omega)\, d_i^e(t,\omega). \tag{4.7}$$

The functions $d(\cdot,\omega)$ are continuously differentiable with

$$d^e(0,\omega) = 0. \tag{4.8}$$

5 Existence of Solutions

The solvability of the system equation (3.1) can be guaranteed in a neighbourhood of a constant stationary solution v_0 that satisfies the equation $S^e(v_0) = 0$.

With boundary controls u that are chosen such that the corresponding values of the solution are sufficiently close to v_0, whose derivatives are sufficiently small and that satisfy compatibility conditions, system (3.1) has a continuously differentiable solution. This follows from results of Cirina in [8].

The linearized system (4.1) has the same system matrix as (3.1) . The characteristic curves are determined by the eigenvalues of this matrix.

Since with smooth boundary values that satisfy compatibility conditions with the initial state, shocks only occur if the characteristic curves intersect, this observation implies that if for the original initial boundary value problem a continuously differentiable solution exists, then also the linearized system (4.1) can be transformed to the same characteristic coordinates.

For the case where S^e is the zero function and in addition $v_x = 0$ or $\nabla_v D = 0$, that is if the source term vanishes, this implies that the solution of the linearized system exists and is continuously differentiable on the same domain as the solution of the original system, since in this case the Riemann invariants remain constant along every characteristic curve (see for example Dafermos, [10], p. 128).

In the sequel we study the existence of a continuous solution of the linearized initial boundary value problem.

Since the linearized problem has the same characteristic curves as the original problem, which by assumption has a continuously differentiable solution, we consider solutions of the linearized problem in the characteristic sense that is solutions that are obtained by integrating along the characteristic curves.

For $e \in E$ and $i \in \{1,...,n\}$ such a curve $(\xi_i^e(s), s)$ satisfies the ordinary differential equation

$$\frac{d}{ds}\xi_i^e(s) = d_{ii}^e(v^e(\xi_i^e(s),s), \xi_i^e(s), s).$$

Along the characteristic curves, the solution H^e of the linearized equation satisfies the following relation:

$$H_i^e(\xi_i^e(t),t) = H_i^e(\xi_i^e(s_0),s_0) + \int_{s_0}^t [[\nabla_v D^e(v^e)H^e]v_x^e + (S^e)'(v^e)H^e]_i \,|_{(\xi_i^e(s),s)}\, ds \tag{5.1}$$

where the notation $[\,]_i$ denotes the i–th component.

The following theorem concerns the existence of continous solutions of the linearized problem.

Theorem 1 Let a control function u be given that generates continuously differentiable solution of equation (3.1) that satisfies the initial condition (3.3) and on the time interval $[0,T]$ the conditions (3.4), (3.5), (3.6), (3.7). Assume that the function d satisfies the compatibility conditions (4.8).

Then on the time interval $[0,T]$ there exists a unique continuous solution of the initial boundary value problem defined by the initial condition $H^e(\cdot,0)=0$ for all $e \in E$, the boundary node conditions (4.4), (4.5), the interior node conditions (4.6), (4.7) and the linearized equation (4.1) in the characteristic sense (5.1).

The Theorem is proved by a fixed–point iteration of Picard–Lindelöf type along the characteristic curves. Using the a priori bounds, the existence of the solution on the time interval $[0,T]$ follows by applying the local existence result on a finite number of time intervals $[0,T_0]$, $[T_0,2T_0]$,... until the time T is reached.

Note that due to the linearity of the problem, we need not assume a bound on the size of the norm of d or its derivative.

Proof.

For $e \in E$ and $(x,t) \in [0,L_e] \times [0,T]$ define the characteristic curves through this point by introducing the functions $\xi_i^e(s) = \xi_i^e(s;x,t)$ for $i = 1,...,n$ that satisfy the conditions

$$\frac{d}{ds}\xi_i^e(s) = d_{ii}^e(v^e(\xi_i^e(s),s),\xi_i^e(s),s),\ s \le t,\ \ \xi_i^e(t) = x.$$

where d_{ii}^e are the eigenvalues of the diagonal system matrix D^e. Then the characteristic curves through the point (x,t) have the form $\{(\xi_i^e(s),s):\ s \in [s_0,t]\}$.

Since by assumption a C^1–solution of the original problem exists, such curves exist and run backwards in time until they reach boundary 0 or L_e at some positive time $s_0 > 0$ or until at the initial time $s_0 = 0$, ξ_i^e has some value in the interval $[0,L_e]$.

Along the characteristic curves, the function H satisfies the following relation:

$$H_i^e(x,t) = H_i^e(\xi_i^e(s_0),s_0) + \int_{s_0}^t [[\nabla_v D^e(v^e)H^e]v_x^e + (S^e)'(v^e)H^e]_i \,|_{(\xi_i^e(s),s)}\ ds \tag{5.2}$$

For a continuous function w such that for all $e \in E$: $w^e : [0, L_e] \times [0, T] \mapsto \mathbb{R}^n$, define the integral operator

$$(Pw)_i^e(x,t) = (Qw)_i^e(\xi_i^e(s_0), s_0)$$

$$+ \int_{s_0}^{t} [[\nabla_v D^e(v^e) w^e] v_x^e + (S^e)'(v^e) w^e]_i \,|_{(\xi_i^e(s),s)} \; ds$$

where $(Qw)_i^e(\xi_i^e(s_0), s_0)$ denotes the boundary values that are determined either by the initial values given in the initial condition (3.3) (if $s_0 = 0$) or by the boundary conditions (4.4), (4.5) or the node conditions (4.6), (4.7) with the control function d and the values of the function Pw determined on the characteristic curves that enter the nodes.

To obtain these boundary values, it is necessary to follow the involved characteristic curves backwards in time until the initial time zero is reached. If the time t is sufficiently small, the initial line can be reached after at most one reflection at the boundary.

Choose $w_0 = 0$. Consider the sequence of functions φ_k with $\varphi_0 = w_0$ and

$$(\varphi_{k+1})_i^e = (P\varphi_k)_i^e.$$

Then the definition of P implies that for all $k \in \mathbb{N}$, the function φ_k is well–defined and continuous.

Let $K \geq 1$ be such that for all $(x,t) \in [0, L_e] \times [0, T]$ and for all continuous functions f^e defined on $[0, L_e] \times [0, T]$ the following inequalities hold:

$$\| \, [\nabla_v D^e(v^e) f^e] v_x^e + (S^e)'(v^e) f^e|_{(x,t)} \| \; \leq K \|f^e(x,t)\|$$

and for an interior node $\omega \in V$, if $f \in (\mathbb{R}^n)^{E_0(\omega)}$ satisfies the node conditions (4.6), (4.7) with $d_i^e = 0$, then for the components f_i^e corresponding to outgoing characteristic curves with respect to this node we have

$$|f_i^e| \leq K \|r^e\|, \tag{5.3}$$

where the vector r^e contains the components of f^e for which the corresponding characteristic curves come from the interior of $[0, L_e] \times [0, T]$ to the node, so r^e contains the components of v^e corresponding to ingoing characteristic curves with respect to the node $\omega \in V$.

