Mauro Garavello

Control of distributed and hybrid systems.

Thesis submitted for the degree of Doctor Philosophiae
ACADEMIC YEAR 2003-2004

Supervisor: Prof. Benedetto Piccoli

To my parents

I am very grateful to my supervisor prof. B. Piccoli for proposing me the study of many interesting problems, for the attention he spent to follow me and also for the hospitality he gave me. I also would like to thank prof. J.–M. Coron and prof. Y. Chitour for their stimulating collaborations during my visiting at the Université Paris–sud and for the hospitality.

I thank my parents, Veronica and my relatives for their constant encouragement and support.

Moreover I thank all my friends, in particular Federico, Giacomo, Massimiliano and Simone, who made my return in Este so nice.

Finally I also thank all my colleagues met at S.I.S.S.A., at Université Paris–sud and at I.A.C. during my Ph.D.

Roma, July 2004 *Mauro Garavello*

Contents

Notation

We collect here a list of notation commonly used in this thesis.

$\mathbb{N}$	the set of natural numbers including 0.
$\mathbb{R}$	the set of real numbers.
$K^{\perp}$	the orthogonal set to K.
X	the state space.
$\mathcal{U}$	the control space.
Dom	the domain of a function.
Id	the identity function.
L^p_{loc}	the set of functions locally L^p integrable.
$\mathcal{B}_{(a,b)}$	the Borel σ–algebra on (a, b).
$\mathfrak{M}_b(a, b; \mathbb{R})$	the set of scalar bounded Radon measures on (a, b).
$\mathfrak{M}_b(a, b; \mathbb{R}^n)$	the set of vector valued bounded Radon measures on (a, b).
Δ	the Laplace operator.
g.c.d.	the great common divisor of two or more natural numbers.
x^{T}	the transpose of the vector x.

Chapter 1

Introduction.

We consider systems whose evolution is given by

$$x(t) = \mathcal{T}(t, u(\cdot))(x_0) \tag{1.0.1}$$

where $t \in \mathcal{I}$, the set of times, $u(\cdot) \in \mathcal{U}$, the space of controls, $x_0 \in X$, the state space, $\mathcal{T}(t, u(\cdot)) : X \to X$, the evolution operator and $\mathcal{T}(0, u(\cdot)) = \mathrm{Id}_X$ for every $u \in \mathcal{U}$. Systems of this type are called *control systems*, since the control function (or control law) $u(\cdot)$ acts modifying evolutions of trajectories. Usually, the control law is assigned in two ways:

- as a function of time (open loop);
- as a function of the state (closed loop or feedback).

Some of the main questions concerning a control system are:

1. **Modelling**. It consists in describing some real processes, in which an exogenous agent, usually human beings, acts on the system modifying its evolution. This approach permits to understand well both the behavior of the system and the effects of agent actions on its evolution.

2. **Controllability**. By controllability we mean the possibility to connect every two distinct states by a trajectory. More precisely controllability holds if, for every $x, y \in X$, there exist a control $u \in \mathcal{U}$ and a time $t \in \mathcal{I}$ such that
$$y = \mathcal{T}(t, u(\cdot))(x).$$

3. **Optimal control**. An optimal control problem consists in assigning a cost functional $J(x(\cdot), u(\cdot))$ to each trajectory–control pair and to minimize J among all the trajectory–control pairs satisfying some constraints, i.e. to find a control $\bar{u} \in \mathcal{U}$ and a corresponding trajectory $\bar{x}(\cdot)$ such that

$$J(\bar{x}(\cdot), \bar{u}(\cdot)) = \min_{\{u \in \mathcal{U} : \mathcal{T}(t, u(\cdot))(x_0) \in S\}} J(x(\cdot), u(\cdot)),$$

where S is a subset of X. If such a trajectory–control pair $(\bar{x}, \bar{u})$ exists, then $(\bar{x}, \bar{u})$ is said an *optimal trajectory–control pair*.

A classical example of control system is the following one. The state space X is equal to $\mathbb{R}^n$ with $n \in \mathbb{N}$ or to a finite dimensional manifold, $\mathcal{I} = [0, +\infty[$, there is a compact set U such that the control space $\mathcal{U}$ is

$$\{u : [0, T] \to U \text{ measurable: } T > 0\}$$

and $\mathcal{T}(t, u(\cdot))$ is equal to $\Phi(t, u(\cdot))$, where $\Phi(t, u(\cdot))$ is the flux of the system

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \\ x(0) = x_0, \end{cases} \tag{1.0.2}$$

with f smooth. In this setting there are a lot of interesting results in the literature regarding 1, 2 and 3 and also for some other questions as stability and observability. We refer for example to [2, 21, 68, 69, 97].

The main purpose of the thesis is to extend some results for control system (1.0.2) to other classes of evolutionary systems.

In Chapters 2, 3 and 4, one considers the evolution system (1.0.1), where the set $\mathcal{I}$ is equal to a compact interval $[0, T]$ with $T > 0$, the state space X is an infinite dimensional Banach space and $\mathcal{T}(t, u(\cdot))$ is the evolution operator of a partial differential equation with boundary conditions both of parabolic and hyperbolic type. For this control system controllability properties are studied.

In Chapters 5 and 6, we construct a model to describe the evolution of traffic flow in a road network. A road network is composed by a finite collection of roads connected by junctions and we model each road by an interval, possibly unbounded, of $\mathbb{R}$. On each road the evolution of some macroscopic variables such as density and speed of cars is considered. So in this setting X is an infinite dimensional space, the set $\mathcal{I}$ is equal to $[0, +\infty[$

and $\mathcal{T}(t, u(\cdot))$ is the evolution operator of hyperbolic conservation laws on a one–dimensional topological graph. We focus on modelling aspects for these systems.

In Chapters 7 and 8 we consider a control system, where both a continuous and discrete evolutions are present. An element x belongs to the state space X if $x = (x_c, x_d)$ and

$$x_c(t) = \mathcal{T}_c(t, u_c(\cdot), u_d(\cdot)) \tag{1.0.3}$$

evolves as in the classical case (1.0.2) with $\mathcal{I}_c = [0, +\infty[$, while

$$x_d(t) = \mathcal{T}_d(t, u_c(\cdot), u_d(\cdot)) \tag{1.0.4}$$

is a discrete evolution and $\mathcal{I}_d$ is isomorphic to a subset of $\mathbb{N}$. For this kind of systems, called *hybrid systems*, optimal control problems and properties of optimal trajectories are studied.

Let us now describe in more details the contents of each chapter.

In the first chapters, one studies exact controllability of some partial differential equations. The problem can be formulated in the following way. Consider an evolution system and a time interval $(0, T)$. Given an initial and final state, the aim is to find a control in such a way the solution of the system coincides with the initial state at $t = 0$ and with the final state at $t = T$. For a general introduction to this topic, the reader can refer to the survey paper by Russell [102]. We consider some controllability problems: steady–state controllability for heat and Saint–Venant equations and controllability for a Burgers viscous equation.

In Chapter 2 we treat the case of the heat equation on an open and bounded subset Ω of $\mathbb{R}^n$. The problem is whether it is possible to drive the initial datum $y_0 \equiv 1$ to the final datum $y_1 \equiv 0$ in time T by choosing appropriately a boundary condition depending only by the time. If this problem has a positive answer, then the Laplace transform of the solution, after reparametrizations, satisfies an eigenvalue problem for the Laplacian with the Dirichlet zero boundary condition. This implies the existence of a holomorphic function $B(s)$, defined on the whole $\mathbb{C}$, which, by Paley–Wiener theorem, satisfies the growth condition

$$|B(s)| \leqslant C e^{T|\operatorname{Re} s|}, \tag{1.0.5}$$

for some $C > 0$ and which is zero on all eigenvalues of the Laplace–Dirichlet operator for which there exists a non–zero eigenfunction with non–zero mean. So the problem becomes the following one:

- find an entire function f, satisfying the growth condition (1.0.5), such that the distribution of the zeroes of f is the same as the distribution of some eigenvalues of the Laplace–Dirichlet operator.

Clearly, if the number of zeroes of f is sufficiently big, then the holomorphic function is subject to many constraints. Hence it is necessarily equal to 0 and so the controllability problem has not a solution.

It turns out that this obstruction depends only on the shape of the set Ω. Therefore if the domain Ω satisfies some assumptions, then the original controllability problem has not solutions. It is also proven that these assumptions on Ω are generic, in the sense that generic domains, with respect to a suitable topology, satisfy them. When Ω is a cube or a parallelepiped, we are able to find all the eigenvalues and to check that the assumptions are valid, hence the controllability problem has not solutions. If instead Ω is a ball in $\mathbb{R}^n$, then it does not satisfy these assumptions and the conjecture is that in this case the problem is controllable.

In Chapter 3, we consider the problem of moving a tank, containing a fluid, from a position to another one in such a way that the fluid is in equilibrium at $t = 0$ and $t = T$. The mathematical description of the problem is given by

$$\begin{cases} \ddot{D}(t) = u(t), & \text{if } t \in (0, T), \\ h_{tt}(t, x) = \Delta h(t, x), & \text{if } (t, x) \in (0, T) \times \Omega, \\ \frac{\partial h}{\partial \nu}(t, x) = -u(t) \cdot \nu(x), & \text{if } (t, x) \in (0, T) \times \partial\Omega. \end{cases} \tag{1.0.6}$$

where D is the position in $\mathbb{R}^2$ of the tank, $\Omega \subseteq \mathbb{R}^2$ is the shape of the tank, h denotes the height of the fluid with respect to an equilibrium configuration, the control u is the acceleration of the tank and $\nu(x)$ denotes the outward unit normal vector to Ω at $x \in \partial\Omega$, see Figure 1.1. The function h satisfies the Saint–Venant equation, that, in this case, reduces to a wave equation with Neumann boundary conditions.

The same technique as in the previous case is used, i.e. applying Laplace transform and translating the problem in the complex analysis setting. In this case the problem is more complicated than the previous one, since we have to consider a holomorphic function on $\mathbb{C}^2$. However, we give conditions on the shape Ω of the tank, that is conditions on eigenvalues and eigenfunctions of the Laplace–Neumann operator, that imply non–controllability.

In Chapter 4 the Burgers viscous equation

$$y_t - y_{xx} + 2yy_x = 0 \tag{1.0.7}$$

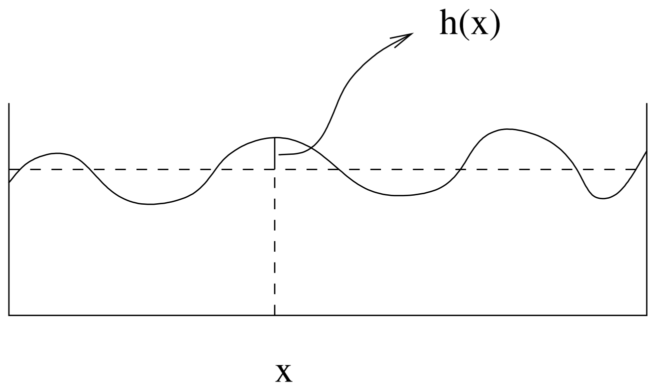


Figure 1.1: Tank containing a fluid. The function h denotes the shift of the fluid from the equilibrium position.

in the domain $(t, x) \in [0, T] \times [0, 1]$ is considered. We suppose that the controls acts on both sides of the domain, that is at $x = 0$ and $x = 1$. If the control acts only on one side, then Fursikov and Imanuvilov in [56] and Diaz in [48] proved that approximated controllability in $L^2(0, 1)$ fails. Therefore a natural question is whether the system has the exact controllability property when the control acts on both sides. Clearly, due to smoothing property of parabolic equations, we can not reach every element of $L^2(0, 1)$ at time $t = T$. The correct definition of exact controllability for this equation is that given by Fursikov and Imanuvilov in [56, 57], i.e. passing from one solution to another solution. Fursikov and Imanuvilov in [56] proved also a local controllability result. They showed that it is possible to drive the system from one solution to another one in any time $T > 0$ if the two solutions are sufficiently near in the norm of the Sobolev space $W^{1,2}(0, 1)$. Here some partial results about global exact controllability are proved, even if the general case remains open. First it is shown that it is possible to drive the system from 0 to an arbitrary big constant, provided the time T is sufficiently big. This result needs the local exact controllability by Fursikov and Imanuvilov and Carleman type estimates on solutions. Then also null controllability is proved for (1.0.7).

In Chapter 5, we propose a model for traffic flow in a road network, based on a first order equation in conservation form, proposed by Lighthill and Whitham [77] in 1955 and independently by Richards [99] in 1956. This formulation permits to reveal in particular congestions of traffic, since these equations may produce discontinuities in finite time even if the initial datum is smooth. On each road the model is given by the equation

$$\rho_t + f(\rho)_x = 0,$$

where ρ is the density of cars, v is the speed of cars and $f(\rho) = \rho v$ is the flux. One usually assumes that v depends only by the density in a decreasing way and f is a strictly concave function of the density with a unique maximum point. Roads are connected by junctions and so a definition of solution at junctions is needed. Conservation of cars is not sufficient to determine a unique solution. So we suppose that

(A) there are some prescribed preferences of drivers, that is the traffic from incoming roads is distributed on outgoing roads according to fixed coefficients;

(B) respecting (A), drivers behave so as to maximize fluxes.

To deal with (A), one assumes the existence for every junction of a distributional matrix A, whose entries are the percentage of the flux from an incoming to an outgoing road. These rules provide the existence of a unique solution to the Riemann Problem at junctions, i.e. to the Cauchy problem with initial data constant on each road. We also prove the existence of entropic solutions to Cauchy problems on the whole network. The main technique is that of wave front tracking approximate solution (see [26]). Unfortunately in the general case the solution does not depend in a Lipschitz way by the initial data, as shown in an example. Some results are proved under special assumptions, while the continuous dependence by initial data is still an open problem for the general case. We also describe the problem from the control point of view, by assuming the matrix A time dependent as in the case of traffic lights at junctions.

In Chapter 6 we study traffic flow using the model proposed by Aw and Rascle [13] in 2000. The model is given by

$$\begin{cases} \partial_t \rho + \partial_x (y - \rho^{\gamma+1}) = 0, \\ \partial_t y + \partial_x (\frac{y^2}{\rho} - y\rho^{\gamma}) = 0, \end{cases}$$

where ρ is the density of cars, y is the “momentum” (see [13]) and $\gamma > 0$ is a constant. It is a second order traffic model in conservation form, where the conserved quantities are the density and the momentum. The first second order models for traffic were proposed by Payne [90] in 1971 and by Whitham [119] in 1974. Some other second order models were proposed in the literature, see [70, 71, 72]. Unfortunately, these models are not reasonable and lead to unrealistic results, since cars may move backwards as shown by Daganzo [44]. Then in 2000 Aw and Rascle in [13] corrected the model and eliminated bad behaviors.

To construct solutions, again a definition of solution to Riemann problems at junctions is needed. In analogy with the Lighthill–Whitham model, let us assume

(a) conservation of cars;

(b) the existence of a matrix A giving the percentage of the flux of the density from an incoming road to an outgoing one;

(c) maximization of the flux of the density.

These rules are sufficient to isolate a unique solution on incoming roads, but in outgoing roads in general there is a one–dimensional manifold of solutions. Hence an additional rule for uniqueness is needed. We propose three different additional rules:

(AR-1) maximize the speed of cars;

(AR-2) maximize the density of cars;

(AR-3) minimize the total variation of the density along the solution.

For each additional rule, we prove that there is a unique solution and we describe in detail stability with respect to L^∞ perturbations.

For a network consisting of just one junction with n incoming roads and m outgoing ones, we prove the existence of an entropic solution to a Cauchy problem when the initial datum is closed to a stable equilibrium configuration and is sufficiently small in total variation. This clearly is the first step to consider control problem in traffic flow with the Aw–Rascle model.

The last part of this thesis deals with hybrid systems. Roughly speaking a hybrid system is a collection of control systems called locations, possible defined on different manifolds, and an automaton that rules the switchings between locations. The definition of hybrid system is that of [59, 93, 110]. The term hybrid indicates the presence of both continuous and discrete dynamics. The continuous part is given by location controlled dynamics, while the discrete one by the automaton. An optimal control problem is obtained assigning Lagrangian running costs on each location and final and switching costs. More precisely a hybrid control system is a 7-tuple

$$\Sigma = (\mathcal{Q}, M, U, f, \mathcal{U}, J, \mathcal{S}) \tag{1.0.8}$$

such that

H1. $\mathcal{Q}$ is a finite set;

H2. $M = \{M_q\}_{q\in\mathcal{Q}}$ is a family of smooth manifolds, indexed by $\mathcal{Q}$;

H3. $U = \{U_q\}_{q\in\mathcal{Q}}$ is a family of sets;

H4. $f = \{f_q\}_{q\in\mathcal{Q}}$ is a family of maps $f_q : M_q \times U_q \mapsto TM_q$ (TM_q is the tangent bundle of M_q), such that $f_q(x,u) \in T_xM_q$ for every $(x,u) \in M_q \times U_q$;

H5. $\mathcal{U} = \{\mathcal{U}_q\}_{q\in\mathcal{Q}}$ is a family of sets $\mathcal{U}_q$ whose members are maps $u : Dom(u) \to U_q$, defined on some interval $Dom(u) \subset \mathbb{R}$;

H6. $J = \{J_q\}_{q\in\mathcal{Q}}$ is a family of subintervals of $\mathbb{R}^+$;

H7. $\mathcal{S}$ is a subset of

$$\{(q, x, q', x', u(\cdot), \tau) : q, q' \in \mathcal{Q}, x \in M_q, x' \in M_{q'}, u(\cdot) \in \mathcal{U}_{q'}, \tau \in J_{q'}\}.$$

The system evolves in a location $q \in \mathcal{Q}$ according to the corresponding controlled dynamic f_q and then switches as prescribed by $\mathcal{S}$. The intervals J_q indicate the lengths of time interval on which the system can stay in location q.

Recently optimization problems for hybrid systems have attracted a lot of attention, see [11, 22, 33, 47, 64, 100]. This is due to the fact that many physical and mechanical systems present both continuous and discrete characteristics. A very elementary example is a car with gears: acceleration and braking constitutes the continuous dynamic, while change of gear represents the discrete dynamic.

For an optimal classical control problem, the main tool toward the construction of an optimal trajectory is the Pontryagin Maximum Principle. For a hybrid system there exists a generalization of PMP, proved by Piccoli [93] in 1998 and by Sussmann [110] in 1999. The key point is the switching mechanism, that permits to pass from one location to another one with possible restrictions on state and time to spend in next location. The strategy to prove the Hybrid Maximum Principle (HMP) by Sussmann is essentially the same of PMP. Some variations are performed on the supposed optimal trajectory and they produce necessary conditions for a trajectory to be optimal. In hybrid setting, it is important to understand how variations propagate after a switching.

A more general case of switching mechanisms for hybrid systems than that of [93, 110] is considered. In particular, we assume that the switching strategy provides some restrictions on the set of admissible controls. These restrictions affect the general strategy of PMP and HMP. In fact, variations in PMP and HMP are generated by "needle variations", that are modifications of the control in a small interval of time and then prolonged after switchings. In our setting, these variations are not admissible, in the sense that they produce a change in the switching strategy, hence we are not allowed to use the same control after switchings. Therefore we introduce a more general kind of variations, according to the fact that the switching strategy affects the choice of the controls, and we define the concept of "map of variations". The basic requests are weak differentiability properties in the space of bounded Radon measure. The problem of producing variations more general than needle variations was extensively considered in the literature for classical setting (see [1, 3, 18, 24, 74, 107]) and also for hybrid setting (see [110]). We prove in this way a Hybrid Necessary Principle (HNP) giving necessary conditions for an optimal hybrid trajectory. Notice that the word "maximum" is eliminated, since the conditions can not be expressed in supremum form.

In Chapter 8 these results are applied to a simple model of a car with gears. We describe this model with a hybrid system, where each location corresponds to a gear of the car. In each location the car is described by the system

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u\, g_i(x_2), \end{cases} \tag{1.0.9}$$

where x_1 is the position of the car, x_2 is its speed, g_i is the gear function and the control u corresponds to acceleration. First we apply HMP to solve some optimality problems, then we show that HNP is appropriate for a hybrid system of this type, since it permits to exclude non–optimal trajectories satisfying HMP.

Part I

Partial Differential Equations.

Chapter 2

Heat equation.

Let Ω be an open, bounded and non empty subset of $\mathbb{R}^n$, with $n \geqslant 2$. For $y_0 \in L^2(\Omega)$ and $T > 0$, consider the heat equation

$$\begin{cases} y_t(t,x) - \Delta y(t,x) = 0, & \text{if } (t,x) \in (0,T) \times \Omega, \\ y(0,x) = y_0, & \text{if } x \in \Omega, \\ y(t,x) = u(t), & \text{if } x \in \partial\Omega, \end{cases} \tag{2.0.1}$$

where $u \in L^2(0,T)$ is the control. Let us first recall classical results about weak solutions of the Cauchy problem (2.0.1). Let $y_0 \in L^2(\Omega)$, $T > 0$ and $u \in L^2(0,T)$. A weak solution of the Cauchy problem (2.0.1) is a function $y \in C^0([0,T], L^2(\Omega))$ such that, for every $\tau \in [0,T]$ and every $\theta \in C^1([0,T]; L^2(\Omega)) \cap C^0([0,T]; H_0^1(\Omega))$ with

$$\theta_t + \Delta\theta = 0 \text{ in } C^1([0,T]; L^2(\Omega)), \tag{2.0.2}$$

$$\tag{2.0.3}$$

one has

$$\int_\Omega y(\tau,x)\theta(\tau,x)dx - \int_\Omega y_0(x)\theta(0,x)dx = \int_0^\tau u(t)\left(\int_\Omega \theta_t(t,x)dx\right)dt. \tag{2.0.4}$$

Of course, every $y \in C^1([0,T]; L^2(\Omega)) \cap C^0([0,T]; H^1(\Omega))$, which is a classical solution of (2.0.1) is a weak solution of (2.0.1). It is also well known that, for every $y_0 \in L^2(\Omega)$, $T > 0$ and $u \in L^2(0,T)$, there exists one and only one weak solution y to (2.0.1). That unique y will be called the solution to the Cauchy problem (2.0.1).

The problem of null controllability associated to (2.0.1) goes as follows. Given $y_0 \in L^2(\Omega)$, does there exist $T > 0$ and $u \in L^2(0,T)$ such that the solution of the Cauchy problem (2.0.1) satisfies $y(T,\cdot) = 0$? The answer to that question is negative, as shown by H. Fattorini in [52, Theorem 2.2] and by S. Avdonin and S. Ivanovin in [12, Theorem IV.2.7, page 187]; see also the articles [82, 83, 84] by S. Micu and E. Zuazua, for even stronger negative results for similar questions.

In this chapter, we look at a particular y_0, namely $y_0 \equiv 1$, and want to see if it is possible to steer that special y_0 to 0 in finite time, that is, again, does there exists $T > 0$ and a control $u \in L^2(0,T)$ such that the solution y of the Cauchy problem (2.0.1) satisfies $y(T,\cdot) = 0$? Of course, a positive answer to that question is equivalent to the steady–state controllability, i.e. given two constant functions $y_0 \equiv C_0$, $y_1 \equiv C_1$, does there exist $T > 0$ and $u \in L^2(0,T)$ such that the solution y of (2.0.1) satisfies $y(0,\cdot) = y_0$ and $y(T,\cdot) = y_1$? As mentioned in the introduction, P. Rouchon shows in [101] that the steady–state controllability holds for $n = 1$ or if Ω is a ball in $\mathbb{R}^n$ and asks what is the answer for general open subsets of $\mathbb{R}^n$, $n \geqslant 2$.

We use $-\Delta_\Omega$ to denote the Laplace–Dirichlet operator defined next,

$$\mathcal{D}\left(-\Delta_\Omega\right) := \{v \in H_0^1(\Omega); \Delta v \in L^2(\Omega)\},$$
$$-\Delta_\Omega v := -\Delta v,\ \forall v \in \mathcal{D}\left(-\Delta_\Omega\right).$$

Let us introduce the definition of Property (A), which will turn out to be an obstruction for steering $y_0 \equiv 1$ to 0 in finite time.

Definition 2.0.1 *The open set Ω has the property (A) if there exists a sequence $(r_k)_{k\in\mathbb{N}^*}$ of distinct eigenvalues of $-\Delta_\Omega$ such that*

(i) One has

$$\sum_{k=1}^{\infty} \frac{1}{r_k} = \infty, \tag{2.0.5}$$

(ii) For every $k \in \mathbb{N}^*$, there exists an eigenfunction w for the eigenvalue r_k and the operator $-\Delta_\Omega$ such that

$$\int_\Omega w dx \neq 0. \tag{2.0.6}$$

We are now able to state the main result of this section.

Theorem 2.0.1 *Let Ω be a bounded, open and non empty subset of $\mathbb{R}^n$, $n \geqslant 2$. If Ω has the property (A), then one cannot steer $y_0 \equiv 1$ to 0 in finite time.*

Proof. Assume that property (A) holds for a bounded, open and non empty subset $\Omega \subset \mathbb{R}^n$, $n \geq 2$. We suppose by contradiction that there exist $T > 0$ and $u \in L^2(0,T)$ such that the solution y of the Cauchy problem (2.0.1) with

$$y_0 \equiv 1, \tag{2.0.7}$$

satisfies

$$y(T,\cdot) = 0. \tag{2.0.8}$$

Let λ be an eigenvalue of $-\Delta_\Omega$ and w be an eigenfunction associated to λ. Consider $\theta \in C^\infty([0,T]; H_0^1(\Omega))$ defined by

$$\theta(t,x) := e^{\lambda t} w(x).$$

Then θ satisfies (2.0.3). Hence, using (2.0.4) with $\tau := T$, (2.0.7) and (2.0.8), one gets

$$B(\lambda) \int_\Omega w dx = 0, \tag{2.0.9}$$

where $B : \mathbb{C} \to \mathbb{C}$ is defined by

$$B(s) := 1 + s \int_0^T u(t) e^{st} dt. \tag{2.0.10}$$

Since property (A) holds for Ω, it results that B vanishes on a sequence $(r_k)_{k \in \mathbb{N}^*}$ of distinct positive real numbers satisfying (2.0.5). By the easy part of the Paley-Wiener theorem, the function B is holomorphic on $\mathbb{C}$ and there exists $C > 0$ such that

$$|B(s)| \leqslant C e^{T \max\{0, Re(s)\}}, \quad \forall s \in \mathbb{C}. \tag{2.0.11}$$

We then apply the following lemma.

Lemma 2.0.1 *Let $f : \mathbb{C} \to \mathbb{C}$ be a holomorphic function such that, for some $C > 0$,*

$$|f(s)| \leqslant Ce^{C|Re(s)|}, \ \forall s \in \mathbb{C}.$$

Let us assume that there exists a sequence a sequence $(r_k)_{k\geqslant 1}$ of distinct positive real numbers such that (2.0.5) holds and

$$f(r_k) = 0, \quad \forall k \geqslant 1. \tag{2.0.12}$$

Then, f is identically equal to 0.

Applying Lemma 2.0.1 with $f := B$, we conclude that B is identically equal to zero. That contradicts the fact that $B(0) = 1$. □

Proof of Lemma 2.0.1. We may suppose that f is symmetric with respect the origin otherwise it is sufficient to consider the function $f(s)f(-s)$. We define

$$g(s) := \prod_{k\in\mathbb{N}} \left[\left(1 - \frac{s}{r_k}\right)\left(1 + \frac{s}{r_k}\right) \right].$$

It is clear that g is well defined since the product is convergent. Therefore the function defined by

$$h(s) := \frac{f(s)}{g(s)}$$

is holomorphic on $\mathbb{C}$. By hypotheses, it is clear that f is bounded on the imaginary axis. Moreover, we have that if $y \in \mathbb{R}$, then

$$\begin{aligned} \ln|g(iy)| &= \sum_{k\in\mathbb{N}} \ln\left[1 + \frac{y^2}{r_k^2}\right] \\ &\geqslant C_1|y| - C_2, \end{aligned}$$

where C_1 and C_2 are positive constants and so

$$|h(iy)| \leqslant Ke^{-C_1|y|}$$

for every $y \in \mathbb{R}$. By [17, Corollary 4.5.7, p. 358], for every $s \in \mathbb{C}$ we have:

$$|h(s)| \leqslant Ce^{A_1|s|}$$

where C and A_1 are positive constants. Moreover, by [61, Théorème 7 and 9, pages 212 and 216] we have that

$$|h(s)| \leqslant C_3 e^{-a|\mathrm{Im}\, y|+b'|\mathrm{Re}\, s|}, \tag{2.0.13}$$

where C_3, a and b' are positive constants. Such a function h must be identically equal to zero ([61, page 225]), and so is f. $\square$

Remark 1 *In Lemma 2.0.1, it is very important that the zeroes of the function f belong to the real axis. Indeed the result is false if the zeroes belong to the imaginary axis. In fact the function*

$$f(s) := \sinh(s)$$

satisfies clearly

$$|f(s)| \leqslant e^{|\mathrm{Re}\, s|}$$

and

$$f(i\pi k) = 0$$

for every $k \in \mathbb{N}$.

2.1 Generic non steady–state controllability

In this section, we want to prove that condition (A) holds for generic bounded open subsets Ω (and therefore, by Theorem 2.0.1, for such generic sets Ω, one cannot steer $y_0 \equiv 1$ to 0 in finite time).

We use here notations and results of [4, 62, 117]. Let $\mathcal{R}(\mathbb{R}^n)$ be the set of all non empty bounded open subsets Ω of class C^3. To state the result, one needs to define a topology on $\mathcal{R}(\mathbb{R}^n)$. We follow a construction closely related to that proposed by by R. Hamilton in [62, pages 86-87]. For $\Omega \in \mathcal{R}(\mathbb{R}^n)$, let $\xi \in C^3(\partial\Omega; \mathbb{R}^n)$ be such that

$$\xi(x) \cdot \nu(x) > 0, \forall x \in \partial\Omega, \tag{2.1.14}$$

where $\nu \in C^2(\partial\Omega, \mathbb{R}^n)$ denotes the outward normal to Ω.

Let $\varepsilon_0 > 0$ be small enough so that the two following properties hold.

(i) For every x in $\mathbb{R}^n$ such that $\mathrm{dist}(x, \partial\Omega) < \varepsilon_0$, there exists a unique $\pi(x) \in \partial\Omega$ such that $x - \pi(x)$ is parallel to $\xi(\pi(x))$.

(ii) The map $x \mapsto \pi(x)$ is of class C^3 on the open set

$$\{x \in \mathbb{R}^n ; \operatorname{dist}(x, \partial\Omega) < \varepsilon_0\}.$$

Let $\varepsilon > 0$ and $\eta \in C^3(\partial\Omega)$ be such that

$$|\eta|_{C^3(\partial\Omega)} < \varepsilon. \tag{2.1.15}$$

Define

$$\Omega_\eta := \{x \in \Omega ; \operatorname{dist}(x, \partial\Omega) \geqslant \varepsilon_0\} \cup \{x \in \mathbb{R}^n ; \operatorname{dist}(x, \partial\Omega) < \varepsilon_0 \\ \text{and } (x - \pi(x)) \cdot \xi(\pi(x)) < \eta(\pi(x))\}.$$

There exists $\varepsilon_1 > 0$ such that, for every $\eta \in C^3(\partial\Omega)$ satisfying such that $|\eta|_{C^3(\partial\Omega)} < \varepsilon_1$, Ω_η is a bounded subset of $\mathbb{R}^n$ of class C^3. Let $\mathcal{V}(\varepsilon)$ be the set of all the Ω_η with $\eta \in C^3(\partial\Omega)$ satisfying (2.1.15). We define a topology on $\mathcal{R}(\mathbb{R}^n)$ by considering the sets $\mathcal{V}(\varepsilon)$, with $\varepsilon \in (0, \epsilon_1)$, as a base of neighborhoods of Ω, i.e. every neighborhood of Ω in $\mathcal{R}(\mathbb{R}^n)$ contains some $\mathcal{V}(\varepsilon)$ for $\varepsilon \in (0, \epsilon_1)$ small enough. (One easily checks that this topology is independent of the choice of ξ and ε_1.) Recall that a topological space is a Baire space if any residual set, i.e. any intersection of denumerable open dense subsets, is dense. Since, for every Ω in $\mathcal{R}(\mathbb{R}^n)$, $C^3(\partial\Omega)$ is a Baire space, it follows from our definition of the topology on $\mathcal{R}(\mathbb{R}^n)$ that $\mathcal{R}(\mathbb{R}^n)$ is also a Baire space. (Proceeding as in [62, 4.4.7], one can also prove that $\mathcal{R}(\mathbb{R}^n)$ with our topology is a C^0-manifold modeled on the Banach spaces $C^3(\partial\Omega)$ with $\Omega \in \mathcal{R}(\mathbb{R}^n)$. But we do not need that property.)

Let us recall that that a property (P) holds for generic $\Omega \in \mathcal{R}(\mathbb{R}^n)$ if there exists a residual subset $\tilde{D} \subset \mathcal{R}(\mathbb{R}^n)$ such that property (P) holds for every $\Omega \in \tilde{D}$.

Remark 2 *The use of the transverse vector field ξ is needed to parameterize the variations of a domain Ω. For that purpose, a simpler choice would have been $\xi := \nu$ but there is a serious difficulty in doing so. Indeed, for an arbitrary Ω of class C^3, ν is of class C^2. More generally, there is always a non zero difference between the regularity of Ω and that of its outward normal vector field ν. To overcome that phenomenon of loss of derivative, one could apply a Nash-Moser type of result in order to get the necessary amount of surjectivity, which is clearly needed at some point of the argument. To avoid all that machinery, we followed, instead, the strategy advocated by D. Bresch and J. Simon in [23], consisting in choosing the transverse vector field ξ, as defined in (2.1.14), which has the same regularity as the domain Ω.*

In this section, we prove the following result.

Theorem 2.1.1 *Condition (A) holds for generic* $\Omega \in \mathcal{R}(\mathbb{R}^n)$*.*

Proof of Theorem 2.1.1. The strategy of proof is standard and goes as follows (cf. [4]). Let $\mathcal{G} \subset \mathcal{R}(\mathbb{R}^n)$ be the set of $\Omega \in \mathcal{R}(\mathbb{R}^n)$ such that

(a) all eigenvalues of $-\Delta_\Omega$ are simple,

(b) $\int_\Omega w dx \neq 0$, for every non zero eigenfunction w of $-\Delta_\Omega$.

Similarly, for every positive integer l, the set $\mathcal{S}_l \subset \mathcal{R}(\mathbb{R}^n)$ (respectively $\mathcal{G}_l \subset \mathcal{R}(\mathbb{R}^n)$) of open sets $\Omega \in \mathcal{R}(\mathbb{R}^n)$ is defined such that property (a) (respectively, and property (b)) holds at least for the first l eigenvalues of $-\Delta_\Omega$. Clearly, $\mathcal{G}$ is the countable intersection of the $\mathcal{G}_l$'s.

We show next that $\mathcal{G}$ is residual, which implies Theorem 2.1.1. Indeed, if property (a) holds for $-\Delta_\Omega$, then, by applying the Weyl formula for $-\Delta_\Omega$ (cf. [105, Theorem 15.2, p.124]), one deduces that $\lambda_k \sim_{k\to\infty} C(\Omega)k^{2/n}$, where $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_j < \lambda_{j+1} < \cdots$ is the ordered sequence of the eigenvalues of the Laplace-Dirichlet operator $-\Delta_\Omega$. Therefore, property (A) holds.

For $l \geqslant 0$, $\mathcal{S}_0 = \mathcal{G}_0 := \mathcal{R}(\mathbb{R}^n)$, $\mathcal{G}_l \subset \mathcal{S}_l$, $\mathcal{S}_{l+1} \subset \mathcal{S}_l$ and $\mathcal{S} := \cap_{l\geqslant 0}\mathcal{S}_l$ and, similarly, $\mathcal{G}_{l+1} \subset \mathcal{G}_l$ and $\mathcal{G} = \cap_{l\geqslant 0}\mathcal{G}_l$. Moreover, for $l \geqslant 0$, it is clear that the sets $\mathcal{S}_l$ and $\mathcal{G}_l$ are open in $\mathcal{R}(\mathbb{R}^n)$ (see [4]). To show that $\mathcal{G}$ is residual, it amounts to establish the next lemma.

Lemma 2.1.1 *For every* $l \geqslant 0$*,* $\mathcal{G}_{l+1}$ *is dense in* $\mathcal{G}_l$*.*

Proof of Lemma 2.1.1. First, recall that, for every $l \geqslant 0$, $\mathcal{S}_l$ is dense in $\mathcal{R}(\mathbb{R}^n)$ (see [117]).

We follow the lines of the argument of Theorem 2 in [4]. Let $\Omega \in \mathcal{G}_l$. It is sufficient to exhibit $\Omega' \in \mathcal{G}_{l+1}$, arbitrarily close to Ω. Since $\mathcal{S}_{l+1}$ is dense, it is enough to establish the previous fact for $\Omega \in \mathcal{G}_l \cap \mathcal{S}_{l+1}$. Let $(\mu_k)_{k\in\mathbb{N}^*}$ be the ordered sequence of the eigenvalues of the Laplace-Dirichlet operator $-\Delta_\Omega$ *repeated according to their multiplicity.* We have

$$\mu_1 < \mu_2 < \cdots \mu_l < \mu_{l+1} < \mu_{l+2} \leqslant \mu_{l+3} \leqslant \cdots .$$

Let w_{l+1} be an eigenfunction of $-\Delta_\Omega$ for the eigenvalue μ_{l+1}. If $\int_\Omega w_{l+1} dx \neq 0$, then $\Omega \in \mathcal{G}_{l+1}$. Otherwise, we may assume that

$$\int_\Omega w_{l+1} dx = 0, \qquad (2.1.16)$$

and we simply use μ and w to denote μ_{l+1} and w_{l+1}. Let $\xi \in C^3(\partial\Omega;\mathbb{R}^n)$ be such that (2.1.14) holds and let $\varepsilon_0 > 0$ be as above (see (i) and (ii) in this subsection). Set

$$\varepsilon_0' := \text{ Min}\,\{\xi(\pi(x))\cdot(\pi(x)-x);\, x\in\Omega,\, \text{dist}(x,\partial\Omega)=\varepsilon_0/2\} > 0.$$

Let $\rho \in C^\infty(\mathbb{R},[0,1])$ be such that

$$\rho = 1 \text{ on a neighborhood of } (-\infty, 0],$$
$$\rho = 0 \text{ on a neighborhood of } [\varepsilon_0', +\infty).$$

We use $C^3_\varepsilon(\partial\Omega)$ to denote the set of $\eta \in C^3(\partial\Omega)$ such that $|\eta|_{C^3(\partial\Omega)} < \varepsilon$. For $\eta \in C^3_\varepsilon(\partial\Omega)$, we consider $h_\eta : \overline{\Omega} \to \mathbb{R}^n$ defined by

$$h_\eta(x) := x,$$

for every $x \in \Omega$ with $\text{dist}(x,\partial\Omega) \geqslant \varepsilon_0/2$ and

$$h_\eta(x) := x + \eta(\pi(x))\,(1-\rho\,(\varepsilon_0' - \xi(\pi(x))\cdot(\pi(x)-x)))\,\xi(\pi(x)),$$

for every $x \in \overline{\Omega}$ with $\text{dist}(x,\partial\Omega) \leqslant \varepsilon_0/2$. We now fix $\varepsilon \in (0,\varepsilon_0)$ small enough so that, for every $\eta \in C^3_\varepsilon(\partial\Omega)$, h_η is a diffeomorphism of class C^3 from $\overline{\Omega}$ into $\overline{\Omega_\eta}$. Let $P : H^2(\Omega) \to H^2(\mathbb{R}^n)$ be a linear continuous map such that

$$P(v) = v \text{ in } \Omega.$$

For $\eta \in C^3_\varepsilon(\partial\Omega)$, let $Q_\eta : H^2(\mathbb{R}^n) \to H^1_0(\Omega_\eta) \cap H^2(\Omega_\eta)$, $\phi \mapsto \psi$, be defined by

$$-\Delta\psi = -\Delta\phi \text{ in } L^2(\Omega_\eta),$$
$$\psi = 0 \text{ on } \partial\Omega_\eta.$$

Consider

$$E := \left\{(v,\eta) \in H^2(\Omega)\times C^3_\varepsilon(\partial\Omega);\, v(x) + \eta(x)\frac{\partial w}{\partial \xi}(x) = 0,\, \forall x \in \partial\Omega\right\},$$

and the following map

$$\begin{aligned} \Phi : E\times\mathbb{R} \quad &\to L^2(\Omega)\times\mathbb{R} \\ ((v,\eta),\chi) \quad &\mapsto \Big(((-\Delta-\chi)(Q_\eta(P(v))))\circ h_\eta, \textstyle\int_{\Omega_\eta} Q_\eta(P(v))dx\Big). \end{aligned}$$

One has $\Phi((w,0),\mu) = (0,0)$ and Lemma 2.1.1 holds if Φ is locally onto at $((w,0),\mu)$. The map Φ is of class C^1 and one has

$$\Phi'((w,0),\mu)((v,\eta),\chi) = (-\Delta v - \mu v - \chi w, \int_\Omega v dx),$$

for every $(v,\eta) \in H^2(\Omega) \times C^3(\partial\Omega)$ such that

$$v(x) + \eta(x)\frac{\partial w}{\partial \xi}(x) = 0, \ \forall x \in \partial\Omega.$$

Using the Fredholm alternative (recall that the eigenvalue μ is assumed to be simple), one easily checks that, for every $f \in L^2(\Omega)$ and every $\eta \in C^3(\partial\Omega)$, there exists one and only one $(v,\chi) \in H^2(\Omega) \times \mathbb{R}$ such that

$$-\Delta v - \mu v - \chi w = f, \tag{2.1.17}$$

$$\int_\Omega vw dx = 0, \tag{2.1.18}$$

$$v(x) + \eta(x)\frac{\partial w}{\partial \xi}(x) = 0, \ \forall x \in \partial\Omega. \tag{2.1.19}$$

For $f = 0$, let us denote by (v_η, χ_η) the corresponding unique solution. We next prove that

$$\text{there exists } \eta_0 \in C^3(\partial\Omega) \text{ such that } \int_\Omega v_{\eta_0} dx \neq 0. \tag{2.1.20}$$

To compute $\int_\Omega v_\eta dx$ in terms of η, we consider the unique solution of the inhomogeneous Dirichlet problem given by

$$\begin{cases} (-\Delta_\Omega - \mu)S = 1, & \text{in } \Omega, \\ S = 0, & \text{on } \partial\Omega, \\ \int_\Omega Sw dx = 0. \end{cases} \tag{2.1.21}$$

Since $\int_\Omega w dx = 0$ and the eigenvalue μ is simple, the Fredholm alternative tell us that such an S exists (and is unique). By applying Stokes formula, one gets, using in particular (2.1.17), (2.1.18), (2.1.19) and (2.1.21),

$$\int_\Omega v_\eta dx = \int_\Omega \Big((-\Delta - \mu)S\Big) v_\eta dx = \int_{\partial\Omega} \eta \frac{\partial S}{\partial \nu}\frac{\partial w}{\partial \xi} d\sigma. \tag{2.1.22}$$

Let us assume that (2.1.20) does not hold. Then, the right hand side of (2.1.22) should be equal to zero for every $\eta \in C^3(\partial\Omega)$ and, therefore,

$$\frac{\partial S}{\partial \nu}\frac{\partial w}{\partial \xi} \equiv 0.$$

By the Holmgren uniqueness theorem (see e.g. [113, Proposition 4.3, p. 433]), since w is a non zero eigenfunction of $-\Delta_\Omega$, $\partial w/\partial\xi$ cannot be equal to zero on any nonempty open subset of $\partial\Omega$. Therefore, for the previous equation to hold, it results that

$$\frac{\partial S}{\partial \nu} = 0 \text{ on } \partial\Omega. \tag{2.1.23}$$

The following lemma tells us that (2.1.23) cannot hold true (and, therefore, yields (2.1.20)).

Lemma 2.1.2 *With the notations above, there is no solution to the following overdetermined eigenvalue problem*

$$\begin{cases} (-\Delta_\Omega + \mu)S = 1, & \textit{in } \Omega, \\ S = 0, & \textit{on } \partial\Omega, \\ \frac{\partial S}{\partial \nu} = 0, & \textit{on } \partial\Omega. \end{cases} \tag{2.1.24}$$

Proof of Lemma 2.1.2. The result is classical but we provide it for sake of completeness. We argue by contradiction. By differentiating (2.1.24), it follows that ∇S is solution of the following partial differential system,

$$\begin{cases} (-\Delta_\Omega + \mu)\nabla S = 0, & \text{in } \Omega, \\ \nabla S = 0, & \text{on } \partial\Omega. \end{cases} \tag{2.1.25}$$

It implies that there exists a non zero constant vector $a \in \mathbb{R}^n$ such that

$$\nabla S = wa, \text{ in } \overline{\Omega}.$$

Indeed, for every constant vector $z \in \mathbb{R}^n$, $w_z := \nabla S \cdot z$ is solution of the Dirichlet problem

$$\begin{cases} (-\Delta_\Omega + \mu)v = 0, & \text{in } \Omega, \\ v = 0, & \text{on } \partial\Omega. \end{cases}$$

Since μ is a simple eigenvalue and w_z is linear in z, there exists $a \in \mathbb{R}^n$ such that $w_z = (a\cdot z)w$ for all $z \in \mathbb{R}^n$. Finally, a is non zero since S is non constant (see (2.1.21)).

Up to a rotation, one can choose $a = \|a\|(1, \cdots, 0)^{\mathrm{tr}}$. Then, S is only function of the variable x_1, with $S = 0$ on $\partial\Omega$. That clearly implies that $S = 0$ on the intersection of $\overline{\Omega}$ with any hyperplane defined by x_1 constant. Therefore $S = 0$ on $\overline{\Omega}$, contradicting $(-\Delta_\Omega + \mu)S = 1$. □

Finally let

$$E_0 := \left\{(v, \tau) \in H^2(\Omega) \times (-1, 1);\ v(x) + \tau\eta_0(x)\frac{\partial w}{\partial \xi}(x) = 0,\ \forall x \in \partial\Omega\right\},$$

and let $\Phi_0 : E_0 \times \mathbb{R} \to L^2(\Omega) \times \mathbb{R}$ be defined by

$$\Phi_0((v, \tau), \chi) := \Phi((v, \tau\eta_0), \chi).$$

Then, Φ_0 is of class C^1, $\Phi_0((w, 0), \mu) = 0$ and $\Phi_0'((w, 0), \mu)$ is onto. Hence, since $E_0 \times \mathbb{R}$ is an open set of the *Hilbert* space

$$\left(\left\{(v, \tau) \in H^2(\Omega) \times \mathbb{R};\ v(x) + \tau\eta_0(x)\frac{\partial w}{\partial \nu}(x) = 0,\ \forall x \in \partial\Omega\right\}\right) \times \mathbb{R},$$

we get that Φ_0 is locally onto at $((w, 0), \mu)$ and the proof of Lemma 2.1.1 is finished. □

Remark 3 *In order to prove Theorem 2.1.1, one can alternatively use the strategy developed by J. Ortega and E. Zuazua to show the simplicity of, on the one hand, the eigenvalues of a plate equation ([87]) and, on the other hand, those of the Stokes system in two space dimensions ([88]) for generic domains. Note that the situation in [87, 88] is much more complicated than ours since, in [87, 88], one cannot apply the Holmgren uniqueness theorem for every* Ω. *Note also that the simplicity of the eigenvalues has already been used in a control problem by J.-L. Lions and E. Zuazua in [78], but in order to get a controllability result instead of a non-controllability one, as in here.*

2.2 Open sets which are rectangles

We consider the domain $\Omega = (0, a) \times (0, b) \subset \mathbb{R}^2$, where a and b are strictly positive real numbers. Our goal is to show the following theorem.

Theorem 2.2.1 *The initial state* $y_0 \equiv 1$ *cannot be steered to zero in finite time.*

The eigenfunctions and eigenvalues of $-\Delta_\Omega$ are respectively

$$u(x_1, x_2) = K \sin(k_1 \pi x_1/a) \sin(k_2 \pi x_2/b),$$

and

$$\lambda = \frac{k_1^2 \pi^2}{a^2} + \frac{k_2^2 \pi^2}{b^2},$$

where $K \in \mathbb{R}$ and $(k_1, k_2) \in \mathbb{N}^2 \setminus \{(0,0)\}$. Note that

$$\int_\Omega \sin(k_1 \pi x_1/a) \sin(k_2 \pi x_2/b) dx_1 dx_2 \neq 0,$$

if and only if both k_1 and k_2 are odd.

Define $m := a^2/b^2$. Let Σ_0 be the set of eigenvalues of $-\Delta_\Omega$ such that there exists a corresponding eigenfunction w satisfying

$$\int_\Omega w dx \neq 0.$$

Then $\lambda \in \Sigma_0$ if and only if there exist two odd positive integers k_1 and k_2 such that

$$\lambda = \frac{\pi^2}{a^2}(k_1^2 + m k_2^2).$$

Therefore, if $N(R)$ denotes the number of $\lambda \in \Sigma_0$ less than or equal to $R > 0$, then $N(R) = N_0(Ra^2/\pi^2)$ where N_0 is the counting function of the set

$$\mathcal{N}_0(R) := \{Y \leqslant R \mid Y = k_1^2 + m k_2^2 \text{ with } k_1, k_2 \text{ odd integers }\}, \tag{2.2.26}$$

i.e. $N_0(R) := \#\mathcal{N}_0(R)$. Theorem 2.2.1 is a consequence of the following lemma.

Lemma 2.2.1 *With the notations above, there exists $C > 0$ such that, for R large enough,*

$$N_0(R) \geqslant C \frac{R}{\ln(R)}. \tag{2.2.27}$$

Assuming the conclusion of the lemma, let us finish the proof of Theorem 2.2.1. We order the real numbers in Σ_0 to get a strictly increasing sequence $(\tilde{\lambda}_n)_{n \geqslant 1}$. Then, it is clear that $N(\tilde{\lambda}_n) = n$, which implies, by Lemma 2.2.1, that there exists $C' > 0$ such that, for n large enough,

$$\tilde{\lambda}_n \leqslant C' n \ln(n).$$

Therefore, Ω has property (A) and consequently, by Theorem 2.0.1, $y_0 \equiv 1$ cannot be steered to 0 in finite time.

It remains to prove Lemma 2.2.1. Let us first assume that $m \notin \mathbb{Q}$. Then, for every $(k_1, k_2, m_1, m_2) \in \mathbb{N}^4$,

$$\left(k_1^2 + mk_2^2 = l_1^2 + ml_2^2\right) \Rightarrow (k_1 = l_1 \text{ and } k_2 = l_2)\,. \tag{2.2.28}$$

Therefore, for $R > 0$, $N_0(R)$ is equal to the number of pairs (k_1, k_2) of odd integers such that $k_1^2 + mk_2^2 \leqslant R$. As one easily checks, there exists $\delta > 0$ depending only on m such that

$$\#\left\{(k_1, k_2) \in \mathbb{N}^2;\, k_1 \text{ and } k_2 \text{ are odd, } k_1^2 + mk_2^2 \leqslant R\right\} \geqslant \delta R,\ \forall R \geqslant 1. \tag{2.2.29}$$

Then, equation (2.2.27) holds.

We now assume that $m = r/q$, where r, q are positive integers with $g.c.d.(r, q) = 1$. By reducing to the same denominator, we have that $N_0(R) = N_1(qR)$ with $N_1(R) := \#\mathcal{N}_1(R)$, where

$$\mathcal{N}_1(R) := \{Y \leqslant R \mid Y = qk_1^2 + rk_2^2 \text{ with } k_1, k_2 \text{ odd integers}\}. \tag{2.2.30}$$

By possibly exchanging q and r, we may assume that q is odd. We will actually use the asymptotic of another counting function, namely $P(R) := \#\mathcal{P}(R)$, where

$$\mathcal{P}(R) := \{3 \leq p \leqslant R \mid p \text{ prime and } p = qk_1^2 + rk_2^2 \text{ for some } (k_1, k_2) \in \mathbb{N}^2\}.$$

Recall that there exists $C_m > 0$ only depending on m such that

$$P(R) \sim_{R\to\infty} C_m \frac{R}{\ln(R)}, \tag{2.2.31}$$

see [75]. For $R > 0$ large enough, let $\mathcal{S}(R)$ be the set of integers $Y \leqslant q(r+q)R$ such that, either $Y = p$ or $Y = q(r+q)p$, where $p \in [3, R]$ is a prime number with $p = qk_1^2 + rk_2^2$ for some $(k_1, k_2) \in \mathbb{N}^2$. Finally define the map $i : \mathcal{P}(R) \to \mathcal{S}(R)$ as follows. For $p \in \mathcal{P}(R)$, then $i(p) = p$ if there exist two odd integers k_1 and k_2 such that $p = qk_1^2 + rk_2^2$. Otherwise $i(p) = q(r+q)p$. It is obvious that i is an injection.

We claim that the image of $\mathcal{P}(R)$ by i is a subset of $\mathcal{N}_1\Big(q(r+q)R\Big)$. From the definition of i, it simply amounts to show that, for a prime $p =$

$qk_1^2+rk_2^2 \leqslant R$ with k_1 and k_2 integers having different parity, then $q(r+q)p \in \mathcal{N}_1\Big(q(p+q)R\Big)$. The latter simply results from the classical identity

$$q(r+q)(qk_1^2+rk_2^2) = q(qk_1-rk_2)^2 + r[q(k_1+k_2)]^2. \tag{2.2.32}$$

Indeed, since q is odd, then $q(k_1+k_2)$ is, and since $qk_1 - rk_2$ has the same parity as $qk_1^2+rk_2^2 = p$, it is also odd. We deduce that

$$N_1\Big(q(r+q)R\Big) \geqslant P(R), \tag{2.2.33}$$

which implies that (2.2.27) follows from (2.2.31).