Moreover, assume that the number K is chosen such that at the boundary nodes $\omega \in V^0$, if $f \in \mathbb{R}^n$ satisfies the node condition

$$f_+^e = B(t, \omega) f_-^e \tag{5.4}$$

we have the inequality

$$\|f_+^e\| \leq K \|f_-^e\| \tag{5.5}$$

and at the boundary nodes $\omega \in V^L$ for all $f \in \mathbb{R}^n$ with

$$f_-^e = B(t, \omega) f_+^e \tag{5.6}$$

we have the inequality

$$\|f_-^e\| \le K\|f_+^e\|. \tag{5.7}$$

Then the following inequality is valid for continuous functions α, β with components α^e, $\beta^e \in C([0,L_e]\times[0,T],I\!R^n)$ for $e \in E$:

$$\max_{x\in[0,L_e]} \|(P\alpha(x,t) - P\beta(x,t))^e\| \le \max_{x\in[0,L_e]} \max_{s\in[0,t]} \|\alpha^e(x,s) - \beta^e(x,s)\| K^{m+1} t.$$

Here m denotes the maximum number of reflections at the boundary until a characteristic curve emanating from the interval $[0,L_e]$ at time zero reaches a point $(x,t) \in [0,L_e]\times[0,T]$. Note that here we use the inequalities (5.3), (5.5), (5.7).

Hence if $T_1 \le T$ is sufficiently small, P is a contraction on the time interval $[0,T_1]$.

Therefore the sequence φ_k converges to a limit v and the limit function v is continuous and satisfies the initial conditions (4.3). Since for all $k \in I\!N$, the boundary conditions (4.4), (4.5), and the node conditions (3.6), (3.7) with the control function d are valid for the functions φ_k by the definition of (Qw), also the function v satisfies these conditions. Moreover, the equation $Pv = v$ implies that v satisfies the integral equation (5.2).

To show the existence of the solution on the whole time interval $[0,T]$, we use the method of continuation (see [20]), namely, starting with the solution at time T_1 as initial condition we show in a similar way the existence of the solution on a time interval $[T_1,T_2]$ that is sufficiently small to yield a contraction and continue in this way on successive time–intervals until we have reached the time T.

So we have shown the existence of a unique continuous solution.

6 Directional Differentiability

In this section we show that if a control function u generates a continuously differentiable solution v of our nonlinear system, then we can expand the solution in the form

$$v(x,t,u+hd) = v(x,t,u) + hH(x,t,u,d) + o(h) \tag{6.1}$$

where H is the solution of the corresponding linearized system and d is a continuously differentiable direction that satisfies the compatibility conditions (4.8).

Theorem 2 Let a control function u be given that generates a continuously differentiable solution of equation (3.1) that satisfies the initial condition (3.3) and on the time interval $[0,T]$ the conditions (3.4), (3.5), (3.6), (3.7).

Assume that the function d satisfies the compatibility conditions (4.8). Assume that for all real numbers h that are sufficiently small, the control functions $u+hd$ also generate continuously differentiable solutions.

Then using the unique continuous solution of the linearized problem (4.3), (4.4), (4.5), (4.6), (4.7) and (4.1) in the characteristic sense (5.1) we can expand the solution of the original system in the form (6.1), that is equation (1.1) holds.

Proof.
For a real number h, define the function

$$R(x,t,u,d,h) = v(x,t,u) + hH(x,t,u,d)$$

where H is a solution of the linearized problem (4.3), (4.4), (4.5), (4.6), (4.7) where the linearized system equation (4.1) holds in the characteristic sense (5.1). Then the following equation holds vor R along the characteristic curves $(\xi_i^e(s), s)$ generated by the control function $u + hd$ (where we use the notation $d_{ii}^e(u + hd, s) = d_{ii}^e(v^e(\xi_i^e(s), s, u + hd), \xi_i^e(s), s)$):

$$R_i^e(\xi_i^e(t), t, u, d, h) = R_i^e(\xi_i^e(s_0), s_0, u, d, h)$$
$$+ \int_{s_0}^t [d_{ii}^e(u + hd, s) - d_{ii}^e(u, s)] \; [R_x(\xi_i^e(s), s, u, d, h)]_i \; ds$$
$$- \int_{s_0}^t [S^e(v^e(\xi_i^e(s), s, u), s) + (S^e)'(v^e(\xi_i^e(s), s, u), s) hH^e(\xi_i^e(s), s, u, d)]_i \; ds$$
$$- \int_{s_0}^t h \, [[\nabla_v D^e H^e] v_x^e(\xi_i^e(s), s, u), s)]_i \; ds$$

where in the matrix

$$[\nabla_v D^e H^e] = \operatorname{diag}\left(\sum_{j=1}^n \partial_{v_j^e} d_{ii}^e(v^e(\xi_i^e(s), s, u), \xi_i^e(s), s) H_j^e(\xi_i^e(s), s, u, d) \right)$$

we have omitted the arguments to obtain a shorter formula.
Define the function

$$r(x,t,u,d,h) = R(x,t,u,d,h) - v(x,t,u + hd).$$

Then along the characteristic curves $(\xi_i^e(s), s)$ generated by the control function $u + hd$, the following equation holds for r:

$$r_i^e(\xi_i^e(t), t, u, d, h) - r_i^e(\xi_i^e(s_0), s_0, u, d, h) \tag{6.2}$$
$$= \int_{s_0}^t [d_{ii}^e(u + hd, s) - d_{ii}^e(u, s)] \; [R_x(\xi_i^e(s), s, u, d, h)]_i \; ds$$
$$+ \int_{s_0}^t [S^e(v^e(\xi_i^e(s), s, u + hd), s) - S^e(v^e(\xi_i^e(s), s, u), s)$$
$$-(S^e)'(v^e(\xi_i^e(s), s, u), s) hH^e(\xi_i^e(s), s, u, d)]_i \; ds$$

$$
\begin{aligned}
& -\int_{s_0}^{t} h\,[[\nabla_v D^e H^e] v_x^e(\xi_i^e(s),s,u),s)]_i \; ds \\
= & \int_{s_0}^{t} [S^e(R^e - r^e) - [S^e(v^e(u)) + (S^e)'(v^e(u))(R^e - v^e(u))]]_i \; ds \\
& -\int_{s_0}^{t} h\,[[\nabla_v D^e H^e] v_x^e(\xi_i^e(s),s,u),s)]_i \; ds \\
& +\int_{s_0}^{t} [d_{ii}^e(u+hd,s) - d_{ii}^e(u,s)]\,[R_x^e(\xi_i^e(s),s,u,d,h)]_i \; ds
\end{aligned}
$$

where in the last equation we used some shorter notation which is self–explaining.