Remark 4 *Theorem 2.2.1 obviously extends to any parallelepiped in any dimension $n \geqslant 2$.*

2.3 Open set which are balls.

Let us consider $\Omega = B(0,1)$ in $\mathbb{R}^2$ and the eigenvalue problems for the Laplacian on Ω

$$\begin{cases} -\Delta u(x) = \lambda u(x), & x \in \Omega, \\ u(x) = 0, & x \in \partial\Omega. \end{cases} \tag{2.3.34}$$

Since the domain is a ball, it is natural to consider polar coordinates in $\mathbb{R}^2$ and trying to solve equation (2.3.34) by the separation of variables method. The first equation in (2.3.34) in polar coordinates (ρ,θ) becomes

$$\frac{\partial^2 v}{\partial \rho^2} + \frac{1}{\rho^2}\frac{\partial^2 v}{\partial \theta^2} + \frac{1}{\rho}\frac{\partial v}{\partial \rho} = -\lambda v. \tag{2.3.35}$$

We look for non zero solutions of the type

$$v(\rho,\theta) = M(\rho)Z(\theta),$$

so that (2.3.35) becomes

$$M''(\rho)Z(\theta) + \frac{1}{\rho^2}M(\rho)Z''(\theta) + \frac{1}{\rho}M'(\rho)Z(\theta) = -\lambda M(\rho)Z(\theta). \tag{2.3.36}$$

Clearly, $Z(\theta)$ must be a periodic function of period 2π, otherwise v is not a single–valued function. Moreover $M(1) = 0$ and $M(0)$ is finite.

We start with the following lemma.

Lemma 2.3.1 *If an eigenfunction v is not radial, i.e. $Z(\theta)$ is not constant, then the mean of v on $B(0,1)$ is equal to 0.*

Proof. By equation (2.3.36) we clearly deduce that the function Z must satisfy the equation

$$\ddot{Z}(\theta) = AZ(\theta).$$

If $A = 0$, then $\ddot{Z}(\theta) = 0$ and so

$$Z(\theta) = \alpha\theta + \beta,$$

for some α and β. The periodicity of Z implies that Z is constant and v radial. Thus we have $A \neq 0$ and

$$\int_0^{2\pi} Z(\theta)d\theta = \frac{1}{A}\int_0^{2\pi} \ddot{Z}(\theta)d\theta = \frac{1}{A}\left[\dot{Z}(2\pi) - \dot{Z}(0)\right] = 0. \tag{2.3.37}$$

Therefore

$$\int_\Omega v\, dxdy = \left[\int_0^1 \rho M(\rho)d\rho\right]\left[\int_0^{2\pi} Z(\theta)d\theta\right] = 0,$$

since (2.3.37) holds. □

Now, we look for radial solutions to (2.3.35), that is solutions where Z is constant. The equation (2.3.36) becomes

$$\rho M''(\rho) + M'(\rho) + \lambda\rho M(\rho) = 0, \tag{2.3.38}$$

with the boundary conditions $M(1) = 0$ and $M(0)$ finite. Multiplying the last equation by ρ and applying the transformation $\rho\sqrt{\lambda} = r$, equation (2.3.38) takes the form

$$M''(r) + \frac{1}{r}M'(r) + M(r) = 0.$$

The last equation is called Bessel equation, see [43, 85, 63]. First let us note that the mean of a radial eigenfunction is different from 0.

Lemma 2.3.2 *If v is a non zero radial eigenfunction of (2.3.34), then the mean of v is different from 0.*

Proof. The mean of v on Ω is given by

$$\int_{\Omega} v\,dxdy = \left[\int_0^1 \rho M(\rho)d\rho\right]\left[\int_0^{2\pi} Z(\theta)d\theta\right].$$

Since v is radial, then Z is a constant different from 0. Moreover

$$\int_0^1 \rho M(\rho)d\rho = -\frac{1}{\lambda}\int_0^1 M'(\rho)d\rho - \frac{1}{\lambda}\int_0^1 \rho M''(\rho)d\rho.$$

Integrating by parts the last integral, we obtain that

$$\int_0^1 \rho M''(\rho)d\rho = M'(1) - \int_0^1 M'(\rho)d\rho,$$

and so

$$\int_0^1 \rho M(\rho)d\rho = -\frac{1}{\lambda}M'(1),$$

which is clearly different from 0, otherwise by uniqueness $M \equiv 0$. □

Now we construct an analytic solution to (2.3.38) in order to have a condition on eigenvalues λ.

Lemma 2.3.3 *A solution to (2.3.38) is given by the analytic function*

$$\varphi_1(\rho) := \sum_{n=0}^{\infty} c_{2n}\rho^{2n}, \tag{2.3.39}$$

where the coefficients c_{2n} satisfy

$$c_{2n} = (-1)^n \frac{\lambda^n}{4^n}\frac{1}{(n!)^2}c_0, \tag{2.3.40}$$

for every $n \in \mathbb{N}$. Moreover an other linear independent solution to (2.3.38) has a singularity for $\rho = 0$.

Proof. Substituting (2.3.39) into (2.3.38), it is clear that (2.3.39) is a solution to (2.3.38). Now, let us consider the Wronskian

$$W(\rho) := \varphi_1(\rho)\varphi_2'(\rho) - \varphi_1'(\rho)\varphi_2(\rho),$$

where $\varphi_2(\rho)$ is a solution to (2.3.38) linear independent by φ_1. By simple computations we get that

$$W' = -\frac{1}{\rho} W,$$

which implies that

$$W(\rho) = \frac{K}{\rho},$$

with K constant. Moreover we have that

$$\frac{d}{dt}\left(\frac{\varphi_2}{\varphi_1}\right) = \frac{K}{\rho\varphi_1^2(\rho)},$$

and so

$$\varphi_2(\rho) = K_1\varphi_1(\rho) + K\varphi_1(\rho)\int_*^\rho \frac{1}{s\varphi_1^2(s)} ds,$$

with K_1 constant. When $\rho \to 0^+$ the last integral blows up, which means that φ_2 has a singularity at $\rho = 0$. □

By the previous lemmata, $\varphi_1(1) = 0$ implies that

$$\sum_{n=0}^{\infty}(-1)^n \frac{\lambda^n}{4^n} \frac{1}{(n!)^2} = 0, \tag{2.3.41}$$

since c_0 must be different from 0, otherwise $M \equiv 0$. There are infinitely many λ solution to (2.3.41), see [43]. In fact every λ is the square of a zero of the Bessel function J_0. Moreover the zeroes of J_0 are countable and can be ordered by the sequence

$$\mu_1 < \mu_2 < \ldots < \mu_k < \ldots$$

with the property that $|\mu_k - \mu_{k-1}| \to \pi$ as $k \to +\infty$, see [63]. The conclusion is that in this case Theorem 2.0.1 cannot be applied, since Ω does not satisfy assumption (A). Th conjecture is that for the ball steady–state controllability holds.

Chapter 3

Saint Venant equation.

Let us consider the controllability problem for a tank containing a fluid. As in [91], we consider an open, bounded and connected subset Ω of $\mathbb{R}^2$, which corresponds to the shape of the tank. The mathematical description of this problem is given by the position D in $\mathbb{R}^2$ of the tank and by the height $h(t,x)$ of the fluid respect to an equilibrium position. The control system is modeled by

$$\begin{cases} \ddot{D}(t) = u(t), & \text{if } t \in (0,T), \\ h_{tt}(t,x) = \Delta h(t,x), & \text{if } (t,x) \in (0,T) \times \Omega, \\ \frac{\partial h}{\partial \nu}(t,x) = -u(t) \cdot \nu(x), & \text{if } (t,x) \in (0,T) \times \partial\Omega. \end{cases} \tag{3.0.1}$$

where the control $u(t) \in \mathbb{R}^2$. Here $\nu(x)$ denotes the outward unit normal vector at $x \in \partial\Omega$. The steady–state control problem is the following one. Let D_0 and D_1 be two arbitrary points in $\mathbb{R}^2$, does there exist $T > 0$ and $u : [0,T] \to \mathbb{R}^2$ such that the solution $D : [0,T] \to \mathbb{R}^2$, $h : [0,T] \times \Omega \to \mathbb{R}$ of (3.0.1) with

$$h(0,\cdot) = 0, h_t(0,\cdot) = 0,\ D(0) = D_0,\ \dot{D}(0) = 0, \tag{3.0.2}$$

satisfies

$$D(T) = D_1,\ \dot{D}(T) = 0,\ h(T,\cdot) = h_t(T,\cdot) = 0? \tag{3.0.3}$$

In [91], N. Petit and P. Rouchon proved that, if the shape Ω of the tank is a rectangle or a circle, then there is a solution to this controllability problem. When Ω has a general form, they address the issue as an open problem. Here, in the spirit of the previous chapter, we propose a necessary condition for that steady–state controllability concerning the behavior of eigenvalues and eigenfunctions of a Neumann problem.

Let us fix $\Omega \subseteq \mathbb{R}^2$ a bounded, open and connected subset of $\mathbb{R}^2$ of class C^2 or a convex polygon. Let us first recall some classical result about weak solution the following Cauchy problem

$$\begin{cases} \ddot{D}(t) = u(t), & \text{if } t \in (0,T), \\ \dot{D}(0) = s_0, & \\ D(0) = D_0, & \\ h_{tt}(t,x) = \Delta h(t,x), & \text{if } (t,x) \in (0,T) \times \Omega, \\ \frac{\partial h}{\partial \nu}(t,x) = -u(t) \cdot \nu(x), & \text{if } (t,x) \in (0,T) \times \partial\Omega, \\ h(0,x) = h_0(x), & \text{if } x \in \Omega, \\ h_t(0,x) = v_0(x), & \text{if } x \in \Omega. \end{cases} \tag{3.0.4}$$

Define

$$H := \{h \in L^2(\Omega); \int_\Omega h dx = 0\},$$
$$V := \{h \in H^1(\Omega); \int_\Omega h dx = 0\},$$

and let V' be the dual space of $V \subset H$. Let $D_0 \in \mathbb{R}^2$, $s_0 \in \mathbb{R}^2$, $(h_0, v_0) \in H \times V'$, $T > 0$ and $u \in L^2(0,T;\mathbb{R}^2)$. A weak solution of the Cauchy problem (3.0.4) is a couple (D,h) such that

$$D \in H^2(0,T;\mathbb{R}^2),\ D(0) = D_0, \dot{D}(0) = s_0,\ \ddot{D} = u \in L^2(0,T), \tag{3.0.5}$$
$$h \in C^0([0,T];H) \cap C^1([0,T];V'), \tag{3.0.6}$$

and such that, for every $\tau \in [0,T]$ and for every $\theta \in C^0([0,T];H^2(\Omega)) \cap C^1([0,T];H^1(\Omega)) \cap C^2([0,T];L^2(\Omega))$ satisfying

$$\theta_{tt} = \Delta\theta, \quad \text{in } C^0([0,T];L^2(\Omega)), \tag{3.0.7}$$
$$\frac{\partial\theta}{\partial\nu} = 0, \quad \text{in } C^0([0,T];H^{1/2}(\partial\Omega)), \tag{3.0.8}$$

one has

$$\begin{aligned} -\int_0^\tau \int_{\partial\Omega} \theta(t,x) u(t) \cdot \nu(x) d\sigma(x) dt + \langle v_0, \theta(0,\cdot)\rangle_{V',V} - \int_\Omega h_0(x)\theta_t(0,x)dx \\ = \langle h_t(\tau,\cdot), \theta(\tau,\cdot)\rangle_{V',V} - \int_\Omega h(\tau,x)\theta_t(\tau,x)dx. \end{aligned} \tag{3.0.9}$$

Of course, for every $D \in H^2(0,T)$ and every $h \in C^0([0,T];H^2(\Omega)) \cap C^1([0,T];H^1(\Omega)) \cap C^2([0,T];L^2(\Omega))$, if (D,h) is a classical solution of (3.0.4), then it is also a weak solution of (3.0.4). Moreover, it is well known that, for every $(D_0,s_0) \in \mathbb{R}^2 \times \mathbb{R}^2$, $(h_0,v_0) \in H \times V'$, $T > 0$ and $u \in L^2(0,T;\mathbb{R}^2)$, there exists one and only one weak solution (D,h) to (3.0.4). This unique (D,h) will be called the solution to the Cauchy problem (3.0.4).

We say that the control system (3.0.1) is steady–state controllable if, for every $(D_0,D_1) \in \mathbb{R}^2 \times \mathbb{R}^2$, there exist $T > 0$ and $u \in L^2(0,T;\mathbb{R}^2)$ with $u(0) = 0$ such that the solution to the Cauchy problem (3.0.4), with $h_0 = v_0 = 0$, $s_0 = 0$, satisfies (3.0.3).

Consider the Laplace–Neumann operator $-\Delta_\Omega^N$ defined as follows:

$$\mathcal{D}\left(-\Delta_\Omega^N\right) := \left\{ v \in H^2(\Omega); \frac{\partial v}{\partial \nu} = 0 \text{ on } \partial\Omega \right\},$$
$$-\Delta_\Omega^N v = -\Delta v, \; \forall v \in \mathcal{D}\left(-\Delta_\Omega^N\right).$$

3.1 Obstructions to controllability.

We next provide property (B), which will turn out to be an obstruction for the steady–state controllability in finite time.

Definition 3.1.1 *The open set Ω has the property (B) if there exists a sequence $(\lambda_k)_{k\in\mathbb{N}^*}$ of distinct eigenvalues of $-\Delta_\Omega^N$ such that*

(i) There exist $\rho \in (0,2)$ and $C > 0$ such that

$$\lambda_k \leqslant C k^\rho, \; \forall k \geqslant 1, \tag{3.1.10}$$

(ii) For every $k \in \mathbb{N}^*$, there exists an eigenfunction w_k for the eigenvalue λ_k and the operator $-\Delta_\Omega^N$ such that

$$\int_{\partial\Omega} w_k \nu dx \neq 0. \tag{3.1.11}$$

We are now able to state the main result of this paragraph.

Theorem 3.1.1 *Let Ω be a bounded, open and non empty subset of $\mathbb{R}^2$. If Ω has property (B), then the control system (3.0.1) is not steady–state controllable.*

Remark 5 *The previous theorem is, in a sense, optimal. Indeed, if Ω is equal to the disc or the rectangle, then steady-state controllability holds true and condition (B) too, except for (i), where ρ is equal to 2.*

Proof. Let us first consider $u \in L^2(0,T;\mathbb{R}^2)$ and let (D,h) be the solution of the Cauchy problem (3.0.4), with

$$D_0 := 0,\ s_0 := 0,\ h_0 := 0,\ v_0 := 0. \tag{3.1.12}$$

We assume that

$$h(T,\cdot) = 0,\ h_t(T,\cdot) = 0. \tag{3.1.13}$$

Let λ be an eigenvalue of $-\Delta_\Omega^N$ and w be an eigenfunction associated to that eigenvalue λ. Let $\theta \in C^\infty([0,T], H^2(\Omega))$ be defined by

$$\theta(t,x) = e^{i\sqrt{\lambda}t} w(x).$$

Then θ satisfies (3.0.7) and (3.0.8). Hence, using (3.0.9) with $\tau = T$, (3.1.12) and (3.1.13), one gets

$$C(i\sqrt{\lambda}) \cdot \int_{\partial\Omega} w(x)\nu(x)d\sigma(x) = 0, \tag{3.1.14}$$

where $C : \mathbb{C} \to \mathbb{C}^2$ is the holomorphic function defined by

$$C(s) := \int_0^T u(t)e^{st}dt. \tag{3.1.15}$$

The proof of Theorem 3.1.1 goes now by contradiction. We suppose that property (B) holds and assume, by contradiction, that the control system (3.0.1) is steady–state controllable. Then, for every $q \in \mathbb{R}^2$, there exists $u(t) \in L^2(0,T;\mathbb{R}^2)$ such that the solution (D,h) to the Cauchy problem (3.0.4), with $D_0 := 0$, $s_0 := 0$, $h_0 := 0$ and $v_0 := 0$, satisfies

$$h(T,\cdot) = 0,\ h_t(T,\cdot) = 0,\ D(T) = q, \dot{D}(T) = 0. \tag{3.1.16}$$

We use u^1, D^1 and u^2, D^2 to denote u, D, for $q := (1,0)^{\mathrm{T}}$ and $q := (0,1)^{\mathrm{T}}$ respectively. Similarly, $C^1 := (C_1^1, C_2^1)^{\mathrm{T}}$ and $C^2 := (C_1^2, C_2^2)^{\mathrm{T}}$ are defined by (see (3.1.15))

$$C^1(s) := \int_0^T u^1(t)e^{st}dt,\ C^2(s) := \int_0^T u^2(t)e^{st}dt.$$

We will derive a contradiction from the existence of both u^1 and u^2. Let $(\lambda_k)_{k\in\mathbb{N}^*}$ and $(w_k)_{k\in\mathbb{N}^*}$ be as in Definition 3.1.1. For $k \in \mathbb{N}^*$, define

$$t_k := \sqrt{\lambda_k},\ v_k := \int_{\partial\Omega} w_k(x)\nu(x)d\sigma(x) \in \mathbb{R}^2 \setminus \{0\}.$$

By (3.1.14), we have, for $j = 1, 2$ and $k \geqslant 1$,

$$C^j(it_k) \cdot v_k = 0. \tag{3.1.17}$$

For every $k \geqslant 1$, v_k is a non zero vector of $\mathbb{C}^2$. Therefore $C^1(it_k)$ and $C^2(it_k)$ must be collinear. We deduce that, for every $k \geqslant 1$,

$$C_1^1(it_k)C_2^2(it_k) - C_2^1(it_k)C_1^2(it_k) = 0. \tag{3.1.18}$$

Introducing the holomorphic function $G : \mathbb{C} \to \mathbb{C}$, $G(s) := C_1^1(s)C_2^2(s) - C_2^1(s)C_1^2(s)$, we reformulate (3.1.18) as

$$G(it_k) = 0. \tag{3.1.19}$$

Let us recall the following classical result.

Lemma 3.1.1 *Let $f : \mathbb{C} \to \mathbb{C}$ be a holomorphic function such that*

$$\exists C_0 > 0,\ \textit{such that}\ \forall s \in \mathbb{C},\ |f(s)| \leqslant C_0 e^{C_0|s|}. \tag{3.1.20}$$

Assume that $f \neq 0$. Let $n : [0, +\infty) \to \mathbb{N}$ be defined by

$$n(R) := \#\{s \in \mathbb{C};\ f(s) = 0\ \textit{and}\ |s| \leqslant R\}.$$

Then,

$$\exists C_1 > 0, \forall R \in (1, +\infty), \int_1^R \frac{n(t)}{t} dt \leqslant C_1 R. \tag{3.1.21}$$

(This lemma follows easily from the Jensen formula, see e.g. [76, Lecture 2, section 2.3, p. 10-11].) We apply this lemma with $f := G$. By the easy part of the Paley-Wiener theorem, G is a holomorphic function which satisfies (3.1.20). By (3.1.10) and (3.1.19), (3.1.21) does not hold. Hence, by Lemma 3.1.1, $G = 0$. On the other hand, we compute $G(s)$ for s small enough. Simple integrations by parts yield, for $j, l = 1, 2$ and $s \in \mathbb{C}^2$,

$$C_l^j(s) = -sD_l^j(T)e^{sT} + s^2 \int_0^T D_l^j(t)e^{st} dt. \tag{3.1.22}$$

As s goes to 0, the previous equation can be written $C_l^j(s) = -sD_l^j(T)e^{sT} + O(s^2)$ and then,

$$G(s) = s^2 e^{2sT}\Big(D_1^1(T)D_2^2(T) - D_2^1(T)D_1^2(T)\Big) + O(s^3) = s^2 e^{2sT} + O(s^3),$$

which implies that $G(s) \neq 0$ for s small enough but nonzero. That contradicts the fact that G is the zero function. Theorem 3.1.1 is proved. □

3.2 Genericity of condition (B).

In this section, we prove that condition (B) holds generically for tank shapes of class C^3, and therefore by Theorem 3.1.1, for such generic tank shapes Ω, steady-state controllability for a water-tank does not hold.

We use here notations and results of [106]. Let $\mathcal{S}_3$ be the set of all non empty open, bounded, connected subsets $\Omega \in \mathbb{R}^2$ of class C^3. The topology on $\mathcal{S}_3$ is defined as follows ([106, p. 7]).

Let $C^3(\mathbb{R}^2)$ be the space of functions $u : \mathbb{R}^2 \to \mathbb{R}^2$ of class C^3. Then $C^3(\mathbb{R}^2)$ equipped with the standard sup norm $\|\cdot\|_3$ is a Banach space. For $\Omega \in \mathcal{S}_3$ and $u \in C^3(\mathbb{R}^2)$, let $\Omega + u := (Id + u)(\Omega)$ be the subset of points $y \in \mathbb{R}^2$ such that $y = x + u(x)$ for some $x \in \Omega$. By simple topological arguments, one easily gets that, for $u \in C^3(\mathbb{R}^2)$ small enough, $\Omega + u$ belongs to $\mathcal{S}_3$.

For $\varepsilon > 0$, let $\mathcal{V}(\varepsilon)$ be the sets of all the $\Omega + u$ with $u \in C^3(\mathbb{R}^2)$ and $\|u\|_3 < \varepsilon$. The topology on $\mathcal{S}_3$ is defined by considering the sets $\mathcal{V}(\varepsilon)$ with ε small enough as a base of neighborhoods of Ω. Then $\mathcal{S}_3$ is a Baire space.

Theorem 3.2.1 *Condition (B) holds for generic $\Omega \in \mathcal{S}_3$.*

Proof. The strategy is entirely similar to that of the argument of Theorem 3.1.1. Let $\mathcal{H} \subset \mathcal{S}_3$ be the set of $\Omega \in \mathcal{S}_3$ such that

(a) all eigenvalues of $-\Delta_\Omega^N$ and $-\Delta_\Omega^D$ are simple;

(b) $\int_{\partial\Omega} w\nu dx \neq 0$, for every non zero eigenfunction w of $-\Delta_\Omega^N$ corresponding to a nonzero eigenvalue.

For $l \geq 1$, define the set $\mathcal{K}_l \in \mathcal{S}_3$ (respectively $\mathcal{H}_l \in \mathcal{S}_3$) of open sets $\Omega \in \mathcal{S}_3$ such that property (a) (respectively, and property (b)) holds at least for the first l eigenvalues of $-\Delta_\Omega^D$ and $-\Delta_\Omega^N$. Then, $\mathcal{H}$ is the countable intersection of the $\mathcal{H}_l$'s.

We show next that $\mathcal{H}$ is residual, which implies Theorem 3.2.1. Similarly to the proof of Theorem 3.1.1, that follows by using a Weyl formula for $-\Delta_\Omega^N$, namely $\lambda_k \sim_{k\to\infty} C(\Omega)k$, (cf. [67, formula (10.2.40), page 505]).

For $l \geqslant 0$, $\mathcal{K}_0 = \mathcal{H}_0 := \mathcal{S}_3$, $\mathcal{H}_l \subset \mathcal{K}_l$, $\mathcal{K}_{l+1} \subset \mathcal{K}_l$ and $\mathcal{K} := \cap_{l\geqslant 0}\mathcal{K}_l$ and, similarly, $\mathcal{H}_{l+1} \subset \mathcal{H}_l$ and $\mathcal{H} = \cap_{l\geqslant 0}\mathcal{H}_l$. Moreover, for $l \geqslant 0$, it is clear that the sets $\mathcal{K}_l$ and $\mathcal{H}_l$ are open in $\mathcal{S}_3$. To show that $\mathcal{H}$ is residual, amounts to establish the next lemma.

Lemma 3.2.1 *For every $l \geqslant 0$, $\mathcal{H}_{l+1}$ is dense in $\mathcal{S}_3$.*

Proof of Lemma 3.2.1.

The argument follows the lines of that of Lemma 2.1.1. Recall that, for $l \geqslant 0$, $\mathcal{K}_l$ is dense in $\mathcal{S}_3$ (which is a trivial consequence of Theorem 3.2 in [89]). In particular, $\mathcal{H}_1$, which is equal to $\mathcal{K}_1$, is dense in $\mathcal{S}_3$. Here, we may assume that $l \geq 1$.

Let $\Omega \in \mathcal{H}_l$. It is sufficient to exhibit $\Omega' \in \mathcal{H}_{l+1}$, arbitrarily close to Ω. Since $\mathcal{K}_{l+1}$ is dense, it is enough to establish the previous fact for $\Omega \in \mathcal{H}_l \cap \mathcal{K}_{l+1}$. Let $(\mu_k)_{k\in\mathbb{N}^*}$ be the ordered sequence of the eigenvalues of $-\Delta_\Omega^D$ repeated according to their multiplicity, and similarly, let $(\lambda_k)_{k\in\mathbb{N}^*}$ be the ordered sequence of the eigenvalues of $-\Delta_\Omega^N$ repeated according to their multiplicity. We have

$$\mu_1 < \mu_2 < \cdots \mu_l < \mu_{l+1} < \mu_{l+2} \leqslant \mu_{l+3} \leqslant \cdots ,$$

and

$$0 = \lambda_1 < \lambda_2 < \cdots \lambda_l < \lambda_{l+1} < \lambda_{l+2} \leqslant \lambda_{l+3} \leqslant \cdots .$$

Let w_{l+1} be a nonzero eigenfunction of $-\Delta_\Omega^N$ for the eigenvalue $\lambda_{l+1} > 0$. If $\int_{\partial\Omega} w_{l+1}\nu dx \neq 0$ then $\Omega \in \mathcal{H}_{l+1}$. Therefore, we may assume that

$$\int_{\partial\Omega} w_{l+1}\nu dx = 0, \tag{3.2.23}$$

and we simply use λ and w to denote λ_{l+1} and w_{l+1}.

Consider the set E of couples $(v, u) \in H^2(\Omega) \times C^3(\mathbb{R}^2)$ such that, on $\partial\Omega$,

$$\frac{\partial v}{\partial \nu} + (u \cdot \nu)\frac{\partial^2 w}{\partial \nu^2} - \frac{\partial (u \cdot \nu)}{\partial \tau}\frac{\partial w}{\partial \tau} = 0,$$

where $\frac{\partial}{\partial\tau}$ denotes the tangential derivative at $x \in \partial\Omega$. Let $\Phi : E \times \mathbb{R} \to L^2(\Omega) \times \mathbb{R}^2$ be the map defined by

$$\Phi(v, u, \chi) := \Big((-\Delta - \chi)v, \int_{\partial\Omega} v\nu dx\Big).$$

One has $\Phi((w, 0), \lambda) = (0, 0)$ and Lemma 3.2.1 holds if Φ is locally onto at $((w, 0), \lambda)$. The map Φ is of class C^1 and one has

$$\Phi'((w, 0), \lambda)((v, u), \chi) = (-\Delta v - \mu v - \chi w, \int_{\partial\Omega} v\nu dx),$$

for every $(v, u) \in E$. To see that, we use equation (1.21) page 33 of [106]. Using the Fredholm alternative (recall that the eigenvalue λ is assumed to be simple), one easily checks that, for every $f \in L^2(\Omega)$ and every $u \in C^3(\mathbb{R}^2)$, there exists one and only one $(v, \chi) \in H^2(\Omega) \times \mathbb{R}$ such that

$$-\Delta v - \lambda v - \chi w = f, \tag{3.2.24}$$

$$\int_\Omega vw dx = 0, \tag{3.2.25}$$

$$(v, u) \in E. \tag{3.2.26}$$

For $f = 0$, let us denote by (v_u, χ_u) the corresponding unique solution. We next prove that

$$\text{there exists } u_0 \in C^3(\mathbb{R}^2) \text{ such that } \int_{\partial\Omega} v_{u_0}\nu dx \neq 0. \tag{3.2.27}$$

To compute $\int_{\partial\Omega} v_u \nu dx$ in terms of u, we consider the unique solution of the inhomogeneous Neumann problem given by

$$\begin{cases} (-\Delta - \lambda)S = 0, & \text{in } \Omega, \\ \frac{\partial S}{\partial\nu} = \nu, & \text{on } \partial\Omega, \\ \int_\Omega Sw dx = 0. \end{cases} \tag{3.2.28}$$

Since $\int_{\partial\Omega} w\nu dx = 0$ and the eigenvalue λ is simple, the Fredholm alternative tells us that such an S exists (and is unique). Then, for every $u \in C^3(\mathbb{R}^2)$, one has, by using (3.2.24), (3.2.26) and (3.2.28), applying Stokes formula and performing an integration by parts,

$$\int_{\partial\Omega} v_u \nu dx = \int_{\partial\Omega} v_u \frac{\partial S}{\partial\nu} dx = \int_{\partial\Omega} (u \cdot \nu)\Big(\lambda w S - \frac{\partial w}{\partial\tau}\frac{\partial S}{\partial\tau}\Big). \tag{3.2.29}$$

Let us assume that (3.2.27) does not hold. Then, the right hand side of (3.2.29) should be equal to zero for every $u \in C^3(\mathbb{R}^2)$ and therefore, the following holds on $\partial\Omega$,

$$\lambda w S - \frac{\partial w}{\partial \tau}\frac{\partial S}{\partial \tau} = 0. \tag{3.2.30}$$

The following lemma tells us that an S subject to (3.2.28) and (3.2.30) does not exist (and, therefore, yields (3.2.27)).

Lemma 3.2.2 *With the above notations, there is no solution to the following overdetermined eigenvalue problem*

$$\begin{cases} (-\Delta - \lambda)S = 0, & \textit{in } \Omega, \\ \frac{\partial S}{\partial \nu} = \nu, & \textit{on } \partial\Omega, \\ \lambda w S - \frac{\partial w}{\partial \tau}\frac{\partial S}{\partial \tau} = 0, & \textit{on } \partial\Omega. \end{cases} \tag{3.2.31}$$

Proof of Lemma 3.2.2. We will show later that the constraints imposed on S actually imply that $S = 0$ on $\partial\Omega$. Assuming that fact, we next conclude. For every vector $a \in \mathbb{R}^2$, $S \cdot a$ is solution of the Dirichlet problem

$$\begin{cases} (-\Delta - \lambda)y = 0, & \text{in } \Omega, \\ y = 0, & \text{on } \partial\Omega. \end{cases} \tag{3.2.32}$$

Since $S \neq 0$, then there exists $a \in \mathbb{R}^2$ such that $S \cdot a$ is a nonzero solution of (3.2.32). Therefore λ is also an eigenvalue of $-\Delta_\Omega^D$. As in the proof of Lemma 2.1.2, one gets that there exists a nonzero vector $a \in \mathbb{R}^2$ such that $S = az$, where z is a nonzero eigenfunction of $-\Delta_\Omega^D$ associated to λ. Up to a rotation, one concludes that $S = (\|a\|z, 0))^{\mathrm{T}}$, and then ν is constant on each connected component of $\partial\Omega$ and equal to $(1,0)^{\mathrm{T}}$ or $-(1,0)^{\mathrm{T}}$. A simple topological argument yields a contradiction with the boundedness of Ω.

It remains to show that, with the hypothesis of Lemma 3.2.2, then $S = 0$ on $\partial\Omega$. Since $\partial\Omega$ is the finite union of regular Jordan curves $\Gamma_1, \cdots, \Gamma_m$, it is enough to show the claim on (let say) Γ_1. We only use the fact that

$$\lambda w S - \frac{\partial w}{\partial \tau}\frac{\partial S}{\partial \tau} = 0, \tag{3.2.33}$$

on Γ_1. Modifying λ if necessary, equation (3.2.33) simply reduces to

$$\lambda w S = w' S',$$

where w and S are seen as C^1 functions on the unit circle S^1.

We first show the following: if $f : S^1 \to \mathbb{R}$ is of class C^1 such that

$$\lambda w f = w' f',$$

then $f \equiv 0$. Indeed, consider $H := f\frac{\lambda w^2}{2}$. Note that H is of class C^1, $H = \frac{ww'f'}{2}$ and, by using the previous equation,

$$H' = f'\Big((w')^2 + \frac{\lambda}{2}w^2\Big). \tag{3.2.34}$$

We first claim that H must be identically equal to zero. Reasoning by contradiction, there exists some $\theta_0 \in S^1$ such that $H'(\theta_0) = 0$ and $H(\theta_0) \neq 0$. By (3.2.34), one has $f'(\theta_0) = 0$ or $w'(\theta_0) = w(\theta_0) = 0$, since $\lambda > 0$. In both cases, that contradicts $H(\theta_0) \neq 0$. We get that $fw \equiv 0$. By the Holmgren uniqueness theorem, w cannot be equal to zero on any nonempty open subset of $\partial\Omega$ since w is a non zero eigenfunction of $-\Delta_\Omega^N$. We conclude that $f \equiv 0$ on S^1.

Applying the previous result to $S \cdot (1,0)^{\mathrm{T}}$ and $S \cdot (0,1)^{\mathrm{T}}$, we conclude that $S \equiv 0$ on $\partial\Omega$. The proof of Lemma 3.2.2 is finished. □

Chapter 4

Burgers viscous equation.

This chapter deals with the exact global controllability of a Burgers viscous equation. Let $\Omega = [0, 1]$ and $T > 0$. For $y_0 \in L^2(\Omega)$, we consider the Burgers viscous equation

$$\begin{cases} y_t(t,x) - y_{xx}(t,x) + 2y(t,x)y_x(t,x) = 0, & \text{if } (t,x) \in (0,T) \times \Omega, \\ y(0,x) = y_0(x), & \text{if } x \in \Omega, \\ y(t,0) = u_0(t), & \text{if } t \in (0,T), \\ y(t,1) = u_1(t), & \text{if } t \in (0,T), \end{cases} \tag{4.0.1}$$

where u_0, u_1, the controls, belong to $L^2(0, T)$. If the control acts only on one side, i.e. if either $u_0 \equiv 0$ or $u_1 \equiv 0$, then the approximated controllability in $L^2(0, 1)$ does not hold, see [42, 48, 56].

Due to smoothing property of parabolic equations, we can not reach every element of $L^2(0, 1)$ at time $t = T$. The correct definition of exact controllability is that given by Fursikov and Imanuvilov in [56, 57], i.e. passing from one solution to another solution. Here some partial results about global exact controllability are proved, even if the general case remains open. In particular we show that it is possible to drive the system from 0 to an arbitrary big constant, provided the time T is sufficiently big. Moreover we prove that it is also possible to drive the system from an arbitrary initial data in $W^{1,2}(0, 1)$ to 0 if the time T is sufficiently big. These results are based on the local exact controllability, proved by Fursikov and Imanuvilov in [56], and on Carleman estimates.

4.1 Controllability at big constants.

The following theorem holds.

Theorem 4.1.1 *If $y_0(x) = 0$ for every $x \in \Omega$, then there exists a time $\bar{T}$ and a constant $C_0 > 0$ such that, for every $T \geqslant \bar{T}$ and for every $C \geqslant C_0$, there exists a solution $y(t,x)$ to (4.0.1) such that $y(T,x) = C$.*

The proof of this theorem is in four steps.

First step: local exact controllability around 0**.**
Let us consider time $0 < T_0 < \frac{T}{3}$. We perform a change of variables, in order to transform the first equation in (4.0.1) into an ordinary differential equation. More precisely, we look for a solution of the type

$$z_1(t,x) = \varphi(x - Ct), \tag{4.1.2}$$

where $\varphi : \mathbb{R} \to \mathbb{R}$ is a function to be determined and C is a real positive constant. Substituting (4.1.2) into (4.0.1) we obtain that φ satisfies the ordinary differential equation

$$\left(-C\varphi + \varphi^2 - \varphi'\right)' = 0.$$

Integrating the last expression and using (4.1.2), we obtain that

$$z_1(t,x) = \frac{C}{1 + e^{C(x - Ct + \eta)}}, \tag{4.1.3}$$

with $\eta > 0$. The claim is that $z_1(0,x)$ converges to the zero function in the Sobolev space $W^{1,2}(0,1)$ as $C \to +\infty$. Indeed

$$\begin{aligned} \int_0^1 |z_1(0,x)|^2 dx &= \int_0^1 \frac{C^2}{\left[1 + e^{C(x+\eta)}\right]^2} dx = \\ = C\left[\log\left(e^C \frac{1 + e^{C\eta}}{1 + e^{C(1+\eta)}}\right)\right. &+ \left.\frac{e^{C\eta} - e^{C(1+\eta)}}{(1 + e^{C(1+\eta)})(1 + e^{C\eta})}\right] \to 0 \end{aligned}$$

as $C \to +\infty$. Moreover

$$\begin{aligned} \int_0^1 \left|\frac{\partial}{\partial x} z_1(0,x)\right|^2 dx &= C^4 \int_0^1 \frac{e^{2C(x+\eta)}}{\left[1 + e^{C(x+\eta)}\right]^4} dx = \\ -\frac{C^3}{2(1 + e^{C(\eta+1)})^2} + \frac{C^3}{3(1 + e^{C(\eta+1)})^3} &+ \frac{C^3}{2(1 + e^{C\eta})^2} - \frac{C^3}{3(1 + e^{C\eta})^3} \to 0 \end{aligned}$$

as $C \to +\infty$. Now, using the local exact controllability theorem in [56, Theorem 5.1], we conclude that there exists a solution y to (4.0.1) such that $y(0,x) = 0$ for every $x \in [0,1]$ and $y(T_0,x) = z_1(T_0,x)$ for every $x \in [0,1]$.

Second step: distance between z_1 and C.

We fix a time T_1 such that $\frac{2T}{3} < T_1 < T$. We evaluate the distance between the function $z_1(T_1,\cdot)$ and the constant function C in the Sobolev space $W^{1,2}(\Omega)$. Using (4.1.3), we easily get

$$\begin{aligned}\int_0^1 |z_1(T_1,x) - C|^2 dx &= C \log \frac{1 + e^{C(1+\eta - CT_1)}}{1 + e^{C(\eta - CT_1)}} + \\ &\quad + C \frac{e^{C(\eta - CT_1)} - e^{C(1+\eta - CT_1)}}{(1 + e^{C(1+\eta - CT_1)})(1 + e^{C(\eta - CT_1)})}\end{aligned}$$

and

$$\begin{aligned}\int_0^1 |\frac{\partial}{\partial x} z_1(T_1,x)|^2 dx = C^3 \frac{e^{3C(\eta - CT_1)} + 3e^{C(\eta - CT_1)} + 3e^{2C(\eta - CT_1)}}{3(1 + e^{C(1+\eta - CT_1)})^3 (1 + e^{C(\eta - CT_1)})^3} \\ + C^3 \frac{2e^{C(1+\eta - CT_1)} + e^{2C(1+\eta - CT_1)} - 2e^{C(\eta - CT_1)} - e^{2C(\eta - CT_1)}}{2(1 + e^{C(1+\eta - CT_1)})^2 (1 + e^{C(\eta - CT_1)})^2} \\ - C^3 \frac{e^{3C(1+\eta - CT_1)} + 3e^{C(1+\eta - CT_1)} + 3e^{2C(1+\eta - CT_1)}}{3(1 + e^{C(1+\eta - CT_1)})^3 (1 + e^{C(\eta - CT_1)})^3}.\end{aligned}$$

Therefore, we obtain that

$$\|z_1(T_1,\cdot) - C\|^2_{W^{1,2}(0,1)} \simeq 2C^3 T_1 e^{C(1+\eta - CT_1)} \tag{4.1.4}$$

when C is big enough. The idea is to apply again Theorem 5.1 in [56] in order to find a function $y(t,x)$ satisfying the first equation in (4.0.1) and such that $y(T_1,x) = z_1(T_1,x)$ for every $x \in \Omega$ and $y(T,x) = C$ for every $x \in \Omega$. If this is possible, then the theorem is proved. Unfortunately Theorem 5.1 in [56] is only a local controllability theorem. Therefore we need to estimate the radius of controllability around the constant C.

Third step: linearization and estimations.

Applying the change of variable $z(t,x) = y(t,x) - C$, the first equation in (4.0.1) becomes

$$z_t(t,x) - z_{xx}(t,x) + 2(z(t,x) + C) z_x(t,x) = 0$$

for every $(t,x) \in (0,T) \times \Omega$. We focus the attention in particular to the time interval (T_1,T). We linearize the last equation and we obtain that

$$z_t(t,x) - z_{xx}(t,x) + 2Cz_x(t,x) + (\bar{y}(t,x)z(t,x))_x = 0 \tag{4.1.5}$$

for every $(t,x) \in (T_1,T) \times \Omega$. By [56, Theorem 4.3], we know that, for every $\bar{y} \in W_2^{1,2}((T_1,T) \times (0,1))$ and for every $y_0 \in W^{1,2}(0,1)$, there exists a solution $z(t,x) \in W_2^{1,2}((T_1,T) \times (0,1))$ to (4.1.5) such that $z(T_1,x) = y_0(x)$ and $z(T,x) = 0$ for every $x \in \Omega$. Therefore, for every $y_0 \in W^{1,2}(\Omega)$, there exists a function

$$\theta : W_2^{1,2}((T_1,T) \times \Omega) \to W_2^{1,2}((T_1,T) \times \Omega) \tag{4.1.6}$$

which maps $\bar{y}$ to a solution z to (4.1.5) such that $z(T_1,x) = y_0(x)$ and $z(T,x) = 0$ for every $x \in \Omega$.

Fursikov and Imanuvilov in [56] proved that θ is a compact function and that there exists a positive constant $r > 0$, depending by C, such that, if $\|y_0\|_{W^{1,2}(0,1)} \leqslant r$, then the map θ has a fixed point. We evaluate now the dependence of r by C.

We take a bigger domain $\widetilde{Q} := (T_1,T) \times (-2,2)$. There exists a linear and continuous operator $E : W^{1,2}(\Omega) \to W^{1,2}(-2,2)$ such that $E\zeta(-2) = E\zeta(2) = 0$ and $E\zeta(x) = \zeta(x)$ for every $x \in \Omega$ and for every $\zeta \in W^{1,2}(\Omega)$. So, in this way, we extend the function y_0 on the interval $(-2,2)$ (for simplicity we denote Ey_0 with y_0 itself) such that $y_0 \in W^{1,2}(-2,2)$, $y_0(-2) = y_0(2) = 0$ and

$$\|y_0\|_{W^{1,2}(-2,2)} \leqslant k\|y_0\|_{W^{1,2}(0,1)}$$

with $k > 0$ not depending by y_0. In the same way we extend $\bar{y}$ to a function belonging to $W_2^{1,2}(\widetilde{Q})$ such that

$$\|\bar{y}\|_{W_2^{1,2}(\widetilde{Q})} \leqslant k\|\bar{y}\|_{W_2^{1,2}(Q)}.$$

We recall that $W_2^{1,2}(\widetilde{Q}) \hookrightarrow L^\infty(\widetilde{Q})$ and so there exists a constant, say k, not depending by $\bar{y}$ such that

$$\|\bar{y}\|_{L^\infty(\widetilde{Q})} \leqslant k\|\bar{y}\|_{W_2^{1,2}(\widetilde{Q})}.$$

We choose $\bar{y}$ in $\overline{B_{W_2^{1,2}(\widetilde{Q})}(\alpha)}$ where $\alpha < \frac{1}{k}\min\{1, \frac{1}{T-T_1}\}$. In this way we may consider equation (4.1.5) in $\widetilde{Q}$ and we may suppose that $y(T_1,x) = y_0(x)$ for every $x \in (-2, 2)$. Let $\chi(t, x)$ be the solution to (4.1.5) such that $\chi(T_1, x) = y_0(x)$ for every $x \in (-2, 2)$ and $\chi(t, -2) = \chi(t, 2) = 0$ for every $t \in (T_1, T)$. Fix $\varphi(t) \in C^\infty(T_1, T)$ such that $\varphi(t) = 1$ for every $t \in (T_1, \frac{2T_1+T}{3})$ and $\varphi(t) = 0$ for every $t \in (\frac{T_1+2T}{3}, T)$. If we denote with $\hat{y}(t, x) = \varphi(t)\chi(t, x)$, then we obtain that

$$\hat{y}_t(t, x) - \hat{y}_{xx}(t, x) + 2C\hat{y}_x(t, x) + (\bar{y}(t, x)\hat{y}(t, x))_x = -f_0(t, x) := \varphi'(t)\chi(t, x) \tag{4.1.7}$$

for every $(t, x) \in (T_1, T) \times (-2, 2)$. Notice that the support of the function f_0 is contained in the set $\left(\frac{2T_1+T}{3}, \frac{T_1+2T}{3}\right) \times (-2, 2)$. We have:

$$\begin{aligned} \|f_0\|^2_{L^2(\widetilde{Q})} &= \|\hat{y}_t - \hat{y}_{xx} + 2C\hat{y}_x + (\bar{y}\hat{y})_x\|^2_{L^2(\widetilde{Q})} \\ &\leqslant \left(\|\hat{y}_t\|^2_{L^2(\widetilde{Q})} + \|\hat{y}_{xx}\|^2_{L^2(\widetilde{Q})} + 4C^2\|\hat{y}_x\|^2_{L^2(\widetilde{Q})} + \|(\bar{y}\hat{y})_x\|^2_{L^2(\widetilde{Q})}\right) \end{aligned}$$

and

$$\begin{aligned} \|(\bar{y}\hat{y})_x\|^2_{L^2(\widetilde{Q})} &= \int_{\widetilde{Q}} |\bar{y}_x\hat{y} + \bar{y}\hat{y}_x|^2 dxdt \\ &\leqslant 2\int_{\widetilde{Q}} |\bar{y}_x\hat{y}|^2 dxdt + 2\int_{\widetilde{Q}} \bar{y}^2|\hat{y}_x|^2 dxdt. \end{aligned}$$

By the Sobolev embedding theorem $W_2^{1,2}(\widetilde{Q}) \hookrightarrow L^\infty(\widetilde{Q})$, (see for example [55]), there exists a positive constant c_1, depending only by $\widetilde{Q}$, such that

$$\begin{aligned} \int_{\widetilde{Q}} |\bar{y}_x\hat{y}|^2 dxdt &\leqslant \|\hat{y}\|^2_{L^\infty(\widetilde{Q})} \int_{\widetilde{Q}} |\bar{y}_x|^2 dxdt \\ &\leqslant c_1^2\|\hat{y}\|^2_{W_2^{1,2}(\widetilde{Q})}\|\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})} \end{aligned}$$

and

$$\int_{\widetilde{Q}} \bar{y}^2|\hat{y}_x|^2 dxdt \leqslant c_1^2\|\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})} \int_{\widetilde{Q}} |\hat{y}_x|^2 dxdt.$$

Therefore we obtain that

$$\|f_0\|^2_{L^2(\widetilde{Q})} \leqslant 8(1 + C^2 + 2c_1^2\|\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})})\|\hat{y}\|^2_{W_2^{1,2}(\widetilde{Q})}. \tag{4.1.8}$$

Consider the change of variable $z = w + \hat{y}$, the problem translates into finding a function $w \in W_2^{1,2}(\widetilde{Q})$ such that

$$\begin{cases} w_t(t, x) - w_{xx}(t, x) + 2Cw_x(t, x) + (\bar{y}(t, x)w(t, x))_x = f_0(t, x), & (t, x) \in \widetilde{Q}, \\ w(T_1, x) = 0, & x \in (-2, 2), \\ w(T, x) = 0, & x \in (-2, 2). \end{cases} \tag{4.1.9}$$

We look for a solution to (4.1.9) minimizing the cost

$$\int_{\widetilde{Q}} w^2(t,x)dxdt. \tag{4.1.10}$$

By [56], if w is a solution to (4.1.9) and minimizes (4.1.10), then there exists a function $p(t,x)$ such that:

$$\begin{cases} p_t(t,x) + p_{xx}(t,x) + 2Cp_x(t,x) + \bar{y}(t,x)p_x(t,x) = w(t,x), & (t,x) \in \widetilde{Q}, \\ p(t,\pm 2) = 0, & t \in (T_1,T), \\ p_x(t,\pm 2) = 0, & t \in (T_1,T). \end{cases} \tag{4.1.11}$$

For a such p, the following Carleman type estimate holds.

Theorem 4.1.2 *Let $\bar{y} \in W_2^{1,2}(\widetilde{Q})$ and $w \in L^2(\widetilde{Q})$. If a function $p(t,x)$ satisfies (4.1.11), then there exist a strictly positive function ψ, defined on $\widetilde{Q}$ and two positive constants c_2 and c_3 such that*

$$\int_{\widetilde{Q}} e^{-2s\psi(t,x)}p^2(t,x)dxdt \leqslant c_3 \int_{\widetilde{Q}} w^2(t,x)dxdt, \tag{4.1.12}$$

for every $s \geqslant \max\left\{\|2C+\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})}, c_2\right\}$.

Proof. Let us consider the function

$$\psi(t,x) = \left[\frac{1}{(t-T_1)^2} + \frac{1}{(T-t)^2}\right]\ln(x+4). \tag{4.1.13}$$

For $s > 0$, we denote

$$u = e^{-s\varphi}p, \quad w_1 = e^{-s\varphi}w. \tag{4.1.14}$$

Using (4.1.11) and (4.1.13), we obtain that

$$u(t,\pm 2) = 0,\ u_x(t,\pm 2) = 0,\ u(0,x) = 0,\ u(T,x) = 0, \tag{4.1.15}$$

for every $t \in (T_1,T)$ and $x \in (-2,2)$. The equation (4.1.11) can be rewritten as

$$(M_1 + M_2)u = w_1 - (2C+\bar{y})u_x - s\psi_t u - s\psi_{xx}u - s(2C+\bar{y})\psi_x u \tag{4.1.16}$$

where $M_1 = \frac{\partial^2}{\partial x^2} + s^2(\psi_x)^2$ and $M_2 = \frac{\partial}{\partial t} + 2s\psi_x \frac{\partial}{\partial x}$. We clearly have

$$\|M_1 u\|^2_{L^2(\widetilde{Q})} + \|M_2 u\|^2_{L^2(\widetilde{Q})} + 2(M_1 u, M_2 u)_{L^2(\widetilde{Q})} = \|w_1 - (2C+\bar{y})u_x - s\psi_t u - s\psi_{xx} u - s(2C+\bar{y})\psi_x u\|^2_{L^2(\widetilde{Q})}.$$

Moreover we have

$$2(M_1 u, M_2 u)_{L^2(\widetilde{Q})} = -\int_{\widetilde{Q}} s^2 \frac{\partial}{\partial t}(\psi_x)^2 u^2 dxdt + \\ + \int_{\widetilde{Q}} (-\psi_{xx})(2s(u_x)^2 + 6s^3(\psi_x)^2 u^2) dxdt.$$

and, using the Sobolev embedding theorem $W_2^{1,2}(\widetilde{Q}) \hookrightarrow L^\infty(\widetilde{Q})$,

$$\|w_1 - (2C+\bar{y})u_x - s\psi_t u - s\psi_{xx} u - s(2C+\bar{y})\psi_x u\|^2_{L^2(\widetilde{Q})} \leqslant \\ k_1(\|w_1\|^2_{L^2(\widetilde{Q})} + \|2C+\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})} \|u_x\|^2_{L^2(\widetilde{Q})} + \\ + s^2(\|\psi_t u\|^2_{L^2(\widetilde{Q})} + \|2C+\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})} \|\psi_x u\|^2_{L^2(\widetilde{Q})} + \|\psi_{xx} u\|^2_{L^2(\widetilde{Q})}))$$

where k_1 is a positive constant. Therefore we obtain that

$$2\int_{\widetilde{Q}} (-\psi_{xx})[3s^3 u^2(\psi_x)^2 + s(u_x)^2] dxdt + \|M_1 u\|^2_{L^2(\widetilde{Q})} + \|M_2 u\|^2_{L^2(\widetilde{Q})} \leqslant \\ k_1(\|w_1\|^2_{L^2(\widetilde{Q})} + \|2C+\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})} \|u_x\|^2_{L^2(\widetilde{Q})} + s^2(\|\psi_t u\|^2_{L^2(\widetilde{Q})} + \\ + \|2C+\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})} \|\psi_x u\|^2_{L^2(\widetilde{Q})} + \|u\psi_{xx}\|^2_{L^2(\widetilde{Q})} + +2\int_{\widetilde{Q}} |\frac{\partial}{\partial t}(\psi_x)^2| u^2 dxdt)).$$

So, there exist two positive constants c_2 and c_3 such that, for every $s \geqslant \max\{\|2C+\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})}, c_2\}$,

$$\int_{\widetilde{Q}} u^2 dxdt \leqslant c_3 \|w_1\|^2_{L^2(\widetilde{Q})}. \tag{4.1.17}$$

Now, using (4.1.14), we obtain

$$\int_{\widetilde{Q}} e^{-2s\psi} p^2 dxdt \leqslant c_3 \int_{\widetilde{Q}} e^{-2s\psi} w^2 dxdt \leqslant c_3 \int_{\widetilde{Q}} w^2 dxdt$$

and so the theorem is proved. □

We also need the following lemma.

Lemma 4.1.1 *Let $\bar{y} \in W_2^{1,2}(\widetilde{Q})$ and $f \in L^2(\widetilde{Q})$. If $w \in L^2(\widetilde{Q})$ is a solution to*

$$\begin{cases} w_t(t,x) - w_{xx}(t,x) + 2Cw_x(t,x) + (\bar{y}(t,x)w(t,x))_x = f(t,x), & (t,x) \in \widetilde{Q}, \\ w(T_1,x) = 0, & x \in (-2,2), \end{cases} \tag{4.1.18}$$

then there exists a constant $c > 0$, not depending by C, such that

$$\sup_{t\in[T_1,T]} \int_{-2}^{2} w^2\rho^2 dx + \int_{\widetilde{Q}} (w_x)^2 \rho^2 dxdt \leqslant$$
$$c(1 + \|2C + \bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C + \bar{y}\|^2_{L^\infty(\widetilde{Q})})\|w\|^2_{L^2(\widetilde{Q})} + 2\int_{\widetilde{Q}} f^2\rho^4 dxdt \tag{4.1.19}$$

and there exists a constant $\gamma > 0$, which depends in a continuous way only by $\|\bar{y}\|_{W_2^{1,2}(\widetilde{Q})}$, such that

$$\int_{\widetilde{Q}} ((w_t)^2 + (w_{xx})^2)\rho^4 dxdt \leqslant \gamma(2 + \|\bar{y}\|^2_{L^\infty(\widetilde{Q})} + C^2)\|f\rho^2\|^2_{L^2(\widetilde{Q})} +$$
$$\gamma(1 + \|\bar{y}\|^2_{L^\infty(\widetilde{Q})} + C^2)(2 + \|2C + \bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C + \bar{y}\|^2_{L^\infty(\widetilde{Q})})\|w\|^2_{L^2(\widetilde{Q})}, \tag{4.1.20}$$

where $\rho(x) = 4 - x^2$.

Proof. Multiplying each side of the equation

$$w_t - w_{xx} + 2Cw_x + (\bar{y}w)_x = f$$

by $w\rho^2$ and integrating in $(-2,2)$, we obtain that

$$\frac{1}{2}\frac{\partial}{\partial t}\int_{-2}^{2} w^2\rho^2 dx + \int_{-2}^{2} (w_x)^2\rho^2 dx + \frac{1}{2}\int_{-2}^{2} (w^2)_x(\rho^2)_x dx$$
$$-\int_{-2}^{2} (2C + \bar{y})ww_x\rho^2 dx - \int_{-2}^{2} (2C + \bar{y})w^2(\rho^2)_x dx = \int_{-2}^{2} fw\rho^2 dx \leqslant$$
$$\int_{-2}^{2} f^2\rho^4 dx + \int_{-2}^{2} w^2 dx.$$

Now, integrating the previous inequality in $(0,t)$, we have that

$$\frac{1}{2}\int_{-2}^{2} w^2(t,x)\rho^2(x)dx + \int_0^t\int_{-2}^{2}(w_x)^2\rho^2 dxdt - \frac{1}{2}\int_0^t\int_{-2}^{2} w^2(\rho^2)_{xx}dxdt \leqslant$$
$$\int_0^t\int_{-2}^{2} f^2\rho^4 dxdt + \int_0^t\int_{-2}^{2} w^2 dxdt + \frac{1}{2}\int_0^t\int_{-2}^{2}(2C+\bar{y})^2 w^2\rho^2 dxdt +$$
$$+\frac{1}{2}\int_0^t\int_{-2}^{2}(w_x)^2\rho^2 dxdt + \int_0^t\int_{-2}^{2}(2C+\bar{y})w^2(\rho^2)_x dxdt.$$

If $\rho(x) = 4 - x^2$, then we obtain that there exists a positive constant $c > 0$ such that

$$\frac{1}{2}\int_{-2}^{2} w^2(t,x)\rho^2(x)dx + \frac{1}{2}\int_0^t\int_{-2}^{2}(w_x)^2\rho^2 dxdt \leqslant$$
$$c(1 + \|2C+\bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C+\bar{y}\|^2_{L^\infty(\widetilde{Q})})\int_{\widetilde{Q}} w^2 dxdt + \int_{\widetilde{Q}} f^2\rho^4 dxdt.$$

So we conclude that

$$\sup_{t\in[0,T]}\int_{-2}^{2} w^2(t,x)\rho^2(x)dx + \int_{\widetilde{Q}}(w_x)^2\rho^2 dxdt \leqslant$$
$$c(1 + \|2C+\bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C+\bar{y}\|^2_{L^\infty(\widetilde{Q})})\int_{\widetilde{Q}} w^2 dxdt + \int_{\widetilde{Q}} f^2\rho^4 dxdt$$

which is the first inequality in the statement of the lemma. We now multiply by ρ^2 each side of

$$w_t - w_{xx} + 2Cw_x + (\bar{y}w)_x = f$$

in order to obtain

$$(w\rho^2)_t - (w\rho^2)_{xx} + \bar{y}_x(w\rho^2) = f\rho^2 - \bar{y}w_x\rho^2 - 2w_x(\rho^2)_x - w(\rho^2)_{xx} - 2Cw_x\rho^2.$$

Since $w\rho^2$ satisfies a parabolic equation, $(w\rho^2)(t,\pm 2) = 0$ for every $t \in (T_1, T)$ and $(w\rho^2)(0,x) = 0$ for every $x \in (-2,2)$, there exists a positive constant γ_1, which depends continuously by $\|\bar{y}\|_{W_2^{1,2}(\widetilde{Q})}$, such that

$$\|w\rho^2\|^2_{W_2^{1,2}(\widetilde{Q})} \leqslant \gamma_1\|f\rho^2 - \bar{y}w_x\rho^2 - 2w_x(\rho^2)_x - w(\rho^2)_{xx} - 2Cw_x\rho^2\|^2_{L^2(\widetilde{Q})}.$$

Therefore we have that

$$\int_{\widetilde{Q}} (w_t)^2 \rho^4 dxdt + \int_{\widetilde{Q}} (w_{xx})^2 \rho^4 dxdt \leqslant -4 \int_{\widetilde{Q}} w_{xx} w_x \rho^2 (\rho^2)_x dxdt$$
$$-2 \int_{\widetilde{Q}} w w_{xx} \rho^2 (\rho^2)_{xx} dxdt +$$
$$+\gamma_1 \| f\rho^2 - \bar{y} w_x \rho^2 - 2w_x(\rho^2)_x - w(\rho^2)_{xx} - 2C w_x \rho^2 \|^2_{L^2(\widetilde{Q})}.$$

Moreover

$$-4 \int_{\widetilde{Q}} w_{xx} w_x \rho^2 (\rho^2)_x dxdt = -2 \int_{\widetilde{Q}} ((w_x)^2)_x \rho^2 (\rho^2)_x dxdt$$
$$= 2 \int_{\widetilde{Q}} (w_x)^2 ((\rho^2)_x)^2 dxdt + 2 \int_{\widetilde{Q}} (w_x)^2 \rho^2 (\rho^2)_{xx} dxdt$$
$$\leqslant \tilde{c} \int_{\widetilde{Q}} (w_x)^2 \rho^2 dxdt$$
$$\leqslant \tilde{c}c(1 + \|2C + \bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C + \bar{y}\|^2_{L^\infty(\widetilde{Q})}) \|w\|^2_{L^2(\widetilde{Q})} + 2\tilde{c} \int_{\widetilde{Q}} f^2 \rho^4 dxdt,$$

where $\tilde{c}$ is a positive constant. Moreover

$$-2 \int_{\widetilde{Q}} w w_{xx} \rho^2 (\rho^2)_{xx} dxdt = 2 \int_{\widetilde{Q}} (w_x)^2 \rho^2 (\rho^2)_{xx} dxdt + \int_{\widetilde{Q}} (w^2)_x (\rho^2)_x (\rho^2)_{xx} dxdt$$
$$+ \int_{\widetilde{Q}} (w^2)_x \rho^2 (\rho^2)_{xxx} dxdt$$
$$\leqslant c_2 \int_{\widetilde{Q}} (w_x)^2 \rho^2 dxdt + c_2 \int_{\widetilde{Q}} w^2 dxdt$$
$$\leqslant c_2 c(2 + \|2C + \bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C + \bar{y}\|^2_{L^\infty(\widetilde{Q})}) \|w\|^2_{L^2(\widetilde{Q})} +$$
$$+2c_2 \int_{\widetilde{Q}} f^2 \rho^4 dxdt,$$

for some positive constant c_2. Now

$$\begin{gathered}
\|f\rho^2 - 2\bar{y}w_x\rho^2 - 2w_x(\rho^2)_x - w(\rho^2)_{xx} - 2Cw_x\rho^2\|^2_{L^2(\widetilde{Q})} \leqslant \\
\leqslant 2\|f\rho^2\|^2_{L^2(\widetilde{Q})} + 4\|\bar{y}w_x\rho^2\|^2_{L^2(\widetilde{Q})} + 4\|w_x(\rho^2)_x\|^2_{L^2(\widetilde{Q})} + \\
+2\|w(\rho^2)_{xx}\|^2_{L^2(\widetilde{Q})} + 4C^2\|w_x\rho^2\|^2_{L^2(\widetilde{Q})} \\
\leqslant 2\|f\rho^2\|^2_{L^2(\widetilde{Q})} + c_3(1+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})} + C^2)\|w_x\rho\|^2_{L^2(\widetilde{Q})} + c_3\|w\|^2_{L^2(\widetilde{Q})} \\
\leqslant cc_3(1+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})} + C^2)(2+\|2C+\bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C+\bar{y}\|^2_{L^\infty(\widetilde{Q})})\|w\|^2_{L^2(\widetilde{Q})} + \\
+cc_3(2+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})} + C^2)\int_{\widetilde{Q}} f^2\rho^4 dxdt,
\end{gathered}$$

where $c_3 > 0$. Therefore we conclude that there exists a constant $\gamma > 0$, which only depends continuously by $\|\bar{y}\|_{W_2^{1,2}(\widetilde{Q})}$, such that

$$\begin{gathered}
\int_{\widetilde{Q}} (w_t)^2\rho^4 dxdt + \int_{\widetilde{Q}} (w_{xx})^2\rho^4 dxdt \leqslant \gamma(2+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})} + C^2)\|f\rho^2\|^2_{L^2(\widetilde{Q})} + \\
+\gamma(1+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})} + C^2)(2+\|2C+\bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C+\bar{y}\|^2_{L^\infty(\widetilde{Q})})\|w\|^2_{L^2(\widetilde{Q})}.
\end{gathered}$$

This concludes the proof of the lemma. □

If $z = \theta(\bar{y})$, where θ is the function defined in (4.1.6), then we have:

$$\begin{aligned}
\|z\|^2_{W_2^{1,2}(Q)} &= \|w+\hat{y}\|^2_{W_2^{1,2}(Q)} \\
&\leqslant 2\|w\|^2_{W_2^{1,2}(Q)} + 2\|\hat{y}\|^2_{W_2^{1,2}(Q)}.
\end{aligned}$$

By definition, we have that

$$\|w\|^2_{W_2^{1,2}(Q)} = \|w\|^2_{L^2(Q)} + \|w_x\|^2_{L^2(Q)} + \|w_{xx}\|^2_{L^2(Q)} + \|w_t\|^2_{L^2(Q)}.$$

By (4.1.19) we obtain that

$$\|w_x\|^2_{L^2(Q)} \leqslant c(1+\|2C+\bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C+\bar{y}\|^2_{L^\infty(\widetilde{Q})})\|w\|^2_{L^2(\widetilde{Q})} + c_4\|f_0\|^2_{L^2(\widetilde{Q})}$$

for some positive constant c_4. Using (4.1.20) we get

$$\begin{gathered}
\|w_{xx}\|^2_{L^2(Q)} + \|w_t\|^2_{L^2(Q)} \leqslant c_5\gamma(2+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})} + C^2)\|f_0\|^2_{L^2(\widetilde{Q})} + \\
+\gamma(1+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})} + C^2)(2+\|2C+\bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C+\bar{y}\|^2_{L^\infty(\widetilde{Q})})\|w\|^2_{L^2(\widetilde{Q})}
\end{gathered}$$

with $c_5 > 0$. Thus we obtain that there exists a positive constant c_6, which depends by $\|\bar{y}\|_{W_2^{1,2}(\widetilde{Q})}$ but not by C, such that

$$\|z\|^2_{W_2^{1,2}(Q)} \leqslant c_6(1+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})}+C^2)(1+\|2C+\bar{y}\|_{L^\infty(\widetilde{Q})}+\|2C+\bar{y}\|^2_{L^\infty(\widetilde{Q})})\|w\|^2_{L^2(\widetilde{Q})}$$
$$+c_6(1+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})}+C^2)\|f_0\|^2_{L^2(\widetilde{Q})}+2\|\hat{y}\|^2_{W_2^{1,2}(\widetilde{Q})}.$$

By (4.1.11) we have that

$$\|w\|^2_{L^2(\widetilde{Q})} = |(f_0, p)_{L^2(\widetilde{Q})}|.$$

Recalling that the support of f_0 is contained in $(\frac{2T_1+T}{3}, \frac{T_1+2T}{3}) \times (-2,2)$, we obtain that

$$\begin{aligned}
\|w\|^2_{L^2(\widetilde{Q})} &= \left| \int_{\frac{2T_1+T}{3}}^{\frac{T_1+2T}{3}} \int_{-2}^{2} f_0(t,x)p(t,x)dxdt \right| \\
&\leqslant \|f_0\|_{L^2(\widetilde{Q})} \left(\int_{\frac{2T_1+T}{3}}^{\frac{T_1+2T}{3}} \int_{-2}^{2} p^2(t,x)dxdt \right)^{1/2} \\
&\leqslant e^{2sk_1}\|f_0\|_{L^2(\widetilde{Q})} \left(\int_{\frac{2T_1+T}{3}}^{\frac{T_1+2T}{3}} \int_{-2}^{2} e^{-2s\psi}p^2(t,x)dxdt \right)^{1/2} \\
&\leqslant e^{2sk_1}\|f_0\|_{L^2(\widetilde{Q})} \left(\int_{\widetilde{Q}} e^{-2s\psi}p^2(t,x)dxdt \right)^{1/2}
\end{aligned}$$

where $s = \|2C+\bar{y}\|^2_{W^{1,2,2}(\widetilde{Q})}$, ψ are that of Theorem 4.1.2 and $k_1 = \frac{45}{4}\frac{\log 6}{(T-T_1)^2}$. Using the Carleman estimate (4.1.12) we obtain

$$\|w\|^2_{L^2(\widetilde{Q})} \leqslant e^{2sk_1}\|f_0\|_{L^2(\widetilde{Q})}\|w\|_{L^2(\widetilde{Q})}$$

and so

$$\|w\|_{L^2(\widetilde{Q})} \leqslant e^{2sk_1}\|f_0\|_{L^2(\widetilde{Q})}.$$

Thus, we obtain that there exists a constant c_7 not depending by C, such that

$$\|z\|^2_{W_2^{1,2}(Q)} \leqslant 2\|\hat{y}\|^2_{W_2^{1,2}(\widetilde{Q})}+$$
$$c_7 e^{4sk_1}(1+\|\bar{y}\|^2_{L^\infty(\widetilde{Q})}+C^2)(1+\|2C+\bar{y}\|_{L^\infty(\widetilde{Q})}+\|2C+\bar{y}\|^2_{L^\infty(\widetilde{Q})})\|f_0\|^2_{L^2(\widetilde{Q})}.$$

and, using (4.1.8), we have

$$\|z\|^2_{W_2^{1,2}(Q)} \leqslant 2\|\hat{y}\|^2_{W_2^{1,2}(\widetilde{Q})}+$$
$$c_8 e^{4sk_1}(1+\|\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})}+C^2)^2(1+\|2C+\bar{y}\|_{L^\infty(\widetilde{Q})}+\|2C+\bar{y}\|^2_{L^\infty(\widetilde{Q})})\|\hat{y}\|^2_{W_2^{1,2}(\widetilde{Q})},$$

where c_8 is a positive constant depending only by $\|\hat{y}\|_{W_2^{1,2}(\widetilde{Q})}$ in a continuous way.

Since $\hat{y}(t,x)=\varphi(t)\chi(t,x)$, there exists a positive constant c_9 such that

$$\|\hat{y}\|_{W_2^{1,2}(Q)} \leqslant c_9\|\chi\|_{W_2^{1,2}(Q)}.$$

We have the following lemmata.

Lemma 4.1.2 *Let* $\bar{y} \in W_2^{1,2}(\widetilde{Q})$ *and* $y_0 \in W^{1,2}(-2,2)$. *If* $\|\bar{y}\|_{L^\infty(\widetilde{Q})} < \min\{2, \frac{1}{T-T_1}\}$, *then*

$$\|\chi\|^2_{L^2(\widetilde{Q})} \leqslant \frac{T-T_1}{1-(T-T_1)\|\bar{y}\|_{L^\infty(\widetilde{Q})}}\|y_0\|^2_{L^2(-2,2)}. \tag{4.1.21}$$

Proof. For every $t \in]T_1,T[$, we clearly have

$$\frac{d}{dt}\int_{-2}^{2}\chi^2(t,x)dx = 2\int_{-2}^{2}\chi(t,x)\chi_{xx}(t,x)dx - 4C\int_{-2}^{2}\chi(t,x)\chi_x(t,x)dx$$
$$-2\int_{-2}^{2}\chi^2(t,x)\bar{y}_x(t,x)dx - 2\int_{-2}^{2}\chi(t,x)\chi_x(t,x)\bar{y}(t,x)dx$$
$$= -2\int_{-2}^{2}(\chi_x(t,x))^2dx - 2\int_{-2}^{2}\chi^2(t,x)\bar{y}_x(t,x)dx - \int_{-2}^{2}(\chi^2(t,x))_x\bar{y}(t,x)dx$$
$$= -\int_{-2}^{2}(\chi_x(t,x))^2dx + \int_{-2}^{2}(\chi^2(t,x))_x\bar{y}(t,x)dx.$$

Moreover we have:

$$\int_{-2}^{2}(\chi^2(t,x))_x\bar{y}(t,x)dx \leqslant 2\|\bar{y}\|_{L^\infty(\widetilde{Q})}\int_{-2}^{2}|\chi(t,x)\chi_x(t,x)|dx$$
$$\leqslant \|\bar{y}\|_{L^\infty(\widetilde{Q})}\int_{-2}^{2}\chi^2(t,x)dx + \|\bar{y}\|_{L^\infty(\widetilde{Q})}\int_{-2}^{2}(\chi_x(t,x))^2dx.$$

Therefore

$$\begin{aligned}\frac{d}{dt}\int_{-2}^{2}\chi^2(t,x)dx &\leqslant (\|\bar y\|_{L^\infty(\widetilde Q)}-2)\int_{-2}^{2}(\chi_x(t,x))^2dx+\\ &\quad +\|\bar y\|_{L^\infty(\widetilde Q)}\int_{-2}^{2}\chi^2(t,x)dx\\ &\leqslant \|\bar y\|_{L^\infty(\widetilde Q)}\int_{-2}^{2}\chi^2(t,x)dx.\end{aligned}$$

Integrating the last inequality on $(-2,2)$ we obtain

$$\int_{-2}^{2}\chi^2(t,x)dx\leqslant\int_{-2}^{2}y_0^2(x)dx+\|\bar y\|_{L^\infty(\widetilde Q)}\|\chi\|^2_{L^2(\widetilde Q)}.$$

Integrating now on (T_1,T) we get

$$\|\chi\|^2_{L^2(\widetilde Q)}\leqslant (T-T_1)\|y_0\|^2_{L^2(-2,2)}+(T-T_1)\|\bar y\|_{L^\infty(\widetilde Q)}\|\chi\|^2_{L^2(\widetilde Q)}$$

and obviously the last inequality implies the claim. □

Lemma 4.1.3 *Let $\bar y\in W_2^{1,2}(\widetilde Q)$ and $y_0\in W^{1,2}(-2,2)$. Then*

$$\|\chi_x\|^2_{L^2(\widetilde Q)}\leqslant\|y_0\|^2_{L^2(-2,2)}+\|\bar y\|^2_{L^\infty(\widetilde Q)}\|\chi\|^2_{L^2(\widetilde Q)}. \tag{4.1.22}$$

Proof. Multiplying by χ the equation

$$\chi_t-\chi_{xx}+2C\chi_x+(\bar y\chi)_x=0$$

and integrating in $(-2,2)$ we obtain that

$$\frac{1}{2}\frac{d}{dt}\int_{-2}^{2}\chi^2dx+\int_{-2}^{2}(\chi_x)^2dx\leqslant\int_{-2}^{2}\bar y\chi\chi_xdx.$$

Integrating now on (T_1,t), we get

$$\begin{aligned}\frac{1}{2}\int_{-2}^{2}\chi^2(t,x)dx+\int_{T_1}^{t}\int_{-2}^{2}(\chi_x)^2dxds &\leqslant \frac{1}{2}\|y_0\|^2_{L^2(\widetilde Q)}+\frac{1}{2}\int_{T_1}^{t}\int_{-2}^{2}(\chi_x)^2dxds\\ &\quad +\frac{1}{2}\|\bar y\|^2_{L^\infty(\widetilde Q)}\|\chi\|^2_{L^2(\widetilde Q)}.\end{aligned}$$

Therefore the lemma is proved. □

Lemma 4.1.4 *Let $\bar{y} \in W_2^{1,2}(\widetilde{Q})$ and $y_0 \in W^{1,2}(-2,2)$. Then*

$$\|\chi\|^2_{W_2^{1,2}(\widetilde{Q})} \leqslant 4\gamma_2 C^2 \|\chi_x\|^2_{L^2(\widetilde{Q})}$$

where γ_2 is a positive constant depending by $\|\bar{y}\|_{W_2^{1,2}(\widetilde{Q})}$ and by $\|y_0\|_{L^2(-2,2)}$ in a continuous way.

Proof. The proof easily follows applying standard estimation on parabolic equations (see for example [55]). □

These lemmata proved that it is possible to control the norm $\|\chi\|_{W_2^{1,2}(\widetilde{Q})}$ with $\|\chi\|_{L^2(\widetilde{Q})}$ and the last one by the $\|y_0\|_{L^2(-2,2)}$. Therefore there exists a constant $k_2 > 0$ which depends by T, T_1 and $\|\bar{y}\|_{W_2^{1,2}}$ such that

$$\|z\|_{W_2^{1,2}(Q)} \leqslant k_2 e^{2sk_1} C(1 + \|\bar{y}\|^2_{W_2^{1,2}(\widetilde{Q})} + C^2)\times$$
$$\times(1 + \|2C + \bar{y}\|_{L^\infty(\widetilde{Q})} + \|2C + \bar{y}\|^2_{L^\infty(\widetilde{Q})})^{\frac{1}{2}} \|y_0\|_{W^{1,2}(0,1)}.$$

Thus, if $\|y_0\|_{W^{1,2}}$ is sufficiently small, it is possible to apply Schauder fixed point theorem to the map θ. First, for C big enough, we may suppose that $\|2C+\bar{y}\|_{L^\infty(\widetilde{Q})} \simeq 2C$ and $\|2C+\bar{y}\|_{W_2^{1,2}(\widetilde{Q})} \simeq 2C$. Therefore, if C is big enough, then there exists a polynomial function $g(C)$ with the following property: if $\|y_0\|_{W^{1,2}}(0,1) \leqslant \frac{1}{e^{8k_1C^2}g(C)}$, then θ has a fixed point.

Fourth step: conclusion of the proof.

By (4.1.4), we have that

$$\|y_0\|_{W^{1,2}(0,1)} \simeq 2^{1/2} C^{3/2} T_1^{1/2} e^{\frac{C(1+\eta-CT_1)}{2}}.$$

From the third step it follows that, if T is big enough, then y_0, given by $z_1(T_1,\cdot) - C$, belongs to the ball in $W^{1,2}(0,1)$ centered at 0 with radius $\frac{1}{e^{8k_1C^2}g(C)}$. This concludes the proof of Theorem 4.1.1.

4.2 Null controllability.

In this section we study null controllability for the Burgers equation (4.0.1). We show that it is possible to drive any initial data in $W^{1,2}(0,1)$ to 0 in finite time. The precisely result is the following.

Theorem 4.2.1 *There exists a time* $T_0 > 0$ *such that, for every* $y_0 \in W^{1,2}(0,1)$, *there exists* $y \in C^0([0,T_0]; W^{1,2}(0,1))$ *such that*

$$\begin{cases} y_t - y_{xx} + 2yy_x = 0, & in\ \mathcal{D}'(]0,T_0[\times]0,1[), \\ y(0,x) = y_0(x), & \forall x \in [0,1], \\ y(T_0,x) = 0, & \forall x \in [0,1]. \end{cases} \tag{4.2.23}$$

Before giving the proof of this theorem, let us consider the following technical lemmata.

Lemma 4.2.1 ***(Maximum Principle)***. *Fix* $T > 0$. *If* y *and* $\bar{y}$ *belong to* $C^0([0,T]; H^1(0,1))$ *and satisfy*

$$\begin{cases} y_t - y_{xx} + 2yy_x \leqslant \bar{y}_t - \bar{y}_{xx} + 2\bar{y}\bar{y}_x, & in\ \mathcal{D}'(]0,T[\times]0,1[), \\ y(t,0) \leqslant \bar{y}(t,0),\ y(t,1) \leqslant \bar{y}(t,1), & \forall t \in [0,1], \\ y(0,x) \leqslant \bar{y}(0,x), & \forall x \in [0,1], \end{cases} \tag{4.2.24}$$

then

$$y(t,x) \leqslant \bar{y}(t,x) \quad \forall x \in [0,1], \forall t \in [0,T]. \tag{4.2.25}$$

For a proof of this lemma, see for instance [98].

Lemma 4.2.2 *Let* $T > 0$, $y \in C^0([0,T]; H^1(0,1))$ *be such that*

$$\begin{cases} y_t - y_{xx} + 2yy_x = 0, & in\ \mathcal{D}'(]0,T_0[\times]0,1[), \\ y(t,0) = y(t,1) = 0, & \forall t \in [0,T]. \end{cases} \tag{4.2.26}$$

Then

$$y(t,x) \leqslant \frac{x}{2t} \tag{4.2.27}$$

for every $t \in]0,T]$ *and every* $x \in [0,1]$.

Proof. Let $\delta > 0$. There exists $\bar{\varepsilon} > 0$ such that, if $\varepsilon \in]0,\bar{\varepsilon}]$, then

$$y(0,x) \leqslant \delta + \frac{x}{\varepsilon} \quad \forall x \in [0,1]. \tag{4.2.28}$$

We define

$$\bar{y}(t,x) := \frac{\delta\varepsilon + \frac{x}{2}}{\varepsilon + t}.$$

Therefore

$$\bar{y}_t - \bar{y}_{xx} + 2\bar{y}\bar{y}_x = 0$$

and

$$y(0,x) \leqslant \bar{y}(0,x),\ 0 = y(t,0) \leqslant \bar{y}(t,0),\ 0 = y(t,1) \leqslant \bar{y}(t,1).$$

Using Lemma 4.2.1, we conclude that

$$y(t,x) \leqslant \bar{y}(t,x) = \frac{\delta\varepsilon + \frac{x}{2}}{\varepsilon + t} \qquad \forall (t,x) \in [0,T] \times [0,1]$$

and thus

$$y(t,x) \leqslant \frac{\delta\varepsilon + \frac{x}{2}}{t} \quad \forall (t,x) \in]0,T] \times [0,1].$$

Passing to the limit as $\varepsilon \to 0$, we finish the proof. □

Lemma 4.2.3 *Fix $T > 0$. If $y \in C^0([0,T]; H^1(0,1))$ satisfies*

$$\begin{cases} y_t - y_{xx} + 2yy_x = 0, & \text{in } \mathcal{D}'(]0,T[\times]0,1[), \\ y(t,0) = y(t,1) = 0, & \forall t \in [0,T], \end{cases}$$

then

$$y(t,x) \geqslant -\frac{1-x}{2t}$$

for every $t \in]0,T]$ and for every $x \in [0,1]$.

Proof. Use Lemma 4.2.2 with $-y(t, 1-x)$. □

Lemma 4.2.4 *Let $T_1 > 0$ and $\eta_1 > 0$. There exists $\eta = \eta(T_1, \eta_1) > 0$ such that, if $y_0 \in H^1_0(0,1)$ satisfies*

$$\|y_0\|_{L^\infty(0,1)} \leqslant \eta,$$

then there exists a unique function $y \in C([0,T_1]; H^1(0,1))$ solution to

$$\begin{cases} y_t - y_{xx} + 2yy_x = 0, & \text{in } \mathcal{D}'(]0,T_1[\times]0,1[), \\ y(t,0) = y(t,1) = 0, & \forall t \in [0,T_1], \\ y(0,x) = y_0(x), & \forall x \in [0,1], \end{cases}$$

and such that

$$\|y(T_1,\cdot)\|_{H^1(0,1)} \leqslant \eta_1.$$

Lemma 4.2.5 ***(Fursikov, Imanuvilov)***. *Let $T_2 > 0$. There exists $\varepsilon = \varepsilon(T_2) > 0$ such that, if $y_0 \in H^1(0,1)$ satisfies*

$$\|y_0\|_{H^1(0,1)} \leqslant \varepsilon,$$

then there exists $y \in C([0,T_2]; H^1(0,1))$ such that

$$\begin{cases} y_t - y_{xx} + 2yy_x = 0 & \text{in } \mathcal{D}'(]0,T_2[\times]0,1[) \\ y(0,x) = y_0(x) & \forall x \in [0,1] \\ y(T_2,x) = 0 & \forall x \in [0,1]. \end{cases}$$

A proof of this lemma can be found in [56].

Lemma 4.2.6 *Let $T > 0$. We suppose that $a \in C^0([0,T])$, $b \in C^0([0,T])$ and $y_0 \in H^1(0,1)$ satisfy*

$$y_0(0) = a(0), \quad y_0(1) = b(0).$$

Then the Cauchy problem

$$\begin{cases} y_t - y_{xx} + 2yy_x = 0, & \text{in } \mathcal{D}'(]0,T[\times]0,1[), \\ y(t,0) = a(t),\ y(t,1) = b(t), & \forall t \in [0,T], \\ y(0,x) = y_0(x,) & \forall x \in [0,1], \end{cases} \tag{4.2.29}$$

admits a unique solution in $C^0([0,T]; H^1(0,1))$.

Proof. The local existence and uniqueness is well known, see for example [114]. For the global existence we argue by contradiction and we suppose that the maximal solution to the Cauchy problem (4.2.29) is defined on $[0,T'[$ with $T' < T$ and

$$\lim_{t\to T', t<T'} \|y(t,\cdot)\|_{L^\infty(0,1)} = +\infty. \tag{4.2.30}$$

Therefore, if $M := \|y_0\|_{L^\infty(0,1)} + \|a\|_{L^\infty(0,T)} + \|b\|_{L^\infty(0,T)}$, then, using Lemma 4.2.1 with $\bar{y} = M$, we obtain

$$y(t,x) \leqslant M \quad \forall (t,x) \in [0,T'[\times[0,1]. \tag{4.2.31}$$

With the same arguments, we have that

$$-M \leqslant y(t,x) \quad \forall (t,x) \in [0,T'[\times[0,1]. \tag{4.2.32}$$

Then the inequalities (4.2.31) and (4.2.32) contradict (4.2.30). □

Proof of Theorem 4.2.1. Fix $T_1 > 0$, $T_2 > 0$ and $T_3 > 0$. We take $\eta_1 := \varepsilon(T_2)$, $\eta = \eta(T_1, \eta_1)$ as in Lemmas 4.2.5 and 4.2.4, and $T_4 = 1/\eta$, $T_0 := T_1 + T_2 + T_3 + T_4$. Let $y_0 \in H^1(0,1)$. On $[0, T_3] \times [0,1]$, let y be the function in $C^0([0,T_3]; H^1(0,1))$ defined by

$$\begin{cases} y_t - y_{xx} + 2yy_x = 0, & \text{in } \mathcal{D}'(]0,T_3[\times]0,1[), \\ y(0,x) = y_0(x), & \forall x \in [0,1], \\ y(t,0) = \frac{T_3 - t}{T_3} y_0(0), & \forall t \in [0,T_3], \\ y(t,1) = \frac{T_3 - t}{T_3} y_1(1), & \forall t \in [0,T_3]. \end{cases}$$

By Lemma 4.2.6, such y exists. On $[T_3, T_3+T_4+T_1]$, we consider the function $y \in C^0([T_3, T_3 + T_4 + T_1]; H^1(0,1))$ such that

$$\begin{cases} y_t - y_{xx} + 2yy_x = 0, & \text{in } \mathcal{D}'(]T_3, T_3 + T_4 + T_1[\times]0,1[), \\ y(t,0) = y(t,1) = 0, & \forall t \in [T_3, T_3 + T_4 + T_1]. \end{cases}$$

By Lemma 4.2.6, such y exists. Using Lemma 4.2.2 and Lemma 4.2.3 we obtain

$$\|y(T_3 + T_4, \cdot)\|_{L^\infty(0,1)} \leqslant \frac{1}{T_4} = \eta(T_1, \eta_1). \tag{4.2.33}$$

Now, using (4.2.33) and Lemma 4.2.4 we get

$$\|y(T_3 + T_4 + T_1, \cdot)\|_{H^1(0,1)} \leqslant \eta_1 = \varepsilon(T_2). \tag{4.2.34}$$

Finally, using (4.2.34) and Lemma 4.2.5, there exists a continuous function $y \in C^0([T_3 + T_4 + T_1, T_3 + T_4 + T_1 + T_2]; H^1(0,1))$ such that

$$y_t - y_{xx} + 2yy_x = 0 \quad \text{in } \mathcal{D}'(]T_3 + T_4 + T_1, T_3 + T_4 + T_1 + T_2[\times]0,1[)$$

and $y(T_3 + T_4 + T_1 + T_2, \cdot) = 0$. So the proof of the theorem is finished. □

Chapter 5

Lighthill–Whitham–Richards traffic model.

This chapter deals with a fluidodynamic model of heavy traffic on a road network. More precisely, we consider the conservation law formulation proposed by Lighthill and Whitham [77] and Richards [99]. This nonlinear framework is based simply on the conservation of cars and is described by the equation:

$$\rho_t + f(\rho)_x = 0, \tag{5.0.1}$$

where $\rho = \rho(t,x) \in [0, \rho_{max}]$, $(t,x) \in \mathbb{R}_+ \times \mathbb{R}$, is the *density* of cars, $v = v(t,x)$ is the *speed* and $f(\rho) = v\,\rho$ is the *flux*. This model is appropriate to reveal shocks formation as it is natural for conservation laws, whose solutions may develop discontinuities in finite time even for smooth initial data (see [26]). In most cases one assumes that v is a function of ρ only and that the corresponding flux is a concave function. We make this assumption, moreover we let f have a unique maximum $\sigma \in]0, \rho_{max}[$ and for notational simplicity we assume $\rho_{max} = 1$.

Here we deal with a network of roads, as in [65]. This means that we have a finite number of roads modeled by intervals $[a_i, b_i]$ (with one of the two endpoints possibly infinite) that meet at some junctions. For endpoints that do not touch a junction (and are not infinite), we assume to have a given boundary data and solve the corresponding boundary problem, as in [6, 9, 16]. The key role is played by junctions at which the system is underdetermined even after prescribing the conservation of cars, that can be written as the

Rankine-Hugoniot relation:

$$\sum_{i=1}^{n} f(\rho_i(t,b_i)) = \sum_{j=n+1}^{n+m} f(\rho_j(t,a_j)), \tag{5.0.2}$$

where ρ_i, $i = 1,\dots,n$, are the car densities on incoming roads, while ρ_j, $j = n+1,\dots,n+m$, are the car densities on outgoing roads. In [65], the Riemann problem, that is the problem with constant initial data on each road, is solved maximizing a concave function of the fluxes and it is proved existence of weak solutions for Cauchy problems with suitable initial data of bounded variation. In this paper we assume that:

(A) there are some prescribed preferences of drivers, that is the traffic from incoming roads is distributed on outgoing roads according to fixed coefficients;

(B) respecting (A), drivers choose so as to maximize fluxes.

To deal with rule (A), we fix a traffic distribution matrix

$$A \doteq \{\alpha_{ji}\}_{j=n+1,\dots n+m,\, i=1,\dots,n} \in \mathbb{R}^{m\times n},$$

such that

$$0 < \alpha_{ji} < 1, \qquad \sum_{j=n+1}^{n+m} \alpha_{ji} = 1, \tag{5.0.3}$$

for each $i = 1,...,n$ and $j = n+1,...,n+m$, where α_{ji} is the percentage of drivers arriving from the $i-$th incoming road that take the $j-$th outgoing road. Notice that with only the rule (A) Riemann problems are still underdetermined. This choice represents a situation in which drivers have a final destination, hence distribute on outgoing roads according to a fixed law, but maximize the flux whenever possible. We are able to solve uniquely Riemann problems, under suitable conditions on the matrix A, and then to construct solutions to Cauchy problems for networks with simple junctions, i.e. junctions with two incoming roads and two outgoing ones. Our main technique is the use of a front tracking algorithm and the control of the total variation of the flux. We refer the reader to [26] for the general theory of conservation laws and for a discussion of wave front tracking algorithms.

The main difficulty in solving systems of conservation laws is the control of the total variation, see [26]. It is easy to see that for a single conservation law the total variation is decreasing, however in our case it may increase due to interaction of waves with junctions.

There is a natural lack of symmetry for *big waves* (i.e. waves crossing the value σ, see Definition 5.4.3) and *bad data* (see Definition 5.4.3) at junctions, since the role of entering roads is different from that of exiting ones. Similarly, for scalar conservation laws with discontinuous coefficients, one has to use a definition of strength for discontinuities of the coefficient, seen as waves, that is not symmetric but depends on the sign of the jump in the solution, see [73, 115, 116]. This is enough to control the total variation in that case, on the contrary our problem is more delicate. In fact, the variation can still increase due to interactions of waves with junctions. The bounded quantity is the total variation of the flux. We prove this fact for junctions with only two incoming roads and two outgoing ones. Unfortunately the total variation of the flux is not equivalent to the total variation of ρ, since $f'(\sigma) = 0$, and so it is not sufficient to prove existence of solutions. Therefore some compactness argument is used together with a bound of big waves near junctions.

Our techniques are quite flexible, so we can deal with time dependent coefficients for the rule (A). In particular, we can model traffic lights and also in this case the control of total variation is extremely delicate. An arbitrarily small change in the coefficients can produce waves whose strength is bounded away from zero. Still it is possible to consider periodic coefficients, a case of particular interest for applications. We can also deal with roads with different fluxes: this can be treated in the same way with the necessary notational modifications.

There is an interesting ongoing discussion on hydrodynamic models for heavy traffic flow. In particular some models using systems of two conservation laws have been proposed, see [13, 40, 60]. We do no treat this aspect.

The chapter is organized as follows. In Section 5.1 we give the definition of weak entropic solution and, following rules (A) and (B), we introduce an admissibility condition at junctions. In Section 5.2 we prove the existence and uniqueness of admissible solutions for the Riemann Problem in a junction, then using this we describe the construction of the approximants for the Cauchy Problem (see Section 5.3). In Section 5.4 we prove the bound on the total variation of the flux and existence of admissible solutions for the Cauchy Problem with suitable initial data. In Section 5.5 we prove with a counterexample that the Lipschitz continuous dependence with respect to initial data does not hold in general, but we also show that this property holds under special assumptions. In Section 5.6 we describe what happens when there are traffic lights and time dependent coefficients. Section 5.7 contains an example of flux variation increase, that does not happen for junctions with only two incoming and two outgoing roads. Finally, in Section 5.8 we show that the interaction of a small wave with a junction can produce a uniformly big wave.

5.1 Basic Definitions.

We consider a network of roads, that is modeled by a finite collection of intervals $I_i = [a_i, b_i] \subset \mathbb{R}$, $i = 1, \ldots, N$, $a_i < b_i$, possibly with either $a_i = -\infty$ or $b_i = +\infty$, on which we consider the equation (5.0.1). Hence the datum is given by a finite collection of functions ρ_i defined on $[0, +\infty[\times I_i$.

On each road I_i we want ρ_i to be a weak entropic solution, that is for every function $\varphi : [0, +\infty[\times I_i \to \mathbb{R}$ smooth with compact support on $]0, +\infty[\times]a_i, b_i[$

$$\int_0^{+\infty}\int_{a_i}^{b_i}\Big(\rho_i\frac{\partial\varphi}{\partial t} + f(\rho_i)\frac{\partial\varphi}{\partial x}\Big)dxdt = 0, \tag{5.1.4}$$

and for every $k \in \mathbb{R}$ and every $\tilde{\varphi} : [0, +\infty[\times I_i \to \mathbb{R}$ smooth, positive with compact support on $]0, +\infty[\times]a_i, b_i[$

$$\int_0^{+\infty}\int_{a_i}^{b_i}\Big(|\rho_i - k|\frac{\partial\tilde{\varphi}}{\partial t} + \text{sgn}\,(\rho_i - k)(f(\rho_i) - f(k))\frac{\partial\tilde{\varphi}}{\partial x}\Big)dxdt \geq 0. \tag{5.1.5}$$

It is well known that, for equation (5.0.1) on $\mathbb{R}$ and for every initial data in L^∞, there exists a unique weak entropic solution depending in a continuous way from the initial data in L^1_{loc}.

We assume that the roads are connected by some junctions. Each junction J is given by a finite number of incoming roads and a finite number of outgoing roads, thus we identify J with $((i_1, \ldots, i_n), (j_1, \ldots, j_m))$ where the first n–tuple indicates the set of incoming roads and the second m–tuple indicates the set of outgoing roads. We assume that each road can be incoming road at most for one junction and outgoing at most for one junction.

Hence the complete model is given by a couple $(\mathcal{I}, \mathcal{J})$, where $\mathcal{I} = \{I_i : i = 1, \ldots, N\}$ is the collection of roads and $\mathcal{J}$ is the collection of junctions.

Fix a junction J with incoming roads, say I_1,. . .,I_n, and outgoing roads, say I_{n+1},. . ., I_{n+m}. A weak solution at J is a collection of functions ρ_l :

$[0,+\infty[\times I_l \to \mathbb{R}$, $l = 1,\ldots,n+m$, such that

$$\sum_{l=0}^{n+m}\left(\int_0^{+\infty}\int_{a_l}^{b_l}\left(\rho_l\frac{\partial\varphi_l}{\partial t}+f(\rho_l)\frac{\partial\varphi_l}{\partial x}\right)dxdt\right)=0, \tag{5.1.6}$$

for every φ_l, $l = 1,\ldots,n+m$, with compact support in $]0,+\infty[\times]a_l,b_l]$ for $l = 1,\ldots,n$ (incoming roads) and in $]0,+\infty[\times[a_l,b_l[$ for $l = n+1,\ldots,n+m$ (outgoing roads), that are also *smooth across the junction*, i.e.

$$\varphi_i(\cdot,b_i)=\varphi_j(\cdot,a_j),\ \frac{\partial\varphi_i}{\partial x}(\cdot,b_i)=\frac{\partial\varphi_j}{\partial x}(\cdot,a_j),\ i=1,...,n,\ j=n+1,...,n+m.$$

Remark 6 *Let $\rho = (\rho_1,\ldots,\rho_{n+m})$ be a weak solution at the junction such that each $x \to \rho_i(t,x)$ has bounded variation. We can deduce that ρ satisfies the Rankine-Hugoniot Condition at the junction J, namely*

$$\sum_{i=1}^{n} f(\rho_i(t,b_i-))=\sum_{j=n+1}^{n+m} f(\rho_j(t,a_j+)), \tag{5.1.7}$$

for almost every $t > 0$.

The rules (A) and (B) can be given explicitly only for solutions with bounded variation as in the next definition.

Definition 5.1.1 *Let $\rho = (\rho_1,\ldots,\rho_{n+m})$ be such that $\rho_i(t,\cdot)$ is of bounded variation for every $t \geqslant 0$. Then ρ is an admissible weak solution of (5.0.1) related to the matrix A, satisfying (5.0.3), at the junction J if and only if the following properties hold:*

(i) ρ is a weak solution at the junction J;

(ii) $f(\rho_j(\cdot,a_j+)) = \sum_{i=1}^{n}\alpha_{ji}f(\rho_i(\cdot,b_i-))$, for each $j = n+1,...,n+m$;

(iii) $\sum_{i=1}^{n} f(\rho_i(\cdot,b_i-))$ is maximum subject to (ii).

For every road $I_i = [a_i,b_i]$, if $a_i > -\infty$ and I_i is not the outgoing road of any junction, or $b_i < +\infty$ and I_i is not the incoming road of any junction, then a boundary data $\psi_i : [0,+\infty[\to \mathbb{R}$ is given. In this case we ask ρ_i to satisfy $\rho_i(t,a_i) = \psi_i(t)$ (or $\rho_i(t,b_i) = \psi_i(t)$) in the sense of [16]. The treatment of boundary data in the sense of [16] can be done in the same way as in [6, 9], thus we treat the case without boundary data. All the stated results hold also for the case with boundary data with obvious modifications.

Our aim is to solve the Cauchy problem on $[0, +\infty[$ for a given initial and boundary data as in next definition.

Definition 5.1.2 *Given $\bar{\rho}_i : I_i \to \mathbb{R}$, $i = 1, \ldots, N$, L^∞ functions, a collection of functions $\rho = (\rho_1, \ldots, \rho_N)$, with $\rho_i : [0, +\infty[\times I_i \to \mathbb{R}$ continuous as functions from $[0, +\infty[$ into L^1_{loc}, is an admissible solution if ρ_i is a weak entropic solution to (5.0.1) on I_i, $\rho_i(0, x) = \bar{\rho}_i(x)$ a.e., at each junction ρ is a weak solution and is an admissible weak solution in case of bounded variation.*

On the flux f we make the following assumption

($\mathcal{F}$) $f : [0, 1] \to \mathbb{R}$ is smooth, strictly concave (i.e. $f'' \leqslant -c < 0$ for some $c > 0$), $f(0) = f(1) = 0$. Therefore there exists a unique $\sigma \in]0, 1[$ such that $f'(\sigma) = 0$ (that is σ is a strict maximum).

5.2 The Riemann Problem.

For a scalar conservation law a Riemann problem is a Cauchy problem for an initial data of Heaviside type, that is piecewise constant with only one discontinuity. One looks for centered solutions, i.e. $\rho(t, x) = \phi(\frac{x}{t})$, which are the building blocks to construct solutions to the Cauchy problem via wave front tracking algorithm. These solutions are formed by continuous waves called rarefactions and by traveling discontinuities called shocks. The speed of waves are related to the values of f', see [26].

Analogously, we call Riemann problem for the road network the Cauchy problem corresponding to an initial data that is piecewise constant on each road. The solutions on each road I_i can be constructed in the same way as for the scalar conservation law, hence it suffices to describe the solution at junctions. Because of finite propagation speed, it is enough to study the Riemann Problem for a single junction.

Consider a junction J in which there are n roads with incoming traffic and m roads with outgoing traffic, and a traffic distribution matrix A. For simplicity we indicate by

$$(t, x) \in \mathbb{R}_+ \times I_i \mapsto \rho_i(t, x) \in [0, 1], \quad i = 1, ..., n, \tag{5.2.8}$$

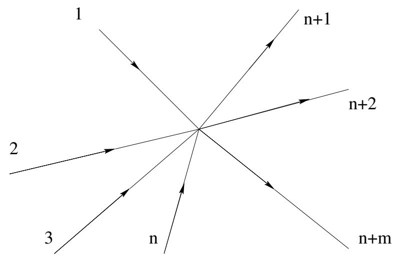


Figure 5.1: a junction.

the densities of the cars on the roads with incoming traffic and

$$(t,x) \in \mathbb{R}_+ \times I_j \mapsto \rho_j(t,x) \in [0,1], \quad j = n+1, ..., n+m \tag{5.2.9}$$

those on the roads with outgoing traffic, see Figure 5.1.

We need some more notation:

Definition 5.2.1 *Let $\tau : [0,1] \to [0,1]$ be the map such that:*

1. *$f(\tau(\rho)) = f(\rho)$ for every $\rho \in [0,1]$;*
2. *$\tau(\rho) \neq \rho$ for every $\rho \in [0,1] \setminus \{\sigma\}$.*

Clearly, τ is well defined and satisfies

$$0 \leqslant \rho \leqslant \sigma \Longleftrightarrow \sigma \leqslant \tau(\rho) \leqslant 1, \qquad \sigma \leqslant \rho \leqslant 1 \Longleftrightarrow 0 \leqslant \tau(\rho) \leqslant \sigma.$$

To state the main result of this section we need some assumption on the matrix A satisfied under generic conditions. Let $\{e_1, \ldots, e_n\}$ be the canonical basis of $\mathbb{R}^n$ and for every subset $V \subset \mathbb{R}^n$ indicate by $V^\perp$ its orthogonal. Define for every $i = 1, \ldots, n$, $H_i = \{e_i\}^\perp$, i.e. the coordinate hyperplane orthogonal to e_i and for every $j = n+1, \ldots, n+m$ let $\alpha_j = (\alpha_{j1}, \ldots, \alpha_{jn}) \in \mathbb{R}^n$ and define $H_j = \{\alpha_j\}^\perp$. Let $\mathcal{K}$ be the set of indices $k = (k_1, ..., k_\ell)$, $1 \leq \ell \leq n-1$, such that $0 \leq k_1 < k_2 < \cdots < k_\ell \leq n+m$ and for every $k \in \mathcal{K}$ set

$$H_k = \bigcap_{h=1}^{\ell} H_{k_h}.$$

Letting $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^n$, we assume

(C) for every $k \in \mathcal{K}$, $\mathbf{1} \notin H_k^{\perp}$.

Remark 7 *Condition (C) is a technical condition, which allows us to have uniqueness to the maximization problem described in Theorem 5.2.1. From (C) we immediately derive* $m \geqslant n$. *Otherwise, since by definition* $\mathbf{1} = \sum_{j=n+1}^{n+m} \alpha_j$, *we get* $\mathbf{1} \in H_k^{\perp}$, *where*

$$H_k = \cap_{j=n+1}^{n+m} H_j.$$

Moreover if $n \geqslant 2$, *then (C) implies that, for every* $j \in \{n+1, \ldots, n+m\}$ *and for every distinct elements* $i, i' \in \{1, \ldots, n\}$, *it holds* $\alpha_{ji} \neq \alpha_{ji'}$. *Otherwise, without loss of generalities, we may suppose that* $\alpha_{n+1,1} = \alpha_{n+1,2}$. *If we consider*

$$H = (\cap_{2<j\leqslant n} H_j) \cap H_{n+1},$$

then, by (C), there exists an element $(x_1, x_2, 0, \ldots, 0) \in H$ *such that* $x_1 + x_2 \neq 0$ *and* $\alpha_{n+1,1}(x_1 + x_2) = 0$.

In the case of a simple junction J *with* 2 *incoming roads and* 2 *outgoing ones, the condition (C) is completely equivalent to the fact that, for every* $j \in \{3, 4\}$, $\alpha_{j1} \neq \alpha_{j2}$.

Remark 8 *Notice that the matrix* A *could have identical lines. For example the matrix*

$$A = \begin{pmatrix} \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{2} & \frac{3}{5} \end{pmatrix}$$

satisfies the condition (C).