Define

$$
\begin{aligned}
F_1^e(r,h) &= S^e(R^e - r^e) - [S^e(v^e(u)) + (S^e)'(v^e(u))(R^e - v^e(u))] \\
&= S^e(R^e - r^e) - S^e(R^e) + S^e(R^e) - [S^e(v^e(u)) + (S^e)'(v^e(u))(R^e - v^e(u))] \\
&= S^e(R^e - r^e) - S^e(R^e) + \int_0^1 [(S^e)'(v^e(u)) - (S^e)'(v^e(u) + \lambda h H^e)]\, h H^e \, d\lambda.
\end{aligned}
$$

Then we have the inequality

$$
\|F_1^e(r,h)\| \le M_1 \|r^e\| + |h| \int_0^1 \|(S^e)'(v^e(u)) - (S^e)'(v^e(u) + \lambda h H^e)\| \|H^e\| \, d\lambda,
$$

where the number $M_1 > 0$ is such that for all $e \in E$ for all $(x,t) \in (0,L_e) \times (0,T)$, $\lambda \in (0,1)$ we have

$$
\|(S^e)'(R^e - r^e + \lambda r^e)(x,t)\| \le M_1.
$$

Hence we have shown that an estimate of the form

$$
\|F_1^e(r,h)\| \le M_1 \|r^e\| + |h|\, \varphi_1^e(h) \tag{6.3}
$$

is valid, with

$$
\lim_{h\to 0} \varphi_1^e(h) = 0
$$

for all $e \in E$.

Define

$$
\begin{aligned}
F_2^e(r,h) &= [D^e(v^e(u+hd)) - D^e(v^e(u))]\, R_x^e - h[\nabla_v D^e(v^e(u^e)) H^e] v_x^e(u^e) \\
&= [D^e(R^e - r^e) - D^e(v^e(u^e)) - [\nabla_v D^e(v^e(u^e)) h H^e]]\, R_x^e \\
&+ h^2 [\nabla_v D^e(v^e(u^e)) H^e] H_x^e \\
&= \left\{ \int_0^1 [(\nabla_v D^e(v^e + \lambda h H^e) - \nabla_v D^e(v^e))\; h H^e]\, d\lambda + D^e(R^e - r^e) - D^e(R^e) \right\} R_x^e \\
&+ h^2 [\nabla_v D^e(v^e(u^e)) H^e] H_x^e \\
&= \left\{ h \int_0^1 [(\nabla_v D^e(v^e + \lambda h H^e) - \nabla_v D^e(v^e))\; H^e] + [\nabla_v D^e(R^e - \lambda r^e) r^e]\; d\lambda \right\} R_x^e \\
&+ h^2 [\nabla_v D^e(v^e(u^e)) H^e] H_x^e.
\end{aligned}
$$

Then in a similar way as above, we obtain an estimate of the form

$$\|F_2^e(r,h)\| \le M_2\|r^e\| + |h|\,\varphi_2^e(h) \tag{6.4}$$

with

$$\lim_{h\to 0}\varphi_2^e(h) = 0$$

for all $e \in E$.

For $t \ge 0$ define the continuous function

$$U(t) = \max\{\|r^e(x,t,u,d,h)\| \,:\, e \in E,\ x \in [0,L_e],\ s \le t\}.$$

The the inequalities (6.3), (6.4) and the integral equation (6.2) imply that an inequality of the following form is valid:

$$0 \le U(t) \le \int_0^t M\,U(s) + |h|\varphi(h)\,ds \tag{6.5}$$

with

$$\lim_{h\to 0}\varphi(h) = 0$$

and a number $M > 0$ that does not depend on h. Then Gronwall's Lemma 1 stated below for the convenience of the reader yields the inequality

$$U(t) \le |h|\varphi(h)\exp(Mt)/M$$

which implies that for $|h| \to 0$, $U(\cdot)/|h|$ converges to zero uniformly on $[0,T]$. The definition of r implies that r satifies approximatively the linearized node conditions (4.4), (4.5), (4.6), (4.7) with $d = 0$ and with an error of order $o(h)$, which yields the assertion.

Lemma 1 (Gronwall) (see for example [27], p.13): Let $C > 0$, $u_0 \ge 0$ and $\varepsilon \ge 0$ be given. Suppose that u is a continuous real–valued function on the interval $[0,T]$ satisfying the inequalities

$$0 \le u(t) \le u_0 + \int_0^t Cu(s) + \varepsilon\,ds.$$

Then the following inequality holds: $u(t) \le u_0 e^{Ct} + \varepsilon\left(e^{Ct} - 1\right)/C$.

7 Evaluation of Derivatives using adjoint solutions

We consider objective or constraint functions of the form

$$\begin{aligned} J(u) \quad &= \quad \sum_{\omega\in V}\sum_{e\in E_0(\omega)}\int_0^T f^{\omega,e}(v^e(x_e(\omega),t),t)\,dt \\ &+ \quad \sum_{e\in E}\int_0^T\int_0^{L_e} f^e(v^e(x,t),x,t)\,dx\,dt. \end{aligned}$$

The functions $f^{\omega,e}$ and f^e are assumed to be continuously differentiable, and v^e (the state of the system) is the solution of the initial boundary value problem (3.1), (3.3), (3.4), (3.5), (3.6), (3.7) generated by the control function u.

Define the matrix function $M^e(v^e)$ with the entries

$$M_{ij}^e(v^e) = \partial_x v_i^e \partial_{v_j} d_{ii}(v^e). \tag{7.1}$$

Then we have for all $i \in \{1, ..., n\}$

$$[M^e(v^e)H^e]_i = \sum_{j=1}^{n} \partial_x v_i^e \partial_{v_j} d_{ii}(v^e) H_j^e = [\sum_{j=1}^{n} \partial_{v_j} d_{ii}(v^e) H_j^e] \partial_x v_i^e = [\ [\nabla_v D^e(v^e) H^e] v_x^e\ ]_i$$

hence $M^e(v^e)H^e = [\nabla_v D^e(v^e)H^e]v_x$.

For $e \in E$, we consider the adjoint equation

$$\mu_t^e + (\mu^e D^e)_x - \mu^e[(S^e)'(v^e)] - \mu^e[M^e(v^e)] = \nabla_v (f^e)^T \tag{7.2}$$

where μ^e is a row–vector with n components and the matrix $(S^e)'(v^e)$ is as in (4.2).

The end conditions for μ^e are

$$\mu^e(x, T) = 0, \text{ for all } e \in E. \tag{7.3}$$

In this section, we want to show that the directional derivative

$$D_d J(u) = \lim_{h \to 0} \frac{J(u + hd) - J(u)}{h}$$

can be represented in a form where the solution of the original nonlinear forward problem and the adjoint solution appear but not the solution of the linearized problem.

Theorem 3 Let the assumptions of Theorem 2 hold. Let μ denote the solution of the terminal value problem that satisfies for all $e \in E$ the end condition (7.3) and for all $(x, t) \in (0, L_e) \times (0, T)$ the adjoint equation (7.2) and the appropriate node conditions given below by (7.4), (7.6), (7.7). Again we consider solutions that satisfy the corresponding integral equations along the characteristic curves.

Then the directional derivative of J is given by the equation

$$\begin{aligned} D_d J(u) \quad = \quad & \sum_{\omega \in V} \sum_{e \in E_0(\omega): x_e(\omega) = L_e} \sum_{i \in I_-^e} \int_0^T \mu_i^e(L_e, t)\ d_{ii}^e(L_e, t)\ \beta(t, \omega)\ d_i^e(t, \omega)\, dt \\ - \quad & \sum_{\omega \in V} \sum_{e \in E_0(\omega): x_e(\omega) = 0} \sum_{i \in I_+^e} \int_0^T \mu_i^e(0, t)\ d_{ii}^e(0, t)\ \beta(t, \omega)\ d_i^e(t, \omega)\, dt \\ + \quad & \sum_{\omega \in V} \sum_{e \in E_0(\omega): x_e(\omega) = L_e} \sum_{i \in I_-^e} \int_0^T \partial_{v_i} f^{\omega,e}(v^e(L_e, t), t)\, \beta(t, \omega)\ d_i^e(t, \omega)\, dt \\ + \quad & \sum_{\omega \in V} \sum_{e \in E_0(\omega): x_e(\omega) = 0} \sum_{i \in I_+^e} \int_0^T \partial_{v_i} f^{\omega,e}(v^e(0, t), t)\, \beta(t, \omega)\ d_i^e(t, \omega)\, dt. \end{aligned}$$

Before we can prove the theorem, we have to define the appropriate node conditions.

Let a node $\omega \in V$ and $t \in [0,T]$ be given. Define the column vectors

$$w_1^\omega(t) = (H_i^e(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_-^e;\ H_i^e(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_+^e)^T,$$

$$w_2^\omega(t) = (H_i^e(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_-^e;\ H_i^e(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_+^e;$$

$$H_i^e(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_+^e;\ H_i^e(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_-^e)^T,$$

$$d^\omega(t) = (d_i^e(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_+^e;\ d_i^e(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_-^e)^T.$$

Since we want to work with matrix vector multiplication in what follows, let the order in which the components appear in the vectors be fixed.

Then we can write the node conditions (4.6), (4.7) in the form

$$w_2^\omega(t) = \begin{pmatrix} I \\ A^\omega(t) \end{pmatrix} w_1^\omega(t) + \begin{pmatrix} 0 \\ \beta(t,\omega)\ d^\omega(t) \end{pmatrix}$$

where I denotes the identity matrix in the space containing $w_1^\omega(t)$ and $A^\omega(t)$ is the appropriate matrix. Let m_1^ω denote the length of $w_1^\omega(t)$ and m_2^ω denote the length of $d^\omega(t)$. Then $A^\omega(t)$ is a $m_2^\omega \times m_1^\omega$ matrix.

Define the row vectors

$$\overline{\mu}_1^\omega(t) = (\mu_i^e(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_+^e;\ \mu_i^e(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_-^e),$$

$$\overline{\mu}_2^\omega(t) = (\mu_i^e(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_-^e;\ \mu_i^e(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_+^e;$$

$$\mu_i^e(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_+^e;\ \mu_i^e(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_-^e).$$

Then for the adjoint node conditions we make the ansatz

$$\overline{\mu}_2^\omega(t) = \overline{\mu}_1^\omega(t)(B^\omega(t), I) + (\alpha^\omega(t), 0) \tag{7.4}$$

where I denotes the identity matrix in the m_2^ω–dimensional space and $B^\omega(t)$ is a $m_2^\omega \times m_1^\omega$ matrix that we have to determine as well as the row vector $\alpha^\omega(t)$ that has m_1^ω components.

Define the diagonal matrices

$$\Delta_1^\omega(t) = \operatorname{diag}(-d_{ii}^e(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_-^e;\ d_{ii}^e(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_+^e),$$

$$\Delta_2^\omega(t) = \operatorname{diag}(-d_{ii}^e(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_+^e;\ d_{ii}^e(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_-^e),$$

$$\Delta^\omega(t) = \begin{pmatrix} \Delta_1^\omega(t) & 0 \\ 0 & \Delta_2^\omega(t) \end{pmatrix}.$$

The matrix $\Delta_1^\omega(t)$ is a $m_1^\omega \times m_1^\omega$ matrix and the matrix $\Delta_2^\omega(t)$ is a $m_2^\omega \times m_2^\omega$ matrix. The following equation is valid:

$$\sum_{i=1}^{n} \sum_{\substack{e \in E_0(\omega): \\ x_e(\omega)=L_e}} \mu_i^e(L_e,t) d_{ii}^e(L_e,t) H_i^e(L_e,t) - \sum_{\substack{e \in E_0(\omega): \\ x_e(\omega)=0}} \mu_i^e(0,t) d_{ii}^e(0,t) H_i^e(0,t) \tag{7.5}$$

$$= \overline{\mu}_2^\omega(t)\, \Delta^\omega(t)\, w_2^\omega(t)$$

$$= [\overline{\mu}_1^\omega(t)(B^\omega(t), I) + (\alpha^\omega(t), 0)] \begin{pmatrix} \Delta_1^\omega(t) & 0 \\ 0 & \Delta_2^\omega(t) \end{pmatrix} \left[\begin{pmatrix} I \\ A^\omega(t) \end{pmatrix} w_1^\omega(t) + \begin{pmatrix} 0 \\ \beta(t,\omega)\, d^\omega(t) \end{pmatrix} \right]$$

$$= \overline{\mu}_1^\omega(t)\, B^\omega(t) \Delta_1^\omega(t)\, w_1^\omega(t) + \overline{\mu}_1^\omega(t)\, \Delta_2^\omega(t)\, A^\omega(t)\, w_1^\omega(t) + \overline{\mu}_1^\omega(t)\, \Delta_2^\omega(t)\, \beta(t,\omega)\, d^\omega(t) + \alpha^\omega(t)\, \Delta_1^\omega(t)\, w_1^\omega(t)$$

$$=: \psi^\omega(t)\,.$$

We define the matrix $B^\omega(t)$ by the equation

$$B^\omega(t) = -\Delta_2^\omega(t)\, A^\omega(t)\, [\Delta_1^\omega(t)]^{-1}\,. \tag{7.6}$$

Then

$$\psi^\omega(t) = \overline{\mu}_1^\omega(t)\, \Delta_2^\omega(t)\, \beta(t,\omega)\, d^\omega(t) + \alpha^\omega(t)\, \Delta_1^\omega(t)\, w_1^\omega(t).$$

Define the column vectors

$$Df_1^\omega(t) = (\partial_{v_i} f^{e,\omega}(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_-^e;\ \partial_{v_i} f^{e,\omega}(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_+^e)^T,$$

$$Df_2^\omega(t) = (\partial_{v_i} f^{e,\omega}(0,t), e \in E_0(\omega), x_e(\omega) = 0, i \in I_+^e;\ \partial_{v_i} f^{e,\omega}(L_e,t), e \in E_0(\omega), x_e(\omega) = L_e, i \in I_-^e)^T.$$