Theorem 5.2.1 *Consider a junction* J, *assume that the flux* $f : [0,1] \to \mathbb{R}$ *satisfies* $(\mathcal{F})$ *and the matrix* A *satisfies condition (C). For every initial datum* $\rho_{1,0}, ..., \rho_{n+m,0} \in [0,1]$, *there exists a unique admissible centered weak solution, in the sense of Definition 5.1.1,* $\rho = \big(\rho_1, ..., \rho_{n+m}\big)$ *of (5.0.1) at the junction* J *such that*

$$\rho_1(0, \cdot) \equiv \rho_{1,0}, ..., \rho_{n+m}(0, \cdot) \equiv \rho_{n+m,0}.$$

Moreover, there exists a unique $(n+m)-$*tuple* $(\hat{\rho}_1, ..., \hat{\rho}_{n+m}) \in [0,1]^{n+m}$ *such that*

$$\hat{\rho}_i \in \begin{cases} \{\rho_{i,0}\} \cup]\tau(\rho_{i,0}), 1], & \textit{if } 0 \leqslant \rho_{i,0} \leqslant \sigma, \\ [\sigma, 1], & \textit{if } \sigma \leqslant \rho_{i,0} \leqslant 1, \end{cases} \qquad i = 1, ..., n, \tag{5.2.10}$$

and

$$\hat{\rho}_j \in \begin{cases} [0,\sigma], & \text{if } 0 \leqslant \rho_{j,0} \leqslant \sigma, \\ \{\rho_{j,0}\} \cup [0,\tau(\rho_{j,0})[, & \text{if } \sigma \leqslant \rho_{j,0} \leqslant 1, \end{cases} \qquad j = n+1,...,n+m. \tag{5.2.11}$$

Fixed $i \in \{1,...,n\}$, *if* $\rho_{i,0} \leqslant \hat{\rho}_i$, *we have*

$$\rho_i(t,x) = \begin{cases} \rho_{i,0}, & \text{if } x < \frac{f(\hat{\rho}_i)-f(\rho_{i,0})}{\hat{\rho}_i-\rho_{i,0}}t + b_i,\ t \geqslant 0, \\ \hat{\rho}_i, & \text{if } x > \frac{f(\hat{\rho}_i)-f(\rho_{i,0})}{\hat{\rho}_i-\rho_{i,0}}t + b_i,\ t \geqslant 0, \end{cases} \tag{5.2.12}$$

and, if $\hat{\rho}_i < \rho_{i,0}$,

$$\rho_i(t,x) = \begin{cases} \rho_{i,0}, & \text{if } x \leqslant f'(\rho_{i,0})t + b_i,\ t \geqslant 0, \\ \left(f'\right)^{-1}\big((x-b_i)/t\big), & \text{if } f'(\rho_{i,0})t + b_i \leqslant x \leqslant f'(\hat{\rho}_i)t + b_i,\ t \geqslant 0, \\ \hat{\rho}_i, & \text{if } x > f'(\hat{\rho}_i)t + b_i,\ t \geqslant 0. \end{cases} \tag{5.2.13}$$

Fixed $j \in \{n+1,...,n+m\}$, *if* $\rho_{j,0} \leqslant \hat{\rho}_j$, *we have*

$$\rho_j(t,x) = \begin{cases} \hat{\rho}_j, & \text{if } x \leqslant f'(\hat{\rho}_j)t + a_j,\ t \geqslant 0, \\ \left(f'\right)^{-1}\big((x-a_j)/t\big), & \text{if } f'(\hat{\rho}_j)t + a_j \leqslant x \leqslant f'(\rho_{j,0})t + a_j,\ t \geqslant 0, \\ \rho_{j,0}, & \text{if } x > f'(\rho_{j,0})t + a_j,\ t \geqslant 0, \end{cases} \tag{5.2.14}$$

and, if $\hat{\rho}_j < \rho_{j,0}$,

$$\rho_j(t,x) = \begin{cases} \hat{\rho}_j, & \text{if } x < \frac{f(\rho_{j,0})-f(\hat{\rho}_j)}{\rho_{j,0}-\hat{\rho}_j}t + a_j,\ t \geqslant 0, \\ \rho_{j,0}, & \text{if } x > \frac{f(\rho_{j,0})-f(\hat{\rho}_j)}{\rho_{j,0}-\hat{\rho}_j}t + a_j,\ t \geqslant 0. \end{cases} \tag{5.2.15}$$

Proof. Define the map

$$E : (\delta_1,...,\delta_n) \in \mathbb{R}^n \longmapsto \sum_{i=1}^{n} \delta_i$$

and the sets

$$\Omega_i \doteq \begin{cases} [0, f(\rho_{i,0})], & \text{if } 0 \leqslant \rho_{i,0} \leqslant \sigma, \\ [0, f(\sigma)], & \text{if } \sigma \leqslant \rho_{i,0} \leqslant 1, \end{cases} \qquad i = 1,...,n,$$

$$\Omega_j \doteq \begin{cases} [0, f(\sigma)], & \text{if } 0 \leqslant \rho_{j,0} \leqslant \sigma, \\ [0, f(\rho_{j,0})], & \text{if } \sigma \leqslant \rho_{j,0} \leqslant 1, \end{cases} \qquad j = n+1,...,n+m,$$

$$\Omega \doteq \Big\{(\delta_1, ..., \delta_n) \in \Omega_1 \times \ldots \times \Omega_n \Big| A \cdot (\delta_1, ..., \delta_n)^T \in \Omega_{n+1} \times \ldots \times \Omega_{n+m}\Big\}.$$

The set Ω is closed, convex and not empty. Moreover, by (C), $\nabla E = \mathbf{1}$ is not orthogonal to any nontrivial subspace contained in a supporting hyperplane of Ω, hence there exists a unique vector $(\hat{\delta}_1, ..., \hat{\delta}_n) \in \Omega$ such that

$$E(\hat{\delta}_1, ..., \hat{\delta}_n) = \max_{(\delta_1,...,\delta_n)\in\Omega} E(\delta_1, ..., \delta_n).$$

For every $i \in \{1, ..., n\}$, we choose $\hat{\rho}_i \in [0, 1]$ such that

$$f(\hat{\rho}_i) = \hat{\delta}_i, \quad \hat{\rho}_i \in \begin{cases} \{\rho_{i,0}\} \cup]\tau(\rho_{i,0}), 1], & \text{if } 0 \leqslant \rho_{i,0} \leqslant \sigma, \\ [\sigma, 1], & \text{if } \sigma \leqslant \rho_{i,0} \leqslant 1. \end{cases}$$

By $(\mathcal{F})$, $\hat{\rho}_i$ exists and is unique. Let

$$\hat{\delta}_j \doteq \sum_{i=1}^{n} \alpha_{ji} \hat{\delta}_i, \qquad j = n + 1, ..., n + m,$$

and $\hat{\rho}_j \in [0, 1]$ be such that

$$f(\hat{\rho}_j) = \hat{\delta}_j, \quad \hat{\rho}_j \in \begin{cases} [0, \sigma], & \text{if } 0 \leqslant \rho_{j,0} \leqslant \sigma, \\ \{\rho_{j,0}\} \cup [0, \tau(\rho_{j,0})[, & \text{if } \sigma \leqslant \rho_{j,0} \leqslant 1. \end{cases}$$

Since $(\hat{\delta}_1, ..., \hat{\delta}_n) \in \Omega$, $\hat{\rho}_j$ exists and is unique for every $j \in \{n+1, \ldots, n+m\}$. Solving the Riemann Problem (see [26, Chapter 6]) on each road, the claim is proved. □

5.3 The Wave Front Tracking Algorithm.

Once the solution to a Riemann problem is provided, we are able to construct piecewise constant approximations via wave-front tracking algorithm. The construction is very similar to that for scalar conservation law, see [26], hence we briefly describe it.

Let $\bar{\rho} = (\rho_1, \ldots, \rho_N)$ be a piecewise constant map defined on the road network. We want to construct a weak solution of (5.0.1) with initial condition $\rho(0, \cdot) \equiv \bar{\rho}$. We begin by solving the Riemann Problems on each road in correspondence of the jumps of $\bar{\rho}$ and the Riemann Problems at junctions determined by the values of $\bar{\rho}$ (see Theorem 5.2.1). We split each rarefaction wave into a rarefaction fan formed by rarefaction shocks, that are discontinuities traveling with the Rankine-Hugoniot speed. We always split rarefaction waves inserting the value σ (if it is in the range of the rarefaction). Moreover, we let any rarefaction shock with endpoint σ have velocity zero.

When a wave interacts with another one we simply solve the new Riemann Problem. Instead, when a wave reaches a junction, we solve the Riemann Problem at the junction. The number of waves may increase only for interactions of waves at junctions. Since the speeds of waves are bounded, there are finitely many waves on the network at each time $t \geqslant 0$. We call the obtained function *an approximate wave front tracking solution*. Given a general initial data, we approximate it by a sequence of piecewise constant functions and construct the corresponding approximate solutions. If they converge in L^1_{loc}, then the limit is a weak entropic solution on each road, see [26] for a proof.

5.4 Estimates on Flux Variation and Existence of Solutions.

This Section is devoted to the estimation of the total variation of the flux along an approximate wave front tracking solution and to the construction of solutions to the Cauchy problem. From now on, we assume that every junction has exactly two incoming roads and two outgoing ones. This hypothesis is crucial, because, as shown in Section 5.7, the presence of more complicate junctions provokes additional increases of the total variation of the flux. The case where junctions have at most two incoming roads and at most two outgoing roads can be treated in the same way. So, for each junction J, the matrix A, defined in the introduction, takes the form

$$A = \begin{pmatrix} \alpha & \beta \\ 1-\alpha & 1-\beta \end{pmatrix}, \tag{5.4.16}$$

where $\alpha, \beta \in]0,1[$ and $\alpha \neq \beta$, so that (C) is satisfied.

From now on we fix an approximate wave front tracking solution ρ, defined on the road network.

Definition 5.4.1 *For every road I_i, $i = 1, \ldots, N$, we indicate by*

$$(\rho_-^\theta, \rho_+^\theta), \quad \theta \in \Theta = \Theta(\rho, t, i), \quad \Theta \text{ finite set},$$

the discontinuities on road I_i at time t, and by $x^\theta(t)$, $\lambda^\theta(t)$, $\theta \in \Theta$, respectively their positions and velocities at time t. We also refer to the wave θ to indicate the discontinuity $(\rho_-^\theta, \rho_+^\theta)$.

For each discontinuity $(\rho_-^\theta, \rho_+^\theta)$ at time $\bar{t}$ on road I_i, we call $y^\theta(t)$, $t \in [\bar{t}, t_\theta]$, the trace of the wave so defined. We start with $y^\theta(\bar{t}) = x^\theta(\bar{t})$ and we continue up to the first interaction with another wave or a junction. If at time $\tilde{t}$ an interaction with a wave or a junction occurs, then either a single new wave $(\rho_-^{\tilde{\theta}}, \rho_+^{\tilde{\theta}})$ on road I_i is produced or no wave is produced. In the latter case we set $t_\theta = \tilde{t}$, otherwise we set $y^\theta(\tilde{t}) = x^{\tilde{\theta}}(\tilde{t})$ and follow $x^{\tilde{\theta}}(t)$ for $t \geqslant t^{\tilde{\theta}}$ up to next interaction and so on.

We start by proving some technical lemmata.

Lemma 5.4.1 *Fix a junction J and an incoming road I_i. Let θ be a wave on road I_i, originated at time $\bar{t}$ from J with a flux decrease, i.e. $x^\theta(\bar{t}) = b_i$, $\lambda^\theta(\bar{t}) < 0$ and $f(\rho_+^\theta) < f(\rho_-^\theta)$. Let y^θ be the traced wave and assume that there exists $\tilde{t}$, the first time of interaction of y^θ with J after $\bar{t}$. Then either y^θ interacts with another junction on $]\bar{t}, \tilde{t}[$ or, letting $\theta_1, \dots, \theta_l$ be the waves interacting with y^θ at times $t_m \in]\bar{t}, \tilde{t}[$, $m = 1, \dots, l$, $(t_1 < t_2 < \dots < t_l)$, we have:*

$$\begin{aligned} & \left|f(\rho(\tilde{t} - \varepsilon, y^\theta(\tilde{t} - \varepsilon)+)) - f(\rho(\tilde{t} - \varepsilon, y^\theta(\tilde{t} - \varepsilon)-))\right| \\ \leqslant & \sum_{m=1}^{l} \left|f(\rho(t_m - \varepsilon, x^{\theta_m}(t_m - \varepsilon)+)) - f(\rho(t_m - \varepsilon, x^{\theta_m}(t_m - \varepsilon)-))\right| \\ & - \left|f(\rho_-^\theta) - f(\rho_+^\theta)\right|, \end{aligned}$$

for $\varepsilon > 0$ small enough. This means that the initial flux variation along y^θ is canceled. The same conclusion holds for an outcoming road I_j.

Proof. Consider the wave $(\rho_-^\theta, \rho_+^\theta)$ as in the statement, then it is a shock with negative velocity and $\rho_+^\theta > \max\{\rho_-^\theta, \tau(\rho_-^\theta)\}$. If y^θ interacts with another junction, then there is nothing to prove. So, we assume that y^θ does not interact with another junction. At time t_1, the wave θ_1 interacts with y^θ. We analyze first the case of interaction from the left of y^θ. We have two possibilities:

1. $\rho_-^{\theta_1} \in [0, \tau(\rho_+^\theta)]$. In this case we have total cancellation of the flux variation and so

$$\left|f(\rho_+^\theta) - f(\rho_-^{\theta_1})\right| = \left|f(\rho_-^{\theta_1}) - f(\rho_-^\theta)\right| - \left|f(\rho_-^\theta) - f(\rho_+^\theta)\right|.$$

Therefore the claim easily follows.

2. $\rho_-^{\theta_1} \in]\tau(\rho_+^\theta), \rho_+^\theta]$. In this case the wave y^θ after the time interaction t_1 is of the same type of y^θ before t_1, i.e.

$$\max\{\rho(t_1, y^\theta(t_1)-), \tau(\rho(t_1, y^\theta(t_1)-))\} < \rho(t_1, y^\theta(t_1)+).$$

We consider now the case of interaction from the right of y^θ. It is clear that $\rho_+^{\theta_1} \in]\rho_-^\theta, 1]$. If moreover $f(\rho_+^{\theta_1}) \geqslant f(\rho_-^\theta)$, then we have total cancellation of the flux and we conclude as before. If instead $f(\rho_+^{\theta_1}) < f(\rho_-^\theta)$, then the wave y^θ after the time t_1 is of the same type of y^θ before t_1.

We repeat this argument at each interaction time t_m. If at some t_m we have total cancellation of the flux, then we conclude. Therefore we may suppose that at each t_m total cancellation of the flux does not occur. Since the type of the wave y^θ does not change, we have

$$\max\{\rho(\tilde{t} - \tilde{\varepsilon}, y^\theta(\tilde{t} - \tilde{\varepsilon})-), \tau(\rho(\tilde{t} - \tilde{\varepsilon}, y^\theta(\tilde{t} - \tilde{\varepsilon})-))\} < \rho(\tilde{t} - \tilde{\varepsilon}, y^\theta(\tilde{t} - \tilde{\varepsilon})+)$$

for $\tilde{\varepsilon} > 0$ small enough and hence the speed $\lambda^\theta(\tilde{t} - \tilde{\varepsilon})$ is negative, which contradicts the fact that y^θ interacts with J at time $\tilde{t}$. □

Lemma 5.4.2 *Fix a junction J and an incoming road I_i. Let θ be a wave on road I_i, originated at time $\bar{t}$ from J by a flux increase, i.e. $x^\theta(\bar{t}) = b_i$, $\lambda^\theta(\bar{t}) < 0$ and $f(\rho_+^\theta) > f(\rho_-^\theta)$. Let y^θ be the traced wave and assume that there exists $\tilde{t}$, the first time of interaction of y^θ with J after $\bar{t}$. Then y^θ interacts with other junctions in $]\bar{t}, \tilde{t}[$ or y^θ cancels the flux variation, or it produces a flux decrease at J at $\tilde{t}$, i.e.*

$$f(\rho(\tilde{t} - \varepsilon, y^\theta(\tilde{t} - \varepsilon)-)) < f(\rho(\tilde{t} - \varepsilon, y^\theta(\tilde{t} - \varepsilon)+)),$$

for $\varepsilon > 0$ small enough. The same holds for outgoing roads.

Proof. Since $\lambda^\theta(\bar{t}) < 0$ and $f(\rho_+^\theta) > f(\rho_-^\theta)$, then $\rho_-^\theta > \sigma$. Moreover the wave $(\rho_-^\theta, \rho_+^\theta)$ is a rarefaction fan, hence $\sigma < \rho_+^\theta < \rho_-^\theta$.

If an interaction on the right with a wave θ_1 happens, then $\rho_+^{\theta_1} \in]\rho_-^\theta, 1]$ and we have total cancellation of the flux variation. Therefore we may suppose that an interaction on the left with a wave θ_1 happens. In this case we have two possibilities:

1. $\rho_-^{\theta_1} \in [0, \tau(\rho_+^\theta)[$;

2. $\rho_-^{\theta_1} \in [\tau(\rho_+^\theta), \rho_+^\theta[$.

In the latter case we have total cancellation of the flux variation and so we conclude. In the first case, instead, the type of the wave changes, since

$$0 < \rho_-^{\theta_1} < \tau(\rho_+^\theta) \leqslant \sigma \leqslant \rho_+^\theta < 1.$$

The speed of the wave y^θ after this interaction is positive and if there are no more interaction, then we have the claim since $f(\rho_-^{\theta_1}) < f(\rho_+^\theta)$. Thus we suppose that an interaction with a wave θ_2 happens. If it is an interaction from the left, then the possibilities are the followings:

1. $\rho_-^{\theta_2} \in [0, \tau(\rho_+^\theta)[$. We do not have total cancellation of the flux variation, but the type of the wave does not change and the situation is identical to the previous one.

2. $\rho_-^{\theta_2} \in [\tau(\rho_+^\theta), \sigma[$. We have total cancellation of the flux variation and so we conclude.

If it is an interaction from the right, then the possibilities are the followings:

1. $\rho_+^{\theta_2} \in [\sigma, \tau(\rho_-^{\theta_1})[$. We do not have total cancellation of the flux variation, but the type of the wave does not change.

2. $\rho_+^{\theta_2} \in [\tau(\rho_-^{\theta_1}), 1]$. We have total cancellation of the flux variation and so we conclude.

The conclusion now easily follows repeating this argument. If at each interaction we do not have total cancellation of the flux variation, then we necessarily have that

$$f(\rho(\tilde{t} - \varepsilon, y^\theta(\tilde{t} - \varepsilon)-)) < f(\rho(\tilde{t} - \varepsilon, y^\theta(\tilde{t} - \varepsilon)+)),$$

for $\varepsilon > 0$ small enough, which concludes the proof. □

Lemma 5.4.3 *Fix a junction J. If a wave interacts with the junction J from an incoming road at time $\bar{t}$, then*

$$Tot.Var.\big(f(\rho(\bar{t}+, \cdot))\big) = Tot.Var.\big(f(\rho(\bar{t}-, \cdot))\big). \tag{5.4.17}$$

Proof. For simplicity let us assume that I_1, I_2 are the incoming roads and I_3, I_4 are the outgoing ones. Let $(\rho_{1,0}, ..., \rho_{4,0})$ be an equilibrium configuration at the junction J. We assume that the wave is coming from the first road and that it is given by the values $(\rho_1, \rho_{1,0})$. Let us define the incoming flux

$$f^{in}(y) \doteq \begin{cases} f(y), & \text{if } 0 \leqslant y \leqslant \sigma, \\ f(\sigma), & \text{if } \sigma \leqslant y \leqslant 1, \end{cases} \tag{5.4.18}$$

and the outgoing flux

$$f^{out}(y) \doteq \begin{cases} f(\sigma), & \text{if } 0 \leqslant y \leqslant \sigma, \\ f(y), & \text{if } \sigma \leqslant y \leqslant 1. \end{cases} \tag{5.4.19}$$

Clearly, since the wave on the first road has positive velocity, we have

$$0 \leqslant \rho_1 < \sigma. \tag{5.4.20}$$

Let $(\hat{\rho}_1, ..., \hat{\rho}_4)$ be the solution of the Riemann Problem in the junction J with initial data $(\rho_1, \rho_{2,0}, \rho_{3,0}, \rho_{4,0})$ (see Theorem 5.2.1). By definition, $\big(f(\rho_{1,0}), f(\rho_{2,0})\big)$ is the maximum point of the map E on the domain

$$\Omega_0 \doteq \Big\{(\delta_1, \delta_2) \in \Omega_{1,0} \times \Omega_{2,0} \Big| A \cdot (\delta_1, \delta_2)^T \in \Omega_{3,0} \times \Omega_{4,0}\Big\},$$

and $\big(f(\hat{\rho}_1), f(\hat{\rho}_2)\big)$ is the maximum point of the map E on the domain

$$\hat{\Omega} \doteq \Big\{(\delta_1, \delta_2) \in \Omega_1 \times \Omega_{2,0} \Big| A \cdot (\delta_1, \delta_2)^T \in \Omega_{3,0} \times \Omega_{4,0}\Big\},$$

where

$$\Omega_{j,0} \doteq \begin{cases} [0, f^{in}(\rho_{j,0})], & \text{if } j = 1, 2, \\ [0, f^{out}(\rho_{j,0})], & \text{if } j = 3, 4, \end{cases}$$

and, by (5.4.20),

$$\Omega_1 \doteq [0, f^{in}(\rho_1)] = [0, f(\rho_1)].$$

It is also clear that

$$\big(f(\rho_{1,0}), f(\rho_{2,0})\big) \in \partial\Omega_0, \qquad \big(f(\hat{\rho}_1), f(\hat{\rho}_2)\big) \in \partial\hat{\Omega}.$$

For simplicity we use the notation (5.4.16).

We distinguish two cases. First we suppose that

$$f(\rho_1) < f(\rho_{1,0}), \tag{5.4.21}$$

(equality can not happen in the previous equation because the wave would have velocity zero). Then $\hat{\Omega} \subset \Omega_0$ and

$$f(\hat{\rho}_1) \leqslant f(\rho_1), \qquad f(\hat{\rho}_1) + f(\hat{\rho}_2) \leqslant f(\rho_{1,0}) + f(\rho_{2,0}). \tag{5.4.22}$$

We claim that

$$f(\rho_{2,0}) \leqslant f(\hat{\rho}_2), \quad f(\hat{\rho}_3) \leqslant f(\rho_{3,0}), \quad f(\hat{\rho}_4) \leqslant f(\rho_{4,0}). \tag{5.4.23}$$

The points $\big(f(\rho_{1,0}), f(\rho_{2,0})\big)$, $\big(f(\hat{\rho}_1), f(\hat{\rho}_2)\big)$ are on the boundaries of Ω_0, $\hat{\Omega}$ respectively, where the function E attains the maximum, hence each one is at least on one of the curves

$$\alpha\delta_1 + \beta\delta_2 = f^{out}(\rho_{3,0}),\ (1-\alpha)\delta_1 + (1-\beta)\delta_2 = f^{out}(\rho_{4,0}),\ \delta_2 = f^{in}(\rho_{2,0}).$$

Let us assume that the two points are on the same curve, the other cases being similar,

$$\alpha\delta_1 + \beta\delta_2 = f^{out}(\rho_{3,0}). \tag{5.4.24}$$

Observe that the map E is increasing on the curve

$$\delta_1 \mapsto \left(\delta_1, \frac{f^{out}(\rho_{3,0})}{\beta} - \frac{\alpha}{\beta}\delta_1\right),$$

otherwise we contradict the maximality of E at $\big(f(\rho_{1,0}), f(\rho_{2,0})\big)$. Thus $\alpha < \beta$, $\hat{\rho}_1 = \rho_1$, the first two inequalities in (5.4.23) hold and

$$f(\hat{\rho}_1) = f(\rho_1), \qquad f(\hat{\rho}_2) > f(\rho_{2,0}), \qquad f(\hat{\rho}_3) = f(\rho_{3,0}) = f^{out}(\rho_{3,0}). \tag{5.4.25}$$

On the other hand, by (5.4.22), we have

$$\begin{aligned} f(\hat{\rho}_4) &= (1-\alpha)f(\hat{\rho}_1) + (1-\beta)f(\hat{\rho}_2) \leqslant \\ &\leqslant (1-\alpha)\big(f(\rho_{1,0}) + f(\rho_{2,0}) - f(\hat{\rho}_2)\big) + (1-\beta)f(\hat{\rho}_2) = \\ &= (1-\alpha)\big(f(\rho_{1,0}) + f(\rho_{2,0})\big) + \big(\alpha-\beta\big)f(\hat{\rho}_2) \leqslant \\ &\leqslant (1-\alpha)\big(f(\rho_{1,0}) + f(\rho_{2,0})\big) + \big(\alpha-\beta\big)f(\rho_{2,0}) = f(\rho_{4,0}). \end{aligned}$$

Thus (5.4.23) holds. Using the Rankine–Hugoniot Condition (5.1.7) at the junction J, and using (5.4.23), and (5.4.25), we get

$$\begin{gathered} \text{Tot.Var.}\big(f(\rho(\bar{t}+,\cdot))\big) = \\ = |f(\hat{\rho}_1) - f(\rho_1)| + |f(\hat{\rho}_2) - f(\rho_{2,0})| + |f(\hat{\rho}_3) - f(\rho_{3,0})| + |f(\hat{\rho}_4) - f(\rho_{4,0})| \\ = \big(f(\hat{\rho}_2) - f(\rho_{2,0})\big) + \big(f(\rho_{3,0}) - f(\hat{\rho}_3)\big) + \big(f(\rho_{4,0}) - f(\hat{\rho}_4)\big) \\ = f(\rho_{1,0}) - f(\hat{\rho}_1) = f(\rho_{1,0}) - f(\rho_1) = \text{Tot.Var.}\big(f(\rho(\bar{t}-,\cdot))\big). \end{gathered}$$

Suppose now that

$$f(\rho_{1,0}) < f(\rho_1),$$

then $\rho_{1,0} < \rho_1 < \sigma$ and $\Omega_0 \subset \hat{\Omega}$. Assuming again that both points of maximum of the function E are on the curve (5.4.24), we have

$$f(\hat{\rho}_1) = f(\rho_1), \quad f(\hat{\rho}_2) \leqslant f(\rho_{2,0}), \quad f(\rho_{3,0}) = f(\hat{\rho}_3), \quad f(\rho_{4,0}) \leqslant f(\hat{\rho}_4).$$

By the Rankine Hugoniot Condition at the junction J (see (5.1.7)), we have

$$\begin{aligned} &\text{Tot.Var.}\big(f(\rho(\bar{t}+,\cdot))\big) = \\ &= |f(\hat{\rho}_1) - f(\rho_1)| + |f(\hat{\rho}_2) - f(\rho_{2,0})| + |f(\hat{\rho}_3) - f(\rho_{3,0})| + |f(\hat{\rho}_4) - f(\rho_{4,0})| \\ &= \big(f(\rho_{2,0}) - f(\hat{\rho}_2)\big) + \big(f(\hat{\rho}_3) - f(\rho_{3,0})\big) + \big(f(\hat{\rho}_4) - f(\rho_{4,0}))\big) \\ &= f(\hat{\rho}_1) - f(\rho_{1,0}) = f(\rho_1) - f(\rho_{1,0}) = \text{Tot.Var.}\big(f(\rho(\bar{t}-,\cdot))\big). \end{aligned}$$

This concludes the proof. □

Lemma 5.4.4 *Consider a network $(\mathcal{I}, \mathcal{J})$. We have*

$$\textit{Tot.Var.}\big(f(\rho(0+,\cdot))\big) \leqslant \textit{Tot.Var.}\big(f(\rho(0,\cdot))\big) + 2Rf(\sigma),$$

where R is the total number of roads of the network.

Proof. At time $t = 0$ we can have an instantaneous increase of the total variation of the flux due to the waves generated by the Riemann problems in the junctions. Clearly, this increase can be estimated by the maximum number of waves generated in the junctions ($\leqslant 2R$) times the maximum variation of the flux on each road ($\leqslant f(\sigma)$). □

We are now ready to prove the following.

Lemma 5.4.5 *Consider a road network $(\mathcal{I}, \mathcal{J})$. For some $K > 0$, we have*

$$\begin{aligned} \textit{Tot.Var.}\big(f(\rho(t+,\cdot))\big) &\leq e^{Kt}\,\textit{Tot.Var.}\big(f(\rho(0+,\cdot))\big) \leqslant \\ &\leqslant e^{Kt}\big(\textit{Tot.Var.}\big(f(\rho(0,\cdot))\big) + 2Rf(\sigma)\big), \end{aligned}$$

for each $t \geqslant 0$.

Proof. Fix a junction J. Notice that there exists a constant C_J, depending on the coefficients of the matrix A at J, so that each interaction of a wave with J provokes an increase of flux variation at most by a factor C_J. More precisely, if $\text{Tot.Var.}_f^{\pm}$ is the flux variation of waves before and after the interaction then $\text{Tot.Var.}_f^{+} \leqslant C_J \text{Tot.Var.}_f^{-}$.

Consider a wave θ interacting with the junction J, then from Lemma 5.4.3 the flux variation can increase only if the wave is coming from an outgoing road. Let $\theta_1, \ldots, \theta_4$ be the waves so produced. Thanks to Lemma 5.4.1 waves produced by a flux decrease can not interact with the junction J without canceling the flux variation or reaching another junction. Moreover, by Lemma 5.4.2, every θ_i can come back to the junction J (without interacting with other junctions) only with a decrease of the flux. Now notice that a wave with decreasing flux interacting with J always produces a flux decrease on outgoing roads. Hence, waves θ_i may come back to the junction only with decreasing flux, thus, by Lemma 5.4.1, producing other waves that can not come back to the junction, unless they cancel their flux variation or interact with other junctions. Finally, each wave flux variation can be magnified just twice by a factor C_J interacting only with junction J and not with other junctions.

Now let η be the minimum length of a road, i.e. $\eta = \min_{i\in\mathcal{I}}(b_i - a_i)$, and $\hat{\lambda}$ be the maximum speed of a wave, i.e. $\hat{\lambda} = \max\{f'(0), |f'(1)|\}$. Then each wave takes at least time $\eta/\hat{\lambda}$ to go from one junction to another.

Finally, recalling that the total variation of the flux may only decrease for interactions on roads, we get that a magnification of flux variation of a factor $C_{\mathcal{J}} = \max_{J\in\mathcal{J}} C_J^2$ may occur only once on each time interval of length $\eta/\hat{\lambda}$. We thus get:

$$\begin{aligned} \text{Tot.Var.}\big(f(\rho(t+,\cdot))\big) &\leqslant C_{\mathcal{J}}^{\frac{t\hat{\lambda}}{\eta}} \text{Tot.Var.}\big(f(\rho(0+,\cdot))\big) = \\ &= e^{Kt}\text{Tot.Var.}\big(f(\rho(0+,\cdot))\big), \end{aligned}$$

where $K = \hat{\lambda}\log(C_{\mathcal{J}})/\eta$. □

Definition 5.4.2 *Consider a road network $(\mathcal{I}, \mathcal{J})$ and an approximate wave front tracking solution ρ. For every road I_i, we define two curves $Y_-^{i,\rho}(t)$, $Y_+^{i,\rho}(t)$, called Boundary of External Flux, briefly BEF, in the following way. We set the initial condition $Y_-^{i,\rho}(0) = a_i$, $Y_+^{i,\rho}(0) = b_i$ (if $a_i = -\infty$,* *then* $Y_-^{i,\rho} \equiv -\infty$ *and if* $b_i = +\infty$*, then* $Y_+^{i,\rho} \equiv +\infty$*). We let* $Y_\pm^{i,\rho}(t)$ *follow the generalized characteristic as defined in [45], letting* $Y_-^{i,\rho}(t) = a_i$ *(resp.* $Y_+^{i,\rho}(t) = b_i$*) if the generalized characteristic reaches the boundary and* $f'(\rho(t,a_i)) < 0$ *(resp.* $f'(\rho(t,b_i)) > 0$*). (In this way* $Y_\pm^{i,\rho}(t)$ *may coincide with* a_i *or* b_i *for some time intervals). Let* $\bar{t}$ *be the first time* $\bar{t}$ *such that* $Y_-^{i,\rho}(\bar{t}) = Y_+^{i,\rho}(\bar{t})$ *(possibly* $\bar{t} = +\infty$*), then we let* $Y_\pm^{i,\rho}$ *be defined on* $[0,\bar{t}]$*. Finally, we define the sets*

$$D_1^i(\rho) = \left\{(t,x) : t \in [0,\bar{t}) : Y_-^{i,\rho}(t) \leqslant x \leqslant Y_+^{i,\rho}(t)\right\},$$

and

$$D_2^i(\rho) = [0,+\infty) \times [a_i,b_i] \setminus D_1^i(\rho).$$

Clearly $Y_\pm^i(t)$ bound the set on which the datum is not influenced by the other roads through the junctions.

Definition 5.4.3 *Fix an approximate wave front tracking solution* ρ*, a road* I_i*,* $i = 1,\dots,N$ *and a junction* J*. A wave* θ *in* I_i *is said a big wave if*

$$sgn(\rho_-^\theta - \sigma) \cdot sgn(\rho_+^\theta - \sigma) \leqslant 0,$$

where $sgn(0) = 0$*.*

We say that an incoming road I_i *has a bad datum at* J *at time* $t > 0$ *if*

$$\rho_i(t, b_i-) \in [0,\sigma[.$$

We say that an outgoing road I_j *has a bad datum at* J *at time* $t > 0$ *if*

$$\rho_j(t, a_j+) \in]\sigma, 1].$$

Lemma 5.4.6 *For every* $t \geqslant 0$*, there exist at most two big waves on*

$$\left\{x : (t,x) \in D_2^i(\rho)\right\} \subseteq [a_i,b_i].$$

Proof. A big wave can originate at time t on road I_i from J only if road I_i has a bad datum at J at time t. If this happens, then road I_i has not a bad datum at J up to the time in which a big wave is absorbed from I_i. Then we reach the conclusion. □

Theorem 5.4.1 *Fix a road network* $(\mathcal{I}, \mathcal{J})$. *Given* $C > 0$ *and* $T > 0$, *there exists an admissible solution defined on* $[0, T]$ *for every initial data* $\bar{\rho} \in cl\{\rho : TV(\rho) \leqslant C\}$, *where cl indicates the closure in* L^1_{loc}.

Proof. We fix a sequence of initial data $\bar{\rho}_\nu$ piecewise constant such that $TV(\bar{\rho}_\nu) \leqslant C$ for every $\nu \geqslant 0$ and $\bar{\rho}_\nu \to \bar{\rho}$ in L^1_{loc} as $\nu \to +\infty$. For each $\bar{\rho}_\nu$ we consider an approximate wave front tracking solution ρ_ν such that $\rho_\nu(0, x) = \bar{\rho}_\nu(x)$ and rarefactions are split in rarefaction shocks of size $\frac{1}{\nu}$.

For every road I_i, we notice that on $D^i_1(\rho_\nu)$, ρ_ν is not influenced by other roads and so the estimates of [26] hold. Since the curves $Y^{i,\rho_\nu}_\pm$ are uniformly Lipschitz continuous, they converge, up to a subsequence, to a limit curve and hence the regions $D^i_1(\rho_\nu)$ "converge" to a limit region D^i_1. Then $\rho_\nu \to \rho$ in L^1_{loc} on D^i_1 with ρ admissible solution to the Cauchy problem.

On $D^i_2 := [0, +\infty[\times [a_i, b_i] \backslash D^i_1$, we have that, up to a subsequence, $\rho_\nu \rightharpoonup^* \rho$ weak* on L^1 and, using Lemma 5.4.5 and Theorem 2.4 of [26], $f(\rho_\nu) \to \bar{f}$ in L^1 for some $\bar{f}$. By Lemma 5.4.6, there are at most two big waves on D^i_2 for every time, hence, splitting the domain D^i_2 in a finite number of pieces where we can invert the function f, getting $\rho_\nu \to f^{-1}(\bar{f})$ in L^1. Together with $\rho_\nu \rightharpoonup^* \rho$ weak* on L^1, we conclude that $\rho_\nu \to \rho$ strongly in L^1.

The other requirements of the definition of admissible solution are clearly satisfied. □

5.5 Lipschitz continuous dependence: a counterexample and two special cases.

In this section we assume that every junction has exactly two incoming roads and two outgoing ones and for every junctions we follow the notation (5.4.16). We present a counterexample to the Lipschitz continuous dependence by initial data with respect to the L^1–norm. The continuous dependence by initial data with respect the L^1–norm remains an open problem. The counterexample is constructed using shifts of waves as in the spirit of [27], to which we refer the reader for general theory.

We show that, for every $C > 0$, it is possible to choose two piecewise constant initial data, which are exactly the same except for a shift ξ of a discontinuity, such that the L^1–distance of the two corresponding solutions increases by the multiplicative factor C. Obviously, the L^1–distance of the

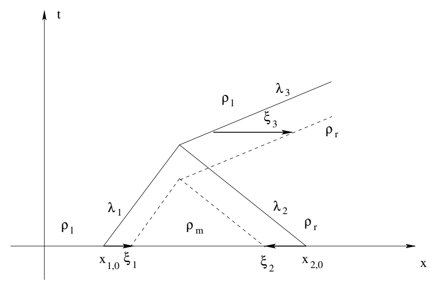


Figure 5.2: Shifts of waves.

initial data is finite and given by $|\xi\,\Delta\rho|$, where ξ is the shift and $\Delta\rho$ is the jump across the corresponding discontinuity. From now on, we consider a junction J, satisfying condition (C), with I_1, I_2 as incoming roads and I_3, I_4 as outgoing ones. Moreover we suppose that the entries of the matrix A satisfy $\alpha < \beta$.

First we need some technical lemmas. The first one is well–known; we report the proof for reader's convenience.

Lemma 5.5.1 *Let us consider in a road two waves, with speeds λ_1 and λ_2 respectively, that interact together at a certain time $\bar{t}$ producing a wave with speed λ_3. If the first wave is shifted by ξ_1 and the second wave by ξ_2, then the shift of the resulting wave is given by*

$$\xi_3 = \frac{\lambda_3 - \lambda_2}{\lambda_1 - \lambda_2}\xi_1 + \frac{\lambda_1 - \lambda_3}{\lambda_1 - \lambda_2}\xi_2. \tag{5.5.26}$$

Moreover we have that

$$\Delta\rho_3\,\xi_3 = \Delta\rho_1\,\xi_1 + \Delta\rho_2\,\xi_2, \tag{5.5.27}$$

where $\Delta\rho_i$ are the signed strengths of the corresponding waves.

Proof. We suppose that ρ_l and ρ_m are the left and the right values of the wave with speed λ_1 and ρ_m and ρ_r are the left and the right values of the wave with speed λ_2, see Figure 5.2.

So $\Delta\rho_1 = \rho_m - \rho_l$, $\Delta\rho_2 = \rho_r - \rho_m$ and $\Delta\rho_3 = \rho_r - \rho_l$. The two wave fronts have respectively equation

$$x = \lambda_1 t + x_{1,0}, \qquad x = \lambda_2 t + x_{2,0},$$

where $x_{1,0}$ and $x_{2,0}$ are the initial positions of the wave fronts with speed λ_1 and λ_2 respectively. Therefore they interact at the point

$$(\bar{x},\bar{t}) = \left(\lambda_1 \frac{x_{1,0}-x_{2,0}}{\lambda_2-\lambda_1} + x_{1,0}, \frac{x_{1,0}-x_{2,0}}{\lambda_2-\lambda_1}\right).$$

If we consider the shifts, then the two wave fronts interact at the point

$$(\tilde{x},\tilde{t}) = \left(x_{1,0}+\xi_1+\lambda_1 \frac{(x_{2,0}+\xi_2)-(x_{1,0}+\xi_1)}{\lambda_1-\lambda_2}, \frac{(x_{2,0}+\xi_2)-(x_{1,0}+\xi_1)}{\lambda_1-\lambda_2}\right),$$

and consequently (5.5.26) holds. Multiplying equation (5.5.26) by $\Delta\rho_3 = \Delta\rho_1 + \Delta\rho_2$, we easily deduce (5.5.27). □

Lemma 5.5.2 *Let us consider a junction J with incoming roads I_1 and I_2 and outgoing roads I_3 and I_4. If a wave on a road I_i ($i \in \{1,\ldots,4\}$) interacts with J without producing waves in the same road I_i and if ξ_i is the shift of the wave in I_i, then the shift ξ_j produced in a different road I_j ($j \in \{1,\ldots,4\} \setminus \{i\}$) satisfies:*

$$\xi_j\left(\rho_j^+ - \rho_j^-\right) = \frac{\Delta\delta_j}{\Delta\delta_i}\xi_i\left(\rho_i^+ - \rho_i^-\right), \tag{5.5.28}$$

where $\Delta\delta_l$ ($l \in \{i,j\}$) represents the variation of the flux in the road I_l and ρ_l^-, ρ_l^+ ($l \in \{i,j\}$) are the states at J in the road I_l respectively before and after the interaction.

Proof. For simplicity let us consider the case $i = 1$ and $j = 3$, the other cases being completely similar. Applying the shift ξ_1 to the wave (ρ_1^+, ρ_1^-), the interaction of the wave with J is shifted in time by

$$-\xi_1 \frac{\rho_1^+ - \rho_1^-}{f(\rho_1^+) - f(\rho_1^-)} = -\xi_1 \frac{\rho_1^+ - \rho_1^-}{\Delta\delta_1}.$$

The shift in time in I_3 must be the same and so

$$\xi_1 \frac{\rho_1^+ - \rho_1^-}{\Delta\delta_1} = \xi_3 \frac{\rho_3^+ - \rho_3^-}{\Delta\delta_3},$$

which concludes the lemma. □

Remark 9 *It is easy to understand that the coefficient of multiplication $\Delta\delta_j/\Delta\delta_i$ in the previous lemma depends by the entries of the matrix A. For example, under the same hypotheses of the previous lemma, if a wave in the I_1 road interacts with J producing a variation of the flux $\Delta\delta_1$ and if no wave is produced in I_1 and I_2, then*

$$\Delta\delta_3 = \alpha\Delta\delta_1, \qquad \Delta\delta_4 = (1-\alpha)\Delta\delta_1.$$

Consequently in this case

$$\frac{\Delta\delta_3}{\Delta\delta_1} = \alpha, \quad \frac{\Delta\delta_4}{\Delta\delta_1} = 1-\alpha.$$

The following lemma is the first step in order to show that the Lipschitz dependence by initial data does not hold in our setting. More precisely, we show that there exists a simple configuration of waves and of shifts, which, after some interactions with J, produces an increase of the L^1-distance, going to a similar configuration.

Lemma 5.5.3 *There exists an initial datum given by $(\rho_{1,0}, \rho_{2,0}, \rho_{3,0}, \rho_{4,0})$, that is an equilibrium configuration at J, a wave $(\bar{\rho}_2, \rho_{2,0})$ on road I_2, waves $(\rho_{3,0}, \rho_3^*)$ with shift $\xi_{3,0}$ and $(\rho_3^*, \bar{\rho}_3)$ on road I_3 such that the followings happen in chronological order:*

1. *the initial distance in L^1 is $\xi_{3,0}\,|\rho_{3,0} - \rho_3^*|$;*
2. *the wave $(\rho_{3,0}, \rho_3^*)$ in I_3 with shift $\xi_{3,0}$ interacts with J;*
3. *waves are produced only in I_2 and I_4;*
4. *the wave on road I_2 interacts with $(\bar{\rho}_2, \rho_{2,0})$ producing a new wave;*
5. *the new wave from road I_2 interacts with J;*
6. *waves are produced only in I_3 and I_4;*
7. *in I_4 the L^1–distance after the interactions, is equal to*

$$2\frac{1-\beta}{\beta}\left|\xi_{3,0}\left(\rho_3^* - \rho_{3,0}\right)\right|,$$

and the L^1–distance on road I_3 is equal to $\xi_{3,0}\,|\rho_{3,0} - \rho_3^*|$.

Proof. Let $(\rho_{1,0}, \rho_{2,0}, \rho_{3,0}, \rho_{4,0})$ be an equilibrium configuration in J such that

$$0 < \rho_{1,0} < \sigma, \quad 0 < \rho_{2,0} < \sigma, \quad 0 < \rho_{3,0} < \sigma, \quad 0 < \rho_{4,0} < \sigma.$$

In road I_3, we consider a wave with negative speed $(\rho_{3,0}, \rho_3^*)$ with shift $\xi_{3,0}$. Since $(\rho_{3,0}, \rho_3^*)$ has negative speed, then $\rho_3^* > \tau(\rho_{3,0})$. Initially the L^1–distance of the two solutions is given by $|\xi_{3,0}(\rho_{3,0} - \rho_3^*)|$. When this wave interacts with J, new waves are produced in I_2 and I_4. It is possible, since $\alpha < \beta$. Therefore the new solution to the Riemann Problem at J is given by

$$(\rho_{1,0}, \hat{\rho}_2, \hat{\rho}_3, \hat{\rho}_4),$$

where $\tau(\rho_{2,0}) < \hat{\rho}_2 < 1$, $0 < \hat{\rho}_4 < \rho_{4,0}$. Moreover some shifts $\hat{\xi}_2$ and $\hat{\xi}_4$ are produced in roads I_2 and I_4 respectively, where obviously $\hat{\xi}_2$ has the same sign of $\xi_{3,0}$ while $\hat{\xi}_4$ has opposite sign. By Lemma 5.5.2, we have

$$\begin{cases} \hat{\xi}_2(\hat{\rho}_2 - \rho_{2,0}) = \frac{1}{\beta}\xi_{3,0}(\rho_3^* - \rho_{3,0}), \\ \hat{\xi}_4(\hat{\rho}_4 - \rho_{4,0}) = \frac{1-\beta}{\beta}\xi_{3,0}(\rho_3^* - \rho_{3,0}). \end{cases}$$

If $0 < \bar{\rho}_2 < \tau(\hat{\rho}_2)$, then the wave $(\bar{\rho}_2, \rho_{2,0})$ in the road I_2 with shift $\bar{\xi}_2 = 0$ interacts with the wave $(\rho_{2,0}, \hat{\rho}_2)$ producing a wave $(\bar{\rho}_2, \hat{\rho}_2)$ with positive speed and with shift $\tilde{\xi}_2$. In this case:

$$\tilde{\xi}_2(\hat{\rho}_2 - \bar{\rho}_2) = \hat{\xi}_2(\hat{\rho}_2 - \rho_{2,0}) = \frac{1}{\beta}\xi_{3,0}(\rho_3^* - \rho_{3,0}).$$

Then, after the interaction of the wave $(\bar{\rho}_2, \hat{\rho}_2)$ with J, the new solution of the Riemann Problem at J is given by

$$(\rho_{1,0}, \bar{\rho}_2, \hat{\rho}_3, \bar{\rho}_4),$$

where $0 < \hat{\rho}_3 < \tau(\rho_3^*)$ and $0 < \bar{\rho}_4 < \hat{\rho}_4$. So in the roads I_3 and I_4 new shifts $\hat{\xi}_3$ and $\bar{\xi}_4$ are created, where:

$$\begin{cases} \hat{\xi}_3(\rho_3^* - \hat{\rho}_3) = \beta\tilde{\xi}_2(\hat{\rho}_2 - \bar{\rho}_2) = \xi_{3,0}(\rho_3^* - \rho_{3,0}), \\ \bar{\xi}_4(\hat{\rho}_4 - \bar{\rho}_4) = (1-\beta)\tilde{\xi}_2(\hat{\rho}_2 - \bar{\rho}_2) = \frac{1-\beta}{\beta}\xi_{3,0}(\rho_3^* - \rho_{3,0}). \end{cases}$$

Now, if $\tau(\hat{\rho}_3) < \bar{\rho}_3 < 1$, then the wave $(\rho_3^*, \bar{\rho}_3)$ with shift $\bar{\xi}_3 = 0$ interacts in I_3 with the wave $(\hat{\rho}_3, \rho_3^*)$ producing a wave $(\hat{\rho}_3, \bar{\rho}_3)$ with negative speed and with shift $\tilde{\xi}_3$ such that

$$\tilde{\xi}_3(\bar{\rho}_3 - \hat{\rho}_3) = \hat{\xi}_3(\rho_3^* - \hat{\rho}_3) = \xi_{3,0}(\rho_3^* - \rho_{3,0}).$$

If the two waves on road I_4 do non interact, and this happens choosing appropriately the position of waves, then in the road I_4 the L^1–distance is

$$2\frac{1-\beta}{\beta}\left|\xi_{3,0}(\rho_3^* - \rho_{3,0})\right|,$$

and so the lemma is proved. □

Applying repeatedly Lemma 5.5.3, we produce a counterexample to the Lipschitz continuous dependence by initial data as the next proposition shows.

Proposition 5.5.1 *Let $C > 0$, J be a junction and let $(\rho_{1,0}, \dots, \rho_{4,0})$ be an equilibrium configuration as in Lemma 5.5.3. There exist two piecewise constant initial data satisfying the equilibrium configuration at J such that the L^1–distance between the corresponding two solutions increases by the multiplication factor C.*

Proof. Let n be big enough so that

$$\left(1 + 2n\frac{1-\beta}{\beta}\right) > C.$$

We want to define an initial data that provides the desired increase. We choose ρ_3^* and two finite sequences $(\bar{\rho}_2^i)$, $(\bar{\rho}_3^i)$, $i = 1, \dots, n$, so that, letting $\hat{\rho}_2^i$, $\hat{\rho}_3^i$ be the states determined as in Lemma 5.5.3, we have:

$$\begin{cases} \rho_3^* \in]\tau(\rho_{3,0}), 1], \\ \bar{\rho}_2^i \in [0, \tau(\hat{\rho}_2^i)[, & i = 1, \dots, n, \\ \bar{\rho}_3^i \in]\tau(\hat{\rho}_3^i), 1], & i = 1, \dots, n. \end{cases}$$

It is easy to check that these sequences can be defined by induction.

The piecewise constant initial data in I_3 is given by

$$\begin{cases} \rho_{3,0}, & \text{if } 0 < x < x^*, \\ \rho_3^*, & \text{if } x^* < x < \hat{x}_1, \\ \hat{\rho}_3^1, & \text{if } \hat{x}_1 < x < \hat{x}_2, \\ \vdots & \dots \\ \bar{\rho}_3^n, & \text{if } \tilde{x}_n < x, \end{cases}$$

where the values x^*, $\hat{x}_1, \ldots, \hat{x}_n$ are to be determined in the sequel. If $\xi_{3,0}$ denotes the shift of the wave $(\rho_{3,0}, \rho_3^*)$ and if no more shifts are present, then the L^1–distance of initial data is given by

$$|\xi_{3,0}|\,(\rho_3^* - \rho_{3,0})\,.$$

The initial data on I_2 is

$$\begin{cases} \rho_{2,0}, & \text{if } \tilde{x}_1 < x < 0, \\ \hat{\rho}_2^1, & \text{if } \tilde{x}_2 < x < \tilde{x}_1, \\ \vdots & \ldots \\ \hat{\rho}_2^n, & \text{if } x < \tilde{x}_n \\ \vdots & \ldots, \end{cases}$$

where $\tilde{x}_1, \ldots, \tilde{x}_n$ are to be chosen appropriately.

The speed of the wave $(\rho_{3,0}, \rho_3^*)$ is given by the Rankine–Hugoniot condition

$$\frac{f(\rho_{3,0}) - f(\rho_3^*)}{\rho_{3,0} - \rho_3^*},$$

and consequently the time needed to go to the junction J is

$$\bar{T} = -\frac{(\rho_{3,0} - \rho_3^*)\,x^*}{f(\rho_{3,0}) - f(\rho_3^*)}.$$

Clearly we adjust $\bar{T}$, choosing x^*. Applying n times Lemma 5.5.3 and adjusting the interaction times by choosing appropriately $\bar{x}_i$, $\tilde{x}_i$, $i \in \{1, \ldots, n\}$, we can create $2n$ waves on road I_4 that do not interact together before the end of these n cycles and so we deduce that, at the end, the L^1–distance of the two solutions is given by

$$\left(1 + 2n\frac{1-\beta}{\beta}\right)|\xi_{3,0}(\rho_3^* - \rho_{3,0})|\,,$$

which concludes the proof. □

Remark 10 *The process described in the proof of Proposition 5.5.1 cannot be infinitely repeated. In fact, the sequences $\bar{\rho}_2^i$, $\bar{\rho}_3^i$ are monotonic and so $\bar{\rho}_3^{i+1} - \bar{\rho}_3^i \sim \frac{\bar{\rho}_3^1}{n}$ as n goes to infinity. Then the corresponding shifts on I_3 tend to infinity, letting waves interact with each other on road I_4. Therefore, with this method, it is not possible to produce a blow–up of the L^1–distance in finite time.*

In some special cases the Lipschitz continuous dependence holds as we show in the next subsections.

5.5.1 Network with only one junction.

We consider a road network with only one junction J and with I_1, I_2 incoming roads and I_3, I_4 outgoing roads. We define

$$\mathcal{D} := \big\{\bar\rho = (\bar\rho_1, \dots, \bar\rho_4) \in L^\infty(I_1 \times \cdots \times I_4) \cap L^1(I_1 \times \cdots \times I_4) : \\ \bar\rho_j \in [0, \sigma], j = 3, 4\big\}\,.$$

The following theorem holds.

Theorem 5.5.1 *There exists a Lipschitz continuous semigroup S, defined on $[0, +\infty[\times\mathcal{D}$ with values in $\mathcal{D}$ so that, for every $\bar\rho \in \mathcal{D}$, $\rho(t, x) = S(t, \bar\rho)(x)$ is an admissible solution with $\rho(0, x) = \bar\rho(x)$.*

Before proving the theorem, we consider the following lemma.

Lemma 5.5.4 *Let $T > 0$ and let $\rho, \tilde\rho$ be two approximate wave front tracking solutions connected by shifts such that $\rho(0, \cdot) \in \mathcal{D}$ and $\tilde\rho(0, \cdot) \in \mathcal{D}$. Then, for every $t \in [0, T]$, we have:*

$$\|\rho(t, \cdot) - \tilde\rho(t, \cdot)\|_{L^1} = \sum_{\theta \in \Theta(t)} \left|\xi^\theta \Delta\rho^\theta\right| = \|\rho(0, \cdot) - \tilde\rho(0, \cdot)\|_{L^1}\,,$$

where $\Theta(t)$ denotes the set of the jumps $\Delta\rho^\theta$ of $\rho(t, \cdot)$ with shifts ξ^θ.

Proof. We note first that $\mathcal{D}$ is invariant with respect approximate wave front tracking solutions. Since $\rho_j \in [0, \sigma]$ for every $j \in \{3, 4\}$, each wave on I_3 and I_4 has positive speed and so shifts on outgoing roads cannot propagate themselves on other roads. The conclusion easily follows from Lemma 5.5.2 and Lemma 5.4.3. □

Proof of Theorem 5.5.1. For every $T > 0$, by Theorem 5.4.1, a solution exists for every initial data in $\mathcal{D}$. Fixed $\rho, \tilde\rho \in \mathcal{D}$, we denote by $\rho_\nu, \tilde\rho_\nu$ two approximate wave front tracking solutions. As in [26, 27], to control the norm $\|\rho_\nu(t, \cdot) - \tilde\rho_\nu(t, \cdot)\|_{L^1}$, $t \in [0, T]$, it is enough to control the lengths of the shifts. Therefore, by Lemma 5.5.4, we obtain

$$\|\rho_\nu(t, \cdot) - \tilde\rho_\nu(t, \cdot)\|_{L^1} \leqslant \|\rho_\nu(0, \cdot) - \tilde\rho_\nu(0, \cdot)\|_{L^1}$$

for every $t \in [0, T]$. Passing to the limit in the last expression, we finish the proof. □

5.5.2 Finite number of big waves, bad data and interactions.

Here we want to show a more general result about the Lipschitz continuity with respect to initial data. We omit the proof of this result, since it can be done with the same techniques as in the last subsection.

Let us consider a road network $(\mathcal{I}, \mathcal{J})$.