We define the row vector $\alpha^\omega(t)$ by the equation

$$\alpha^\omega(t) = \left[-Df_1^\omega(t)^T - Df_2^\omega(t)^T A^\omega(t) \right] [\Delta_1^\omega(t)]^{-1}\,. \tag{7.7}$$

Then we have

$$\psi^\omega(t) = \overline{\mu}_1^\omega(t)\, \Delta_2^\omega(t)\, \beta(t,\omega)\, d^\omega(t) - Df_1^\omega(t)^T\, w_1^\omega(t) - Df_2^\omega(t)^T\, A^\omega(t)\, w_1^\omega(t)$$

$$= \overline{\mu}_1^\omega(t)\, \Delta_2^\omega(t)\, \beta(t,\omega)\, d^\omega(t) - \begin{pmatrix} Df_1^\omega(t) \\ Df_2^\omega(t) \end{pmatrix}^T \left(w_2^\omega(t) - \begin{pmatrix} 0 \\ \beta(t,\omega)\, d^\omega(t) \end{pmatrix} \right). \tag{7.8}$$

Proof of Theorem 3. Consider the directional derivative of J. By Theorem 1 a continuous solution H of the linearized problem exists, and by Theorem 2 and integration by parts we obtain

$$
\begin{aligned}
D_d J(u) &= \sum_{\omega\in V}\sum_{e\in E_0(\omega)}\int_0^T \sum_{i=1}^n \partial_{v_i} f^{\omega,e}(v^e(x_e(\omega),t),t)\, H_i^e(x_e(\omega),t)\,dt \\
&+ \sum_{e\in E}\int_0^T\int_0^{L_e}\sum_{i=1}^n \partial_{v_i} f^e(v^e(x,t),x,t)\, H_i^e(x,t) \\
&+ \mu^e\left\{H_t^e + D^e(v^e)H_x^e + [\nabla_v D^e(v^e)H^e]v_x^e + (S^e)'(v^e)H^e\right\}|_{(x,t)}\;dx\;dt \\
&= \sum_{\omega\in V}\sum_{e\in E_0(\omega)}\int_0^T \sum_{i=1}^n \partial_{v_i} f^{\omega,e}(v^e(x_e(\omega),t),t)\, H_i^e(x_e(\omega),t)\,dt \\
&+ \sum_{e\in E}\int_0^T\int_0^{L_e}\sum_{i=1}^n \partial_{v_i} f^e(v^e(x,t),x,t)\, H_i^e(x,t) \\
&- \left\{\mu_t^e + (\mu^e\, D^e(v^e))_x - \mu^e M^e(v^e) - \mu^e(S^e)'(v^e)\right\} H^e|_{(x,t)}\;dx\;dt \\
&+ \int_0^T \mu^e(x,t)\, D^e(v^e(x,t),x,t)\, H^e(x,t)|_{x=0}^{L_e}\;dt \\
&= \sum_{\omega\in V}\sum_{e\in E_0(\omega)}\int_0^T \sum_{i=1}^n \partial_{v_i} f^{\omega,e}(v^e(x_e(\omega),t),t)\, H_i^e(x_e(\omega),t)\,dt \\
&+ \sum_{\substack{e\,\in\, E_0(\omega)\,:\\ x_e(\omega)\,=\,L_e}} \int_0^T \sum_{i=1}^n \mu_i^e(L_e,t)d_{ii}^e(L_e,t)H_i^e(L_e,t)\,dt \\
&- \sum_{\substack{e\,\in\, E_0(\omega)\,:\\ x_e(\omega)\,=\,0}} \int_0^T \sum_{i=1}^n \mu_i^e(0,t)d_{ii}^e(0,t)H_i^e(0,t)\,dt.
\end{aligned}
$$

Now we have represented $D_d J(u)$ in a form that requires only the values of μ and H at the nodes of the network. Up to now, for μ only the adjoint equation (7.2) and the end condition (7.3) have been used. It remains to be shown that in fact, due to the definition of the adjoint node conditions only the values of the solution μ of the adjoint backwards problem are necessary and the values of the linearized problem H are not needed. In fact the definition of $\psi^\omega(t)$ (7.5) and equation (7.8) imply that

$$\begin{aligned}
D_d J(u) &= \sum_{\omega\in V} \sum_{e\in E_0(\omega)} \int_0^T \sum_{i=1}^n \partial_{v_i} f^{\omega,e}(v^e(x_e(\omega),t),t)\, H_i^e(x_e(\omega),t)\, dt \\
&+ \sum_{\omega\in V} \int_0^T \psi^\omega(t)\, dt \\
&= \sum_{\omega\in V} \int_0^T \begin{pmatrix} Df_1^\omega(t) \\ Df_2^\omega(t) \end{pmatrix}^T w_2^\omega(t) + \overline{\mu}_1^\omega(t)\, \Delta_2^\omega(t)\, \beta(t,\omega)\; d^\omega(t)\, dt \\
&- \int_0^T \begin{pmatrix} Df_1^\omega(t) \\ Df_2^\omega(t) \end{pmatrix}^T \left(w_2^\omega(t) - \begin{pmatrix} 0 \\ \beta(t,\omega)\; d^\omega(t) \end{pmatrix} \right) dt \\
&= \int_0^T Df_2^\omega(t)^T\; \beta(t,\omega)\; d^\omega(t) \; + \overline{\mu}_1^\omega(t)\, \Delta_2^\omega(t)\, \beta(t,\omega)\; d^\omega(t)\, dt
\end{aligned}$$

and the assertion follows.

Remark 1 The numerical evaluation of gradients is also possible by automatic differentiation. For a comparison between both approaches see [14].

8 Numerical Results

In this section, we present a numerical solution for the problem defined in Section 2 to illustrate that the representation of the directional derivatives given in Theorem 3 is useful for numerical purposes.

Our optimal control problem is a problem of global optimization, where local minima can appear. In general, for problems of this type it is only possible to find local minima. For our problem, we have the special situation that it is known that the optimal value is zero, which allows to verify wether the objective value of a computed approximation is close to the optimal value. In general this is not possible, so usually one must be content with an approximation of a local minimum where the norm of the gradient is small.

In our computations, we use the steepest descent method (see [7]) that tries to find a point where the gradient vanishes. In each step of the method, a line–search (that is an approximate one–dimensional minimization) is performed in the direction of steepest descent, which is the negative gradient. For the evaluation of the gradient, we use the results from Theorem 3. We choose the steplength in the line–search according to Armijo's sufficient decrease condition introduced in [1]. A detailed description of the Algorithm can be found in [17], Chapter 3. In fact, we have implemented the following variation of Algorithm 3.1.1 in [17]:

- For $k = 1, ..., kmax$

  (a) Compute $J(u)$ and the gradient g of J with respect to u.
  Test for termination.

(b) Find the least integer $m \geq 0$ such that the sufficient decrease condition

$$J(u - \lambda g) - J(u) < -\alpha\lambda \|g\|^2$$

holds for $\lambda = \delta_k \beta^m$. Let $\gamma_k = \lambda$.