Definition 5.5.1 *Let us fix an approximate wave front tracking solution ρ. For every junction J and for every incoming road I_i, the function $b_\rho(J,i,\cdot)$ is defined on $[0,T]$ by*

$$b_\rho(J,i,t) = \begin{cases} 0, & \textit{if } \rho_i(t, b_i-) \in [\sigma, 1], \\ 1, & \textit{if } \rho_i(t, b_i-) \in [0, \sigma[. \end{cases}$$

If ρ_ν is a sequence of approximate wave front tracking solutions (briefly AWFTS), then we say that the sequence ρ_ν has the property (H) if:

H1. there exists $M \in \mathbb{N}$ such that the function $b_{\rho_\nu}(J,i,\cdot)$ has at most M discontinuities for every $J \in \mathcal{J}$, for every $i \in \{1, \ldots, N\}$ and for every $\nu \geqslant 0$;

H2. there exists $\tilde{\delta} > 0$ such that

$$|\rho_\nu(t, a_i+) - \sigma| > \tilde{\delta}$$

and

$$|\rho_\nu(t, b_i-) - \sigma| > \tilde{\delta}$$

for every $J \in \mathcal{J}$, for every $i \in \{1, \ldots, N\}$, for every $\nu \geqslant 0$ and for every $t \in [0,T]$.

The following proposition holds.

Proposition 5.5.2 *Fixed $T > 0$, we consider a solution ρ defined on $[0,T]$ such that, for every $t \in [0,T]$, $\rho(t,\cdot)$ is a bounded variation function. Given $\eta > 0$, $\tilde{\delta} > 0$ and $M \in \mathbb{N}$, we define*

$$\begin{aligned} \mathcal{D}^\eta_\rho(\tilde{\delta}, M) := \{\bar{\rho} \in L^1_{loc} \ : \ & \exists (\rho_\nu)_{\nu\in\mathbb{N}} \textit{ sequence of AWFTS satisfying (H)} \\ & \textit{with parameters } \tilde{\delta} \textit{ and } M, \\ & \rho_\nu(0,\cdot) \to \bar{\rho}(\cdot) \textit{ in } L^1_{loc}, \\ & \textit{Tot.Var.}(\rho_\nu(0,\cdot) - \rho(0,\cdot)) < \eta\}. \end{aligned}$$

If there exist $0 < \eta' < \eta$, $\tilde{\delta} > 0$ *and* $M \in \mathbb{N}$ *such that*

$$\mathcal{D} := cl\,\{\tilde{\rho} : Tot.Var.(\rho - \tilde{\rho}) < \eta'\} \subseteq \mathcal{D}^{\eta}_{\rho}(\tilde{\delta}, M),$$

then there exists a Lipschitz continuous semigroup S *of solutions defined on* $[0,T] \times \mathcal{D}$*.*

Remark 11 *We expect the existence of* $\eta, \eta', \tilde{\delta}, M$ *as in Proposition 5.5.2, if we have* $\eta < \tilde{\delta}$ *and if we assume that big waves of* ρ *have velocity bounded away from zero.*

5.6 Time Dependent Traffic.

In this section we consider a model of traffic including traffic lights and time dependent traffic. The latter means that the choice of drivers at junctions may depend on the period of the day, for instance during the morning the traffic flows towards some specific parts of the network and during the evening it may flow back. This means that the matrix A may depend on time t.

Consider a single junction J as in Section 5.2 with two incoming roads I_1, I_2 and two outgoing ones I_3 and I_4. Let $\alpha = \alpha(t)$, $\beta = \beta(t)$ be two piecewise constant functions such that

$$0 < \alpha(t) < 1, \qquad 0 < \beta(t) < 1, \qquad \alpha(t) \neq \beta(t), \tag{5.6.29}$$

for each $t \geqslant 0$. Moreover let $\chi_1 = \chi_1(t), \chi_2 = \chi_2(t)$ be piecewise constant maps such that

$$\chi_1(t) + \chi_2(t) = 1, \qquad \chi_i(t) \in \{0,1\}, \qquad i = 1,2,$$

for each $t \geqslant 0$. The two maps represent traffic lights, the value 0 corresponding to red light and the value 1 to green light.

Definition 5.6.1 *Consider* $\rho = \big(\rho_1, ..., \rho_4\big)$ *with bounded variation. We say that* ρ *is a solution at the junction* J *if it satisfies (i), (iii) of Definition 5.1.1 and the following property holds:*
(iv) $f(\rho_3(t, a_3+)) = \alpha(t)\chi_1(t) f(\rho_1(t, b_1-)) + \beta(t)\chi_2(t) f(\rho_2(t, b_2-))$ *and* $f(\rho_4(t, a_4+)) = \big(1-\alpha(t)\big)\chi_1(t) f(\rho_1(t, b_1+)) + \big(1-\beta(t)\big)\chi_2(t) f(\rho_2(t, b_2+))$ *for each* $t > 0$*.*

The construction of the solution can be done as in Section 5.4. However, the total variation of $f(\rho)$ does not depend continuously on the total variation of the maps $\alpha(\cdot)$, $\beta(\cdot)$. Indeed, let us suppose that there are no traffic lights, i.e. $\chi_i \equiv 1$, and let

$$\alpha(t) = \begin{cases} \eta_1 & \text{if } 0 \leqslant t \leqslant \bar{t}, \\ \eta_2 & \text{if } \bar{t} \leqslant t \leqslant T, \end{cases} \qquad \beta(t) = \begin{cases} \eta_2 & \text{if } 0 \leqslant t \leqslant \bar{t}, \\ \eta_1 & \text{if } \bar{t} \leqslant t \leqslant T, \end{cases}$$

where $0 < \eta_2 < \eta_1 < \frac{1}{2}$ and $0 < \bar{t} < T$. Consider the initial data $(\rho_{1,0},\ \rho_{2,0},\ \rho_{3,0},\ \rho_{4,0})$, where

$$f(\rho_{1,0}) = f(\rho_{4,0}) = f(\sigma), \qquad f(\rho_{2,0}) = f(\rho_{3,0}) = \frac{\eta_1}{1-\eta_2} f(\sigma),$$

and

$$\sigma < \rho_{2,0} < 1, \qquad 0 < \rho_{3,0} < \sigma.$$

This is an equilibrium configuration and hence the solution of the Riemann Problem for $0 \leqslant t \leqslant \bar{t}$. At time $t = \bar{t}$ we have to solve a new Riemann Problem. Let $(\hat{\rho}_1,\ \hat{\rho}_2,\ \hat{\rho}_3,\ \hat{\rho}_4)$ be the new solution. We have:

$$f(\hat{\rho}_2) = f(\hat{\rho}_4) = f(\sigma), \qquad f(\hat{\rho}_1) = f(\hat{\rho}_3) = \frac{\eta_1}{1-\eta_2} f(\sigma).$$

Now, if $\eta_1 \to \eta_2$, then

$$\text{Tot.Var.}\big(\alpha; [0,T]\big) \longrightarrow 0, \qquad \text{Tot.Var.}\big(\beta; [0,T]\big) \longrightarrow 0,$$

but

$$\big(f(\rho_{1,0}),\ f(\rho_{2,0})\big) \to \left(f(\sigma),\ \frac{\eta_2}{1-\eta_2} f(\sigma)\right)$$

and

$$\big(f(\hat{\rho}_1),\ f(\hat{\rho}_2)\big) \to \left(\frac{\eta_2}{1-\eta_2} f(\sigma),\ f(\sigma)\right),$$

hence $\text{Tot.Var.}\big(f(\rho); [0,T]\big)$ is bounded away from zero.

5.7 Total Variation of the Fluxes.

Let J be a junction with 3 incoming roads and 3 outgoing ones. We show, with an example, that the total variation of the flux may increase if a wave arrives to J from an incoming road. Let us suppose that the matrix A is given by

$$A \doteq \begin{pmatrix} \frac{1}{2}-\varepsilon & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{2}+\varepsilon \\ \frac{1}{6}+\varepsilon & 0 & \frac{1}{6}-\varepsilon \end{pmatrix}, \tag{5.7.30}$$

with $\varepsilon > 0$. Notice that the matrix A satisfies condition (C) for every $\varepsilon > 0$ small enough.

Let us choose $\rho_1,\ \rho_{1,0}, ...,\ \rho_{6,0} \in [0,1]$ such that

$$\rho_{1,0} = \rho_{4,0} = \rho_{5,0} = \sigma,\ \sigma < \rho_{2,0} < 1,\ \sigma < \rho_{3,0} < 1,\ 0 < \rho_{6,0} < \sigma,\ 0 < \rho_1 < \sigma,$$

$$f(\rho_{2,0}) = \frac{1+36\varepsilon+36\varepsilon^2}{3(1+6\varepsilon)}, \quad f(\rho_{3,0}) = \frac{1-6\varepsilon}{1+6\varepsilon}, \quad f(\rho_{6,0}) = \frac{1}{6} + \varepsilon + \frac{(1-6\varepsilon)^2}{6(1+6\varepsilon)}.$$

Assuming that $f(\sigma) = 1$, $(\rho_{1,0}, ...,\ \rho_{6,0})$ is an equilibrium configuration and ρ given by

$$\rho_1(0,x) = \begin{cases} \rho_{1,0} & \text{if } x_1 \leqslant x \leqslant b_1, \\ \rho_1 & \text{if } x < x_1, \end{cases} \qquad \rho_i(0,\cdot) \equiv \rho_{i,0}, \quad i = 2, ..., 6,$$

is a solution (see Figure 5.3). Moreover the point $\big(f(\rho_{1,0}), f(\rho_{2,0}), f(\rho_{3,0})\big)$ is given by the intersection of the planes

$$\left(\frac{1}{2} - \varepsilon\right)\delta_1 + \frac{1}{2}\delta_2 + \frac{1}{3}\delta_3 = 1, \quad \frac{1}{3}\delta_1 + \frac{1}{2}\delta_2 + \left(\frac{1}{2} + \varepsilon\right)\delta_3 = 1, \quad \delta_1 = 1.$$

At some time, say $\bar{t}$, the wave $(\rho_1, \rho_{1,0})$ interacts with the junction. Let $(\hat{\rho}_1, ..., \hat{\rho}_6)$ be the solution of the Riemann Problem at the junction for the data $(\rho_1,\ \rho_{2,0}, ...,\ \rho_{6,0})$. If $f(\rho_1)$ is sufficiently near to 1, then we have:

$$\begin{array}{llll} f(\hat{\rho}_1) & = f(\rho_1), & f(\hat{\rho}_2) & = 2 - \frac{5-36\varepsilon^2}{3(1+6\varepsilon)} f(\rho_1), \\ f(\hat{\rho}_3) & = \frac{1-6\varepsilon}{1+6\varepsilon} f(\rho_1), & f(\hat{\rho}_4) & = f(\hat{\rho}_5) = 1, \\ f(\hat{\rho}_6) & = \frac{1+36\varepsilon^2}{3(1+6\varepsilon)} f(\rho_1). & & \end{array}$$

Therefore

$$\text{Tot.Var.}\big(f(\rho(\bar{t}-,\cdot))\big) = 1 - f(\rho_1),$$

and

$$\text{Tot.Var.}\big(f(\rho(\bar{t}+,\cdot))\big) = \frac{3(1-2\varepsilon)}{1+6\varepsilon}\big(1 - f(\rho_1)\big) > 2\,\text{Tot.Var.}\big(f(\rho(\bar{t}-,\cdot))\big).$$

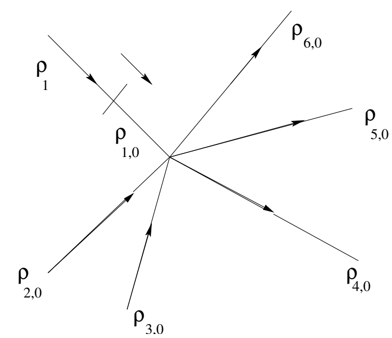


Figure 5.3: Configuration at J.

5.8 Total Variation of the Densities.

Consider a junction J with two incoming roads and two outgoing ones that we parameterize with the intervals $]-\infty, b_1]$, $]-\infty, b_2]$, $[a_3, +\infty[$, $[a_4, +\infty[$ respectively. We suppose that $0 < \beta < \alpha < 1/2$, where α and β are the entries of the matrix A as in (5.4.16).

Define a solution ρ by

$$\rho_1(0,x) = \begin{cases} \rho_{1,0} & \text{if } x_1 \leqslant x \leqslant b_1, \\ \rho_1 & \text{if } x < x_1, \end{cases} \qquad \rho_2(0,x) = \rho_{2,0},$$
$$\rho_3(0,x) = \rho_{3,0}, \quad \rho_4(0,x) = \rho_{4,0}, \tag{5.8.31}$$

where $\rho_1,\ \rho_{1,0},\ \rho_{2,0},\ \rho_{3,0},\ \rho_{4,0}$ are constants such that

$$\sigma < \rho_{2,0} < 1, \quad \sigma < \rho_{3,0} < 1, \quad 0 \leqslant \rho_1 < \sigma, \quad \rho_{1,0} = \rho_{4,0} = \sigma, \tag{5.8.32}$$
$$f(\rho_{1,0}) = f(\rho_{4,0}) = f(\sigma), \quad f(\rho_{2,0}) = f(\rho_{3,0}) = \frac{\alpha}{1-\beta} f(\sigma),$$

so $(\rho_{1,0},\ \rho_{2,0},\ \rho_{3,0},\ \rho_{4,0})$ is an equilibrium configuration.

After some time the wave $(\rho_1, \rho_{1,0})$ interacts with the junction. Let $(\hat\rho_1,\ \hat\rho_2,\ \hat\rho_3,\ \hat\rho_4)$ be the solution of the Riemann Problem in the junction for the data $(\rho_1,\ \rho_{2,0},\ \rho_{3,0},\ \rho_{4,0})$, see Figure 5.4. By (5.8.31) and (5.8.32),

$$f(\hat\rho_1) = f(\rho_1), \quad f(\hat\rho_2) = \frac{f(\sigma) - (1-\alpha) f(\rho_1)}{1-\beta},$$

$$f(\hat\rho_3) = \frac{\alpha - \beta}{1-\beta} f(\rho_1) + \frac{\beta}{1-\beta} f(\sigma), \quad f(\hat\rho_4) = f(\sigma),$$

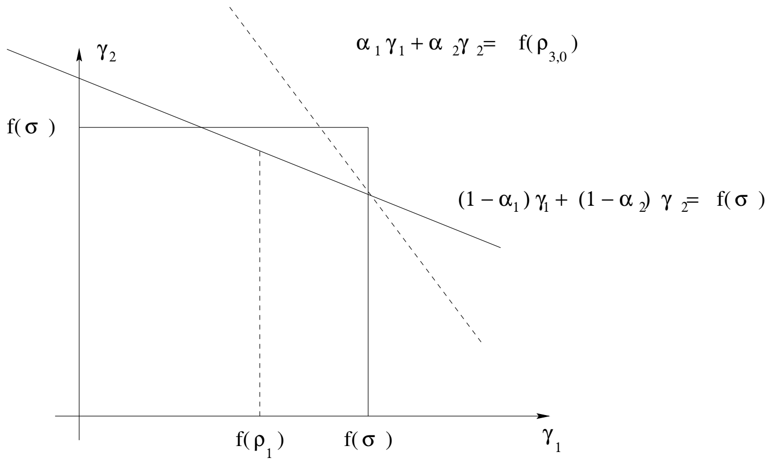


Figure 5.4: Solution to the Riemann problem at J.

and

$$0 < \hat{\rho}_3 < \sigma \leqslant \hat{\rho}_2 < 1. \tag{5.8.33}$$

Therefore, if $\rho_1 \to \rho_{1,0} = \sigma$, then

$$f(\hat{\rho}_3) \longrightarrow \frac{\alpha}{1-\beta} f(\sigma) = f(\rho_{3,0}),$$

and, by (5.8.33), (5.8.32), we have $\hat{\rho}_3 \to \tau(\rho_{3,0})$. Then, we are able to create on the third road a wave with strength bounded away from zero using an arbitrarily small wave on the first one.

Chapter 6

Aw–Rascle traffic model.

In this chapter, traffic flow is described by the second order model, proposed in 2000 by Aw and Rascle, see [13]. The first prototype of a second order traffic model was the so called Payne–Whitham model, see [90, 119]. Some other second order models were proposed in the literature, see [70, 71, 72]. In [44], Daganzo showed that these second order models are not satisfactory to describe the behavior of traffic flow. Therefore Aw and Rascle in [13] proposed a correction to the Payne–Whitham model overcoming the bad behavior of previous models.

We consider a network of roads connected by junctions and, on each road, the traffic flow is described by two equations, written in conservation form. The macroscopic variables are the density ρ of the cars and the momentum y, which can be expressed as function of the density, of the speed of cars and of the pressure.

At junctions, instead, we provide some rules to describe how the flux evolves. Similar to the previous chapter, we assume at each junction the existence of a distributional matrix A, whose entries determine the percentage of the flux from an incoming road to an outgoing one. Moreover we assume that drivers want to maximize the flux. Unfortunately these rules are not sufficient to isolate a unique solution for the Riemann Problem at junctions. More precisely, the Riemann Problem at junctions is underdetermined in outgoing roads, hence the necessity to give an additional rule. We propose three different additional rules: (AR-1) maximization of the speed v of cars, (AR-2) maximization of the density of cars and (AR-3) minimization of the total variation of the density along the solution. The first two rules are motivated by model aspects, while the third one is introduced only for a mathematical reason. These rules permit to solve in unique way the Riemann problem at junctions.

We study in detail solutions, equilibria and stability for Riemann problems at junctions. Equilibria are constant solutions to Riemann Problem at junctions, while stability means that small perturbations of an equilibrium produce a small variation of the equilibrium itself. We prove a result about the existence of solutions to a Cauchy problem when the road network has only one junction and the initial data are sufficiently near to an equilibrium configuration in total variation. More general results about existence and Lipschitz continuous dependence from initial data are still open in this framework.

The chapter is organized in the following way. In Section 6.1, we give the basic definitions and notations. In Section 6.2, we note that the system is strictly hyperbolic if the state $(\rho, y) \neq (0,0)$ and we evaluate the speed of the first and second family of waves. In Section 6.3, we define an invariant domain $\mathcal{D}$ for the Riemann Problem in the roads. In Section 6.4, we solve in details the Riemann Problem at junctions and in Section 6.5 we analyze the stability of equilibria. Section 6.6 is devoted to prove the existence of a Cauchy problem on a road network with only one junction.

6.1 Basic definitions and notations.

We consider a network of roads, that is modeled by a finite collection of connected intervals $I_i = [a_i, b_i] \subseteq \mathbb{R}$, $i = 1, \ldots, N$, possibly with either $a_i = -\infty$ or $b_i = +\infty$, on which the dynamic is governed by the system:

$$\begin{cases} \partial_t \rho + \partial_x (y - \rho^{\gamma+1}) = 0 \\ \partial_t y + \partial_x (\frac{y^2}{\rho} - y\rho^{\gamma}) = 0 \end{cases} \tag{6.1.1}$$

where $\gamma > 0$, ρ is the density of the cars and $y = \rho v + \rho^{\gamma+1}$ is the momentum (v is the velocity of the cars). Thus, on each road, the datum is given by two functions $\rho_i, y_i : [0, +\infty[\times I_i \to \mathbb{R}$.

On each road I_i, we say that $U_i := (\rho_i, y_i) : [0, +\infty[\times I_i \to \mathbb{R}$ is a weak solution to (6.1.1) if, for every C^∞–function $\varphi : [0, +\infty[\times I_i \to \mathbb{R}^2$ with compact support in $]0, +\infty[\times]a_i, b_i[$,

$$\int_0^{+\infty} \int_{a_i}^{b_i} \left(U_i \cdot \frac{\partial \varphi}{\partial t} + f(U_i) \cdot \frac{\partial \varphi}{\partial x} \right) dx\, dt = 0 \tag{6.1.2}$$

where

$$f(U_i) = \begin{pmatrix} y_i - \rho_i^{\gamma+1} \\ \frac{y_i^2}{\rho_i} - y_i \rho_i^{\gamma} \end{pmatrix}, \tag{6.1.3}$$

is the flux of the system (6.1.1). For the definition of entropic solution, we refer to [24].

As in [39], we assume that the roads connect together at junctions and that each road could be an incoming road for at most one junction and an outgoing road for at most one junction, that is in each road cars can run in a unique direction.

With *flux of the density* we indicate the first component of the flux f and precisely the quantity $y - \rho^{\gamma+1}$, while *flux of momentum* stands for the second component of the flux, i.e. $y^2/\rho - y\rho^{\gamma}$.

6.2 Characteristic fields.

We observe that the system (6.1.1) is strictly hyperbolic when $\rho > 0$. In fact the Jacobian matrix of the flux of the system (6.1.1) is given by

$$\begin{pmatrix} -(\gamma+1)\rho^{\gamma} & 1 \\ -\frac{y^2}{\rho^2} - \gamma y \rho^{\gamma-1} & 2\frac{y}{\rho} - \rho^{\gamma} \end{pmatrix} \tag{6.2.4}$$

whose eigenvalues are

$$\lambda_1 = \frac{y}{\rho} - (\gamma+1)\rho^{\gamma}, \qquad \lambda_2 = \frac{y}{\rho} - \rho^{\gamma}. \tag{6.2.5}$$

Therefore, if $\rho > 0$, then $\lambda_1 < \lambda_2$ and the system is strictly hyperbolic. Notice that the second eigenvalue λ_2 is equal to the velocity v of the cars.

It is easy to see that the first characteristic field is genuinely nonlinear, while the second characteristic field is linearly degenerate, see [13, 24]. Moreover the rarefaction curves of the first family are lines passing through the origin. Since the rarefaction curves are lines, also the shock curves of the first family are lines and they have the same expression.

Instead, the curves of the second family through (ρ_0, y_0) are given by

$$y = \frac{y_0}{\rho_0}\rho + \rho^{\gamma+1} - \rho_0^{\gamma}\rho. \tag{6.2.6}$$

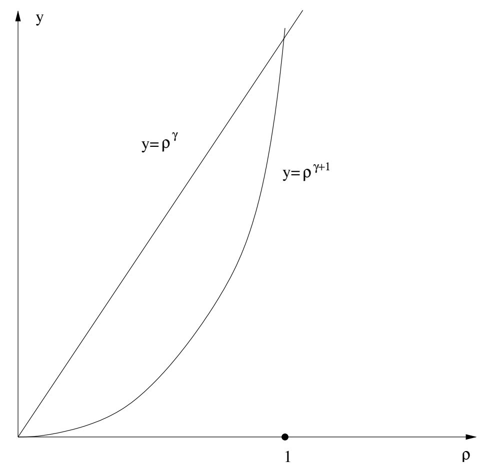


Figure 6.1: Domain of invariance.

6.3 Invariant domain.

It is natural to assume that in each road the density ρ is positive and bounded by a constant ρ_{max}, which for simplicity we assume to be 1.

Also the velocity v of cars must be positive and bounded. In particular we suppose that the maximum velocity of cars is decreasing with respect to the density ρ and it has the following expression:

$$v_{max}(\rho) = 1 - \rho^{\gamma}.$$

Thus we obtain that $\rho^{\gamma+1} \leqslant y \leqslant \rho$; see Figure 6.1. Thus (ρ, y) takes value in the domain

$$\mathcal{D} = \left\{(\rho, y) \in \mathbb{R}^{+} \times \mathbb{R}^{+} : \rho^{\gamma+1} \leqslant y \leqslant \rho\right\}. \tag{6.3.7}$$

We show that the region $\mathcal{D}$ is invariant for the system (6.1.1). To this purpose, it is enough to show that the solution to every Riemann Problem with data in $\mathcal{D}$, remains in $\mathcal{D}$. Consider a road I, modeled by $\mathbb{R}$ and the following Riemann problem:

$$\begin{cases} \partial_t \rho + \partial_x (y - \rho^{\gamma+1}) = 0, \\ \partial_t y + \partial_x (\frac{y^2}{\rho} - y\rho^{\gamma}) = 0, \\ (\rho(0,x), y(0,x)) = (\rho_-, y_-), & \text{if } x < 0, \\ (\rho(0,x), y(0,x)) = (\rho_+, y_+), & \text{if } x > 0. \end{cases} \tag{6.3.8}$$

As pointed out in [13], there are some different cases.

1. The points (ρ_-, y_-) and (ρ_+, y_+) belong either to a curve of the first family or to a curve of the second family. In this case the two points can be connected either by a wave of the first family or by a wave of the second family. Notice that (ρ_-, y_-) or (ρ_+, y_+) can be equal to $(0, 0)$.

2. $\rho_- > 0$, $\rho_+ > 0$ and the curve of the first family through (ρ_-, y_-) intersects the curve of the second family through (ρ_+, y_+) in a point of $\mathcal{D}$ different from $(0, 0)$. We call (ρ_0, y_0) this point.

   If $\rho_0 < \rho_-$ then $\lambda_1(\rho_-, y_-) < \lambda_1(\rho_0, y_0) < \lambda_2(\rho_0, y_0) = \lambda_2(\rho_+, y_+)$. So it is possible to connect (ρ_-, y_-) with (ρ_0, y_0) by a wave of the first family with maximum speed $\lambda_1(\rho_0, y_0)$ and then (ρ_0, y_0) with (ρ_+, y_+) by a wave of the second family with speed $\lambda_2(\rho_0, y_0)$.

   If instead $\rho_0 > \rho_-$, then it is possible to connect (ρ_-, y_-) with (ρ_0, y_0) by a shock wave of the first family with speed

$$\frac{(y_- - \rho_-^{\gamma+1}) - (y_0 - \rho_0^{\gamma+1})}{\rho_- - \rho_0}$$

   and then (ρ_0, y_0) with (ρ_+, y_+) by a wave of the second family with speed

$$\lambda_2(\rho_0, y_0) = \frac{y_0}{\rho_0} - \rho_0^\gamma.$$

   Clearly this process is admissible if and only if

$$\frac{(y_- - \rho_-^{\gamma+1}) - (y_0 - \rho_0^{\gamma+1})}{\rho_- - \rho_0} < \frac{y_0}{\rho_0} - \rho_0^\gamma. \tag{6.3.9}$$

   Since (ρ_-, y_-) and (ρ_0, y_0) belong to the same line $y = c\rho$ with $c > 0$, (6.3.9) is valid if and only if

$$\frac{c(\rho_- - \rho_0) - (\rho_-^{\gamma+1} - \rho_0^{\gamma+1})}{\rho_- - \rho_0} < c - \rho_0^\gamma$$

   which is equivalent to

$$\frac{\rho_-^{\gamma+1} - \rho_0^{\gamma+1}}{\rho_- - \rho_0} > \rho_0^\gamma.$$

   Multiplying by $(\rho_- - \rho_0)$ the last inequality, it results $\rho_-^{\gamma+1} - \rho_0^{\gamma+1} < \rho_0^\gamma \rho_- - \rho_0^{\gamma+1}$ and so (6.3.9) is equivalent to $\rho_-^\gamma < \rho_0^\gamma$ which is clearly true. Thus the analysis of this case is completed.

3. $\rho_- > 0$, $\rho_+ > 0$ and the curve of the first family through (ρ_-, y_-) intersects in $\mathcal{D}$ the curve of the second family through (ρ_+, y_+) only at $(0,0)$. Let $y = c_1\rho$ be the curve of the first family through (ρ_-, y_-) and let $y = c_2\rho + \rho^{\gamma+1}$ be the curve of the second family through (ρ_+, y_+). In this case it is easy to see that $c_1 \leqslant c_2$. It is possible to connect (ρ_-, y_-) to $(0,0)$ by a wave of the first family whose maximum speed is
$$\lim_{\rho\to 0^+} \lambda_1(\rho, c_1\rho) = \lim_{\rho\to 0^+} c_1 - (\gamma+1)\rho^\gamma = c_1$$
and then $(0,0)$ to (ρ_+, y_+) by a wave of the second family with speed c_2. The conclusion follows from the fact that $c_2 \geqslant c_1$.

6.4 Riemann problems at junctions.

To construct solutions on the network, we need to define a solution to Riemann problems at junctions, that is a solution to the Cauchy problem with initial data constant on each road.

In the whole section, we consider a fixed junction J with n incoming roads (say $I_1, \ldots, I_n$) and m outgoing roads (say $I_{n+1}, \ldots, I_{n+m}$) and we assume that $((\rho_{1,0}, y_{1,0}), \ldots, (\rho_{n+m,0}, y_{n+m,0}))$ are the initial data on the roads. It is natural to impose to solutions to the Riemann problem at J the following rules:

(R-1) the waves produced must have negative speed in incoming roads and positive speed in outgoing roads;

(R-2) the first component of the flux (i.e. the flux of the density) must be conserved;

(R-3) there exist some fixed coefficients describing the preferences of the drivers. Each of them determines the percentage of the flux of the density which passes from an incoming road to an outgoing one;

(R-4) the sum of the first components of the flux in incoming roads is maximized.

The first rule means that the waves produced by solving a Riemann problem at a junction travel in the right direction in each road. The second one is conservation of car density, i.e. cars cannot be created or destroyed at junctions. The third one requires that each driver knows her destination and then she chooses the direction according to it. The last rule implies the maximization of the number of cars passing through the junction.

In next sections it is shown that these four rules are not sufficient to solve in a unique way the Riemann problem at junctions. More precisely, these rules are sufficient to isolate a unique solution to the Riemann problem only for incoming roads, but for outgoing roads there exist, in general, infinitely many solutions respecting rules (R-1)–(R-4). Therefore we need an additional rule and we propose three different ones:

(AR-1) maximize the velocity v of cars in outgoing roads;

(AR-2) maximize the density ρ of cars in outgoing roads;

(AR-3) minimize the total variation of ρ along the solution of the Riemann problem in outgoing roads.

Remark 12 *The solution to the Riemann problem at junctions implies the conservation of the density of car, but does not imply the conservation of the momentum. This means that the solution of the Riemann problem at junctions is not a weak solution to (6.1.1), that is it is not a solution to (6.1.1) in integral sense.*

Remark 13 *Rules (AR-1) and (AR-2) are given for model reason, assuming that drivers prefer to maximize ρ or v. On the other side rule (AR-3) is motivated mathematically to control the BV norm.*

To satisfy rule (R-3), we fix an $m \times n$ matrix

$$A = \begin{pmatrix} \alpha_{n+1,1} & \alpha_{n+1,2} & \cdots & \alpha_{n+1,n} \\ \alpha_{n+2,1} & \alpha_{n+2,2} & \cdots & \alpha_{n+2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n+m,1} & \alpha_{n+m,2} & \cdots & \alpha_{n+m,n} \end{pmatrix} \tag{6.4.10}$$

where every $\alpha_{j,i}$ represents the percentage of flux of the density of the cars of the I_i incoming road which goes to the I_j outgoing road and

(A1) $0 < \alpha_{j,i} < 1$ for every $j \in \{n+1, \ldots, n+m\}$ and for every $i \in \{1, \ldots, n\}$;

(A2) $\sum_{j=n+1}^{n+m} \alpha_{j,i} = 1$ for every $i \in \{1, \ldots, n\}$;

(A3) denoting with $\{e_1, \ldots, e_n\}$ the canonical basis of $\mathbb{R}^n$, with $H_i = \{e_i\}^\perp$ the orthogonal hyperplane to e_i, with H_j the orthogonal hyperplane to $\alpha_j = (\alpha_{j,1}, \ldots, \alpha_{j,n})$ ($j \in \{n+1, \ldots, n+m\}$), it holds

$$(1, \ldots, 1) \notin H^\perp$$

for every H defined as the intersection of l distinct hyperplanes H_h, where $l \in \{1, \ldots, n-1\}$ and $h \in \{1, \ldots, n+m\}$.

We want to determine $((\hat{\rho}_1, \hat{y}_1), \ldots, (\hat{\rho}_{n+m}, \hat{y}_{n+m}))$ such that,

- for every $i \in \{1, \ldots, n\}$, waves generated by $((\rho_{i,0}, y_{i,0}), (\hat{\rho}_i, \hat{y}_i))$ have negative velocity;
- for every $j \in \{n+1, \ldots, n+m\}$, waves obtained by $((\hat{\rho}_j, \hat{y}_j), (\rho_{j,0}, y_{j,0}))$ have positive velocity;
- for every $j \in \{n+1, \ldots, n+m\}$, we have $\hat{y}_j - \hat{\rho}_j^{\gamma+1} = \sum_{i=1}^n \alpha_{j,i}(\hat{y}_i - \hat{\rho}_i^{\gamma+1})$;
- the summation $\sum_{i=1}^n (\hat{y}_i - \hat{\rho}_i^{\gamma+1})$ is maximum subject to the previous constrains.
- one of (AR-1), (AR-2), (AR-3) holds.

First of all, we have to calculate all the admissible final states for incoming roads and for outgoing ones, that is we want to find all the final states $((\hat{\rho}_1, \hat{y}_1), \ldots, (\hat{\rho}_{n+m}, \hat{y}_{n+m}))$ such that the first two conditions hold.

In the following analysis, some curves in the domain $\mathcal{D}$ play a crucial role:

1. the curves of the first family;
2. the curves of the second family;
3. the curve $y = (\gamma+1)\rho^{\gamma+1}$.

We call the last one *curve of maxima*, since the first component of the flux restricted to a curve of the first family has the maximum point at the intersection with such curve. Moreover $y = (\gamma+1)\rho^{\gamma+1}$ divides the domain $\mathcal{D}$ into two subdomains $\mathcal{D}_1$ and $\mathcal{D}_2$:

$$\mathcal{D}_1 := \left\{(\rho, y) \in \mathcal{D} : \; y \geqslant (\gamma+1)\rho^{\gamma+1}\right\} \tag{6.4.11}$$

and

$$\mathcal{D}_2 := \left\{(\rho, y) \in \mathcal{D} : \; y \leqslant (\gamma+1)\rho^{\gamma+1}\right\}. \qquad (6.4.12)$$

We use the symbols $\mathring{\mathcal{D}}_1$ and $\mathring{\mathcal{D}}_2$ to denote the sets:

$$\mathring{\mathcal{D}}_1 := \left\{(\rho, y) \in \mathcal{D}_1 : \; y > (\gamma+1)\rho^{\gamma+1}\right\} \qquad (6.4.13)$$

and

$$\mathring{\mathcal{D}}_2 := \left\{(\rho, y) \in \mathcal{D}_2 : \; y < (\gamma+1)\rho^{\gamma+1}\right\}. \qquad (6.4.14)$$

6.4.1 Admissible states in incoming roads.

Fix an incoming road I_i with an initial state $(\rho_{i,0}, y_{i,0})$. We want to find all the possible states $(\hat{\rho}_i, \hat{y}_i)$ such that waves generated by the Riemann problem with data $(\rho_{i,0}, y_{i,0})$ and $(\hat{\rho}_i, \hat{y}_i)$ have negative speed.

Proposition 6.4.1 *Let $(\rho_{i,0}, y_{i,0}) \neq (0,0)$ be the initial value in an incoming road. The admissible states $(\hat{\rho}_i, \hat{y}_i)$ generated by the Riemann problem at the junction must belong to the curve of the first family through $(\rho_{i,0}, y_{i,0})$. More precisely, we have the following cases:*

1. *$(\rho_{i,0}, y_{i,0}) \in \mathcal{D}_1$. In this case, the two states are connected by a shock wave of the first family. There exists a unique point $(\bar{\rho}, \bar{y}) \in \mathcal{D}_2$ on the curve of the first family through $(\rho_{i,0}, y_{i,0})$ with the properties:*
   - *(a) $y_{i,0} - \rho_{i,0}^{\gamma+1} = \bar{y} - \bar{\rho}^{\gamma+1}$;*
   - *(b) $(\hat{\rho}_i, \hat{y}_i)$ is admissible if and only if $\hat{\rho}_i > \bar{\rho}$; see Figure 6.2.*
2. *$(\rho_{i,0}, y_{i,0}) \in \mathcal{D}_2$. In this case all the admissible final states belong to $\mathcal{D}_2$; see Figure 6.3.*

If instead $(\rho_{i,0}, y_{i,0}) = (0,0)$ then the only admissible final state is the same point $(0,0)$.

Proof. If we connect two states with a wave of the second family, then the speed of the wave is greater or equal to 0. Therefore, to obtain waves with negative speed one has to restrict to waves of the first family. First, consider the case $(\rho_{i,0}, y_{i,0}) \neq (0,0)$.

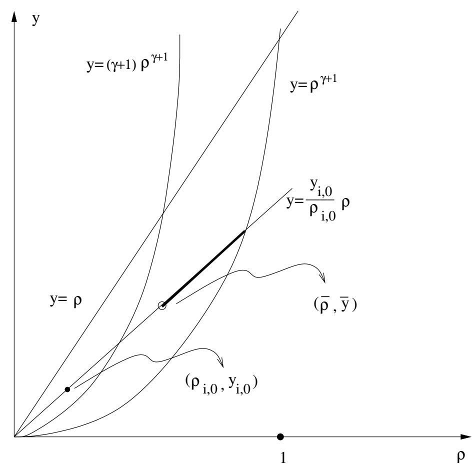


Figure 6.2: Admissible states in an incoming road I_i when $y_{i,0} > (\gamma+1)\rho_{i,0}^{\gamma+1}$. The final state either is $(\rho_{i,0}, y_{i,0})$ or belongs to the part in bold of the line $y = \frac{y_{i,0}}{\rho_{i,0}}\rho$.

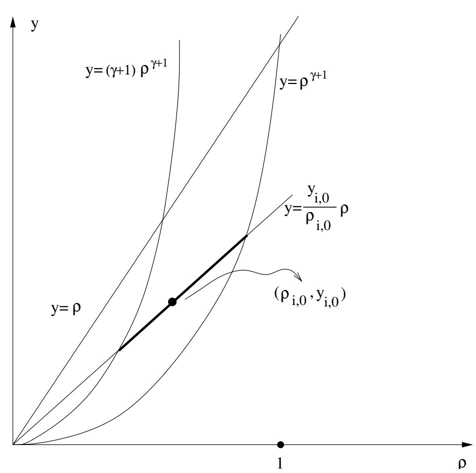


Figure 6.3: Admissible states in an incoming road I_i when $y_{i,0} < (\gamma+1)\rho_{i,0}^{\gamma+1}$. The final state belongs to the part in bold of the line $y = \frac{y_{i,0}}{\rho_{i,0}}\rho$.

If $\hat{\rho}_i < \rho_{i,0}$ then there exists a rarefaction wave of the first family connecting $(\rho_{i,0}, y_{i,0})$ to $(\hat{\rho}_i, \hat{y}_i)$. The maximum speed of the wave is given by

$$\lambda_1(\hat{\rho}_i, \hat{y}_i) = \frac{\hat{y}_i}{\hat{\rho}_i} - (\gamma+1)\hat{\rho}_i^{\gamma}.$$

Since we need $\lambda_1(\hat{\rho}_i, \hat{y}_i) \leqslant 0$, then

$$\hat{\rho}_i^{\gamma+1} \leqslant \hat{y}_i \leqslant (\gamma+1)\hat{\rho}_i^{\gamma+1}.$$

If $\hat{\rho}_i > \rho_{i,0}$ then there exists a shock wave of the first family connecting $(\rho_{i,0}, y_{i,0})$ to $(\hat{\rho}_i, \hat{y}_i)$. Since the speed of the wave, given by the Rankine-Hugoniot condition, must be negative, it results

$$\hat{\rho}_i^{\gamma+1} - \frac{y_{i,0}}{\rho_{i,0}}\hat{\rho}_i + y_{i,0} - \rho_{i,0}^{\gamma+1} > 0.$$

The previous inequality can also be written in the form

$$\frac{y_{i,0}}{\rho_{i,0}} < \frac{\rho_{i,0}^{\gamma+1} - \hat{\rho}_i^{\gamma+1}}{\rho_{i,0} - \hat{\rho}_i}. \tag{6.4.15}$$

If $y_{i,0} \leqslant (\gamma+1)\rho_{i,0}^{\gamma+1}$, then all the points on the curve of the first family through $(\rho_{i,0}, y_{i,0})$ with $\hat{\rho}_i > \rho_{i,0}$ satisfy the last inequality. In fact $y_{i,0}/\rho_{i,0}$ is the slope of the curve of the first family through $(\rho_{i,0}, y_{i,0})$, while

$$\frac{\rho_{i,0}^{\gamma+1} - \hat{\rho}_i^{\gamma+1}}{\rho_{i,0} - \hat{\rho}_i}$$

is strictly greater than the minimum of the derivative of $\rho^{\gamma+1}$ when ρ belongs to the interval

$$\left[\left(\frac{1}{\gamma+1}\right)^{\frac{1}{\gamma}}\left(\frac{y_{i,0}}{\rho_{i,0}}\right)^{\frac{1}{\gamma}}, \left(\frac{y_{i,0}}{\rho_{i,0}}\right)^{\frac{1}{\gamma}}\right],$$

which is exactly $y_{i,0}/\rho_{i,0}$.

Instead, if $y_{i,0} > (\gamma+1)\rho_{i,0}^{\gamma+1}$, then there exists a unique point $(\bar{\rho}, \bar{y})$ on the curve of the first family through $(\rho_{i,0}, y_{i,0})$ with $\bar{\rho} > \rho_{i,0}$ such that

$$\frac{y_{i,0}}{\rho_{i,0}} = \frac{\rho_{i,0}^{\gamma+1} - \bar{\rho}^{\gamma+1}}{\rho_{i,0} - \bar{\rho}}.$$

In fact, since the function $\rho \mapsto \rho^{\gamma+1}$ is convex, then the function

$$\rho \mapsto \frac{\rho_{i,0}^{\gamma+1} - \rho^{\gamma+1}}{\rho_{i,0} - \rho}$$

is strictly increasing when $\rho \geqslant \rho_{i,0}$; moreover

$$\lim_{\rho \to (\frac{1}{\gamma+1})^{1/\gamma} (\frac{y_{i,0}}{\rho_{i,0}})^{1/\gamma}} \frac{\rho_{i,0}^{\gamma+1} - \rho^{\gamma+1}}{\rho_{i,0} - \rho} < \frac{y_{i,0}}{\rho_{i,0}} \quad \text{and} \quad \frac{\rho_{i,0}^{\gamma+1} - \left(\frac{y_{i,0}}{\rho_{i,0}}\right)^{\frac{\gamma+1}{\gamma}}}{\rho_{i,0} - \left(\frac{y_{i,0}}{\rho_{i,0}}\right)^{\frac{1}{\gamma}}} > \frac{y_{i,0}}{\rho_{i,0}},$$

gives the existence of $(\bar{\rho}, \bar{y}) \in \mathring{\mathcal{D}}_2$. Notice that the points $(\rho_{i,0}, y_{i,0})$ and $(\bar{\rho}, \frac{y_{i,0}}{\rho_{i,0}} \bar{\rho})$ have the same first component of the flux.

Now, it remains the case $(\rho_{i,0}, y_{i,0}) = (0, 0)$. In this case no point $(\hat{\rho}_i, \hat{y}_i)$ is admissible, since the speed of the wave of the first family connecting $(0, 0)$ to $(\hat{\rho}_i, \hat{y}_i)$ is given by

$$\frac{\hat{y}_i - \hat{\rho}_i^{\gamma+1}}{\hat{\rho}_i},$$

which is clearly positive. Therefore the proof is finished. □

By the previous proposition, the first component of the flux in an incoming road I_i, may take values in the set

$$\Omega_i = \begin{cases} \left[0, \gamma \left(\frac{1}{\gamma+1}\right)^{\frac{\gamma+1}{\gamma}} \left(\frac{y_{i,0}}{\rho_{i,0}}\right)^{\frac{\gamma+1}{\gamma}}\right], & \text{if } (\rho_{i,0}, y_{i,0}) \in \mathcal{D}_2, \\ \left[0, y_{i,0} - \rho_{i,0}^{\gamma+1}\right], & \text{if } (\rho_{i,0}, y_{i,0}) \in \mathcal{D}_1. \end{cases} \tag{6.4.16}$$

6.4.2 Admissible states in outgoing roads.

Consider an outgoing road I_j, with an initial state $(\rho_{j,0}, y_{j,0})$. We describe the solutions given by an intermediate state $(\bar{\rho}, \bar{y})$ and a final state $(\hat{\rho}_j, \hat{y}_j)$.

Proposition 6.4.2 *Any state $(\bar{\rho}, \bar{y})$ on a curve of the second family through the point $(\rho_{j,0}, y_{j,0})$ can be connected to $(\rho_{j,0}, y_{j,0})$ by a contact discontinuity wave of the second family with speed greater than or equal to* 0.

Proof. The proof follows from the fact that the second eigenvalue λ_2 is greater than or equal to 0 in $\mathcal{D}$. □

Proposition 6.4.3 *A state* $(\hat{\rho}_j, \hat{y}_j) \neq (0,0)$ *is connectible to a given state* $(\bar{\rho}, \bar{y})$ *by a wave of the first family with strictly positive speed if and only if* $\bar{y} = \frac{\hat{y}_j}{\hat{\rho}_j}\bar{\rho}$ *and one of the followings holds:*

1. $\bar{y} < (\gamma+1)\bar{\rho}^{\gamma+1}$. *In this case there exists* $\tilde{\rho} < \bar{\rho}$ *such that all the possible final states* $(\hat{\rho}_j, \hat{y}_j)$ *are those with* $\hat{\rho}_j < \tilde{\rho}$.

2. $\bar{y} \geqslant (\gamma+1)\bar{\rho}^{\gamma+1}$. *In this case we have that*

$$0 \leqslant \hat{\rho}_j \leqslant \left(\frac{1}{\gamma+1}\right)^{1/\gamma} \left(\frac{\hat{y}_j}{\hat{\rho}_j}\right)^{1/\gamma}.$$

If $\hat{\rho}_j < \bar{\rho}$, *then the wave of the first family connecting* $(\hat{\rho}_j, \hat{y}_j)$ *to* $(\bar{\rho}, \bar{y})$ *is a shock wave, while, if* $\hat{\rho}_j > \bar{\rho}$, *then the wave of the first family connecting* $(\hat{\rho}_j, \hat{y}_j)$ *to* $(\bar{\rho}, \bar{y})$ *is a rarefaction wave.*

Proof. First, we note that, if $(\hat{\rho}_j, \hat{y}_j)$ is connectible to $(\bar{\rho}, \bar{y})$ with a wave of the first family, then $\bar{y} = \frac{\hat{y}_j}{\hat{\rho}_j}\bar{\rho}$.

If $\bar{\rho} < \hat{\rho}_j$, then the minimum speed of the wave of the first family connecting $(\hat{\rho}_j, \hat{y}_j)$ to $(\bar{\rho}, \bar{y})$ is given by

$$\lambda_1(\hat{\rho}_j, \hat{y}_j) = \frac{\hat{y}_j}{\hat{\rho}_j} - (\gamma+1)\hat{\rho}_j^{\gamma}.$$

Therefore the speed is positive if and only if

$$\hat{y}_j \geqslant (\gamma+1)\hat{\rho}_j^{\gamma+1}.$$

Instead, if $\bar{\rho} > \bar{\rho}_j$, then the speed of the wave of the first family connecting $(\hat{\rho}_j, \hat{y}_j)$ to $(\bar{\rho}, \bar{y})$ is positive if and only if

$$\frac{(\hat{y}_j - \hat{\rho}_j^{\gamma+1}) - (\bar{y} - \bar{\rho}^{\gamma+1})}{\hat{\rho}_j - \bar{\rho}} > 0,$$

which is equivalent to

$$\frac{\bar{y}}{\bar{\rho}} > \frac{\bar{\rho}^{\gamma+1} - \hat{\rho}_j^{\gamma+1}}{\bar{\rho} - \hat{\rho}_j}. \tag{6.4.17}$$

If $\bar{y} \geqslant (\gamma+1)\bar{\rho}^{\gamma+1}$, then the supremum of the second member of (6.4.17) when $0 < \hat{\rho}_j < \bar{\rho}$ is equal to $(\gamma+1)\bar{\rho}^{\gamma}$, which is lower than or equal to $\bar{y}/\bar{\rho}$.

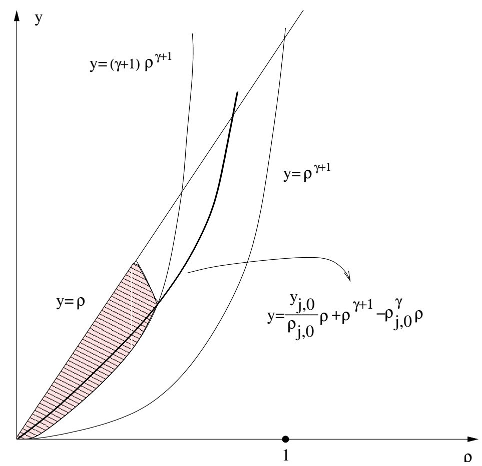


Figure 6.4: Admissible states in an outgoing road I_j when the curve of the second family through $(\rho_{j,0}, y_{j,0})$ (in bold) is not completely in $\mathcal{D}_1$. The admissible final states $(\hat{\rho}_j, \hat{y}_j)$ belongs to that curve or to the drawn region.

If instead $\bar{y} < (\gamma+1)\bar{\rho}^{\gamma+1}$, then as in the previous subsection there exists $\tilde{\rho} < \bar{\rho}$ with the desired property. □

Putting together the results of the last two propositions, we obtain the set of admissible states in an outgoing road; see Figures 6.4 and 6.5.

The possible values of the first component of the flux in an outgoing road I_j is given by

$$\Omega_j = \left[0, \gamma\left(\frac{1}{\gamma+1}\right)^{\frac{\gamma+1}{\gamma}}\right] \tag{6.4.18}$$

if the curve of the second family through $(\rho_{j,0}, y_{j,0})$ is completely inside $\mathcal{D}_1$, while

$$\Omega_j = \left[0, \frac{1}{\rho_{j,0}}(y_{j,0} - \rho_{j,0}^{\gamma+1})\left(1 + \rho_{j,0}^{\gamma} - \frac{y_{j,0}}{\rho_{j,0}}\right)^{\frac{1}{\gamma}}\right] \tag{6.4.19}$$

in the other case.

Remark 14 *Notice that if $(\rho_{j,0}, y_{j,0}) \neq (0,0)$ satisfies $y_{j,0} = \rho_{j,0}^{\gamma+1}$, then the final state $(\hat{\rho}_j, \hat{y}_j)$ must be equal to $(\rho_{j,0}, y_{j,0})$. In fact, if $(\hat{\rho}_j, \hat{y}_j)$ belongs to the curve of the second family through $(\rho_{j,0}, y_{j,0})$, then the wave connecting the*

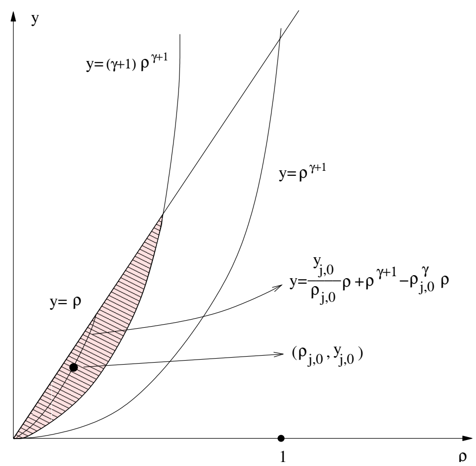


Figure 6.5: Admissible states in an outgoing road I_j when the curve of the second family through $(\rho_{j,0}, y_{j,0})$ is completely in $\mathcal{D}_1$. The admissible final states $(\hat{\rho}_j, \hat{y}_j)$ belongs to the drawn region.

two states has zero speed and so it is not admissible, while if $(\hat{\rho}_j, \hat{y}_j)$ belongs to the curve of the first family through $(\rho_{j,0}, y_{j,0})$, then the speed of the wave is negative.

6.4.3 Riemann problem with rules (R-1)–(R-4).

By the analysis of the previous subsections, (R-1) gives the possible density fluxes in each road of J, (R-3) individuates the set

$$\Omega := \{(\delta_1, \ldots, \delta_n) \in \Omega_1 \times \cdots \times \Omega_n | A \cdot (\delta_1, \ldots, \delta_n) \in \Omega_{n+1} \times \cdots \times \Omega_{n+m}\}, \tag{6.4.20}$$

(R-4) prescribes the maximization of the function

$$E : (\delta_1, \ldots, \delta_n) \mapsto \sum_{i=1}^{n} \delta_i, \tag{6.4.21}$$

while (R-2) is granted once (R-3) is satisfied. The set Ω is closed, convex and non empty. Moreover, by (A3), ∇E is not orthogonal to any nontrivial subspace contained in a supporting hyperplane to Ω, hence there exists a unique vector $\hat{\delta} = (\hat{\delta}_1, \ldots, \hat{\delta}_n) \in \Omega$ such that

$$E(\hat{\delta}_1, \ldots, \hat{\delta}_n) = \max_{(\delta_1, \ldots, \delta_n) \in \Omega} E(\delta_1, \ldots, \delta_n). \tag{6.4.22}$$

With this procedure we find uniquely the values of the fluxes of density of the solution of the Riemann problem at the junction J. More precisely, $\hat{\delta}_i$ gives the value of density fluxes in incoming roads, while density fluxes in outgoing roads are defined by

$$(\hat{\delta}_{n+1}, \ldots, \hat{\delta}_{n+m})^T = A \cdot (\hat{\delta}_1, \ldots, \hat{\delta}_n)^T.$$

For every $i \in \{1, \ldots, n\}$, we have to choose an element $(\hat{\rho}_i, \hat{y}_i)$, which is an admissible state as discussed in Subsection 6.4.1 and such that the flux $\hat{y}_i - \hat{\rho}_i^{\gamma+1}$ is equal to $\hat{\delta}_i$. In order to do this, we need to solve the system

$$\begin{cases} y = \frac{y_{i,0}}{\rho_{i,0}}\rho, \\ y = \rho^{\gamma+1} + \hat{\delta}_i. \end{cases} \tag{6.4.23}$$

This system in general admits two solutions in $\mathcal{D}$, but only one is admissible. So we take

$$(\hat{\rho}_i, \hat{y}_i) = (\rho_{i,0}, y_{i,0}) \tag{6.4.24}$$

if $y_{i,0} = \rho_{i,0}^{\gamma+1} + \hat{\delta}_i$, otherwise $(\hat{\rho}_i, \hat{y}_i)$ is the unique solution in $\mathcal{D}_2$ of the system (6.4.23).