(c) $u = u - \gamma_k g$

In our implementation we have chosen $\alpha = 10^{-7}$, $\beta = 1/2$ and $\delta_k = 10^{-3}$ for $k \leq 50$, $\delta_k = 10^{-2}$ for $k > 50$.

The iteration was started with constant control functions (u_1, u_2, u_3) that generated the constant initial state where for all $i \in \{1, 2, 3\}$

$$U_i = 0.2, \ h_i = 0.1.$$

The lengths are $L_1 = L_2 = L_3 = 400$ and the widths $b_1 = b_2 = 1$ and $b_3 = 2$. We work with $T = 5000$.

For $i \in \{1, 2, 3\}$ let $u_i(t)$ denote the control corresponding to channel i, that is we have the boundary conditions:

$$\begin{aligned} u_1(t) &= R_+^1(0, t), \\ u_2(t) &= R_+^2(0, t), \\ u_3(t) &= R_-^3(L, t). \end{aligned}$$

Let $u_i^{(k)}$ denote the current iterate for u_i. Then Step (c) of the algorithm has the following form due to the representation of the directional derivatives given in Theorem 3, for $t \in [0, T]$:

$$\begin{aligned} u_1^{(k+1)}(t) &= u_1^{(k)}(t) + \gamma_k \ \mu_+^1(0, t) \ d_{11}^1(0, t) \\ u_2^{(k+1)}(t) &= u_2^{(k)}(t) + \gamma_k \ \mu_+^2(0, t) \ d_{11}^2(0, t) \\ u_3^{(k+1)}(t) &= u_3^{(k)}(t) - \gamma_k \ \mu_-^3(L, t) \ d_{22}^3(L, t) \end{aligned}$$

Here μ denotes the solution of the adjoint system defined in Theorem 3. The steplength $\gamma_k \geq 0$ has to be chosen sufficiently short such that a subcritical state is generated.

The corresponding gradients are computed by an upwind/downwind finite difference discretization of the representation given in Theorem 3 with a flux–vector splitting (see [26]) in the sense that the space–derivatives in the forwards equation corresponding to positive eigenvalues are replaced by an upwind discretization and the space–derivatives in the forwards equation corresponding to negative eigenvalues are replaced by a downwind discretization. For the adjoint backwards equation, in the equations for the components corresponding to positive eigenvalues a downwind discretization is used and for the other components an upwind discretization. The discretization is described in detail in [11].

In our computations we have $T_1 = T/3$ and $\alpha(t) = 0.4(t - T_1)_+/(T - T_1)$.

Figure 2 shows the flow rates $Q_i = b_i \ U_i \ h_i$ corresponding to the computed control functions. The dotted line is the flow rate in channel 1 and channel 2 where due to symmetry the controls are identical and the top line is the flow rate in channel 3. At the end of channel 3, where outflow occurs, the reduction of the flow rate starts earlier than in the other nodes, where the water flows in.

The following figures illustrate the state of the system generated by the computed control. Figure 3 shows snapshots of the velocity U_i and the water height h_i. On the left hand side are channel 1 and channel 2 where the water flows into the system. Due to the symmetry of the system the states in these channels are identical. On the right hand side of the figure is channel 3, where outflow from the system occurs. At $t = 0$, the top line gives the positive initial velocity. The lower line gives the initial water height, which is plotted at the z–level zero in order to make it possible to include it in the same figure. It can be seen that the control steers the system to a state where the water has zero velocity. The water height for the final state is greater than for the initial state, namely at time T we have $h = 0.1069$. Since during the control process, the water height rises, in the figure at time T the water height is given by the upper line and the lower line is the velocity that has been steered to zero.

9 Analytical solution of the example problem

We are looking for control functions that generate a state with $U_i(x,t) = 0$ for $t \geq T_1,\ x \in [0,L],\ i \in \{1,2,3\}$. The objective function does not determine the water height of the terminal state, so we can choose $h_0 = 0.1$ as the water height for our final constant state. Since our system is symmetric we can find an optimal control with $u_1 = u_2$. In fact, if we assume that $u_1(t) = u_2(t)$ for all t, we can consider our system as a single channel of length $2L$ and width b_3, which allows to use the results from [12]. Now we present the control constructed in [12].

Let

$$r_+^0 = 0.2 + 2\sqrt{gh_0}, \qquad r_-^0 = 0.2 - 2\sqrt{gh_0},$$
$$r_+^1 = 2\sqrt{gh_0}, \qquad r_-^1 = -2\sqrt{gh_0}.$$

Let $T_2 \in (0, T_1)$. We choose continuously differentiable decreasing control funtions $u_1 = u_2$ and u_3 with $u_1(0) = r_+^0$ and $u_1(t) = r_+^1$ if $t \geq T_2$, and $u_3(0) = r_-^0$ and $u_3(t) = r_-^1$ if $t \geq T_2$.

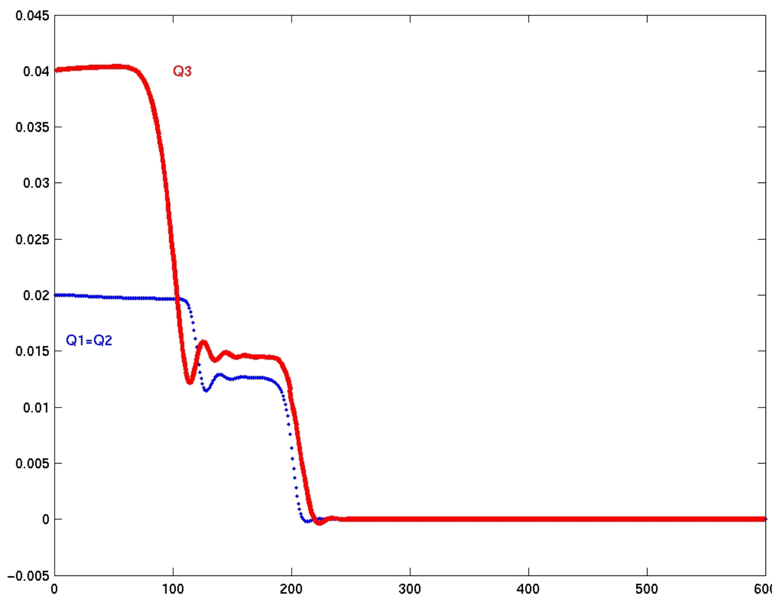
If T_2 is sufficiently large, we can choose control functions whose derivatives are sufficiently close to zero such that a continuously differentiable state is generated. If T_2 is sufficiently small, we have for $t \geq T_1$

$$U_i(x,t) = 0,\ h_i(x,t) = 0.1,\ x \in [0,L],\ i \in \{1,2,3\}. \tag{9.1}$$

For the corresponding flow rates at the inflow of channel i ($i \in \{1,2\}$ we obtain

$$Q_i(0,t) = U_i(0,t)b_ih_i(0,t) = \left(\frac{u_i(t) + R_-^i(0,t)}{2}\right)\frac{b_i}{g}\left(\frac{u_i(t) - R_-^i(0,t)}{4}\right)^2,$$

Figure 2: The computed flow rates Q_i: After finite time, the flow rates are steered to zero (that is the ends of the system are closed) and the system has reached a stationary state with velocity zero.



so to evaluate $Q_i(0,t)$, we have to solve our state equation.