The situation is more complicated in outgoing roads. In fact, by the analysis of subsection 6.4.2, it is evident that, in general, there are infinitely many solutions satisfying rules (R-1)–(R-4), since the intersection between the level curve of the flux $y = \rho^{\gamma+1} + \hat{\delta}_j$ with the region of admissible states is an one dimensional manifold.

6.4.4 (AR-1): maximize the speed.

Proposition 6.4.4 *The rule (AR-1) determines a unique solution of the Riemann problem at the junction J. The final state $(\hat{\rho}_j, \hat{y}_j)$ belongs to the line $y = \rho$.*

Proof. First of all, we recall that the second characteristic field is linearly degenerate and the second eigenvalue λ_2 is equal to the velocity v of the cars. So the curves of the second family are the level sets for the speed v of the cars. On the contrary, the speed v is monotone decreasing in ρ on level curves of density flux. Therefore, the solution is given by

$$\begin{cases} y = \rho^{\gamma+1} + \hat{\delta}_j, \\ y = \rho. \end{cases} \tag{6.4.25}$$

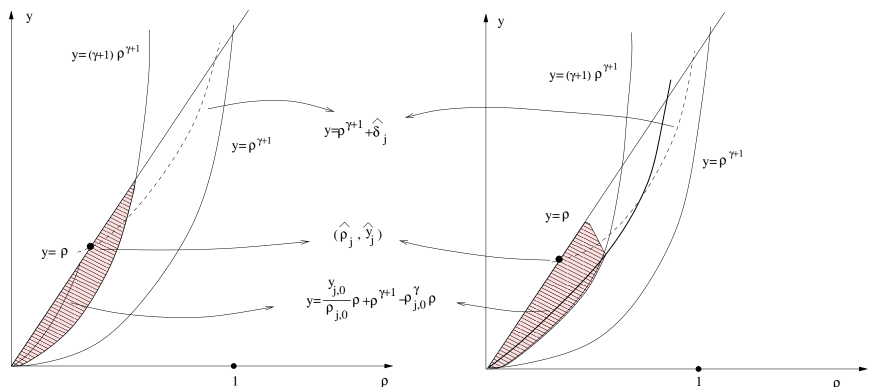


Figure 6.6: Solution $(\hat{\rho}_j, \hat{y}_j)$ to the Riemann problem on an outgoing road I_j in the case 1 with the additional rule (AR-1). In the first picture it is drawn the case in which the curve $y = \rho^{\gamma+1} + \hat{\delta}_j$ does not intersect in $\mathcal{D}$ the curve of the second family through $(\rho_{j,0}, y_{j,0})$, while in the second picture the other case.

There are some different cases.

1. $\hat{\delta}_j < \sup \Omega_j$. In this case $(\hat{\rho}_j, \hat{y}_j)$ is the solution to the system (6.4.25), that is in $\mathcal{D}_1$. In general, to connect $(\hat{\rho}_j, \hat{y}_j)$ with $(\rho_{j,0}, y_{j,0})$ we use a wave of the first family with positive speed and a wave of the second family; see Figure 6.6.

Remark 15 *If we recalculate Ω_j using $(\hat{\rho}_j, \hat{y}_j)$ instead of $(\rho_{j,0}, y_{j,0})$, then the obtained set $\hat{\Omega}_j$ may be bigger than Ω_j. Thus it seems that the solution can be found in two steps, but this is not the case, since $\hat{\delta}_j < \sup(\Omega_j)$ implies that the maximization problem (6.4.22) has the same solution.*

2. $\hat{\delta}_j = \sup(\Omega_j)$, $y_{j,0} < \rho_{j,0}$, and the curve of the second family through $(\rho_{j,0}, y_{j,0})$ lies completely in the region $\mathcal{D}_1$; see Figure 6.7. By the analysis of subsection 6.4.2, the set Ω_j is given by (6.4.18) and so it is the maximum possible. Hence there exists only one point in $\mathcal{D}$ with the first component of the flux equal to $\hat{\delta}_j$ and this point is precisely given by the intersection between the line $y = \rho$ and the curve of maxima. Thus

$$(\hat{\rho}_j, \hat{y}_j) = \left(\left(\frac{1}{\gamma+1} \right)^{\frac{1}{\gamma}}, \left(\frac{1}{\gamma+1} \right)^{\frac{1}{\gamma}} \right),$$

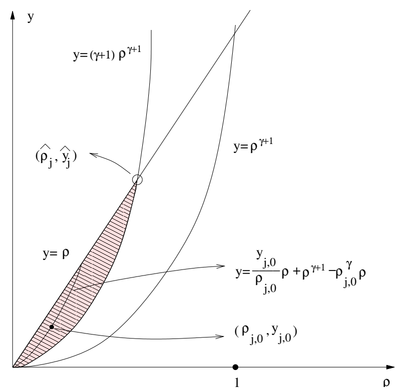


Figure 6.7: Solution $(\hat{\rho}_j, \hat{y}_j)$ to the Riemann problem on an outgoing road I_j in the case 2 with the additional rule (AR-1).

and to connect $(\hat{\rho}_j, \hat{y}_j)$ with $(\rho_{j,0}, y_{j,0})$ we may use a wave of the first family with positive speed and a wave of the second family.

3. $\hat{\delta}_j = \sup(\Omega_j)$, $y_{j,0} < \rho_{j,0}$ and the curve of the second family through $(\rho_{j,0}, y_{j,0})$ is not completely contained in the region $\mathcal{D}_1$. In this case $(\hat{\rho}_j, \hat{y}_j)$ is given by the solution of the system

$$\begin{cases} y = \rho \\ y = \rho^{\gamma+1} + \hat{\delta}_j \\ (\rho, y) \in \mathcal{D}_2. \end{cases}$$

since as in the previous case the intersection between the region of admissible final states and the curve $y = \rho^{\gamma+1} + \hat{\delta}_j$ consists of a single point; see Figure 6.8. To connect $(\hat{\rho}_j, \hat{y}_j)$ with $(\rho_{j,0}, y_{j,0})$ we use only a wave of the second family.

4. $\hat{\delta}_j = \sup(\Omega_j)$ and $y_{j,0} = \rho_{j,0}$. If $(\rho_{j,0}, y_{j,0}) \in \mathcal{D}_1$, then as in case 2 the set Ω_j is the maximum possible and so the solution is given by

$$(\hat{\rho}_j, \hat{y}_j) = \left(\left(\frac{1}{\gamma+1} \right)^{\frac{1}{\gamma}}, \left(\frac{1}{\gamma+1} \right)^{\frac{1}{\gamma}} \right),$$

and to connect $(\hat{\rho}_j, \hat{y}_j)$ with $(\rho_{j,0}, y_{j,0})$ we use only a wave of the first family with positive speed. If instead $(\rho_{j,0}, y_{j,0}) \in \mathcal{D}_2$, then, as in case 3, the solution is

$$(\hat{\rho}_j, \hat{y}_j) = (\rho_{j,0}, y_{j,0}),$$

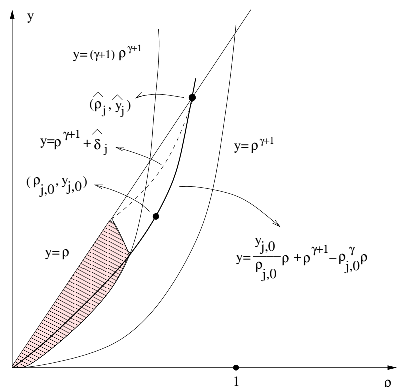


Figure 6.8: Solution $(\hat{\rho}_j, \hat{y}_j)$ to the Riemann problem on an outgoing road I_j in the case 3 with the additional rule (AR-1).

and no wave is produced.

So the proof is finished. □

Remark 16 *Notice that the problem of finding the solution to the Riemann problem with the additional rule (AR-1) is completely equivalent to using the additional rule: minimize the density of the cars in outgoing roads.*

6.4.5 (AR-2): maximize the density.

In this case we have to find the admissible point $(\hat{\rho}_j, \hat{y}_j)$ belonging to the curve $y = \rho^{\gamma+1} + \hat{\delta}_j$ with the maximum ρ.

Proposition 6.4.5 *The solution of the Riemann problem at the junction J with the additional rule (AR-2) is unique and the final state $(\hat{\rho}_j, \hat{y}_j)$ belongs to the region $\mathcal{D}_2$. Moreover, $(\hat{\rho}_j, \hat{y}_j)$ belongs to the curve of the second family through $(\rho_{j,0}, y_{j,0})$ or $(\hat{\rho}_j, \hat{y}_j)$ belongs to the curve of maxima.*

Proof. We have two different possibilities.
1. The curve of the second family through $(\rho_{j,0}, y_{j,0})$ is completely in the region $\mathcal{D}_1$. In this case the admissible final states are exactly all the points of $\mathcal{D}_1$ and so the solution $(\hat{\rho}_j, \hat{y}_j)$ belongs to the part of the curve $y = \rho^{\gamma+1} + \hat{\delta}_j$

which lies in $\mathcal{D}_1$. If we want to maximize the density we have to choose the point $(\hat{\rho}_j, \hat{y}_j)$ given by

$$\begin{cases} y = \rho^{\gamma+1} + \hat{\delta}_j, \\ y = (\gamma+1)\rho^{\gamma+1}, \end{cases}$$

as we clearly see in figure 6.9.a. To connect $(\hat{\rho}_j, \hat{y}_j)$ with $(\rho_{j,0}, y_{j,0})$, we have to use in general a wave of the first family with positive speed and a wave of the second family.

2. The curve of the second family through $(\rho_{j,0}, y_{j,0})$ is not completely in the region $\mathcal{D}_1$. If the curve of the second family through $(\rho_{j,0}, y_{j,0})$ intersects the curve $y = \rho^{\gamma+1} + \hat{\delta}_j$ only in the region $\mathcal{D}_1$, then the solution $(\hat{\rho}_j, \hat{y}_j)$ is given as in the previous case by the system

$$\begin{cases} y = \rho^{\gamma+1} + \hat{\delta}_j, \\ y = (\gamma+1)\rho^{\gamma+1}, \end{cases}$$

and to connect the two states we use a wave of the first family with positive speed and a wave of the second family. Otherwise $(\hat{\rho}_j, \hat{y}_j)$ is given solving

$$\begin{cases} y = \rho^{\gamma+1} + \hat{\delta}_j, \\ y = \frac{y_{j,0}}{\rho_{j,0}}\rho + \rho^{\gamma+1} - \rho_{j,0}^{\gamma}\rho, \end{cases}$$

as we see in figure 6.9.b. To connect $(\hat{\rho}_j, \hat{y}_j)$ with $(\rho_{j,0}, y_{j,0})$, we use only a wave of the second family.

Thus the proof is finished. □

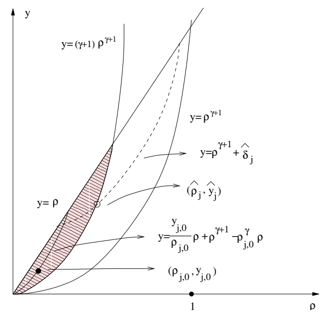
6.4.6 (AR-3): minimize the total variation.

We start by proving the following lemmata.

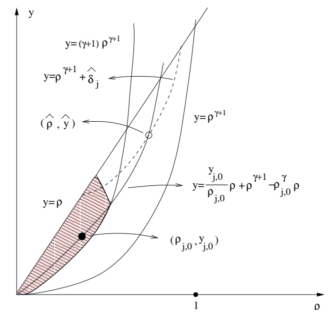
Lemma 6.4.1 *If $(\rho_{j,0}, y_{j,0}) = (0,0)$ then the point $(\hat{\rho}_j, \hat{y}_j)$ belongs to the line $y = \rho$ and to the region $\mathcal{D}_1$.*

Proof. Here the set of admissible states is the whole $\mathcal{D}$ and in this case minimizing the total variation of ρ along a solution is equivalent to choose the point of the curve $y = \rho^{\gamma+1} + \hat{\delta}_j$ with minimum ρ. □

From Remark 14, we have immediately



(a)



(b)

Figure 6.9: Solution $(\hat{\rho}_j, \hat{y}_j)$ to the Riemann problem on an outgoing road I_j with the additional rule (AR-3). The first figure shows the case where the curve of the second family through $(\rho_{j,0}, y_{j,0})$ is completely in the region $\mathcal{D}_1$, while the second figure shows the other case.

Lemma 6.4.2 *Let $(\rho_{j,0}, y_{j,0}) \neq (0,0)$ and $y_{j,0} = \rho_{j,0}^{\gamma+1}$. In this case the solution $(\hat{\rho}_j, \hat{y}_j)$ is equal to $(\rho_{j,0}, y_{j,0})$.*

Lemma 6.4.3 *Assume $y_{j,0} > \rho_{j,0}^{\gamma+1}$. If the curve $y = \rho^{\gamma+1} + \hat{\delta}_j$ intersects the curve of the second family through $(\rho_{j,0}, y_{j,0})$, then the solution $(\hat{\rho}_j, \hat{y}_j)$ is given by the unique intersection of those curves.*

Proof. It is easy to see that the intersection between the curve of the second family through $(\rho_{j,0}, y_{j,0})$ and the curve $y = \rho^{\gamma+1} + \hat{\rho}_j$ consists of at most one point. Then $(\hat{\rho}_j, \hat{y}_j)$ is such intersection and the total variation of the density ρ along the solution is simply given by $|\hat{\rho}_j - \rho_{j,0}|$. In order to prove that the solution attains the minimum of variation in ρ, let us consider an other admissible point $(\bar{\rho}, \bar{y})$ such that

$$\begin{cases} \bar{y} = \bar{\rho}^{\gamma+1} + \hat{\delta}_j, \\ \bar{\rho} \neq \hat{\rho}_j. \end{cases}$$

We must have $\min\{\hat{\rho}_j, \rho_{j,0}\} \leqslant \bar{\rho} \leqslant \max\{\hat{\rho}_j, \rho_{j,0}\}$. If such a point $(\bar{\rho}, \bar{y})$ exists, then to connect $(\bar{\rho}, \bar{y})$ with $(\rho_{j,0}, y_{j,0})$ we need to use first a wave of the first family until a point $(\tilde{\rho}, \tilde{y})$ and then a wave of the second family. Thus the total variation of the density ρ along this solution is given by $|\bar{\rho} - \tilde{\rho}| + |\tilde{\rho} - \rho_{j,0}|$, where $\tilde{\rho}$ satisfies either $\tilde{\rho} < \min\{\hat{\rho}_j, \rho_{j,0}\}$ or $\tilde{\rho} > \max\{\hat{\rho}_j, \rho_{j,0}\}$ and so the proof is finished. □

Lemma 6.4.4 *Assume $y_{j,0} > \rho_{j,0}^{\gamma+1}$. If the curve $y = \rho^{\gamma+1} + \hat{\delta}_j$ does not intersect the curve of the second family through $(\rho_{j,0}, y_{j,0})$, then the solution $(\hat{\rho}_j, \hat{y}_j)$ is given by*

$$\begin{cases} y = \rho^{\gamma+1} + \hat{\delta}_j, \\ y = \rho, \\ (\rho, y) \in \mathcal{D}_1. \end{cases} \tag{6.4.26}$$

Proof. The only possibility is that the curve of the second family through $(\rho_{j,0}, y_{j,0})$ is completely inside the region $\mathcal{D}_1$. Let us call $(\hat{\rho}_j, \hat{y}_j)$ the solution to (6.4.26). To connect $(\hat{\rho}_j, \hat{y}_j)$ with $(\rho_{j,0}, y_{j,0})$ we have to use first a rarefaction wave of the first family until a state $(\bar{\rho}, \bar{y})$ with $\rho_{j,0} < \bar{\rho} < \hat{\rho}_j$ and then a wave of the second family. The total variation of the density ρ along this solution is equal to $|\hat{\rho}_j - \rho_{j,0}|$. Any other point of $y = \rho^{\gamma+1} + \hat{\delta}_j$ generates a variation in ρ strictly bigger than $|\hat{\rho}_j - \rho_{j,0}|$ and so the lemma is proved. □

6.5 Stability of solutions to Riemann problems at junctions.

The aim of this section is to investigate stability of constant (on each road) solutions to Riemann problem, called equilibria. Stability simply means that small perturbations of the data in L^∞ norm, that may be produced by waves arriving at junctions, produce small variations of the equilibrium in L^∞ norm. As in the previous section, we have to consider different cases according to the additional rules (AR-1), (AR-2) or (AR-3).

In the whole section, we consider a fixed junction J with n incoming roads (say $I_1, \ldots, I_n$) and m outgoing roads (say $I_{n+1}, \ldots, I_{n+m}$) and we assume that $((\rho_{1,0}, y_{1,0}), \ldots, (\rho_{n+m,0}, y_{n+m,0}))$ is an equilibrium at J.

We want to remark that waves of the second family have always positive speed. Moreover waves of the first family connecting two states in the region $\mathcal{D}_1$ have positive speed, while waves of the first family connecting two states in the region $\mathcal{D}_2$ have negative speed. The consequences of this fact are the followings.

Claim 1. In an outgoing road only waves of the first family can reach the junction. Therefore if $(\rho_{j,0}, y_{j,0}) \in \mathring{\mathcal{D}}_1$, then it can not be perturbed by waves connecting $(\rho_{j,0}, y_{j,0})$ with an other state $(\bar{\rho}, \bar{y}) \in \mathring{\mathcal{D}}_1$; in fact, in this case, also waves of the first family have positive speed.

Claim 2. Assume $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_1$. If a wave on a road different from I_i produces a variation of the solution of the Riemann problem at the junction, then the new solution $(\hat{\rho}_i, \hat{y}_i)$ in the incoming road I_i either is equal to $(\rho_{i,0}, y_{i,0})$ or $(\hat{\rho}_i, \hat{y}_i)$ belongs to $\mathring{\mathcal{D}}_2$. In the latter case the distance between $(\rho_{i,0}, y_{i,0})$ and $(\hat{\rho}_i, \hat{y}_i)$ is proportional to the distance between $(\rho_{i,0}, y_{i,0})$ and the curve of maxima. Thus, such configuration is unstable.

6.5.1 (AR-1): maximize the speed.

Recall that, by Proposition 6.4.4, all equilibria for outgoing roads must belong to the line $y = \rho$. The analysis of all equilibria is very complicated, hence we prefer to treat in detail only some significant cases. We also consider all the general case when $n = m = 2$.

We have some different possibilities.

1. $(\rho_{j,0}, y_{j,0}) \in \mathring{\mathcal{D}}_1$ for every $j \in \{n+1, \ldots, n+m\}$. Therefore, the maximization problem (6.4.22) implies that $(\rho_{i,0}, y_{i,0}) \in \mathcal{D}_1$ for every $i \in \{1, \ldots, n\}$. In this case, by (6.4.16) and (6.4.18), we deduce that: for incoming roads

$$\Omega_i = \left[0, y_{i,0} - \rho_{i,0}^{\gamma+1}\right],$$

while for outgoing ones

$$\Omega_j = \left[0, \gamma\left(\frac{1}{\gamma+1}\right)^{\frac{\gamma+1}{\gamma}}\right].$$

If we denote by $\delta_{i,0} := y_{i,0} - \rho_{i,0}^{\gamma+1}$ for every $i \in \{1, \ldots, n\}$ and by $\delta_{j,0} := y_{j,0} - \rho_{j,0}^{\gamma+1}$ for every $j \in \{n+1, \ldots, n+m\}$, then clearly $(\delta_{1,0}, \ldots, \delta_{n,0})$ is the solution of the maximization problem (6.4.22) and

$$(\delta_{n+1,0}, \ldots, \delta_{n+m,0})^T = A \cdot (\delta_{1,0}, \ldots, \delta_{n,0})^T.$$

The hypothesis $y_{j,0} > (\gamma+1)\rho_{j,0}^{\gamma+1}$ for every $j \in \{n+1, \ldots, n+m\}$ has the following two consequences. Firstly, $\delta_{j,0} < \sup \Omega_j$ and hence the outgoing roads give no constraint for the maximization problem (6.4.22). Secondly, by claim 1, the outgoing roads cannot be perturbed by waves with negative speed. Consider a perturbation produced by a wave of the first or second family from an incoming road I_i connecting $(\tilde{\rho}_i, \tilde{y}_i)$ with $(\rho_{i,0}, y_{i,0})$. The possible density fluxes are in the set

$$\widetilde{\Omega}_i = \left[0, \tilde{y}_i - \tilde{\rho}_i^{\gamma+1}\right]$$

if $(\tilde{\rho}_i, \tilde{y}_i) \in \mathcal{D}_1$, while

$$\widetilde{\Omega}_i = \left[0, \gamma \left(\frac{1}{\gamma+1}\right)^{\frac{\gamma+1}{\gamma}} \left(\frac{\tilde{y}_i}{\tilde{\rho}_i}\right)^{\frac{\gamma+1}{\gamma}}\right]$$

in the other case. Since the outgoing roads are not constraints for the maximization problem (6.4.22), we may suppose the following, provided the perturbation is sufficiently small:

(a) the new maximum point for (6.4.22) is

$$(\hat{\delta}_1, \ldots, \hat{\delta}_n) := \left(y_{1,0} - \rho_{1,0}^{\gamma+1}, \ldots, \tilde{y}_i - \tilde{\rho}_i^{\gamma+1}, \ldots, y_{n,0} - \rho_{n,0}^{\gamma+1}\right)$$

if $(\tilde{\rho}_i, \tilde{y}_i) \in \mathcal{D}_1$, while

$$(\hat{\delta}_1, \ldots, \hat{\delta}_n) := \left(y_{1,0} - \rho_{1,0}^{\gamma+1}, \ldots, \gamma \left(\frac{1}{\gamma+1}\right)^{\frac{\gamma+1}{\gamma}} \left(\frac{\tilde{y}_i}{\tilde{\rho}_i}\right)^{\frac{\gamma+1}{\gamma}}, \ldots, y_{n,0} - \rho_{n,0}^{\gamma+1}\right).$$

in the other case;

(b) the solution $(\hat{\delta}_{n+1}, \ldots, \hat{\delta}_{n+m})$ defined by

$$(\hat{\delta}_{n+1}, \ldots, \hat{\delta}_{n+m})^T = A \cdot (\hat{\delta}_1, \ldots, \hat{\delta}_n)^T$$

satisfies

$$\hat{\delta}_j < \sup \widetilde{\Omega}_j$$

for every $j \in \{n+1, \ldots, n+m\}$ (the outgoing roads do not become constraints for the maximization problem (6.4.22));

(c)

$$|(\hat{\rho}_i, \hat{y}_i) - (\tilde{\rho}_i, \tilde{y}_i)| + \sum_{j=n+1}^{n+m} |(\hat{\rho}_j, \hat{y}_j) - (\rho_{j,0}, y_{j,0})| < C \, |(\rho_i, y_i) - (\tilde{\rho}_i, \tilde{y}_i)| \, ,$$

where C is a positive constant.

Moreover in outgoing roads waves of the first family are produced, while in incoming roads no waves are produced except in the I_i road.

The conclusion is that this kind of equilibrium is stable under small perturbations.

2. $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_1$ for every $i \in \{1, \ldots, n\}$ and $(\rho_{j,0}, y_{j,0}) \in \mathcal{D}_2$ for some $j \in \{n+1, \ldots, n+m\}$. This is an unstable equilibrium. In fact, let I_{j_1} be the outgoing road with the property $y_{j_1,0} \leqslant (\gamma+1)\rho_{j_1,0}^{\gamma+1}$. It is possible to consider a perturbation generated by a wave of the first family connecting $(\rho_{j_1,0}, y_{j_1,0})$ with $(\tilde{\rho}_{j_1}, \tilde{y}_{j_1})$ such that

$$\sup \widetilde{\Omega}_{j_1} < \sup \Omega_{j_1},$$

where $\widetilde{\Omega}_{j_1}$ is defined as in (6.4.19) for the state $(\tilde{\rho}_{j_1}, \tilde{y}_{j_1})$. In this case the maximization problem (6.4.22) produces a flux in an incoming road I_i, which is strictly lower than $\sup \Omega_i$, hence the final state jumps into the region $\mathring{\mathcal{D}}_2$.

3. $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_2$ for every $i \in \{1, \ldots, n\}$ and $(\rho_{j,0}, y_{j,0}) \in \mathring{\mathcal{D}}_2$ for every $j \in \{n+1, \ldots, n+m\}$. The fact that $y_{i,0} < (\gamma+1)\rho_{i,0}^{\gamma+1}$ for every $i \in \{1, \ldots, n\}$ implies that $\delta_{i,0} := y_{i,0} - \rho_{i,0}^{\gamma+1} < \sup \Omega_i$ for every $i \in \{1, \ldots, n\}$ and hence the incoming roads are not constraints for the maximization problem (6.4.22). Therefore we have stability for perturbations by waves from incoming roads.

Instead the perturbation of an outgoing road in general produces a variation of the maximization problem (6.4.22), since by hypotheses $\delta_{j,0} := y_{j,0} - \rho_{j,0}^{\gamma+1} = \sup \Omega_j$ for every $j \in \{n+1, \ldots, n+m\}$ (all the outgoing roads are constraints for the maximization problem (6.4.22)).

First of all, let us consider the case $m > n$. The maximum for (6.4.22) is determined only by n constraints. Consider a wave of the first family in an outgoing road I_j connecting $(\rho_{j,0}, y_{j,0})$ with $(\tilde{\rho}_j, \tilde{y}_j) \in \mathring{\mathcal{D}}_2$. We denote by $\widetilde{\Omega}_j$ the set defined by (6.4.19) where $(\tilde{\rho}_j, \tilde{y}_j)$ is the initial state. If $\sup \widetilde{\Omega}_j > \sup \Omega_j$, then the maximum for (6.4.22) does not vary, but the I_j road is no more an active constraint since $\delta_j < \sup \widetilde{\Omega}_j$. Then the final state $(\hat{\rho}_j, \hat{y}_j) \in \mathring{\mathcal{D}}_1$ and the equilibrium is unstable.

Now let us consider the case $m = n$. We consider a perturbation in an outgoing road I_j by a wave of the first family connecting $(\rho_{j,0}, y_{j,0})$ with $(\tilde{\rho}_j, \tilde{y}_j)$. If the perturbation is sufficiently small, then we may suppose the following:

(a) the new solution $(\hat{\delta}_1, \ldots, \hat{\delta}_n)$ of the maximization problem (6.4.22) satisfies $\hat{\delta}_i < \sup \Omega_i$ for every $i \in \{1, \ldots, n\}$ (the incoming roads are not constraints for the maximization problem (6.4.22)) and the final states $(\hat{\rho}_i, \hat{y}_i)$ in the incoming roads belong to $\mathring{\mathcal{D}}_2$;

(b) $\hat{\delta}_j = \tilde{y}_j - \tilde{\rho}_j^{\gamma+1}$, the fluxes for the other outgoing roads remain the same and the final state $(\hat{\rho}_j, \hat{y}_j)$ in the I_j outgoing road coincides with $(\tilde{\rho}_j, \tilde{y}_j)$;

(c)
$$\sum_{i=1}^{n} |(\rho_{i,0}, y_{i,0}) - (\hat{\rho}_i, \hat{y}_i)| < C\,|(\rho_{j,0}, y_{j,0}) - (\tilde{\rho}_j, \tilde{y}_j)|\,,$$
where C is a positive constant.

Therefore the equilibrium is stable.

We may summarize all these results in the following.

Theorem 6.5.1 *If $(\rho_{j,0}, y_{j,0}) \in \mathring{\mathcal{D}}_1$ for every $j \in \{n+1, \ldots, n+m\}$, then the equilibrium is stable.*

If $m = n$, $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_2$ for every $i \in \{1, \ldots, n\}$ and $(\rho_{j,0}, y_{j,0}) \in \mathring{\mathcal{D}}_2$ for every $j \in \{n+1, \ldots, 2n\}$, then the equilibrium is stable.

Consider now the case $m = n = 2$. We discuss here the generic cases. For generic we mean that the active constraints are given exactly by two roads and the states belong to the interior of the admissible regions.

1. $(\rho_{1,0}, y_{1,0}) \in \mathring{\mathcal{D}}_1$, $(\rho_{2,0}, y_{2,0}) \in \mathring{\mathcal{D}}_1$, $(\rho_{3,0}, y_{3,0}) \in \mathring{\mathcal{D}}_1$, $(\rho_{4,0}, y_{4,0}) \in \mathring{\mathcal{D}}_1$. This case is covered by the previous theorem.

2. $(\rho_{1,0}, y_{1,0}) \in \mathring{\mathcal{D}}_1$, $(\rho_{2,0}, y_{2,0}) \in \mathring{\mathcal{D}}_2$, $(\rho_{3,0}, y_{3,0}) \in \mathring{\mathcal{D}}_2$, $(\rho_{4,0}, y_{4,0}) \in \mathring{\mathcal{D}}_1$. In this case the active constraints are given by the roads I_1 and I_3. By claim 1, we know that the datum $(\rho_{4,0}, y_{4,0})$ can not be perturbed. Consider a perturbation produced by a wave of the second family connecting $(\tilde{\rho}_2, \tilde{y}_2)$ with $(\rho_{2,0}, y_{2,0})$. If the strength of the wave is sufficiently small, then the maximization problem (6.4.22) admits the same maximum point. Therefore no change happens in road I_1 and I_3, and

$$|(\tilde{\rho}_2, \tilde{y}_2) - (\hat{\rho}_2, \hat{y}_2)| + |(\rho_{4,0}, y_{4,0}) - (\hat{\rho}_4, \hat{y}_4)| \leqslant C\,|(\tilde{\rho}_2, \tilde{y}_2) - (\rho_{2,0}, y_{2,0})|\,,$$

where C is a positive constant and $(\hat{\rho}_2, \hat{y}_2)$ and $(\hat{\rho}_4, \hat{y}_4)$ are the final states respectively in roads I_2 and I_4.

Consider now a perturbation produced by a wave connecting $(\tilde{\rho}_1, \tilde{y}_1)$ with $(\rho_{1,0}, y_{1,0})$. We may suppose the followings, provided the perturbation is small:

(a) the active constraints remain the roads I_1 and I_3;

(b) the final state in I_3 is $(\hat{\rho}_3, \hat{y}_3) = (\rho_{3,0}, y_{3,0})$, while in I_1 is $(\hat{\rho}_1, \hat{y}_1) = (\tilde{\rho}_1, \tilde{y}_1)$;

(c)

$$|(\hat{\rho}_2, \hat{y}_2) - (\rho_{2,0}, y_{2,0})| + |(\rho_{4,0}, y_{4,0}) - (\hat{\rho}_4, \hat{y}_4)| \leqslant C \left|(\tilde{\rho}_1, \tilde{y}_1) - (\rho_{1,0}, y_{1,0})\right|,$$

where C is a positive constant and $(\hat{\rho}_2, \hat{y}_2)$ and $(\hat{\rho}_4, \hat{y}_4)$ are the final states respectively in roads I_2 and I_4.

The case of a perturbation in I_3 is completely similar. Therefore this equilibrium is stable. The other cases, in which the active constraints are given by an incoming road and an outgoing road, are equal to this one and so stable.

3. $(\rho_{1,0}, y_{1,0}) \in \mathring{\mathcal{D}}_2$, $(\rho_{2,0}, y_{2,0}) \in \mathring{\mathcal{D}}_2$, $(\rho_{3,0}, y_{3,0}) \in \mathring{\mathcal{D}}_2$, $(\rho_{4,0}, y_{4,0}) \in \mathring{\mathcal{D}}_2$. This case is covered by the previous theorem.

We conclude with the following.

Theorem 6.5.2 *Let J be a junction with 2 incoming and 2 outgoing roads. A generic equilibrium is stable.*

6.5.2 (AR-2): maximize the density.

By Proposition 6.4.5, we know that all equilibria in outgoing roads must be in the region $\mathcal{D}_2$. We notice that the instability for the equilibrium for the Riemann problem at J happens when there is a jump in incoming roads from the region $\mathring{\mathcal{D}}_1$ to the region $\mathring{\mathcal{D}}_2$.

We have some possibilities.

1. $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_1$ for every $i \in \{1, \ldots, n\}$ and $y_{j,0} = \rho_{j,0}$ for some $j \in \{n+1, \ldots, n+m\}$. This implies that $\delta_{i,0} := y_{i,0} - \rho_{i,0}^{\gamma+1} = \sup \Omega_i$ for every $i \in \{1, \ldots, n\}$ and $\delta_{j,0} := y_{j,0} - \rho_{j,0}^{\gamma+1} \leqslant \sup \Omega_j$ for every $j \in \{n+1, \ldots, n+m\}$. Moreover there exists $j_1 \in \{n+1, \ldots, n+m\}$ such that $\delta_{j_1,0} = \sup \Omega_{j_1}$. This means that all the incoming roads and at least one outgoing road give a constraint for the maximization problem (6.4.22). This fact implies that the equilibrium is unstable. Indeed consider an incoming road I_i and a wave of the first family connecting $(\tilde{\rho}_i, \tilde{y}_i)$ with $(\rho_{i,0}, y_{i,0})$ such that the set $\widetilde{\Omega}_i$, defined as in (6.4.16) for the state $(\tilde{\rho}_i, \tilde{y}_i)$, strictly contains Ω_i. There are at least

n active constraints, so the point of maximum does not change and, if the perturbation is sufficiently small, then we produce a jump on the road I_i.

2. $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_1$ for every $i \in \{1, \ldots, n\}$ and $y_{j,0} < \rho_{j,0}$ for every $j \in \{n+1, \ldots, n+m\}$. Define η as

$$\eta := \min_{j \in \{n+1,\ldots,n+m\}} \left\{ \sup \Omega_j - \left(y_j - \rho_j^{\gamma+1} \right) \right\},$$

then, by hypotheses we have that $\eta > 0$. Assume that a wave of the first family on an outgoing road I_j connecting $(\rho_{j,0}, y_{j,0})$ with $(\tilde{\rho}_j, \tilde{y}_j)$ arrives to J. If the perturbation is sufficiently small, then the new set Ω_j defined as in (6.4.19) with the new state $(\tilde{\rho}_j, \tilde{y}_j)$ satisfies

$$\left| \sup \tilde{\Omega}_j - \sup \Omega_j \right| \leqslant \frac{\eta}{2}$$

and this implies that the maximization problem (6.4.22) remains unchanged. Then only a wave of the second family on I_j connecting $(\hat{\rho}_j, \hat{y}_j)$ with $(\tilde{\rho}_j, \tilde{y}_j)$ is created. Moreover if the perturbation is sufficiently small, then

$$|(\hat{\rho}_j, \hat{y}_j) - (\tilde{\rho}_j, \tilde{y}_j)| \leqslant C \, |(\rho_{j,0}, y_{j,0}) - (\tilde{\rho}_j, \tilde{y}_j)|,$$

where C is a positive constant. Now, suppose that a wave connecting $(\tilde{\rho}_i, \tilde{y}_i)$ with $(\rho_{i,0}, y_{i,0})$ arrives at J. Assume first $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_1$. If the perturbation is sufficiently small, then:

(a) the new solution of the maximization problem (6.4.22) is given by

$$(\delta_{1,0} := y_{1,0} - \rho_{1,0}^{\gamma+1}, \ldots, \hat{\delta}_i, \ldots, \delta_{n,0} := y_{n,0} - \rho_{n,0}^{\gamma+1})$$

with $\hat{\delta}_i := \tilde{y}_i - \tilde{\rho}_i^{\gamma+1}$ and the final state $(\hat{\rho}_i, \hat{y}_i)$ is equal to $(\tilde{\rho}_i, \tilde{y}_i)$;

(b) the solution

$$(\hat{\delta}_{n+1}, \ldots, \hat{\delta}_{n+m})^T = A \cdot (\delta_{1,0}, \ldots, \hat{\delta}_i, \ldots, \delta_{n,0})^T$$

satisfies $\hat{\delta}_j < \sup \Omega_j$ for every $j \in \{n+1, \ldots, n+m\}$, the final states $(\hat{\rho}_j, \hat{y}_j)$ are such that $\hat{y}_j < \hat{\rho}_j$ for every $j \in \{n+1, \ldots, n+m\}$ (the outgoing roads are not constraints for the maximization problem (6.4.22)) and

$$\sum_{j=n+1}^{n+m} |(\hat{\rho}_j, \hat{y}_j) - (\rho_{j,0}, y_{j,0})| < C \, |(\tilde{\rho}_i, \tilde{y}_i) - (\rho_{i,0}, y_{i,0})|$$

for some C positive constant.

If, on the contrary, $(\rho_{i,0}, y_{i,0})$ is on the curve of maxima, then $\left|\hat{\delta}_i - \delta_{i,0}\right|$ is proportional to the incoming wave, $(\hat{\rho}_i, \hat{y}_i)$ is on the curve of maxima, (b) holds and we conclude similarly. So the equilibrium is stable.

3. $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_2$ for every $i \in \{1, \dots, n\}$, and $y_{j,0} = \rho_{j,0}$ for at least n indices $j \in \{n+1, \dots, n+m\}$. Define

$$\eta := \min_{i \in \{1,\dots,n\}} \left\{ \sup \Omega_i - \left(y_i - \rho_i^{\gamma+1} \right) \right\}.$$

If from an incoming road I_i a wave connecting $(\tilde{\rho}_i, \tilde{y}_i)$ with $(\rho_{i,0}, y_{i,0})$ arrives at J, then the new set $\tilde{\Omega}_i$, defined as in (6.4.16) for the state $(\tilde{\rho}_i, \tilde{y}_i)$, satisfies

$$\left| \sup \tilde{\Omega}_i - \sup \Omega_i \right| \leqslant \frac{\eta}{2}$$

provided that the perturbation is sufficiently small. Thus the maximization problem (6.4.22) remains unchanged and only a wave of the first family connecting $(\tilde{\rho}_i, \tilde{y}_i)$ with $(\hat{\rho}_i, \hat{y}_i)$ is created. Moreover,

$$\left| (\hat{\rho}_i, \hat{y}_i) - (\tilde{\rho}_i, \tilde{y}_i) \right| \leqslant C \left| (\rho_i, y_i) - (\tilde{\rho}_i, \tilde{y}_i) \right|,$$

where C is a positive constant.

A similar case happens if the perturbation is on an outgoing road I_j with $y_{j,0} < \rho_{j,0}$.

Now, consider a wave connecting $(\rho_{j,0}, y_{j,0})$ with $(\tilde{\rho}_j, \tilde{y}_j)$ on an outgoing road I_j with $y_{j,0} = \rho_{j,0}$. For the maximization problem (6.4.22), the active constraints remain the same. Waves are produced only in incoming roads and on outgoing roads that give no active constraints. Then

$$\sum_{i=1}^{n+m} \left| (\hat{\rho}_i, \hat{y}_i) - (\rho_{i,0}, y_{i,0}) \right| \leqslant C \left| (\hat{\rho}_j, \hat{y}_j) - (\rho_{j,0}, y_{j,0}) \right|,$$

where C is a positive constant. Thus the equilibrium is stable.

Putting together all the previous results, we obtain the following.

Theorem 6.5.3 *If $(\rho_{i,0}, y_{i,0}) \in \mathcal{D}_1$ for every $i \in \{1, \dots, n\}$ and $y_{j,0} < \rho_{j,0}$ for every $j \in \{n+1, \dots, n+m\}$, then the equilibrium is stable.*

If $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_2$ for every $i \in \{1, \dots, n\}$ and $y_{j,0} = \rho_{j,0}$ for at least n outgoing roads, then the equilibrium is stable.

Consider now the generic case when $m = n = 2$. Generically the states in outgoing roads belong to the region $y < \rho$, hence the outgoing roads are not constraints. Therefore there is only one generic case: the incoming roads are constraints for the maximization problem (6.4.22). So this is a stable equilibrium by Theorem 6.5.3. We have the following.

Theorem 6.5.4 *Let J be a junction with 2 incoming and 2 outgoing roads. A generic equilibrium is stable.*

6.5.3 (AR-3): minimize the total variation.

Notice that, in this case, the instability for the equilibrium happens when there is a jump in incoming roads from the region $\mathring{\mathcal{D}}_1$ to the region $\mathring{\mathcal{D}}_2$. We have some possibilities.

1. For every index $j \in \{n+1, \ldots, n+m\}$, $(\rho_{j,0}, y_{j,0}) \in \mathcal{D} \setminus \{(\rho, y) : \rho = y, \rho \geqslant (\frac{1}{\gamma+1})^{1/\gamma}\}$. Then $(\rho_{i,0}, y_{i,0}) \in \mathcal{D}_1$, $y_{i,0} - \rho_{i,0}^{\gamma+1} = \sup \Omega_i$ for every $i \in \{1, \ldots, n\}$ and $y_{j,0} - \rho_{j,0}^{\gamma+1} < \sup \Omega_j$ for every $j \in \{n+1, \ldots, n+m\}$. Define

$$\eta := \min_{j \in \{n+1, \ldots, n+m\}} \left\{ \sup \Omega_j - \left(y_{j,0} - \rho_{j,0}^{\gamma+1} \right) \right\}.$$

Assume that a wave connecting $(\rho_{j,0}, y_{j,0})$ with $(\tilde{\rho}_j, \tilde{y}_j)$ reaches J. This may happen only if $(\rho_{j,0}, y_{j,0}) \in \mathring{\mathcal{D}}_2$. If the wave is sufficiently small, then

$$\left| \sup \widetilde{\Omega}_j - \sup \Omega_j \right| \leqslant \frac{\eta}{2}$$

which implies that the maximization problem (6.4.22) remains unchanged. Only a wave connecting $(\hat{\rho}_j, \hat{y}_j)$ with $(\tilde{\rho}_j, \tilde{y}_j)$ is created. Moreover, if the perturbation is small, then

$$\left| (\hat{\rho}_j, \hat{y}_j) - (\tilde{\rho}_j, \tilde{y}_j) \right| \leqslant C \left| (\rho_{j,0}, y_{j,0}) - (\tilde{\rho}_j, \tilde{y}_j) \right|,$$

with C positive constant. Now, consider a wave connecting $(\tilde{\rho}_i, \tilde{y}_i)$ with $(\rho_{i,0}, y_{i,0})$ on the incoming road I_i. If the perturbation is sufficiently small, then the maximization problem (6.4.22) has the following solution:

$$(\delta_{1,0}, \ldots, \tilde{\delta}_i, \ldots, \delta_{n,0}),$$

with $\tilde{\delta}_i := \sup \widetilde{\Omega}_i$ and $\delta_{l,0} := y_{l,0} - \rho_{l,0}^{\gamma+1}$. Moreover the fluxes of the density in the outgoing roads change in a continuous way with respect the strength of the perturbation. Thus this equilibrium is stable.

2. $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_1$ for every $i \in \{1, \ldots, n\}$ and $y_{j,0} = \rho_{j,0}$, $(\rho_{j,0}, y_{j,0}) \in \mathcal{D}_2$ for some $j \in \{n+1, \ldots, n+m\}$. This is an unstable case. In fact, if a wave on an incoming road I_i reaches J, in such a way the set Ω_i increases, then the maximization problem (6.4.22) admits the same point of maximum (at least one outgoing road is an active constraint) and a jump happens in the incoming road I_i.

3. $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_2$ for every $i \in \{1, \ldots, n\}$ and $y_{j,0} = \rho_{j,0}$, $(\rho_{j,0}, y_{j,0}) \in \mathcal{D}_2$ for at least n indices $j \in \{n+1, \ldots, n+m\}$. The active constraints are given by the outgoing roads. For small perturbations these are again the only active constraints. Thus the equilibrium is stable.

Putting together all the previous results we have:

Theorem 6.5.5 *If $(\rho_{i,0}, y_{i,0}) \in \mathcal{D}_1$ for every $i \in \{1, \ldots, n\}$ and if, for every $j \in \{n+1, \ldots, n+m\}$, $(\rho_{j,0}, y_{j,0}) \in \mathcal{D} \setminus \{(\rho, y) : \rho = y, \rho \geqslant (\frac{1}{\gamma+1})^{1/\gamma}\}$, then the equilibrium is stable.*

If $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_2$ for every $i \in \{1, \ldots, n\}$ and $y_{j,0} = \rho_{j,0}$, $(\rho_{j,0}, y_{j,0}) \in \mathcal{D}_2$ for at least n indices $j \in \{n+1, \ldots, n+m\}$, then the equilibrium is stable.

Consider now the generic case when $m = n = 2$. As in the previous subsection, the outgoing roads are not constraints. Therefore there is only one generic case: the incoming roads are constraints for the maximization problem (6.4.22). So this is a stable equilibrium by Theorem 6.5.5. We have the following.

Theorem 6.5.6 *Let J be a junction with 2 incoming and 2 outgoing roads. A generic equilibrium is stable.*

6.6 Existence of solution at a junction.

Fix a road network with only one junction J with n incoming and m outgoing roads and fix $((\rho_{1,0}, y_{1,0}), \ldots, (\rho_{n+m,0}, y_{n+m,0}))$, a stable equilibrium for the Riemann Problem at J with one of the additional rules (AR-1),(AR-2) or (AR-3). Assume the following hypothesis:

(H) there exist $k_1, k_2 \in \{1, 2\}$ such that $(\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_{k_1}$ for every $i \in \{1, \ldots, n\}$ and $(\rho_{j,0}, y_{j,0}) \in \mathring{\mathcal{D}}_{k_2}$ for every $j \in \{n+1, \ldots, n+m\}$.

By the analysis of the previous section, the following proposition holds.

Proposition 6.6.1 *There exists a positive constant C such that, if a wave in an incoming road I_i connecting $(\tilde{\rho}_i, \tilde{y}_i)$ with $(\rho_{i,0}, y_{i,0})$ arrives at J and if the wave has sufficiently small total variation, then the solution to the Riemann Problem at J $((\hat{\rho}_1, \hat{y}_1), \ldots, (\hat{\rho}_{n+m}, \hat{y}_{n+m}))$ has the following properties:*

1. *every $(\hat{\rho}_l, \hat{y}_l)$ ($l \in \{1, \ldots, n+m\}$) belongs to the same region ($\mathring{\mathcal{D}}_1$ or $\mathring{\mathcal{D}}_2$) of $(\rho_{l,0}, y_{l,0})$;*

2. $$\sum_{l=1, l\neq i}^{n+m} \left|(\hat{\rho}_l, \hat{y}_l) - (\rho_{l,0}, y_{l,0})\right| + \left|(\hat{\rho}_i, \hat{y}_i) - (\tilde{\rho}_i, \tilde{y}_i)\right| \leqslant C \left|(\tilde{\rho}_i, \tilde{y}_i) - (\rho_{i,0}, y_{i,0})\right|.$$

The same holds for a perturbation on an outgoing road.

Theorem 6.6.1 *There exists $\varepsilon > 0$ such that the following holds. For every initial datum $((\rho_{1,0}(x), y_{1,0}(x)), \ldots, (\rho_{n+m,0}(x), y_{n+m,0}(x)))$ with*

$$\left\|(\rho_{l,0}(x), y_{l,0}(x))\right\|_{BV} \leqslant \varepsilon$$

and

$$\sup_{x\in(a_l,b_l)} |\rho_{l,0}(x) - \rho_{l,0}| + \sup_{x\in(a_l,b_l)} |y_{l,0}(x) - y_{l,0}| \leqslant \varepsilon$$

for every $l \in \{1, \ldots, n+m\}$, there exists a solution

$$((\rho_1(t,x), y_1(t,x)), \ldots, (\rho_{n+m}(t,x), y_{n+m}(t,x))),$$

defined for every $t \geqslant 0$, such that

1. *$(\rho_l(0,x), y_l(0,x)) = (\rho_{l,0}(x), y_{l,0}(x))$ for a.e. $x \in I_l$ and for every $l \in \{1, \ldots, n+m\}$;*

2. *$(\rho_l(t,x), y_l(t,x))$ is an entropic solution to (6.1.1) on each road I_l;*

3. *for a.e. $t > 0$,*

$$((\rho_1(t,b_1-), y_1(t,b_1-)), \ldots, (\rho_{n+m}(t,a_{n+m}+), y_{n+m}(t,a_{n+m}+)))$$

provides an equilibrium at J.

Proof. We consider a wave front tracking approximate solution, see [24]. For every $t > 0$, we denote by (x_k^i, σ_k^i) and (z_l^i, θ_l^i) the positions and strengths in the road I_i of all waves respectively of the first family and of the second family, where k and l belong to some finite sets of indices. For every road I_i, we consider as in [24] the two functionals

$$V_i(t) := \sum_k \left|\sigma_k^i\right| + \sum_l \left|\theta_l^i\right|$$

and

$$Q_i(t) := \sum_{z_l^i < x_k^i} \left|\sigma_k^i \theta_l^i\right| + \sum_{\sigma_k^i < 0} \left|\sigma_k^i \sigma_{k'}^i\right|,$$

which are the classical components of the Glimm functional. We introduce also a functional $\widetilde{V}$ measuring the strength of waves approaching J. If $i \in \{1, \ldots, n\}$, then define

$$\widetilde{V}_i(t) := \begin{cases} V_i(t), & \text{if } (\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_1, \\ \sum_l \left|\theta_l^i\right|, & \text{if } (\rho_{i,0}, y_{i,0}) \in \mathring{\mathcal{D}}_2. \end{cases}$$

For an outgoing road I_j, we put

$$\widetilde{V}_j(t) := \begin{cases} 0, & \text{if } (\rho_{j,0}, y_{j,0}) \in \mathring{\mathcal{D}}_1, \\ \sum_k \left|\sigma_k^j\right|, & \text{if } (\rho_{j,0}, y_{j,0}) \in \mathring{\mathcal{D}}_2. \end{cases}$$

Define $V(t) := \sum_{i=1}^{n+m} V_i(t)$, $Q(t) := \sum_{i=1}^{n+m} Q_i(t)$ and $\widetilde{V}(t) := \sum_{i=1}^{n+m} \widetilde{V}_i(t)$. We claim that there exist two positive constants C_1 and C_2 such that the functional

$$\Upsilon(t) := V(t) + C_1 \widetilde{V}(t) + C_2 Q(t)$$

is decreasing in time.

For a moment we suppose that the claim holds. Then, for every $t > 0$,

$$\begin{aligned} \Upsilon(t) &\leqslant \Upsilon(0) = V(0) + C_1 \widetilde{V}(0) + C_2 Q(0) \\ &\leqslant V(0) + C_1 V(0) + C_2 V^2(0) \end{aligned}$$

and, since Υ is equivalent as norm to the total variation, then the total variation of the approximate wave front tracking solution remains bounded for every $t > 0$, hence we have the conclusion by standard compactness arguments.

We prove now that Υ is decreasing in time. Clearly Υ changes only at times where two waves interact or a wave approaches J. If at a time $\tau > 0$ two waves interact in a road I_i, then, by standard evaluations (see [24]), we have

$$\begin{aligned}\Delta V_i(\tau) &\leqslant C \cdot \text{ product of strength of waves,}\\ \Delta \widetilde{V}_i(\tau) &\leqslant C \cdot \text{ product of strength of waves,}\\ \Delta Q_i(\tau) &\leqslant -\frac{\text{product of strength of waves}}{2},\end{aligned}$$

for some $C > 0$. If

$$\frac{C_2}{2} \geqslant C(1 + C_1), \tag{6.6.27}$$

then $\Delta\Upsilon \leqslant 0$ when waves interact in the roads. Consider now an interaction of a wave with J. For simplicity we assume that a wave of the second family (z_l^1, θ_l^1) arrives at J from the incoming road I_1 at time τ. The other cases are completely similar. By Proposition 6.6.1 , we have that:

$$\Delta V(\tau) \leqslant C \left|\theta_l^1\right|, \quad \Delta Q(\tau) \leqslant C \left|\theta_l^1\right| V(\tau-),$$

and

$$\Delta \widetilde{V}_1(\tau) = -\left|\theta_k^1\right|, \quad \Delta \widetilde{V}_i(\tau) = 0 \quad \text{for } i \neq 1.$$

Therefore $\Delta \widetilde{V}(\tau) = -\left|\theta_k^1\right|$ and

$$\Delta\Upsilon(\tau) \leqslant C \left|\theta_k^1\right| - C_1 \left|\theta_k^1\right| + C_2 C \left|\theta_k^1\right| T.V.(\tau-).$$

If

$$C_1 \geqslant C + C C_2 T.V., \tag{6.6.28}$$

then $\Delta\Upsilon \leqslant 0$ when a wave interacts with J.

Fix $C_1 \geqslant C$. Then it is possible to take C_2 satisfying (6.6.27). As long as (6.6.28) holds, i.e. as long as the total variation of the solution is bounded by a constant δ depending by C_1 and C_2, the functional Υ is decreasing in time. Therefore as long as the total variation is bounded by δ, then we have:

$$\begin{aligned}T.V.(t) \leqslant \Upsilon(t) \leqslant \Upsilon(0) &= V(0) + C_1 \widetilde{V}(0) + C_2 Q(0)\\ &\leqslant (1 + C_1)V(0) + C_2 V^2(0) \leqslant C_3 \cdot T.V.(0),\end{aligned}$$

for some constant $C_3 > 1$. Choosing $\varepsilon = \frac{\delta}{(n+m)C_3}$, we have that $T.V.(0) \leqslant \frac{\delta}{C_3} \leqslant \delta$ and $T.V.(t) \leqslant \delta$ for every $t > 0$. So we conclude the proof. □

Appendix: total variation of the flux.

In the case of a road network where the Lighthill–Whitham–Richards scalar model is considered in each road, if every junction has exactly 2 incoming and 2 outgoing roads, then an increment of the total variation of the flux can happen only when a wave on an outgoing road interacts with the junction; see [39] and Chapter 5. Here the situation is different since there are cases in which the total variation of the flux of the density strictly increases after an interaction of a wave from an incoming road, even if we are considering a junction with 2 incoming and 2 outgoing roads. In fact, consider a junction J with I_1 and I_2 incoming roads and I_3 and I_4 outgoing roads. Moreover suppose $\gamma = 1$ and the matrix A defined by

$$A = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{pmatrix}.$$

Consider the point

$$(\delta_{1,0}, \delta_{2,0}, \delta_{3,0}, \delta_{4,0}) = (1/8, 1/8, 5/48, 7/48).$$

It is clear that

$$\begin{pmatrix} \delta_{3,0} \\ \delta_{4,0} \end{pmatrix} = A \cdot \begin{pmatrix} \delta_{1,0} \\ \delta_{2,0} \end{pmatrix}.$$

We show that there exists an equilibrium configuration with $\delta_{i,0}$ as density fluxes. In I_1 we consider the point on the curve of maxima

$$(\rho_{1,0}, y_{1,0}) = \left(\frac{1}{2\sqrt{2}}, \frac{1}{4} \right)$$

so that $\Omega_1 = [0, 1/8]$ and $y_{1,0} - \rho_{1,0}^2 = 1/8$. In road I_2 we consider a point $(\rho_{2,0}, y_{2,0})$ such that $y_{2,0} - \rho_{2,0}^2 = 1/8$, $y_{2,0} < 2\rho_{2,0}^2$ and $\frac{1}{8} < \sup \Omega_2$. In road I_3 we consider the point

$$(\rho_{3,0}, y_{3,0}) = \left(\frac{1 + \sqrt{\frac{7}{12}}}{2}, \frac{1 + \sqrt{\frac{7}{12}}}{2} \right)$$

and so $\frac{5}{48} = y_{3,0} - \rho_{3,0}^2 = \sup \Omega_3$. Finally in I_4 we consider a point $(\rho_{4,0}, y_{4,0})$ such that $y_{4,0} - \rho_{4,0}^2 = \frac{7}{48} < \sup \Omega_4$. Notice that for every additional rule, it is possible to choose $(\rho_{4,0}, y_{4,0})$ such that $((\rho_{1,0}, y_{1,0}), \dots, (\rho_{4,0}, y_{4,0}))$ is an equilibrium for the Riemann problem at J. For this equilibrium, the active constraints are given by roads I_1 and I_3.

We perturb the equilibrium with a wave of the second family connecting $(\tilde{\rho}_1, \tilde{y}_1)$ with $(\rho_{1,0}, y_{1,0})$ such that the set $\widetilde{\Omega}_1$, defined as in (6.4.16) for the state $(\tilde{\rho}_1, \tilde{y}_1)$, is equal to $[0, 1/8 + \varepsilon]$, where ε is a small positive parameter. This is possible by taking

$$(\tilde{\rho}_1, \tilde{y}_1) = \left(\frac{\frac{1}{4} + 2\varepsilon}{\sqrt{\frac{1}{8} + \varepsilon}} - \frac{\sqrt{2}}{4}, \frac{\left(\frac{1}{4} + 2\varepsilon\right)^2}{\frac{1}{8} + \varepsilon} - \frac{\sqrt{2}}{4} \cdot \frac{\frac{1}{4} + 2\varepsilon}{\sqrt{\frac{1}{8} + \varepsilon}} \right) \in \mathring{\mathcal{D}}_2.$$

The new solution of (6.4.22) is given by

$$(\hat{\delta}_1, \hat{\delta}_2, \hat{\delta}_3, \hat{\delta}_4) = (1/8 + \varepsilon, 1/8 - 2\varepsilon/3, 5/48, 7/48 + \varepsilon/3).$$

Therefore the total variation of the first component of the flux after the interaction is given by

$$\begin{aligned} &\left|\hat{\delta}_1 - \delta_1\right| + \left|\hat{\delta}_2 - \delta_{2,0}\right| + \left|\hat{\delta}_3 - \delta_{3,0}\right| + \left|\hat{\delta}_4 - \delta_{4,0}\right| = \\ &= \hat{\delta}_1 - \delta_1 + \frac{2}{3}\varepsilon + \frac{\varepsilon}{3} = \frac{1}{8} - \delta_1 + 2\varepsilon, \end{aligned}$$

where $\delta_1 = y_1 - \rho_1^2 < \hat{\delta}_1$. Instead, the total variation of the first component of the flux before the interaction is given by

$$|\delta_1 - \delta_{1,0}| = \delta_1 - \frac{1}{8},$$

and so an increment of the total variation of the density flux happens, since $\delta_1 < \hat{\delta}_1 = 1/8 + \varepsilon$.

Part II

Hybrid systems.

Chapter 7

Hybrid Necessary Principle.

This chapter deals with hybrid systems. Roughly speaking a hybrid system is a collection of control systems called locations, possible defined on different manifolds, and an automaton that rules the switchings between locations. The definition of hybrid system is that of [59, 93, 110]. The term hybrid indicates the presence of both continuous and discrete dynamics. The continuous part is given by location controlled dynamics, while the discrete one by the automaton. An optimal control problem is obtained assigning Lagrangian running costs on each location and final and switching costs.

For an optimal classical control problem, the main tool toward the construction of an optimal trajectory is the Pontryagin Maximum Principle. For a hybrid system there exists a generalization of PMP, proved by Piccoli [93] in 1998 and by Sussmann [110] in 1999. The key point is the switching mechanism, that permits to pass from one location to another one with possible restrictions on state and time to spend in next location. The strategy to prove the Hybrid Maximum Principle (HMP) by Sussmann is essentially the same of PMP. Some variations are performed on the supposed optimal trajectory and they produce necessary conditions for a trajectory to be optimal. In hybrid setting, it is important to understand how variations propagate after a switching.

A more general case of switching mechanisms for hybrid systems than that of [93, 110] is considered. In particular, we assume that the switching strategy provides some restrictions on the set of admissible controls. These restrictions affect the general strategy of PMP and HMP. In fact, variations in PMP and HMP are generated by "needle variations", that are modifications of the control in a small interval of time and then prolonged after switchings.

In our setting, these variations are not admissible, in the sense that they produce a change in the switching strategy, hence we are not allowed to use the same control after switchings. Therefore we introduce a more general kind of variations, according to the fact that the switching strategy affects the choice of the controls, and we define the concept of “map of variations”. The basic requests are weak differentiability properties in the space of bounded Radon measure. We prove in this way a Hybrid Necessary Principle (HNP) giving necessary conditions for an optimal hybrid trajectory.

7.1 Basic Definitions and HMP

We start introducing the definition of hybrid system.

Definition 7.1.1 A *hybrid control system* is a 7-tuple

$$\Sigma = (\mathcal{Q}, M, U, f, \mathcal{U}, J, \mathcal{S}) \tag{7.1.1}$$

such that

H1. $\mathcal{Q}$ is a finite set;

H2. $M = \{M_q\}_{q\in\mathcal{Q}}$ is a family of smooth manifolds, indexed by $\mathcal{Q}$;

H3. $U = \{U_q\}_{q\in\mathcal{Q}}$ is a family of sets;

H4. $f = \{f_q\}_{q\in\mathcal{Q}}$ is a family of maps $f_q : M_q \times U_q \mapsto TM_q$ (TM_q is the tangent bundle of M_q), such that $f_q(x,u) \in T_xM_q$ for every $(x,u) \in M_q \times U_q$;

H5. $\mathcal{U} = \{\mathcal{U}_q\}_{q\in\mathcal{Q}}$ is a family of sets $\mathcal{U}_q$ whose members are maps $u : Dom(u) \to U_q$, defined on some interval $Dom(u) \subset \mathbb{R}$;

H6. $J = \{J_q\}_{q\in\mathcal{Q}}$ is a family of subintervals of $\mathbb{R}^+$;

H7. $\mathcal{S}$ is a subset of $Switch(\Sigma)$, where $Switch(\Sigma)$ is equal to

$$\{(q,x,q',x',u(\cdot),\tau) : q,q' \in \mathcal{Q}, x \in M_q, x' \in M_{q'}, u(\cdot) \in \mathcal{U}_{q'}, \tau \in J_{q'}\}. \tag{7.1.2}$$

The members of $\mathcal{Q}$ are called *locations* and represent the states of the automaton. The families M, U, are, respectively, the *family of state spaces* and the *family of control spaces* of Σ. For each q, the manifold M_q, the set U_q, the map f_q and the set $\mathcal{U}_q$ are, respectively, the *state space*, the *control space*, the *controlled dynamical law* and the *class of admissible controls* at location q.

The system evolves in a location q according to the corresponding controlled dynamic and then switches as prescribed by $\mathcal{S}$. The intervals J_q indicate the lengths of time intervals on which the system can stay in location q. So, for example, if $J_q = [0, +\infty[$ then the system can evolve in location q on every interval of time.

For $q, q' \in \mathcal{Q}$, the set $\mathcal{S}_{q,q'}$ is defined by

$$\{(x, x') \in M_q \times M_{q'} : (q, x, q', x', u(\cdot), \tau) \in \mathcal{S} \text{ for some } u(\cdot) \in \mathcal{U}_{q'} \text{ and } \tau \in J_{q'}\}.$$

The sets $\mathcal{S}_{q,q'}$ are called the *switching sets* of Σ from location q to location q'. Moreover, for $q, q' \in \mathcal{Q}$ and $x \in M_q$, $x' \in M_{q'}$, we write

$$\mathcal{U}_{q,x,q',x'} \stackrel{\text{def}}{=} \{u(\cdot) \in \mathcal{U}_{q'} : (q, x, q', x', u(\cdot), \tau) \in \mathcal{S} \text{ for some } \tau \in J_{q'}\}. \qquad (7.1.3)$$

The set $\mathcal{U}_{q,x,q',x'}$ is formed by the controls we can use at location q' if there is a switching from the point x of M_q to the point x' of $M_{q'}$.

Definition 7.1.2 *A hybrid state is a triplet (q, x, τ), where $q \in \mathcal{Q}$ is the location, $x \in M_q$ is the state of the control system and $\tau \in [0, \sup J_q)$ is the time since last switching. We denote by $\mathcal{HS}$ the set of all hybrid states.*

The evolution of the hybrid system is as follows. Given a hybrid initial state $(q_1, x_0, 0)$, at time t_0, on some time interval $[t_0, t_1[$, with $t_1 - t_0 \in J_{q_1}$, the system evolves solving:

$$\begin{cases} q(t) \equiv q_1, & \\ \dot{x}(t) = f_{q_1}(x(t), u_1(t)), & x(t_0) = x_0, \\ \dot{\tau}(t) = 1, & \tau(t_0) = 0, \end{cases} \qquad (7.1.4)$$

for some $u_1(\cdot) \in \mathcal{U}_{q_1}$ such that $Dom(u_1) \supset [t_0, t_1]$. This means that the system remains in location q_1 until $\tau = t_1 - t_0$ and it evolves on M_{q_1} according to the dynamic $f_{q_1}(x(t), u_1(t))$ for the control $u_1(\cdot) \in \mathcal{U}_{q_1}$. If the solution to the previous system can be prolonged on the whole interval $[t_0, t_1]$, then we can choose another hybrid state $(q_2, x_1, 0)$, a control $u_2(\cdot) \in \mathcal{U}_{q_2}$ and t_2 such that $(q_1, x(t_1), q_2, x_1, u_2(\cdot), t_2 - t_1) \in \mathcal{S}$ and let the system evolve in location q_2 following the corresponding controlled dynamics on the interval $[t_1, t_2]$:

$$\begin{cases} q(t) \equiv q_2, & \\ \dot{x}(t) = f_{q_2}(x(t), u_2(t)), & x(t_1) = x_1, \\ \dot{\tau}(t) = 1, & \tau(t_1) = 0. \end{cases} \tag{7.1.5}$$

We say that a location switching from q_1 to q_2 occurs at time t_1. Then we can proceed in the same way with a location switching and so on. Notice that the time t_1 (t_2 and so on) can be chosen freely in J_{q_1} (respectively J_{q_2} and so on), hence it represents a control for the hybrid system.

We assume that if $u \in \mathcal{U}_q$ then every time translation of u is in $\mathcal{U}_q$, more precisely we assume

(A1) If $u \in \mathcal{U}_q$ for some $q \in \mathcal{Q}$, then for every $\sigma \in \mathbb{R}$ the control $\tilde{u}(t) = u(t + \sigma)$ satisfies $\tilde{u} \in \mathcal{U}_q$.

Hence we can always assume that $t_0 = 0$.

Let us now give a precise definition of trajectories, cost functionals and optimal control problems.

Definition 7.1.3 *A trajectory is a map $\boldsymbol{X} : [0, T] \to \mathcal{HS}$ such that $\boldsymbol{X}(t) = (q(t), x(t), \tau(t))$ and the following holds. There exist $0 = t_0 < t_1 < \ldots < t_\nu = T$ such that, if $i \in \{1, \ldots, \nu\}$, then $q(\cdot)$ is constant in $[t_{i-1}, t_i[$ and equal to $q_i \in \mathcal{Q}$, $\tau(t) = t - t_{i-1}$ on $[t_{i-1}, t_i[$, $t_i - t_{i-1} \in J_{q_i}$. Moreover, for every $i \in \{1, \ldots, \nu\}$, there exists $u_i \in \mathcal{U}_{q_i}$ such that:*

- *$x_i(\cdot) := x|_{]t_{i-1}, t_i[}(\cdot)$ is an absolutely continuous function in $]t_{i-1}, t_i[$, continuously prolongable to $[t_{i-1}, t_i]$;*
- *$\frac{d}{dt} x_i(t) = f_{q_i}(x_i(t), u_i(t))$ for a.e. $t \in]t_{i-1}, t_i[$;*
- *$(x_i(t_i), x_{i+1}(t_i)) \in \mathcal{S}_{q_i, q_{i+1}}$ if $i = 1, \ldots, \nu - 1$;*
- *$u_{i+1} \in \mathcal{U}_{q_i, x_i(t_i), q_{i+1}, x_{i+1}(t_i)}$ if $i = 1, \ldots, \nu - 1$.*

Remark 17 *In this setting, for a Cauchy type problem, it is not appropriate to choose first a sequence of controls and then determine the trajectory associated to it, because a priori the sequence could not be admissible, in the sense that there could exist no trajectory corresponding to it. This is due to the fact that in every location q, it is possible to use, as controls, only a subset of $\mathcal{U}_q$, depending on the switching strategy.*

Definition 7.1.4 *If Σ is a hybrid system, then a Lagrangian for Σ is a family $L = \{L_q\}_{q\in\mathcal{Q}}$, $L_q : M_q \times U_q \to \mathbb{R}$ such that, for every trajectory $\boldsymbol{X}$, for every $i \in \{1,\dots,\nu\}$ and for every control u_i associated to x_i, the function $t \mapsto L_{q_i}(x_i(t), u_i(t))$ is integrable in $]t_{i-1}, t_i[$.*

Definition 7.1.5 *If Σ is a hybrid system, then a switching cost function is a family $\Phi = \{\Phi_{q,q'}\}_{(q,q')\in\mathcal{Q}\times\mathcal{Q}}$ such that each $\Phi_{q,q'}$ is a real valued function defined on $\mathcal{S}_{q,q'}$.*

Definition 7.1.6 *If Σ is a hybrid system, then an endpoint cost function is a family $\varphi = \{\varphi_{q,q'}\}_{(q,q')\in\mathcal{Q}\times\mathcal{Q}}$ such that each $\varphi_{q,q'}$ is a real valued function defined on $M_q \times M_{q'}$.*

If $L = \{L_q\}_{q\in\mathcal{Q}}$ is a Lagrangian, $\Phi = \{\Phi_{q,q'}\}_{(q,q')\in\mathcal{Q}\times\mathcal{Q}}$ is a switching cost function, $\varphi = \{\varphi_{q,q'}\}_{(q,q')\in\mathcal{Q}\times\mathcal{Q}}$ is an endpoint cost function for the hybrid control system Σ, then we can define the corresponding *cost functional* C, by letting

$$\begin{aligned} C(\mathbf{X}) \;=\; & \sum_{j=1}^{\nu} \int_{t_{j-1}}^{t_j} L_{q_j}(x_j(t), u_j(t))\, dt + \sum_{j=1}^{\nu-1} \Phi_{q_j,q_{j+1}}(x_j(t_j), x_{j+1}(t_j)) + \\ & + \varphi_{q_1,q_\nu}(x_1(t_0), x_\nu(t_\nu)), \end{aligned}$$

where $\mathbf{X}$ is a trajectory for Σ.

Definition 7.1.7 *Given a hybrid control system Σ, a cost functional C and two non empty subsets $\mathcal{N}_{in}, \mathcal{N}_{fin}$ of $\mathcal{HS}$, we call with $\mathcal{P}$ the problem of minimizing $C(\boldsymbol{X})$ over all trajectories $\boldsymbol{X}$ for Σ such that:*

i) $(q_1, x_1(t_0), 0) \in \mathcal{N}_{in}$ *;*

ii) $(q_\nu, x_\nu(t_\nu), t_\nu - t_{\nu-1}) \in \mathcal{N}_{fin}$.

Remark 18 *Note that there could be no trajectory satisfying boundary data. However, we expect that in many applications the set $\mathcal{N}_{fin}$ should be chosen so to impose restriction only on the final location q and point x. So if $(q, x, t) \in \mathcal{N}_{fin}$ then $\mathcal{N}_{fin}$ should contain also all the points (q, x, s) with $s \le \sup J_{q_\nu}$ (with possible equality only if $\sup J_{q_\nu} \in J_{q_\nu}$).*

The *Maximum Principle* gives a necessary condition for a trajectory $\mathbf{X}$ to be a solution of $\mathcal{P}$. The set of variations involves trajectories having the same *history* (see [93]) of the candidate optimal one, that is having the same switching strategy. As suggested in [93], if there is a finite number of possible switching strategies, for the optimization problem $\mathcal{P}$, then the Maximum Principle can sometimes single out the optimal trajectory.

Definition 7.1.8 *If Σ is a hybrid system and L is a Lagrangian for Σ, then we say that (ψ, ψ_0) is an adjoint pair along a trajectory $\boldsymbol{X}$ if:*

1. *$\psi = (\psi_1, \ldots, \psi_\nu)$ is such that, for every $i \in \{1, \ldots, \nu\}$, $\psi_i : [t_{i-1}, t_i] \to T^* M_{q_i}$ is an absolutely continuous function, $\psi_i(t) \in T^*_{x_i(t)} M_{q_i}$ and*

$$\dot{\psi}_i(t) = - < \psi_i(t), \frac{\partial}{\partial x} f_{q_i}(x_i(t), u_i(t)) > + \psi_0 \frac{\partial}{\partial x} L_{q_i}(x_i(t), u_i(t))$$

for a.e. $t \in [t_{i-1}, t_i]$;

2. *$\psi_0 \in \mathbb{R}^+$.*

In order to state the switching condition, we need a concept of a tangent cone. In this thesis, as in [110], we use the notion of a Boltyanskii approximating cone.

Definition 7.1.9 *Let S be a subset of a smooth manifold $\mathcal{X}$ and let $\bar{s} \in S$. A Boltyanskii approximating cone to S at $\bar{s}$ is a closed convex cone K in the tangent space $T_{\bar{s}}\mathcal{X}$ such that there exists a neighborhood W of 0 in $T_{\bar{s}}\mathcal{X}$ and a continuous map $\omega : W \cap K \to S$ with the property that $\omega(0) = \bar{s}$ and $\omega(w) = \bar{s} + w + o(\|w\|)$ as $w \to 0$ via values in $W \cap K$.*

Definition 7.1.10 *If Σ is a hybrid system, L is a Lagrangian and Φ is a switching cost function, then we say that an adjoint pair (ψ, ψ_0) along a trajectory $\boldsymbol{X}$ satisfies the switching condition if*

$$(-\psi_i(t_i), \psi_{i+1}(t_i)) - \psi_0 \nabla \Phi_{q_i, q_{i+1}}(x_i(t_i), x_{i+1}(t_i)) \in K_i^\perp$$

for every $i \in \{1, \ldots, \nu - 1\}$, where K_i is a Boltyanskii approximating cone to the set $\mathcal{S}_{q_i, q_{i+1}}$ at the point $(x_i(t_i), x_{i+1}(t_i))$ and $K_i^\perp$ is its polar cone.

Definition 7.1.11 *If (ψ, ψ_0) is an adjoint pair along $\boldsymbol{X}$, and*

$$H_i := \sup\{< \psi_i(t), f_{q_i}(x_i(t), u) > -\psi_0 L_{q_i}(x_i(t), u) : u \in U_{q_i}\},$$

then we say that (ψ, ψ_0) satisfies the Hamiltonian value condition *if, for every $i \in \{1,\ldots, \nu - 1\}$,*

- *if $t_i - t_{i-1} \in \mathrm{Int}(J_{q_i})$, then $H_i = H_\nu = 0$;*
- *if $t_i - t_{i-1}$ is the left endpoint of J_{q_i}, but J_{q_i} is nontrivial, then $H_i \leqslant 0$;*
- *if $t_i - t_{i-1}$ is the right endpoint of J_{q_i}, but J_{q_i} is nontrivial, then $H_i \geqslant 0$.*

As explained in the introduction for "simple" switching constraints a Hybrid Maximum Principle is valid. The condition ensuring this is precisely the following:

Assumption (H). For every fixed $q, q' \in \mathcal{Q}$, $x \in M_q$, $x' \in M_{q'}$, we have $\mathcal{U}_{q,x,q',x'} = \mathcal{U}_{q'}$.

Assumption (H) says that in every location $q \in \mathcal{Q}$ we can use always all the controls which are in $\mathcal{U}_q$. Thus the admissible controls do not depend on the location switchings. In particular, the classical "needle variations" are still admissible variations.

Before stating the Hybrid Maximum Principle, we need the definitions of *Hamiltonian maximization*, *nontriviality* and *transversality*, see [95, 109, 110].

Definition 7.1.12 *If Σ is a hybrid system, L is a Lagrangian and Φ is a switching cost function, then we say that an adjoint pair (ψ, ψ_0) along a trajectory $\boldsymbol{X}$ satisfies the Hamiltonian maximization condition if, for every $i \in \{1, \ldots, \nu\}$, the identity*

$$< \psi_i(t), f_{q_i}(x_i(t), u_i(t)) > -\psi_0 L_{q_i}(x_i(t), u_i(t)) = H_i(x_i(t), \psi_i(t), \psi_0)$$

holds for almost every $t \in [t_{i-1}, t_i]$.

Definition 7.1.13 *If Σ is a hybrid system, L is a Lagrangian and Φ is a switching cost function, then we say that an adjoint pair (ψ, ψ_0) along a trajectory $\boldsymbol{X}$ satisfies the nontriviality condition if either $\psi_0 \neq 0$ or there exists $i \in \{1, \ldots, \nu\}$ such that the function ψ_i is not identically zero.*

Definition 7.1.14 *If Σ is a hybrid system, L is a Lagrangian and Φ is a switching cost function, then we say that an adjoint pair (ψ, ψ_0) along a trajectory $\boldsymbol{X}$ satisfies the transversality condition if*

$$(-\psi_\nu(t_\nu), \psi_1(t_0)) - \psi_0 \nabla \tilde{\varphi}_{q_1,q_\nu}(x_\nu(t_\nu), x_1(t_0)) \in K_e^\perp,$$

where K_e is a Boltyanskii approximating cone to the projection of $\mathcal{N}_{fin} \times \mathcal{N}_{in}$ on $M_{q_\nu} \times M_{q_1}$ at $(x_\nu(t_\nu), x_1(t_0))$ and $\tilde{\varphi}_{q_1,q_\nu}(x, x') := \varphi_{q_1,q_\nu}(x', x)$ for every $x' \in M_{q_1}$, $x \in M_{q_\nu}$.

Hybrid Maximum Principle. *Consider the problem $\mathcal{P}$ and assume (H). Let $\boldsymbol{X}$ be a solution for $\mathcal{P}$. Then, under suitable assumptions, there exists an adjoint pair (ψ, ψ_0) along $\boldsymbol{X}$ that satisfies the switching condition, the Hamiltonian maximization, nontriviality, transversality, and Hamiltonian value conditions for $\mathcal{P}$.*

A proof of this result can be found in [110].

7.2 Simple necessary conditions

In this section we present some introductory results about necessary conditions for optimality for hybrid systems that do not satisfy assumption (H). We postpone to the next section the statement and the proof of the *Hybrid Necessary Principle*, the main result of this chapter. So this section is intended as a clarifying introduction to the subject of the next section.

Assumption (H) is restrictive in many mechanical systems. For example, to describe a car with gears, a hybrid system, where each location corresponds to a gear of the car, can be used. In this case it is clear that, when a switch from a low gear to the next one happens, not all the controls (for example strength of accelerations) can be used. For instance, if the change of a gear happens at low speed, then a strong acceleration may cause the stop of the engine of the car. Hence the necessity to consider hybrid systems without assumption (H). In this case we produce a weaker result that that of [110], but we have results in more general and complicate situations. In Section 8.3 we present a simple model of a car with two gears, which does not satisfy assumption (H).

In order to avoid too many technicalities, we prefer to consider simplified hypotheses about the manifolds, the vector fields and the lagrangians. However it is possible to prove all the results of this paper in a similar way using weaker assumptions. Therefore, we suppose that every M_q is equal to $\mathbb{R}^{d_q}$ for some $d_q \in \mathbb{N}$, $d_q \geqslant 1$ and that every U_q is a compact subset of $\mathbb{R}^l$ for some $l \in \mathbb{N}$, $l \geqslant 1$. So $f_q : \mathbb{R}^{d_q} \times U_q \to \mathbb{R}^{d_q}$ and we assume that

$$f_q \in C^2(\mathbb{R}^{d_q} \times U_q; \mathbb{R}^{d_q}), \tag{7.2.6}$$

hence, for every compact $K \subseteq \mathbb{R}^{d_q}$ there exists a constant $\Gamma_K > 0$ such that

$$\begin{cases} |f_q(x,u) - f_q(y,u)| \leqslant \Gamma_K \, |x-y| & \forall x, y \in K \quad \forall u \in U_q \\ |f_q(x,u) - f_q(x,v)| \leqslant \Gamma_K \, |u-v| & \forall x \in K \quad \forall u, v \in U_q. \end{cases} \tag{7.2.7}$$

Besides, we consider the case $\mathcal{U}_q = L^{p_q}_{loc}(\mathbb{R}; U_q)$ for some $1 \leqslant p_q \leqslant +\infty$ and $L_q \in C^2(\mathbb{R}^{d_q} \times U_q; \mathbb{R})$. Obviously we are in the situation of local existence and uniqueness for every Cauchy problem.

Needle variations are the basic tool to prove the *Pontryagin Maximum Principle* in non–hybrid setting and the *Hybrid Maximum Principle* in hybrid setting. Needle variations consist in modifying the supposed optimal control in a small interval of times and to understand how the trajectory and the cost vary in this way. Unfortunately, under our hypotheses, since the choice of admissible controls depends by the switching strategy, needle variations do not produce admissible trajectories. Therefore, it is necessary to modify the notion of needle variation.

For the aim of simplicity, we consider only admissible needle variations of the following type: the control is the same of the candidate optimal trajectory until a certain time $\bar{\tau}$, then we produce a constant variation for a small interval of times and finally, in the following locations, we consider controls satisfying the switching conditions and some continuity and differentiability properties. More precisely:

Definition 7.2.1 *Let us fix a trajectory* $\boldsymbol{X}$ *and* $i \in \{1, \ldots, \nu\}$*. We say that the family of trajectories* $\boldsymbol{X}^\varepsilon = (q, x^\varepsilon, \tau)$*,* $\boldsymbol{X}^\varepsilon : [0,T] \to \mathcal{HS}$ *(*$\varepsilon > 0$*) is an* ***admissible needle variation*** *at location* i *if*

1. $\boldsymbol{X}^0 \equiv \boldsymbol{X}$;

2. $\boldsymbol{X}^\varepsilon(t) = \boldsymbol{X}(t)$ *for every* $t \in [0, t_{i-1}]$;

3. *the curves* $\varepsilon \mapsto x_j^\varepsilon(t_{j-1})$ *are differentiable at* $\varepsilon = 0^+$ *for every* $j \in \{1, \ldots, \nu\}$;

4. *there exists a time* $\bar{\tau} \in [t_{i-1}+\varepsilon, t_i]$ *such that*

$$u_i^{\varepsilon}(t) = \begin{cases} u_i(t) & t \in [t_{i-1}, \bar{\tau}-\varepsilon[\\ \omega & t \in [\bar{\tau}-\varepsilon, \bar{\tau}[\\ u_i(t) & t \in [\bar{\tau}, t_i] \end{cases} \tag{7.2.8}$$

   for some $\omega \in U_{q_i}$*, where the symbol* u_j^{ε} *(*$j \in \{1, \dots, \nu\}$*) denotes the control at location* j *of* x_j^{ε}*;*

5. *for every* $j \in \{i+1, \dots, \nu\}$*,* $u_j^{\varepsilon} \to u_j$ *strongly in* $L^1([t_{j-1}, t_j])$ *as* $\varepsilon \to 0^+$ *and* $\frac{u_j^{\varepsilon}-u_j}{\varepsilon} \rightharpoonup \theta_j$ *weakly in* $L^1([t_{j-1}, t_j])$ *as* $\varepsilon \to 0^+$ *for some* $\theta_j \in L^1([t_{j-1}, t_j])$*.*

Remark 19 *Notice that, in Definition 7.2.1, we require that* $\boldsymbol{X}^{\varepsilon}$*, when* $\varepsilon > 0$*, is a family of trajectories. This means that, for a fixed* $\varepsilon > 0$*,* $\boldsymbol{X}^{\varepsilon}$ *is a trajectory and hence, by Definition 7.1.3,*

$$(x_j^{\varepsilon}(t_j), x_{j+1}^{\varepsilon}(t_j)) \in \mathcal{S}_{q_j, q_{j+1}}$$

for every $j \in \{1, \dots, \nu-1\}$ *and*

$$u_{j+1}^{\varepsilon} \in \mathcal{U}_{q_j, x_j(t_j), q_{j+1}, x_{j+1}(t_j)}$$

for every $j \in \{1, \dots, \nu-1\}$*.*

Moreover we require the existence of a location q_i*,* $i \in \{1, \dots, \nu\}$*, in which a variation originates. In particular we demand that, in the fixed location* q_i*, the variation is a classical needle variation and so the expression of the control* u_i^{ε} *is given in (7.2.8). In another location* q_j*,* $j \in \{1, \dots, \nu\}$*,* $j \neq i$*, we have the following possibilities.*

1. *If* $j < i$*, then* $u_j^{\varepsilon} = u_j$ *and* $x_j^{\varepsilon} = x_j$ *since the variation originates in location* q_i*.*

2. *If* $j > i$*, then we need some regularity properties of the control with respect to the parameter* ε*. These properties are described in 5 of Definition 7.2.1 and they imply that* u_j^{ε} *(*$j > i$*) cannot have an expression similar to that of equation (7.2.8), since otherwise the limit of* u_j^{ε} *as* $\varepsilon \to 0^+$ *does not belong to* $L^1([t_{j-1}, t_j])$*.*

For a needle variation $\mathbf{X}^\varepsilon$ we define:

$$v_j(t) = \frac{d}{d\varepsilon} x_j^\varepsilon(t)\big|_{\varepsilon=0}. \tag{7.2.9}$$

We have the following lemmata:

Lemma 7.2.1 *Let us assume (7.2.7). Let $\boldsymbol{X}^\varepsilon$ be an admissible needle variation. Then x^ε converges to x uniformly as ε goes to 0.*

Proof. It is sufficient to prove that, for every $j \in \{1, \ldots, \nu\}$, x_j^ε converges uniformly to x_j in $[t_{j-1}, t_j]$ as $\varepsilon \to 0$.
Obviously, if $1 \leqslant j < i$, then $x_j^\varepsilon = x_j$ and so the conclusion is true. Therefore we can treat the case $j \geqslant i$. For $t \in [t_{j-1}, t_j]$, we have, for some $\Gamma > 0$,

$$\begin{aligned} \left|x_j^\varepsilon(t) - x_j(t)\right| &\leqslant \left|x_j^\varepsilon(t_{j-1}) - x_j(t_{j-1})\right| + \\ &+ \left|\int_{t_{j-1}}^t \left[f_{q_j}(x_j^\varepsilon(s), u_j^\varepsilon(s)) - f_{q_j}(x_j(s), u_j(s))\right] ds\right| \\ &\leqslant \left|x_j^\varepsilon(t_{j-1}) - x_j(t_{j-1})\right| + \Gamma \int_{t_{j-1}}^t \left|x_j^\varepsilon(s) - x_j(s)\right| ds + \\ &+ \Gamma \int_{t_{j-1}}^t \left|u_j^\varepsilon(s) - u_j(s)\right| ds. \end{aligned}$$

Now, using Gronwall lemma, we obtain

$$\begin{aligned} \left|x_j^\varepsilon(t) - x_j(t)\right| &\leqslant \left|x_j^\varepsilon(t_{j-1}) - x_j(t_{j-1})\right| e^{\Gamma(t_j - t_{j-1})} + \\ &+ \Gamma \int_{t_{j-1}}^{t_j} \left|u_j^\varepsilon(s) - u_j(s)\right| ds\, e^{\Gamma(t_j - t_{j-1})}. \end{aligned}$$

Obviously, by definition of admissible needle variation, $\left|x_j^\varepsilon(t_{j-1}) - x_j(t_{j-1})\right|$ tends to 0 as $\varepsilon \to 0$ and also the last term goes to 0 as $\varepsilon \to 0$. So the proof is finished. □

Lemma 7.2.2 *Let us assume (7.2.6). Let $\boldsymbol{X}^\varepsilon$ be an admissible needle variation. Then $v_j \equiv 0$ if $j < i$,*

$$\begin{cases} \dot{v}_i(t) = D_x f_{q_i}(x_i(t), u_i(t)) v_i(t), \\ v_i(\bar{\tau}) = f_{q_i}(x_i(\bar{\tau}), \omega) - f_{q_i}(x_i(\bar{\tau}), u_i(\bar{\tau})), \end{cases} \tag{7.2.10}$$

in the i–location, while

$$\begin{cases} \dot{v}_j(t) = D_u f_{q_j}(x_j(t), u_j(t))\theta_j(t) + D_x f_{q_j}(x_j(t), u_j(t))v_j(t) \\ v_j(t_{j-1}) = \frac{d}{d\varepsilon} x_j^\varepsilon(t_{j-1})\big|_{\varepsilon=0} \end{cases} \tag{7.2.11}$$

if $j > i$.

Proof. Clearly, if $j < i$ then $v_j \equiv 0$. The case $j = i$ is well known, so we consider only the case $j > i$. In particular we prove the case $j = i + 1$, the other cases being similar.

For simplicity, let us denote with f, x, u respectively $f_{q_{i+1}}$, x_{i+1}, u_{i+1}. So, it is sufficient to prove that, if z in $[t_i, t_{i+1}]$ is the solution to

$$\begin{cases} \dot{z}(t) = D_u f(x(t), u(t))\theta_{i+1}(t) + D_x f(x(t), u(t))z(t) \\ z(t_i) = \frac{d}{d\varepsilon} x_{i+1}^\varepsilon(t_i)\big|_{\varepsilon=0} \end{cases}$$

then $z(t) = \frac{d}{d\varepsilon} x_{i+1}^\varepsilon(t)\big|_{\varepsilon=0}$ for almost every $t \in [t_i, t_{i+1}]$. In order to prove this, for $t \in [t_i, t_{i+1}]$, we estimate

$$\begin{aligned} \left| \frac{x_{i+1}^\varepsilon(t) - x(t)}{\varepsilon} - z(t) \right| &\leqslant \left| \frac{x_{i+1}^\varepsilon(t_i) - x(t_i)}{\varepsilon} - z(t_i) \right| + \\ &+ \frac{1}{\varepsilon} \left| \int_{t_i}^t \left[f(x_{i+1}^\varepsilon(s), u_{i+1}^\varepsilon(s)) - f(x(s), u(s)) \right] ds - \right. \\ &\left. -\varepsilon \int_{t_i}^t D_u f(x(s), u(s))\theta_{i+1}(s) ds - \varepsilon \int_{t_i}^t D_x f(x(s), u(s)) z(s) ds \right|. \end{aligned}$$

Fix $\delta > 0$. Then for ε sufficiently small (depending on δ), the first term of the right hand side is less than δ. Using Taylor's expansion

$$\begin{aligned} \left| \frac{x_{i+1}^\varepsilon(t) - x(t)}{\varepsilon} - z(t) \right| &\leqslant \delta + \left| \int_{t_i}^t D_x f(x(s), u(s)) \cdot \left[\frac{x_{i+1}^\varepsilon(s) - x(s)}{\varepsilon} - z(s) \right] ds \right| \\ &+ \left| \int_{t_i}^t D_u f(x(s), u(s)) \cdot \left[\frac{u_{i+1}^\varepsilon(s) - u(s)}{\varepsilon} - \theta_{i+1}(s) \right] ds \right| + \\ &\frac{c}{\varepsilon} \int_{t_i}^t \left[\left| x_{i+1}^\varepsilon(s) - x(s) \right| + \left| u_{i+1}^\varepsilon(s) - u(s) \right| \right]^2 ds \end{aligned} \tag{7.2.12}$$

where c is a positive constant, depending on the second derivatives of f. Now $x^{\varepsilon}_{i+1} \to x$ uniformly, so for ε sufficiently small it holds that

$$\begin{aligned} \frac{c}{\varepsilon}\int_{t_i}^{t} \left|x^{\varepsilon}_{i+1}(s) - x(s)\right|^2 ds &\leqslant c\delta \int_{t_i}^{t} \frac{\left|x^{\varepsilon}_{i+1}(s) - x(s)\right|}{\varepsilon} ds \\ &\leqslant c\delta \int_{t_i}^{t} \left|\frac{x^{\varepsilon}_{i+1}(s) - x(s)}{\varepsilon} - z(s)\right| ds + c\delta \int_{t_i}^{t} |z(s)|\, ds \\ &\leqslant c\delta \int_{t_i}^{t} \left|\frac{x^{\varepsilon}_{i+1}(s) - x(s)}{\varepsilon} - z(s)\right| ds + c_1\delta \end{aligned}$$

with c_1 positive constant. Moreover

$$\frac{c}{\varepsilon}\int_{t_i}^{t} \left|u^{\varepsilon}_{i+1}(s) - u(s)\right|^2 ds \leqslant c_2\delta$$

with c_2 positive constant, since $\frac{u^{\varepsilon}_{i+1}-u}{\varepsilon}$ converges weakly in $L^1([t_i, t_{i+1}])$ and $u^{\varepsilon}_{i+1} - u$ converges strongly to 0 in $L^1([t_i, t_{i+1}])$. For a detailed proof of this fact see Section 7.4. Analogously

$$\frac{2c}{\varepsilon}\int_{t_i}^{t} \left|u^{\varepsilon}_{i+1}(s) - u(s)\right| \left|x^{\varepsilon}_{i+1}(s) - x(s)\right| ds \leqslant c_2\delta.$$

The third addendum of the right hand side of (7.2.12) is estimated similarly, since $\frac{u^{\varepsilon}_{i+1}-u}{\varepsilon} \rightharpoonup \theta_{i+1}$ weakly in $L^1([t_i, t_{i+1}])$. Thus:

$$\left|\frac{x^{\varepsilon}_{i+1}(t) - x(t)}{\varepsilon} - z(t)\right| \leqslant M_1\delta + (M_2 + c\delta)\int_{t_i}^{t} \left|\frac{x^{\varepsilon}_{i+1}(s) - x(s)}{\varepsilon} - z(t)\right| ds,$$

where M_1, M_2 are positive constants, depending on f and $U_{q_{i+1}}$. Using Gronwall lemma we conclude that

$$\left|\frac{x^{\varepsilon}_{i+1}(t) - x(t)}{\varepsilon} - z(t)\right| \leqslant M_1\delta e^{(M_2+c\delta)(t_{i+1}-t_i)} \tag{7.2.13}$$

and so the lemma is proved, by the arbitrariness of $\delta > 0$. □

Remark 20 *From the last result, we note that the evolution equation for v_j in general is an affine equation, since a term depending by θ_j appears. This is due by definition of admissible needle variation. For hybrid systems with assumption (H), we may consider usual needle variations and so the resulting equation for v_j is linear, without the term containing θ_j.*

Remark 21 *It is useful to recall that equation (7.2.11) is valid only if $j > i$, i.e. only if the variation is originated in a previous location. Therefore, to prove equation (7.2.11) we do not consider expression (7.2.8), but we use properties 3 and 5 of Definition 7.2.1.*

Now, we want to evaluate how the Lagrangian cost varies for an admissible needle variation. If we define

$$G_\varepsilon(t) := \sum_{h=1}^{j-1} \int_{t_{h-1}}^{t_h} L_{q_h}(x_h^\varepsilon(s), u_h^\varepsilon(s))ds + \int_{t_{j-1}}^{t} L_{q_j}(x_j^\varepsilon(s), u_j^\varepsilon(s))ds$$

when $t_{j-1} \leqslant t < t_j$, and set $w(t) := \frac{d}{d\varepsilon} G_\varepsilon(t)\big|_{\varepsilon=0^+}$, then we can get the following result:

Lemma 7.2.3 *Let $\bar{\tau} \in]t_{i-1}, t_i[$ be the time at which an admissible needle variation originates. If $t \in]\bar{\tau}, t_i[$, then w satisfies the following differential equation:*

$$\begin{cases} \dot{w}(t) = \frac{\partial}{\partial x} L_{q_i}(x_i(t), u_i(t)) v_i(t) \\ w(\bar{\tau}) = L_{q_i}(x_i(\bar{\tau}), \omega) - L_{q_i}(x_i(\bar{\tau}), u_i(\bar{\tau})). \end{cases}$$

Moreover if $i < j \leqslant \nu$, then we have:

$$\begin{cases} \dot{w}(t) = \frac{\partial}{\partial x} L_{q_j}(x_j(t), u_j(t)) v_j(t) + \frac{\partial}{\partial u} L_{q_j}(x_j(t), u_j(t)) \theta_j(t), \ \ t_{j-1} < t < t_j \\ w(t_{j-1}) = \lim_{t \to t_{j-1}^-} w(t). \end{cases}$$

Proof. First we evaluate $w(\bar{\tau})$. By definition

$$w(\bar{\tau}) = \frac{d}{d\varepsilon} G_\varepsilon(\bar{\tau})\big|_{\varepsilon=0^+} = \lim_{\varepsilon=0^+} \frac{G_\varepsilon(\bar{\tau}) - G_0(\bar{\tau})}{\varepsilon}$$

and, by the fact that x^ε and u^ε coincide respectively with x and u before $\bar{\tau} - \varepsilon$, we conclude that

$$\begin{aligned} w(\bar{\tau}) &= \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \int_{\bar{\tau}-\varepsilon}^{\bar{\tau}} \left[L_{q_i}(x_i^\varepsilon(s), \omega) - L_{q_i}(x_i(s), u_i(s)) \right] ds \\ &= L_{q_i}(x_i(\bar{\tau}), \omega) - L_{q_i}(x_i(\bar{\tau}), u_i(\bar{\tau})). \end{aligned}$$

Now suppose that $\bar{\tau} < t < t_i$. In this case we have

$$\begin{aligned} w(t) = \lim_{\varepsilon\to 0^+} \Bigg\{ \frac{1}{\varepsilon} \int_{\bar{\tau}-\varepsilon}^{\bar{\tau}} \left[L_{q_i}(x_i^\varepsilon(s), \omega) - L_{q_i}(x_i^\varepsilon(s), u_i(s)) \right] ds + \\ + \frac{1}{\varepsilon} \int_{\bar{\tau}}^{t} \left[L_{q_i}(x_i^\varepsilon(s), u_i(s)) - L_{q_i}(x_i(s), u_i(s)) \right] ds \Bigg\} \\ = L_{q_i}(x_i(\bar{\tau}), \omega) - L_{q_i}(x_i(\bar{\tau}), u_i(\bar{\tau})) + \\ + \lim_{\varepsilon\to 0^+} \Bigg\{ \int_{\bar{\tau}}^{t} \frac{\partial}{\partial x} L_{q_i}(x_i(s), u_i(s)) \frac{x_i^\varepsilon(s) - x_i(s)}{\varepsilon} ds + \\ + \int_{\bar{\tau}}^{t} \frac{\partial^2}{\partial x^2} L_{q_i}(\tilde{x}_i(s), \tilde{u}(s)) \frac{(x_i^\varepsilon(s) - x_i(s))^2}{2\varepsilon} ds \Bigg\}. \end{aligned}$$

Using the estimate (7.2.13) we conclude by Lebesgue theorem that

$$w(t) = L_{q_i}(x_i(\bar{\tau}), \omega) - L_{q_i}(x_i(\bar{\tau}), u_i(\bar{\tau})) + \int_{\bar{\tau}}^{t} \frac{\partial}{\partial x} L_{q_i}(x_i(s), u_i(s)) v_i(s) ds.$$

So the first part of the lemma is proved.

In order to prove the last statement, note that, if $t_{j-1} < t < t_j$, then $w(t)$ is equal to:

$$\begin{aligned} w(t) = \left[L_{q_i}(x_i(\bar{\tau}), \omega) - L_{q_i}(x_i(\bar{\tau}), u_i(\bar{\tau})) \right] + \\ \int_{\bar{\tau}}^{t_i} \frac{\partial}{\partial x} L_{q_i}(x_i(s), u_i(s)) v_i(s) ds + \int_{t_i}^{t_{i+1}} \frac{\partial}{\partial x} L_{q_{i+1}}(x_{i+1}(s), u_{i+1}(s)) v_{i+1}(s) ds + \\ \int_{t_i}^{t_{i+1}} \frac{\partial}{\partial u} L_{q_{i+1}}(x_{i+1}(s), u_{i+1}(s)) \theta_{i+1}(s) ds + \ldots + \\ \int_{t_{j-1}}^{t} \frac{\partial}{\partial x} L_{q_j}(x_j(s), u_j(s)) v_j(s) ds + \int_{t_{j-1}}^{t} \frac{\partial}{\partial u} L_{q_j}(x_j(s), u_j(s)) \theta_j(s) ds. \end{aligned}$$

Indeed, if $l \in \{i+1, \ldots, j\}$, then we have for $t \in [t_{l-1}, t_l]$,

$$\lim_{\varepsilon \to 0^+} \int_{t_{l-1}}^{t} \frac{L_{q_l}(x_l^\varepsilon(s), u_l^\varepsilon(s)) - L_{q_l}(x_l(s), u_l(s))}{\varepsilon} ds =$$

$$\lim_{\varepsilon \to 0^+} \left\{ \int_{t_{l-1}}^{t} \frac{L_{q_l}(x_l^\varepsilon(s), u_l^\varepsilon(s)) - L_{q_l}(x_l(s), u_l^\varepsilon(s))}{\varepsilon} ds + \right.$$

$$\left. \int_{t_{l-1}}^{t} \frac{L_{q_l}(x_l(s), u_l^\varepsilon(s)) - L_{q_l}(x_l(s), u_l(s))}{\varepsilon} ds \right\}$$

$$= \int_{t_{l-1}}^{t} \frac{\partial}{\partial x} L_{q_l}(x_l(s), u_l(s)) v_l(s) ds +$$

$$+ \lim_{\varepsilon \to 0^+} \int_{t_{l-1}}^{t} \frac{L_{q_l}(x_l(s), u_l^\varepsilon(s)) - L_{q_l}(x_l(s), u_l(s))}{\varepsilon} ds.$$

Now the last term is equal to

$$\lim_{\varepsilon \to 0^+} \int_{t_{l-1}}^{t} \frac{L_{q_l}(x_l(s), u_l^\varepsilon(s)) - L_{q_l}(x_l(s), u_l(s))}{\varepsilon} ds =$$

$$\lim_{\varepsilon \to 0^+} \left\{ \int_{t_{l-1}}^{t} \frac{\partial}{\partial u} L_{q_l}(x_l(s), u_l(s)) \frac{u_l^\varepsilon(s) - u_l(s)}{\varepsilon} ds + \right.$$

$$\left. \int_{t_{l-1}}^{t} \frac{\partial^2}{\partial u^2} L_{q_l}(\tilde{x}_l(s), \tilde{u}_l(s)) \frac{u_l^\varepsilon(s) - u_l(s)}{\varepsilon} (u_l^\varepsilon(s) - u_l(s)) ds \right\} =$$

$$\int_{t_{l-1}}^{t} \frac{\partial}{\partial u} L_{q_l}(x_l(s), u_l(s)) \theta_l(s) ds +$$

$$\lim_{\varepsilon \to 0^+} \int_{t_{l-1}}^{t} \frac{\partial^2}{\partial u^2} L_{q_l}(\tilde{x}_l(s), \tilde{u}_l(s)) \frac{u_l^\varepsilon(s) - u_l(s)}{\varepsilon} (u_l^\varepsilon(s) - u_l(s)) ds$$

by definition of admissible needle variation. The last integral converges to 0 since $\frac{u_l^\varepsilon(s) - u_l(s)}{\varepsilon}$ converges weakly in $L^1([t_{l-1}, t_l])$ and $u_l^\varepsilon(s) - u_l(s)$ converges to 0 strongly in $L^1([t_{l-1}, t_l])$ and so the product converges to 0 weakly in $L^1([t_{l-1}, t_l])$, see Section 7.4. $\square$

Putting together all the previous results we have the following proposition.

Proposition 7.2.1 *Let $\boldsymbol{X}$ be a trajectory and let $\boldsymbol{X}^\varepsilon$ be an admissible needle variation. Then, for every adjoint pair (ψ, ψ_0) along $\boldsymbol{X}$ and for every $j \in \{1, \ldots, \nu\}$ the function*

$$\psi_j(t) \cdot v_j(t) - \psi_0 w(t) + q_j(t) \tag{7.2.14}$$

is constant in $[t_{j-1}, t_j]$, where q_j is any function defined by

$$\dot{q}_j(t) = -\psi_j(t)\frac{\partial}{\partial u} f_{q_j}(x_j(t), u_j(t))\theta_j(t) + \psi_0 \frac{\partial}{\partial u} L_{q_j}(x_j(t), u_j(t))\theta_j(t) \tag{7.2.15}$$

if $j > i$, while $q_j \equiv 0$ otherwise.

Proof. It is a simple consequence of Lemma 7.2.2 and Lemma 7.2.3. If $j < i$, then $v_j \equiv 0$, $q_j \equiv 0$ and $w(t) = 0$ for every $t \in [0, t_j]$.

If $j = i$, where q_i is the location at which the admissible needle variation originates, then we have

$$\begin{gathered}
\frac{d}{dt}\left[\psi_i(t) \cdot v_i(t) - \psi_0 w(t)\right] \\
= \dot{\psi}_i(t) \cdot v_i(t) + \psi_i(t) \cdot \dot{v}_i(t) - \psi_0 \dot{w}(t) \\
= -\psi_i(t) D_x f_{q_i}(x_i(t), u_i(t)) v_i(t) + \psi_0 \frac{\partial}{\partial x} L_{q_i}(x_i(t), u_i(t)) v_i(t) \\
+\psi_i(t) D_x f_{q_i}(x_i(t), u_i(t)) v_i(t) - \psi_0 \frac{\partial}{\partial x} L_{q_i}(x_i(t), u_i(t)) v_i(t) = 0
\end{gathered}$$

and so we have the thesis when $j = i$.

Now if $j > i$, then

$$\begin{gathered}
\frac{d}{dt}\left[\psi_j(t) \cdot v_j(t) - \psi_0 w(t) + q_j(t)\right] \\
= \dot{\psi}_j(t) \cdot v_j(t) + \psi_j(t) \cdot \dot{v}_j(t) - \psi_0 \dot{w}(t) + \dot{q}_j(t) \\
= -\psi_j(t) D_x f_{q_j}(x_j(t), u_j(t)) v_j(t) + \psi_0 \frac{\partial}{\partial x} L_{q_j}(x_j(t), u_j(t)) v_j(t) \\
+\psi_j(t) D_u f_{q_j}(x_j(t), u_j(t))\theta_j(t) + \psi_j(t) D_x f_{q_j}(x_j(t), u_j(t)) v_j(t) \\
-\psi_0 \frac{\partial}{\partial x} L_{q_j}(x_j(t), u_j(t)) v_j(t) - \psi_0 \frac{\partial}{\partial u} L_{q_j}(x_j(t), u_j(t))\theta_j(t) \\
-\psi_j(t) \cdot D_u f_{q_j}(x_j(t), u_j(t))\theta_j(t) + \psi_0 \frac{\partial}{\partial u} L_{q_j}(x_j(t), u_j(t))\theta_j(t) = 0
\end{gathered}$$

So, the proof is finished. □

Now, we study how to deduce some necessary conditions from the previous analysis. For clarity, we consider optimal control problems where the cost is formed only by the lagrangian part, that is the switching cost and the endpoint cost vanish. We suppose that $\mathbf{X}$ is an optimal trajectory and we consider an admissible needle variation $\mathbf{X}^\varepsilon$. Clearly, by optimality, $C(\mathbf{X}) \leqslant C(\mathbf{X}^\varepsilon)$. This implies that $w(T) \geqslant 0$. Let us consider an adjoint pair (ψ, ψ_0) along $\mathbf{X}$ with the properties that, for every $j \in \{1, \dots, \nu\}$,

$$\psi_j(t_j) \cdot v_j(t_j) \leqslant 0. \tag{7.2.16}$$

Thus

$$\psi_\nu(t_\nu) \cdot v_\nu(t_\nu) - \psi_0 w(t_\nu) \leqslant 0. \tag{7.2.17}$$

This implies that, for every $q_\nu(\cdot)$ defined as in Proposition 7.2.1 with $q_\nu(t_\nu) \leqslant 0$, it holds:

$$\psi_\nu(t) \cdot v_\nu(t) - \psi_0 w(t) + q_\nu(t) \leqslant 0 \tag{7.2.18}$$

for every $t \in [t_{\nu-1}, t_\nu]$. Therefore in the $\nu - 1$ location we have

$$\psi_{\nu-1}(t_{\nu-1}) \cdot v_{\nu-1}(t_{\nu-1}) - \psi_0 w(t_{\nu-1}) + q_\nu(t_{\nu-1}) \leqslant 0 \tag{7.2.19}$$

and so

$$\psi_{\nu-1}(t) \cdot v_{\nu-1}(t) - \psi_0 w(t) + q_\nu(t_{\nu-1}) + q_{\nu-1}(t) \leqslant 0 \tag{7.2.20}$$

for every $q_{\nu-1}$ with $q_{\nu-1}(t_{\nu-1}) \leqslant 0$ and for every $t \in [t_{\nu-2}, t_{\nu-1}]$.