If (9.1) holds, the solution of the adjoint problem satisfies the equations

$$0 = \mu_+^1(0,t) = \mu_+^2(0,t) = \mu_-^3(L,t).$$

Theorem 3 implies that for all feasible directions d we have

$$D_d J(u) = \int_0^T -\mu_+^1(0,t)\lambda_+^1(0,t)d_+^1(t) - \mu_+^2(0,t)\lambda_+^2(0,t)d_+^2(t) + \mu_-^3(L,t)\lambda_-^3(L,t)d_-^3(t)\,dt.$$

Hence for our optimal control u, for all feasible directions d we have

$$D_d J(u) = 0.$$

The author thanks an anonymous referee for valuable suggestions.

References

[1] L. Armijo. Minimization of functions having lipschitz-continuous first partial derivatives. *Pacific J. Math.*, 16:1–3, 1966.

[2] C. Bardos, A.Y. Leroux, and J.C. Nedelec. First order quasilinear equations with boundary conditions. *Comm. in Partial Differential Equations*, 4:1017–1034, 1979.

[3] C. Bardos and O. Pironneau. A formalism for the differentiation of conservation laws. *Comptes Rendus Mathematique*, 335:839–845, 2002.

[4] C. Bardos and O. Pironneau. Derivatives and control in the presence of shocks. *Computational Fluid Dynamics Journal*, 12, 2003.

[5] J. Borggaard, J. Burns, E. Cliff, and M. Gunzburger. Sensitivity calculations for a 2d, inviscid, supersonic forebody problem. In *Identification and Control in Systems Governed by Partial Differential Equations*, pages 14–25. SIAM, Philadelphia, 1993.

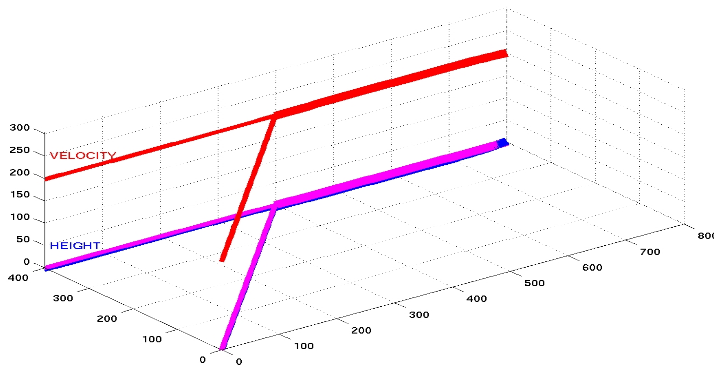
[6] A. Bressan. *Hyperbolic Systems of conservation laws. The one–dimensional Cauchy Problem.* Oxford University Press, 2000.

[7] A. Cauchy. Methode generale pour la resolution des systemes d'equations simultanees. *Comp. Rend. Acad. Sci. Paris*, pages 536–538, 1847.

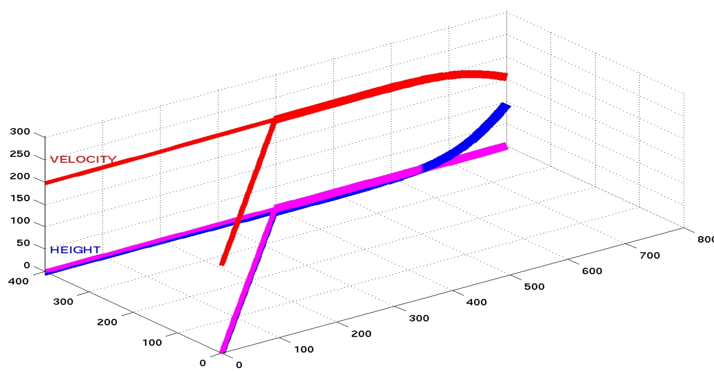
[8] M. Cirina. Nonlinear hyperbolic problems with solutions on preassigned sets. *Michigan Math. J.*, 17:193–209, 1970.

[9] J. M. Coron. Local controllability of a 1–d tank containing a fluid modelled by the shallow water equations. *ESAIM:COCV*, 8:513–554, 2002.

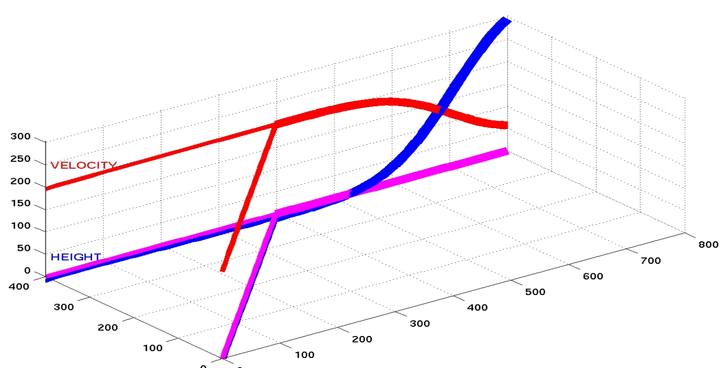
Figure 3: Velocity and water height with the computed control



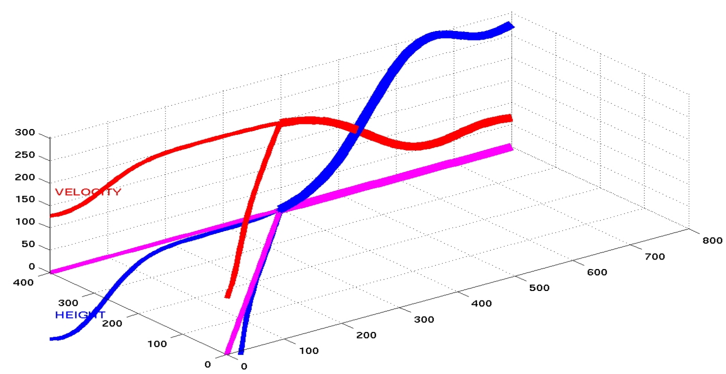
(a) t=7T/60



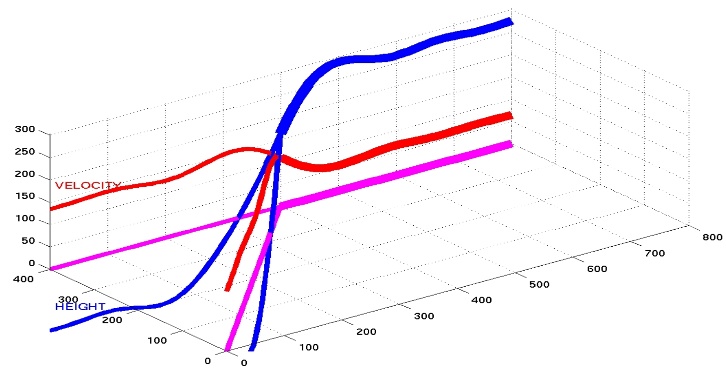
(b) t=9T/60



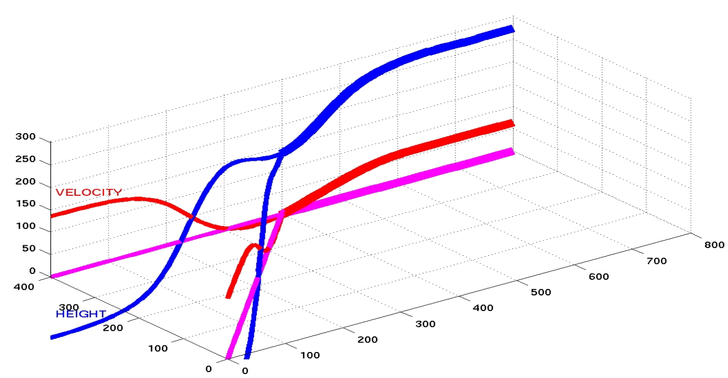
(c) t=11T/60



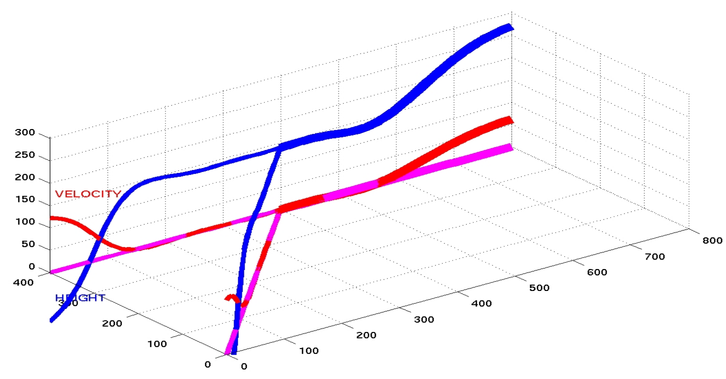
(d) t=13T/60



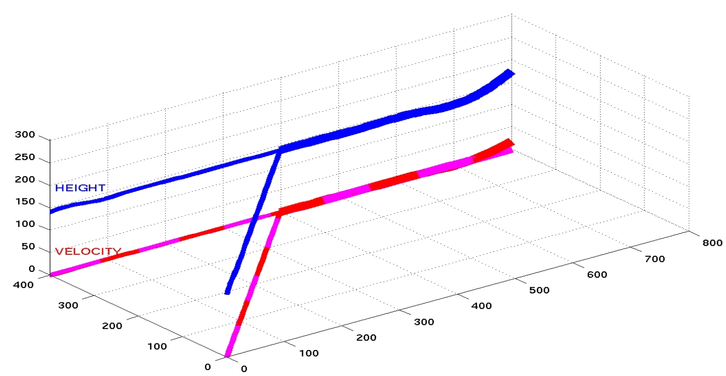
(e) t=15T/60



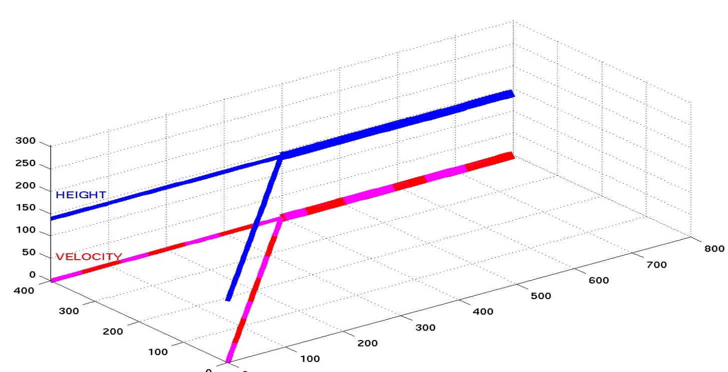
(f) t=17T/60



(g) t=19T/60



(h) t=21T/60



(i) t=23T/60

[10] C. M. Dafermos. *Hyperbolic Conservation Laws in Continuum Physics*. Springer, Berlin, 2000.

[11] M. Gugat. Optimal boundary control in flood management. *submitted to Computational Optimization and Applications*, 2004.

[12] M. Gugat and G. Leugering. Global boundary controllability of the de st. venant equations between steady states. *Annales de l'Institut Henri Poincaré, Nonlinear Analysis*, 20:1–11, 2003.

[13] M. Gugat, G. Leugering, and E.J.P.G. Schmidt. Global controllability between steady supercritical flows in channel networks. *Mathematical Methods in the Applied Sciences*, 27:781–802, 2004.

[14] M. Hinze and T. Slawig. Adjoint gradients compared to gradients form algorithmic differentiation in instantaneous control of the navier stokes equations. *Optimization Methods and Software*, 18:299–315, 2003.

[15] H. Holden and N. H. Risebro. A mathematical model of traffic flow on a network of unidirectional roads. *SIAM J. Math. Analysis*, 26:999–1017, 1995.

[16] H. Holden and N. H. Risebro. *Front tracking for hyperbolic conservation laws*. Springer, Berlin, 2002.

[17] C. T. Kelley. *Iterative Methods for Optimization*. SIAM, Philadelphia, 1999.

[18] P. G. LeFloch. *Hyperbolic Systems of conservation laws. The Theory of classical and nonclassical shock waves*. Birkhaeuser, Basel, 2002.

[19] G. Leugering and E. J. P. Georg Schmidt. On the modelling and stabilisation of flows in networks of open canals. *SIAM J. on Control and Optimization*, 41:164–180, 2002.

[20] P. Linz. *Analytical and Numerical Methods for Volterra Equations*. SIAM, Philadelphia, 1985.

[21] B. F. Sanders and N. D. Katapodes. Adjoint sensitivity analysis for shallow water wave control. *Journal of Irrigation and Drainage Engineering*, 126:909–919, 2000.

[22] Li Ta-tsien. *Global Classical solutions for quasilinear hyperbolic systems*. Masson, Paris, 1994.

[23] Ta tsien LI, Bopeng RAO, and Yi JIN. Semi–global c^1 solution and exact boundary controllability for reducible quasilinear hyperbolic systems. *Math. Modell. and Num. Anal.*, 33:399–408, 2000.

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

[25] S. Ulbrich. Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws. *Systems and Control Letters*, 48:309–324, 2003.

[26] R. Le Veque. *Finite Volume Methods for Hyperbolic Problems.* Cambridge University Press, Cambridge, 2002.

[27] W. Walter. *Differential and Integral Inequalities.* Springer Verlag Berlin, 1970.

[28] G. B. Whitham. *Linear and Nonlinear Waves.* John Wiley, New York, 1974.