Iterating this argument we conclude that

$$\psi_j(t) \cdot v_j(t) - \psi_0 w(t) + \sum_{l=j+1}^{\nu} q_l(t_{l-1}) + q_j(t) \leqslant 0 \tag{7.2.21}$$

for every $j \in \{1, \dots, \nu\}$, $t \in [t_{j-1}, t_j]$ and for every function q_l with $q_l(t_l) \leqslant 0$.

Equation (7.2.21) gives a necessary condition for optimality when the hybrid system does not satisfy assumption (H). In next section we generalize this approach.

7.3 Hybrid Necessary Principle

This section is dedicated to prove the Hybrid Necessary Principle if assumption (H) does not hold. Our result generalizes the Hybrid Maximum Principle proved in [110]. In this section we suppose that the assumptions (7.2.6) and (7.2.7) hold.

We recall first some basic facts about measure theory and in particular about Radon measures, see [54, Chapter 7, page 204].

Definition 7.3.1 *Let $a, b \in \mathbb{R}$, $a < b$. If $\mathcal{B}_{(a,b)}$ denotes the Borel σ–algebra on (a, b), then a signed measure $\mu : \mathcal{B}_{(a,b)} \to \overline{\mathbb{R}}$ is called a Radon measure if:*

1. $|\mu(K)| < +\infty$ *for every K, compact set in (a, b);*
2. $\mu(E) = \inf\{\mu(U) : E \subseteq U, U$ *open set in* $(a, b)\}$ *for every $E \in \mathcal{B}_{(a,b)}$;*
3. $\mu(E) = \sup\{\mu(K) : K \subseteq E, K$ *compact set in* $(a, b)\}$ *for every $E \in \mathcal{B}_{(a,b)}$.*

Moreover, if μ is finite, then μ is said bounded. We denote by $\mathfrak{M}_b(a, b; \mathbb{R})$ the set of bounded Radon measures on (a, b).

If $a, b \in \mathbb{R}$, $a < b$, then $L^1(a, b)$ is contained in $\mathfrak{M}_b(a, b; \mathbb{R})$. The inclusion is to be intended in the following way: to every function $f \in L^1(a, b)$ we associate a measure μ defined by

$$\mu(A) := \int_A f(t)dt \tag{7.3.22}$$

where A is a Borel subset of (a, b). The space $\mathfrak{M}_b(a, b; \mathbb{R})$, equipped with the norm $\|\mu\| := |\mu|(a, b)$, is equal to $(C(a, b))'$ (see [54]). So in $\mathfrak{M}_b(a, b; \mathbb{R})$ we may consider the weak* topology. We have that $\mu_n \rightharpoonup^* \mu$ as $n \to +\infty$ if and only if, for every $g \in C(a, b)$,

$$\int_a^b g(t)d\mu_n(t) \to \int_a^b g(t)d\mu(t) \qquad \text{as } n \to +\infty.$$

In the same way, we consider the space of bounded Radon measures on (a, b) with values in $\mathbb{R}^n$ and we indicate it by $\mathfrak{M}_b(a, b; \mathbb{R}^n)$.

Remark 22 *The measure μ defined in (7.3.22) is clearly a signed measure. In fact if the function f is strictly negative in a Borel set E, whose Lebesgue measure is strictly positive, then $\mu(E) < 0$.*

All the properties of Radon measures are clearly satisfied by any measure defined as in (7.3.22). Moreover, equation (7.3.22) says that $\mu \ll m$, that is μ is an absolutely continuous measure with respect to the Lebesgue measure.

We recall a result about the differentiability of a trajectory with respect to a parameter, used to prove the main result.

Definition 7.3.2 *Let P be a normed space, $\{\bar{x}^p\}_{p\in P}$ be a family in $\mathbb{R}^n$ and $\boldsymbol{f} = \{\boldsymbol{f}_p\}_{p\in P}$ be a family of time–varying vector fields on $\mathbb{R}^n$. For every $p \in P$, we denote with $x^p : [a,b] \to \mathbb{R}^n$ a solution to*

$$\begin{cases} \dot{x}^p(t) = \boldsymbol{f}_p(t, x^p(t)) & a.e.\ t \in [a,b] \\ x^p(a) = \bar{x}^p. \end{cases}$$

Let us fix an element $p_0 \in P$. We say that $\boldsymbol{f}$ is weakly differentiable at $p_0 \in P$ along $x^{p_0}(\cdot)$ if there exists $\bar{\varepsilon} > 0$ such that:

1. *for every p such that $\|p - p_0\| \leqslant \bar{\varepsilon}$ and for every $\xi : [a,b[\to \mathbb{R}^{d_q}$, solution to $\dot{\xi}(t) = \boldsymbol{f}_p(t, \xi(t))$, with $\|\xi(t) - x^{p_0}(t)\| \leqslant \bar{\varepsilon}$ for every $t \in [a,b[$, the limit $\lim_{t\to b^-} \xi(t)$ exists;*

2. *for every p such that $\|p - p_0\| \leqslant \bar{\varepsilon}$ and for every x such that $\|x - x^{p_0}(t)\| < \bar{\varepsilon}$ for some $a \leqslant t < b$, there exists a local forward solution to*

$$\begin{cases} \dot{\xi} = \boldsymbol{f}_p(t, \xi) \\ \xi(t) = x; \end{cases}$$

3. *there exists $A \in L^1([a,b]; \mathbb{R}^{n\times n})$ and positive functions $\zeta_\varepsilon \in L^1([a,b]; \mathbb{R}^+)$ for every $\varepsilon \in]0, \bar{\varepsilon}]$, such that*

$$\|\boldsymbol{f}_p(t,x) - \boldsymbol{f}_p(t, x^{p_0}(t)) - A(t)(x - x^{p_0}(t))\| \leqslant \zeta_\varepsilon(t)(\|x - x^{p_0}(t)\| + \|p - p_0\|)$$

whenever $a \leqslant t \leqslant b$, $\|x - x^{p_0}(t)\| \leqslant \varepsilon$, $\|p - p_0\| \leqslant \varepsilon$ and

$$\lim_{\varepsilon\to 0^+} \int_a^b \zeta_\varepsilon(t)dt = 0;$$

4. *the function* $\tilde{w}_p(t) := \boldsymbol{f}_p(t, x^{p_0}(t)) - \boldsymbol{f}_{p_0}(t, x^{p_0}(t))$ *is integrable in* $[a, b]$ *for every* $\|p - p_0\| \leqslant \bar{\varepsilon}$*;*

5. *for every* $\alpha \in C([a, b]; \mathbb{R}^n)$ *the map*

$$p \mapsto \int_{[a,b]} \alpha(t)\, d\left(\int_a^t \tilde{w}_p(s) ds\right)$$

is Fréchet differentiable at p_0*.*

Proposition 7.3.1 *We use here the notations of Definition 7.3.2. Fix* $p_0 \in P$*, let* $v \in P$ *and assume that*

$$\lim_{\varepsilon \to 0^+} \frac{\bar{x}^{p_0+\varepsilon v} - \bar{x}^{p_0}}{\varepsilon}$$

exists in $\mathbb{R}^n$ *and moreover that there exists* $\alpha_{p_0,v}(\cdot) \in \mathfrak{M}_b(a, b; \mathbb{R}^n)$ *such that*

$$\frac{\boldsymbol{f}_{p_0+\varepsilon v}(t, x^{p_0}(t)) - \boldsymbol{f}_{p_0}(t, x^{p_0}(t))}{\varepsilon} \rightharpoonup^* \alpha_{p_0,v}(t) \tag{7.3.23}$$

as $\varepsilon \to 0^+$*. If we denote with* $y_{p_0,v}(t)$ *the derivative of* $x^{p_0}(t)$ *in the* v*-direction evaluated at* p_0*, then we have that*

$$\dot{y}_{p_0,v}(t) = \frac{\partial}{\partial x}\boldsymbol{f}_{p_0}(t, x^{p_0}(t)) \cdot y_{p_0,v}(t) + \alpha_{p_0,v}(t) \tag{7.3.24}$$

where the last equation is to be intended in integral sense. Moreover, if $\boldsymbol{f}$ *is weakly differentiable at* p_0 *along* $x^{p_0}(\cdot)$*, then* $x^p(\cdot) \to x^{p_0}(\cdot)$ *uniformly on* $[a, b]$ *when* $p \to p_0$ *in* P*.*

For a proof, see [96, Appendix A].

Remark 23 *The previous proposition gives us a tool to evaluate the evolution of the derivative of a trajectory of an ODE with respect to a parameter. In particular, it is useful in order to understand the behavior of modifications due to "variations" on a supposed optimal trajectory. We apply the last proposition in the case where the time–varying vector fields are the coupled dynamic–lagrangian functions evaluated on controls produced by a variation. Notice that the limit in (7.3.23) is in the weak star topology of Radon measures.*

Definition 7.3.3 *Let X, Y be finite dimensional vector spaces on $\mathbb{R}$ and Λ a cone in X. Consider a function $f : x + \Lambda \to Y$ for some $x \in X$. We say that f is differentiable at x in the direction Λ if there exists a linear map $D_\Lambda f(x) : X \to Y$ such that*

$$f(x+\lambda) = f(x) + D_\Lambda f(x) \cdot \lambda + o(\|\lambda\|) \quad \text{as } \lambda \to 0,\ \lambda \in \Lambda. \tag{7.3.25}$$

Obviously the map $D_\Lambda f(x)$ is uniquely determined on $span\{\Lambda\}$.

Let $\mathbf{X}$ be an optimal trajectory for the problem $\mathcal{P}$ and let $\bar{\varepsilon} > 0$. We denote with K a cone in $\mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_\nu}$, with $v = (v_1, \ldots, v_\nu)$ an element of K and with $(u_1, \ldots, u_\nu)$ the controls of the candidate optimal hybrid trajectory $\mathbf{X}$. The next definition needs many technicalities and hence it could be difficult to understand well. The aim is to give a rigorous description of all variations we are able to consider. In analogy with [110], we treat variations depending by two parameters: ε and v. ε is a real positive number, while v belongs to a cone in a finite dimensional manifold. The reader can think v as the parameter responsible for the variation of the initial points of each trajectory x_j, $j \in \{1, \ldots, \nu\}$, and ε as the parameter influencing the control of the trajectory. Unfortunately the roles of ε and v are more complicated than these ones, but this is the basic idea.

Definition 7.3.4 ***(Map of variations)****. A map V defined on $[0, \bar{\varepsilon}] \times K$, $V(\varepsilon, v) = (x_1^{(\varepsilon,v)}, u_1^{(\varepsilon,v)}, \ldots, x_\nu^{(\varepsilon,v)}, u_\nu^{(\varepsilon,v)})$, is called a map of variations if, for every $(\varepsilon, v) \in [0, \bar{\varepsilon}] \times K$, the following conditions hold:*

1. *for every $i \in \{1, \ldots, \nu\}$, $u_i^{(\varepsilon,v)} \in \mathcal{U}_{q_i}$ and $u_i^{(\delta\varepsilon,\delta v)} \to u_i$ in $L^1(t_{i-1}, t_i)$ as $\delta \to 0^+$;*

2. *for every $i \in \{1, \ldots, \nu\}$, $x_i^{(\varepsilon,v)} :]t_{i-1}, t_i[\to \mathbb{R}^{d_i}$ is an absolutely continuous function continuously prolongable to $[t_{i-1}, t_i]$ such that*

$$\frac{d}{d\delta} x_i^{(\delta\varepsilon,\delta v)}(t_{i-1})|_{\delta=0} = v_i$$

and

$$\frac{d}{dt} x_i^{(\varepsilon,v)}(t) = f_{q_i}(x_i^{(\varepsilon,v)}(t), u_i^{(\varepsilon,v)}(t)) \text{ for a.e. } t \in [t_{i-1}, t_i]; \tag{7.3.26}$$

3. *for every $i \in \{1, \ldots, \nu - 1\}$, $u_{i+1}^{(\varepsilon,v)} \in \mathcal{U}_{q_i, x_i^{(\varepsilon,v)}(t_i), q_{i+1}, x_{i+1}^{(\varepsilon,v)}(t_i)}$;*

4. *the map*

$$\widetilde{C}_V : [0,\bar{\varepsilon}] \times K \quad \to \quad (\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \times \cdots \times (\mathbb{R}^{d_\nu} \times \mathbb{R}^{d_1}) \times \mathbb{R} \qquad (7.3.27)$$

defined by

$$\widetilde{C}_V(\varepsilon, v) = ((x_1^{(\varepsilon,v)}(t_1), x_2^{(\varepsilon,v)}(t_1)),..,(x_\nu^{(\varepsilon,v)}(t_\nu), x_1^{(\varepsilon,v)}(t_0)), \gamma(\varepsilon, v)),$$

where

$$\gamma(\varepsilon, v) = \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} L_{q_i}(x_i^{(\varepsilon,v)}(t), u_i^{(\varepsilon,v)}(t))dt,$$

is differentiable at 0 *in the direction* $\mathbb{R}^+ \times K$*;*

5. *for every* $i \in \{1, \ldots, \nu\}$*, there exist two Radon measures* $\alpha_{i,f,V}^{(\varepsilon,v)} \in \mathfrak{M}_b(t_{i-1}, t_i; \mathbb{R}^{d_i})$ *and* $\alpha_{i,L,V}^{(\varepsilon,v)} \in \mathfrak{M}_b(t_{i-1}, t_i; \mathbb{R})$ *such that*

$$\frac{f_{q_i}(x_i(t), u_i^{(\delta\varepsilon,\delta v)}(t)) - f_{q_i}(x_i(t), u_i(t))}{\delta} \rightharpoonup^* \alpha_{i,f,V}^{(\varepsilon,v)}(t) \qquad (7.3.28)$$

and

$$\frac{L_{q_i}(x_i(t), u_i^{(\delta\varepsilon,\delta v)}(t)) - L_{q_i}(x_i(t), u_i(t))}{\delta} \rightharpoonup^* \alpha_{i,L,V}^{(\varepsilon,v)}(t) \qquad (7.3.29)$$

as $\delta \downarrow 0$*.*

We denote by $\mathcal{V}$ *the set of all maps of variations.*

Remark 24 *The idea of the proof in order to obtain necessary conditions is classical. Indeed all the admissible variations for an optimal trajectory produce a finite–dimensional cone, which is separated by the cone produced by the profitable directions. Therefore it is possible to find a covector, which separates these two cones. From these considerations we deduce the necessary conditions.*

In Definition 7.2.1 (map of variations), we require various assumptions. In particular, the assumptions 1*,* 2 *and* 3 *are necessary in order that the map of variations produces admissible trajectories for our hybrid system. Moreover assumption* 4 *implies the existence of the cone generated by the variations and, finally, assumption* 5 *is necessary in order to apply Proposition 7.3.1.*

Remark 25 *Notice that, in order to have the differentiability of the function* $\widetilde{C}_V$*, the Radon measures* $\alpha^{(\varepsilon,v)}_{i,f,V}$ *and* $\alpha^{(\varepsilon,v)}_{i,L,V}$ *must depend continuously on the parameters* (ε, v)*. This is guaranteed if* $\alpha^{(\varepsilon,v)}_{i,f,V}$ *and* $\alpha^{(\varepsilon,v)}_{i,L,V}$ *are linear with respect to the parameters* (ε, v) *and if* $(\varepsilon_n, v_n) \in \mathbb{R}^+ \times K$ *with* $(\varepsilon_n, v_n) \to (\varepsilon, v) \in \mathbb{R}^+ \times K$*, then* $\alpha^{(\varepsilon_n,v_n)}_{i,f,V} \rightharpoonup^* \alpha^{(\varepsilon,v)}_{i,f,V}$ *in* $\mathfrak{M}_b(t_{i-1}, t_i; \mathbb{R}^{d_i})$ *for every* $i = 1, \dots, \nu$ *and* $\alpha^{(\varepsilon_n,v_n)}_{i,L,V} \rightharpoonup^* \alpha^{(\varepsilon,v)}_{i,L,V}$ *in* $\mathfrak{M}_b(t_{i-1}, t_i; \mathbb{R})$ *for every* $i = 1, \dots, \nu$*.*

Now, if $V \in \mathcal{V}$, then we have that

$$D\widetilde{C}_V(0) : \mathbb{R}^+ \times K \to (\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \times \cdots \times (\mathbb{R}^{d_\nu} \times \mathbb{R}^{d_1}) \times \mathbb{R}$$

and, if $\varepsilon \in [0, \bar{\varepsilon}]$, then

$$D\widetilde{C}_V(0)(\varepsilon, v) = ((w_1, v_2), \dots, (w_\nu, v_1), \beta(\varepsilon, v)) \tag{7.3.30}$$

where

$$\begin{aligned}
\beta(\varepsilon, v) &= \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} \frac{\partial}{\partial x} L_{q_i}(x_i(s), u_i(s)) M_i(s, t_{i-1}) v_i ds + \\
&\quad + \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} \frac{\partial}{\partial x} L_{q_i}(x_i(s), u_i(s)) \int_{t_{i-1}}^{s} M_i(s, r)\, d\alpha^{(\varepsilon,v)}_{i,f,V}(r)\, ds \\
&\quad + \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} d\alpha^{(\varepsilon,v)}_{i,L,V}(s), \qquad (7.3.31) \\
w_i &= M_i(t_i, t_{i-1}) v_i + \int_{t_{i-1}}^{t_i} M_i(t_i, s) d\alpha^{(\varepsilon,v)}_{i,f,V}(s) \qquad (7.3.32)
\end{aligned}$$

and $M_i(t, s)$ $(i = 1, \dots, \nu)$ is the fundamental matrix solution for the linear system

$$\dot{y}(t) = \frac{\partial}{\partial x} f_{q_i}(x_i(t), u_i(t)) y(t).$$

Indeed, the differential of the components $x_i^{(\varepsilon,v)}(t_{i-1})$ is equal to v_i by hypothesis 2 of the definition of Map of Variations. For the differential of the components $x_i^{(\varepsilon,v)}(t_i)$ we use Proposition 7.3.1, while for the differential of $\gamma(\varepsilon, v)$ we have to estimate, for every $i = 1, \dots, \nu$, the limit as $\delta \to 0^+$ of the expression

$$\int_{t_{i-1}}^{t_i} \frac{L_{q_i}(x_i^{(\delta\varepsilon,\delta v)}(t), u_i^{(\delta\varepsilon,\delta v)}(t)) - L_{q_i}(x_i(t), u_i(t))}{\delta} dt =$$
$$\int_{t_{i-1}}^{t_i} \frac{L_{q_i}(x_i^{(\delta\varepsilon,\delta v)}(t), u_i^{(\delta\varepsilon,\delta v)}(t)) - L_{q_i}(x_i(t), u_i^{(\delta\varepsilon,\delta v)}(t))}{\delta} dt +$$
$$+ \int_{t_{i-1}}^{t_i} \frac{L_{q_i}(x_i(t), u_i^{(\delta\varepsilon,\delta v)}(t)) - L_{q_i}(x_i(t), u_i(t))}{\delta} dt.$$

For the last addendum we use hypothesis 5 of the definition of Map of Variations, while for the other term we have to use Proposition 7.3.1 and Lemma 7.4.2 of Section 7.4.

Let us denote by $\mathbf{X}^{(\varepsilon,v)}(\cdot)$ the candidate hybrid trajectory obtained piecing together $x_i^{(\varepsilon,v)}$, $i = 1, \ldots, \nu$. Then $\mathbf{X}^{(\varepsilon,v)}(\cdot)$ is a trajectory if and only if

- $(x_i^{(\varepsilon,v)}(t_i), x_{i+1}^{(\varepsilon,v)}(t_i)) \in \mathcal{S}_{q_i,q_{i+1}}$ for $i = 1, \ldots, \nu - 1$;
- $(q_1, x_1^{(\varepsilon,v)}(t_0), 0) \in \mathcal{N}_{in}$;
- $(q_\nu, x_\nu^{(\varepsilon,v)}(t_\nu), t_\nu - t_{\nu-1}) \in \mathcal{N}_{fin}$.

Since $\mathbf{X}$ is optimal we have that

$$C(\mathbf{X}^{(\varepsilon,v)}) \geqslant C(\mathbf{X}) \tag{7.3.33}$$

whenever the previous conditions hold.

In the sequel we identify, for simplicity, $x_{\nu+1}(t_\nu)$ with $x_1(t_0)$ and $\mathbb{R}^{d_{\nu+1}}$ with $\mathbb{R}^{d_1}$. Moreover

$$\mathcal{S}_{q_\nu,q_{\nu+1}} := \{(z, z') \in \mathbb{R}^{d_\nu} \times \mathbb{R}^{d_1} : (q_1, z', 0) \in \mathcal{N}_{in}, (q_\nu, z, t_\nu - t_{\nu-1}) \in \mathcal{N}_{fin}\}.$$

Now, fix smooth functions $\sigma_i : \mathbb{R}^{d_i} \times \mathbb{R}^{d_{i+1}} \to \mathbb{R}$, $(i = 1, \ldots, \nu)$ such that

$$\sigma_i(x_i(t_i), x_{i+1}(t_i)) = 0$$

and $\sigma_i(z_i, z_i') > 0$ if $(z_i, z_i') \in \mathbb{R}^{d_i} \times \mathbb{R}^{d_{i+1}} \setminus \{(x_i(t_i), x_{i+1}(t_i))\}$. Let P be the set of points $((z_1, z_1'), \ldots, (z_\nu, z_\nu'), r)$ of $(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \times \cdots \times (\mathbb{R}^{d_\nu} \times \mathbb{R}^{d_1}) \times \mathbb{R}$ such that $(z_i, z_i') \in \mathcal{S}_{q_i,q_{i+1}}$ $(i = 1, \ldots, \nu)$ and

$$r \leqslant C(\mathbf{X}) - \sum_{i=1}^{\nu-1} \Phi_{q_i,q_{i+1}}(z_i, z_i') - \widetilde{\varphi}_{q_1,q_\nu}(z_\nu, z_\nu') - \sum_{i=1}^{\nu} \sigma_i(z_i, z_i'),$$

where $\widetilde{\varphi}_{q_1,q_\nu}(z_\nu, z'_\nu) = \varphi_{q_1,q_\nu}(z'_\nu, z_\nu)$. Notice that if $\widetilde{C}_V(\varepsilon, v) \in P$ then $\mathbf{X}^{(\varepsilon,v)}$ is a hybrid trajectory and $C(\mathbf{X}^{(\varepsilon,v)}) \leqslant C(\mathbf{X})$ with strict inequality if $\mathbf{X}^{(\varepsilon,v)} \not\equiv \mathbf{X}$, then from (7.3.33), we get $\widetilde{C}_V([0,\bar{\varepsilon}] \times K) \cap P = \{p_*\}$ where

$$p_* = ((x_1(t_1), x_2(t_1)), \ldots, (x_\nu(t_\nu), x_1(t_0)), C_L(\mathbf{X}))\,.$$

Define

$$K_V = D\widetilde{C}_V(0)([0,\bar{\varepsilon}] \times K). \tag{7.3.34}$$

Let K_P be the set of all $((z_1, z'_1), \ldots, (z_\nu, z'_\nu), r)$ such that (z_i, z'_i), for $i = 1, \ldots, \nu$, belongs to a Boltyanskii approximating cone to $\mathcal{S}_{q_i,q_{i+1}}$ at the point $(x_i(t_i), x_{i+1}(t_i))$ and

$$r \leqslant -\sum_{i=1}^{\nu-1} \nabla\Phi_{q_i,q_{i+1}}(x_i(t_i), x_{i+1}(t_i)) \cdot (z_i, z'_i) - \nabla\widetilde{\varphi}_{q_1,q_\nu}(x_\nu(t_\nu), x_1(t_0)) \cdot (z_\nu, z'_\nu).$$

Then K_P is a Boltyanskii approximating cone to P at p_* and is not a linear subspace. By a general separation theorem (see [108]), for every convex cone $\widehat{K} \subseteq \bigcup_{V\in\mathcal{V}} K_V$, there exists an element $\psi \in (\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \times \cdots \times (\mathbb{R}^{d_\nu} \times \mathbb{R}^{d_1}) \times \mathbb{R}$ such that it weakly separates K_P from $\widehat{K}$. In particular we may suppose that

$$(\psi, k) \geqslant 0 \quad \forall k \in K_P$$

and

$$(\psi, k') \leqslant 0 \quad \forall k' \in \widehat{K},$$

where $(\cdot,\cdot)$ denotes the usual scalar product in $\mathbb{R}^n$. We may write

$$\psi = ((\psi_1^+, \psi_2^-), \ldots, (\psi_\nu^+, \psi_1^-), -\psi_0).$$

We note that $(0, \ldots, 0, -1) \in K_P$ and so $\psi_0 \geqslant 0$.

If $V \in \mathcal{V}$ and $(\varepsilon, v) \in [0,\bar{\varepsilon}] \times K$ is such that $D\widetilde{C}_V(0)(\varepsilon, v) \in \widehat{K}$, then

$$\sum_{i=1}^{\nu} \left(\psi_i^- \cdot v_i + \psi_i^+ \cdot M_i(t_i, t_{i-1}) \cdot v_i + \psi_i^+ \int_{t_{i-1}}^{t_i} M_i(t_i, s) d\alpha_{i,f,V}^{(\varepsilon,v)}(s) \right)$$
$$-\psi_0 \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} \frac{\partial}{\partial x} L_{q_i}(x_i(s), u_i(s)) M_i(s, t_{i-1}) v_i ds$$
$$-\psi_0 \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} \frac{\partial}{\partial x} L_{q_i}(x_i(s), u_i(s)) \int_{t_{i-1}}^{s} M_i(s, r) d\alpha_{i,f,V}^{(\varepsilon,v)}(r) ds$$
$$-\psi_0 \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} d\alpha_{i,L,V}^{(\varepsilon,v)}(s) \leqslant 0.$$

Hence, if we define, for every $i = 1, \dots, \nu$, $\psi_i : [t_{i-1}, t_i] \to \mathbb{R}^{d_i}$ to be the Caratheodory solution to

$$\begin{cases} \dot{\psi}_i(t) = -\psi_i(t)\frac{\partial}{\partial x} f_{q_i}(x_i(t), u_i(t)) + \psi_0 \frac{\partial}{\partial x} L_{q_i}(x_i(t), u_i(t)) \\ \psi_i(t_i) = \psi_i^+ \end{cases}$$

then we obtain that

$$\begin{aligned} \sum_{i=1}^{\nu} \left(\psi_i^- \cdot v_i + \psi_i(t_i) \int_{t_{i-1}}^{t_i} M_i(t_i, s) d\alpha_{i,f,V}^{(\varepsilon,v)}(s) \right) \\ + \sum_{i=1}^{\nu} \psi_i(t_{i-1}) v_i - \psi_0 \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} d\alpha_{i,L,V}^{(\varepsilon,v)}(s) \\ -\psi_0 \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} \frac{\partial}{\partial x} L_{q_i}(x_i(t), u_i(t)) \int_{t_{i-1}}^{s} M_i(s,r) d\alpha_{i,f,V}^{(\varepsilon,v)}(r) ds \leqslant 0. \end{aligned} \tag{7.3.35}$$

Fix $i \in \{1, \dots, \nu - 1\}$ and a point $(z, z') \in K_i$, where K_i is a Boltyanskii approximating cone to $\mathcal{S}_{q_i, q_{i+1}}$ at the point $(x_i(t_i), x_{i+1}(t_i))$. Let us consider

$$\mathbf{z} = ((0,0), \dots, (z, z'), \dots, (0,0), r)$$

with $r = -\nabla\Phi_{q_i, q_{i+1}}(x_i(t_i), x_{i+1}(t_i)) \cdot (z, z')$. Obviously $\mathbf{z} \in K_P$ and so

$$\psi_i^+ \cdot z + \psi_{i+1}^- \cdot z' - \psi_0 r \geqslant 0$$

that is

$$((\psi_i(t_i), \psi_{i+1}^-) + \psi_0 \nabla\Phi_{q_i, q_{i+1}}(x_i(t_i), x_{i+1}(t_i))) \cdot (z, z') \geqslant 0$$

and so

$$((-\psi_i(t_i), -\psi_{i+1}^-) - \psi_0 \nabla\Phi_{q_i, q_{i+1}}(x_i(t_i), x_{i+1}(t_i))) \in K_i^{\perp}, \tag{7.3.36}$$

where $K_i^{\perp}$ is the polar of the cone K_i.

Now take a point $(z, z') \in K_\nu$, where K_ν is a Boltyanskii approximating cone to $\mathcal{S}_{q_\nu, q_{\nu+1}}$ at the point $(x_\nu(t_\nu), x_1(t_0))$. Let us consider

$$\mathbf{z} = ((0,0), \dots, (0,0), (z, z'), r)$$

with $r = -\nabla\widetilde{\varphi}_{q_1, q_\nu}(x_\nu(t_\nu), x_1(t_0)) \cdot (z, z')$. Analogously we obtain that

$$((-\psi_\nu(t_\nu), -\psi_1^-) - \psi_0 \nabla\widetilde{\varphi}_{q_\nu, q_1}(x_\nu(t_\nu), x_1(t_0))) \in K_\nu^{\perp}, \tag{7.3.37}$$

where $K_\nu^{\perp}$ is the polar of the cone K_ν.

So we have just proved the following theorem:

Theorem 7.3.1 ***(Hybrid Necessary Principle)***. *Let $\boldsymbol{X}$ be an optimal trajectory for problem $\mathcal{P}$. For every convex cone $\widehat{K}$ contained in $\cup_{V\in\mathcal{V}}K_V$, where K_V is defined in (7.3.34), there exist an adjoint pair (ψ,ψ_0) along $\boldsymbol{X}$ and $(\psi_1^-,\ldots,\psi_\nu^-)\in\mathbb{R}^{d_1}\times\ldots\times\mathbb{R}^{d_\nu}$ such that (7.3.35) holds for every $V\in\mathcal{V}$, $(\varepsilon,v)\in[0,\bar{\varepsilon}]\times K$ such that $D\widetilde{C}_V(0)(\varepsilon,v)\in\widehat{K}$. Moreover the equations (7.3.36) and (7.3.37) hold.*

Remark 26 *Notice that in the previous theorem we have implicitly supposed that the times of switchings are fixed. Obviously it is possible to consider variations of these times as in [110], using a more complicate covector. We obtain analogous necessary conditions that are more complicated and less readable.*

Remark 27 *If assumption (H) holds and if we can take $(0,\ldots,v_i,\ldots,0)\in K$, then from the previous theorem we can obtain the same result of [110, Theorem 1.4.1].*

7.4 A lemma on integrable functions

Lemma 7.4.1 *Let I be a compact interval of $\mathbb{R}$, f_n,f functions of $L^\infty(I)$ such that $\|f_n\|_\infty\leqslant c<+\infty$ and $f_n\to f$ strongly in $L^1(I)$. Moreover let g_n be a sequence s.t. $g_n\rightharpoonup g$ weakly in $L^1(I)$. Then $f_ng_n\rightharpoonup fg$ weakly in $L^1(I)$.*

Proof. Let $\varphi\in L^\infty(I)$. We have to prove that

$$\int_I\varphi(s)(f_n(s)g_n(s)-f(s)g(s))ds\to 0$$

as $n\to+\infty$. We have that

$$\int_I\varphi(s)(f_n(s)g_n(s)-f(s)g(s))ds=\int_I\varphi(s)(f_n(s)-f(s))g_n(s)ds+$$
$$+\int_I\varphi(s)f(s)(g_n(s)-g(s))ds.$$

The last integral goes to 0 since $g_n\rightharpoonup g$. Fix $\varepsilon>0$. Then we can write

$$\int_I\varphi(s)(f_n(s)-f(s))g_n(s)ds=$$
$$\int_{|f_n-f|\geqslant\varepsilon}\varphi(s)(f_n(s)-f(s))g_n(s)ds+\int_{|f_n-f|<\varepsilon}\varphi(s)(f_n(s)-f(s))g_n(s)ds.$$

Moreover

$$\left|\int_{|f_n-f|<\varepsilon}\varphi(s)(f_n(s)-f(s))g_n(s)ds\right| \leqslant \varepsilon\|\varphi\|_\infty \int_I |g_n(s)|\,ds \leqslant M\varepsilon$$

where M is a positive constant. Besides:

$$\left|\int_{|f_n-f|\geqslant\varepsilon}\varphi(s)(f_n(s)-f(s))g_n(s)ds\right| \leqslant \|\varphi\|_\infty M_1 \int_{|f_n-f|\geqslant\varepsilon} |g_n(s)|\,ds$$

with M_1 positive constant. Since $f_n \to f$ strongly in $L^1(I)$, f_n converges to f in measure. Moreover g_n is equiintegrable by Dunford Pettis theorem (see [30, Théorème IV.29]), hence we can find $\bar{n} \in \mathbb{N}$ such that

$$\left|\int_I \varphi(s)(f_n(s)-f(s))g_n(s)ds\right| \leqslant (M + M_1\|\varphi\|_\infty)\varepsilon$$

for every $n \geqslant \bar{n}$ and we conclude by the arbitrariness of ε. □

With an analogous proof, that we omit, we can generalize the previous lemma to the case of Radon measures in the following way:

Lemma 7.4.2 *Let I be a compact interval of $\mathbb{R}$, f_n, f functions of $C(I)$ such that $f_n \to f$ uniformly on I as $n \to +\infty$. Moreover let g_n be a sequence in $C(I)$ such that $g_n \rightharpoonup^* g$ in $\mathfrak{M}_b(I)$, where $g \in \mathfrak{M}_b(I)$. Then $f_n g_n \rightharpoonup^* fg$ in $\mathfrak{M}_b(I)$.*

Chapter 8

Application to a car with gears.

We want now to present some examples of applicability of Hybrid Maximum Principle by Sussmann (see [110]) and of Hybrid Necessary Principle (see [59]). In particular, we introduce a simple model of a car with gears and at this model we apply HMP if we assume that assumption (H) holds, and we apply HNP in the other case.

8.1 A car with gears and brakes.

We model a car with gears and brakes introducing first a discrete variable S that takes values in the set $\{A, B\}$. This corresponds to the fact that the action of the driver is to accelerate or to use brakes. Then the taken action has different effect depending on another discrete variable G that takes values in a finite set $\{1, 2, ..., n\}$ and corresponds to the used gear. We thus get a discrete variable $q = (S, G) \in \mathcal{Q}$, where the set $\mathcal{Q}$ is composed by $2n$ elements. Regarding the changes of locations, there are essentially two reasonable choices:

1. it is possible to jump from a location to another arbitrary one;
2. from a location (S, G) ($S \in \{A, B\}$, $G \in \{1, \ldots, n\}$), it is possible to jump to one of the following locations:
   (a) (S', G) with $S' \in \{A, B\}$;
   (b) $(S', G+1)$ with $S' \in \{A, B\}$ if $G < n$;
   (c) $(S', G-1)$ with $S' \in \{A, B\}$ if $G > 1$.

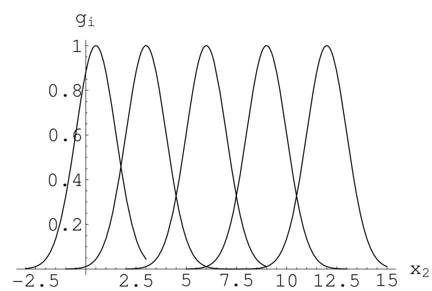


Figure 8.1: Graphs of the gear functions g_i.

i.e. we let the pilot accelerate or brake changing gears only by one.

The first choice can be convenient to consider the problem in most generality, however it is clear that the second one is not very restrictive on the behaviour of the pilot. Therefore from now on we assume only the second possibilities.

We now describe the continuous controlled dynamics for each location $q \in \mathcal{Q}$. For every $q \in \mathcal{Q}$, we describe the car by its position x_1 and its speed x_2. Therefore each manifold M_q is a subset of $\mathbb{R}^2$. If $S = A$ and $G = i$ then the control system evolves according to

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u\, g_i(x_2) \end{cases}$$

where the control u takes values in the set $U_q = [0, 1]$. The function g_i describes the value of the acceleration as a function of the velocity when the pilot uses the maximum control. We choose the function g_i to behave as Gaussians with centre that depends on the gear i, see Figure 8.1.

When $S = B$ and $G = i$ the control system evolves according to

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u\, h_i(x_2) \end{cases}$$

where the control u takes value in the set $U_q = [-1, 0]$. The function h_i describes the value of the deceleration as a function of the velocity when the pilot uses the maximum braking. Thus the control in these locations takes only negative values. Taking into account the braking effect of the engine we choose the function h_i to be increasing and only partially defined, that is on a bounded interval $[0, a_i]$. (Notice that for simplicity we restrict to the

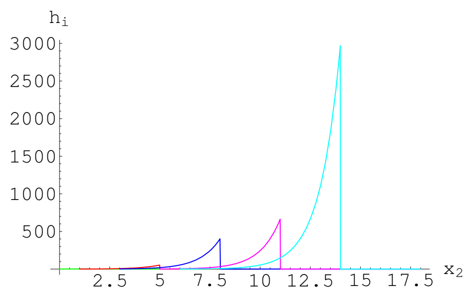


Figure 8.2: Graphs of the braking functions h_i.

set where $x_2 \geq 0$.) Moreover, it seems natural to assume $a_i < a_j$, $h_j < h_i$ when $i < j$ and both functions are defined. See Figure 8.2 representing the braking functions h_i.

Therefore, according to Definition 7.1.1, we have:

$$\mathcal{Q} = \{A, B\} \times \{1, 2, ..., n\}$$

and

$$f_{(A,i)} \begin{pmatrix} x_1 \\ x_2 \\ u \end{pmatrix} = \begin{pmatrix} x_2 \\ u\, g_i(x_2) \end{pmatrix}, \quad f_{(B,i)} \begin{pmatrix} x_1 \\ x_2 \\ u \end{pmatrix} = \begin{pmatrix} x_2 \\ u\, h_i(x_2) \end{pmatrix}$$

We have, $M_q = \mathbb{R}^2$ if the first component of q is A and $M_q = \mathbb{R} \times [0, a_i]$ otherwise, $U_q = [0, 1]$ if the first component of q is A and $U_q = [-1, 0]$ otherwise. All measurable functions with values in U_q are admissible controls, $J_q =]0, +\infty[$ (that is we can use a gear for any time length). Finally, $\mathcal{S}$ contains some 6-tuples of the type

$$((S, G), x, (S', G'), x, u(\cdot), \tau)$$

such that if $S' = B$ and $x = (x_1, x_2)$ then x_2 is in the domain of definition of $h_{G'}$. This correspond to the fact that the position and velocity of the car do not jump, but a location switching to a braking situation can be done only if the velocity is not too big.

8.2 HMP for a car with gears and brakes.

We treat, first, the case in which assumption (H) holds. Therefore, given the sets

$$\mathcal{S}_1 := \{((A,i),x,(B,i),x,u(\cdot),\tau) \; : \; i \in \{1,\dots,n\}, x \in \mathbb{R}\times[0,a_i], \; u(\cdot) \in \mathcal{U}_{(B,i)}, \tau > 0\},$$

$$\mathcal{S}_2 := \{((B,i),x,(A,i),x,u(\cdot),\tau) \; : \; i \in \{1,\dots,n\}, x \in \mathbb{R}\times[0,a_i], \; u(\cdot) \in \mathcal{U}_{(A,i)}, \tau > 0\},$$

$$\mathcal{S}_3 := \{((S,i),x,(S',i+1),x,u(\cdot),\tau) : i \in \{1,\dots,n-1\}, S,S' \in \{A,B\}, \; x \in M_{(S,i)} \cap M_{(S',i+1)}, u(\cdot) \in \mathcal{U}_{(S',i+1)}, \tau > 0\},$$

$$\mathcal{S}_4 := \{((S,i),x,(S',i-1),x,u(\cdot),\tau) : i \in \{2,\dots,n\}, S,S' \in \{A,B\}, \; x \in M_{(S,i)} \cap M_{(S',i-1)}, u(\cdot) \in \mathcal{U}_{(S',i-1)}, \tau > 0\},$$

the switching set is equal to

$$\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4.$$

In next subsections, we apply HMP to this hybrid system for some different problems.

8.2.1 Minimum time problems for fixed final position.

As first example, we consider a minimum time problem to reach a fixed final position $x_1 = M > 0$ (no condition on final velocity) from the origin, that is for initial position $x_1 = 0$ and initial velocity $x_2 = 0$. This amounts to consider the Lagrangian L such that $L_q \equiv 1$ for any $q \in \mathcal{Q}$, the switching cost function and the endpoint cost function constantly equal to zero.

Proposition 8.2.1 *The optimal trajectory $\boldsymbol{X}$ for the problem $\mathcal{P}$ has the following properties:*

- *every location is of the type* (A, i) *with* $i \in \{1, \dots, n\}$*;*
- *the optimal switching strategy is* $(A, 1)$, $(A, 2)$, $\dots$, (A, ν), *where* ν *is in* $\{1, \dots, n\}$*;*
- *the control in each location is always equal to* 1*;*
- *if* t_i *is the switching time between* (A, i) *and* $(A, i+1)$ *(*$i \in \{1, \dots, \nu - 1\}$*), then we have that*

$$g_i(x_i^{(2)}(t_i)) = g_{i+1}(x_{i+1}^{(2)}(t_i)).$$

where $x_l^{(j)}$ *stands for the* j*-th component of* x_l*.*

Proof. Let us consider the last location q_ν for the optimal trajectory $\mathbf{X}$. Obviously either $q_\nu = (A, i)$ or $q_\nu = (B, i)$ for some $i \in \{1, \dots, n\}$. We suppose by contradiction that $q_\nu = (B, i)$ for some $i \in \{1, \dots, n\}$. Therefore for every $t \in [t_{\nu-1}, t_\nu[$, $x_\nu(\cdot)$ satisfies:

$$\dot{x}_\nu(t) = f_{(B,i)} \begin{pmatrix} x_\nu(t) \\ u(t) \end{pmatrix}.$$

For every $t \in [t_{\nu-1}, t_\nu[$, we have that

$$\begin{cases} \dot{x}_\nu^{(1)}(t) = x_\nu^{(2)}(t) \\ \dot{x}_\nu^{(2)}(t) = u(t) h_i(x_\nu^{(2)}(t)) \end{cases}$$

and $x_\nu^{(2)} > 0$. Clearly we assume $x_\nu^{(1)}(t_\nu) = M$ and $x_\nu^{(1)}(t) < M$ for every $t \in [t_{\nu-1}, t_\nu[$. Now, we define $\tilde{x}(t) = (\tilde{x}_1(t), \tilde{x}_2(t))$ $(t \geqslant t_{\nu-1})$ as the solution to

$$\begin{cases} \frac{d}{dt}\tilde{x}_1(t) = \tilde{x}_2(t) \\ \frac{d}{dt}\tilde{x}_2(t) = g_i(\tilde{x}_2(t)) \\ \tilde{x}_1(t_{\nu-1}) = x_\nu^{(1)}(t_{\nu-1}) \\ \tilde{x}_2(t_{\nu-1}) = x_\nu^{(2)}(t_{\nu-1}). \end{cases}$$

We have that

$$\dot{x}_\nu^{(2)}(t) \leqslant 0 < \frac{d}{dt}\tilde{x}_2(t), \quad x_\nu^{(2)}(t_{\nu-1}) = \tilde{x}_2(t_{\nu-1})$$

which implies

$$x_\nu^{(2)}(t) < \tilde{x}_2(t)$$

for every $t \in]t_{\nu-1}, t_\nu]$. Thus

$$\dot{x}_\nu^{(1)}(t) < \frac{d}{dt}\tilde{x}_1(t)$$

for $t \in]t_{\nu-1}, t_\nu]$, which implies that

$$x_\nu^{(1)}(t) < \tilde{x}_1(t)$$

for every $t \in]t_{\nu-1}, t_\nu]$. Let $\bar{t} \in]t_{\nu-1}, t_\nu[$ be such that $\tilde{x}_1(\bar{t}) = M$. Therefore, if we replace the optimal trajectory $\mathbf{X}$ with $\tilde{\mathbf{X}}$ obtained substituting $x_\nu(\cdot)$ with $\tilde{x}(\cdot)$, then the cost of $\tilde{\mathbf{X}}$ is strictly less than the cost of $\mathbf{X}$, a contradiction. Thus q_ν must be equal to (A, i) for some $i \in \{1, \ldots, \nu\}$. With the same argument, we conclude that, for every $l \in \{1, \ldots, \nu\}$, $q_l = (A, m)$ for some $m = 1, \ldots, n$.

We consider an adjoint pair (ψ, ψ_0) along $\mathbf{X}$. We indicate by $\psi_\nu^{(1)}$ and $\psi_\nu^{(2)}$ the two scalar components of ψ_ν. Therefore we have that

$$\begin{cases} \dot{\psi}_\nu^{(1)} = 0, \\ \dot{\psi}_\nu^{(2)} = -\psi_\nu^{(1)} - \psi_\nu^{(2)} u g_i'(x_2). \end{cases}$$

Thus $\psi_\nu^{(1)}$ is constant. Notice that the first component of every ψ_l ($l = 1, \ldots, \nu$) is constant in each location. The transversality condition implies that $\psi_\nu^{(2)}(t_\nu) = 0$, while, since the trajectory does not jump at every switching times, the switching condition gives the continuity of the covector ψ at switching times. In fact the tangent cone to $\mathcal{S}_{(A,i-1),(A,i)}$ or to $\mathcal{S}_{(A,i+1),(A,i)}$ is equal to $\{(v, w, v, w) : v, w \in \mathbb{R}\}$. This clearly implies that $\psi_{\nu-1}^{(1)}(t_{\nu-1}) = \psi_\nu^{(1)}(t_{\nu-1})$ and $\psi_{\nu-1}^{(2)}(t_{\nu-1}) = \psi_\nu^{(2)}(t_{\nu-1})$. The same argument proves the continuity of the covector at the other switching times.

We have three possibilities:

1. $\psi_\nu^{(1)} > 0$. Therefore we obtain that $\psi_\nu^{(2)}$ is strictly positive for every $t_{\nu-1} \leqslant t < t_\nu$. By the Hamiltonian maximization we have that the control in the last location must be equal to 1.

2. $\psi_\nu^{(1)} = 0$. Therefore $\psi_\nu^{(2)} \equiv 0$ and so, by the Hamiltonian value condition, we have that also $\psi_0 = 0$. Clearly $\psi_i^{(2)} \equiv 0$ for every $i = 1, \ldots, \nu$ and this is not possible by nontriviality.

3. $\psi_\nu^{(1)} < 0$. Therefore $\psi_\nu^{(2)}$ is strictly negative. So the Hamiltonian maximization implies that the control u must be constantly equal to 0 in the last location. By the same arguments, it follows that the control must be 0 in all locations, but in this case $x_i^{(1)}(\cdot) \equiv 0$ for every $i \in \{1, \ldots, \nu\}$ and so the trajectory $\mathbf{X}$ is not admissible.

Thus only the first possibility occurs. Therefore $\psi_i^{(1)} > 0$ for every $i \in \{1, \ldots, \nu\}$ and in each location the control must be equal to 1.

Recalling the definition of $\mathcal{S}$, it follows that $q_{\nu-1} = (A, i-1)$ or $q_{\nu-1} = (A, i+1)$. First suppose that $q_{\nu-1} = (A, i-1)$. The Hamiltonian value condition gives

$$\psi_{\nu-1}^{(2)}(t_{\nu-1}) g_{i-1}(x_{\nu-1}^{(2)}(t_{\nu-1})) = \psi_\nu^{(2)}(t_{\nu-1}) g_i(x_\nu^{(2)}(t_{\nu-1})).$$

The continuity of the covector ψ and the fact that $\psi_j^{(2)}(t) > 0$ for every $t \in [t_{j-1}, t_j]$ and for every $j = 1, \ldots, \nu$, imply that

$$g_{i-1}(x_{\nu-1}^{(2)}(t_{\nu-1})) = g_i(x_\nu^{(2)}(t_{\nu-1})).$$

Assume now $q_{\nu-1} = (A, i+1)$. Also in this case the Hamiltonian value condition implies

$$g_{i+1}(x_{\nu-1}^{(2)}(t_{\nu-1})) = g_i(x_\nu^{(2)}(t_{\nu-1})).$$

On a left neighborhood of $t_{\nu-1}$ we get $x_{\nu-1}^{(2)}(t) < x_{\nu-1}^{(2)}(t_{\nu-1})$. This implies that $g_{i+1}(x_{\nu-1}^{(2)}(t)) < g_i(x_{\nu-1}^{(2)}(t))$. Replacing the location $q_{\nu-1} = (A, i+1)$ with (A, i), we get, as before, a trajectory with lower cost, contradicting the optimality of $\mathbf{X}$. So, all the statements of the proposition are proved. □

The result of Proposition 8.2.1 means that the change of gear should happen at a velocity at which the acceleration is the same for the two gears. Similar conditions can be found for other optimization problems with convex Lagrangians.

We then verify these results through computer simulations. We assume $g_i(z) = \exp(-0.5(z - k_i)^2)$ for some $k_1 < k_2 < ... < k_5$ and for every $z \in \mathbb{R}$. We consider a car with three gears with $k_1 = 0.5, k_2 = 3, k_3 = 6$, and final position $M = 20$. Therefore, the first two gear functions coincide at $z = 1.75$, while the second and the third at $z = 4.5$. Computing along the corresponding solutions one gets that the Hybrid Maximum Principle prescribes the gear

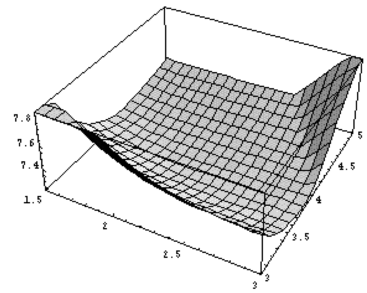


Figure 8.3: Time as function of gear changes: the first axis displays the switching time between gear 1 and gear 2, the second one the switching time between gear 2 and gear 3, while the last one the final time.

changes at times 2.2 and 3.99 respectively. This is confirmed by simulations: Figure 8.3 shows the time in which the car reaches the final position for first gear change between 1.5 and 3 and second gear change between 3 and 5. The minimum is attained at the expected position.

8.2.2 Minimum time problems for fixed final position and velocity.

Now we consider the problem of starting from the origin and reaching a position $(M, 0)$, that is with final velocity equal to zero, in minimum time. In this case the supposed optimal strategy is $(A, 1), ..., (A, n), (B, n), ..., (B, 1)$ (obviously this strategy can be optimal supposing that the distance M to be run is big enough). For the first part of the trajectory we find conditions completely similar to those obtained above. On the contrary on the second part of the trajectory, the braking one, the maximisation condition depends on the set of gears for which it is possible to apply the corresponding braking.

We first need a new definition:

Definition 8.2.1 *A switching from location (S, G) to location (S', G'), happening at time t, is said to be fictitious if the control used in (S, G) is zero on a left neighborhood of t and the control used in (S', G') is zero on a right neighborhood of t.*

It is clear that a location switching is fictitious if the change in location do not affect at all the continuous dynamics. Every optimal trajectory, for which the control is zero on an interval of positive measure, can be modified inserting an arbitrary number of fictitious switchings without changing the trajectory performance. In particular this implies the presence of possible Zeno behaviour. More precisely:

Proposition 8.2.2 *Let $\boldsymbol{X}$ be an optimal trajectory and t a fictitious location switching. Then for every n there exists an optimal trajectory $\boldsymbol{X}_n$, satisfying the same boundary data of $\boldsymbol{X}$ and having n fictitious location switchings.*

Therefore we are interested in determining an optimal trajectory that has no fictitious switching. We get:

Proposition 8.2.3 *Every optimal trajectory $\boldsymbol{X}$ can be replaced by an optimal trajectory satisfying the same boundary conditions, with no fictitious location switching and such that the following holds.*

1. *At a location switching time t between location (A, i) and location $(A, i+1)$ it holds: $g_i(x_i^{(2)}(t)) = g_{i+1}(x_{i+1}^{(2)}(t))$.*

2. *At a location switching time t between (A, n) and (B, n) either $|u| = 1$ on a neighborhood of t and $x_n^{(2)}(t) < a_n$ or $u = 0$ on a left neighborhood of t, $u = -1$ on a right neighborhood of t and $x_n^{(2)}(t) = a_n$.*

3. *At a location switching time t between $(B, i+1)$ and (B, i) it holds $x_{2n-i}^{(2)}(t) = a_i$.*

Proof. It is clear that we can replace **X** by an optimal trajectory satisfying the same boundary conditions and presenting no fictitious location switching.

The first assertion on location switchings is proved in Proposition 8.2.1.

Let us prove the second claim. We denote with $\psi_{(S,n)}^{(i)}$ the i-th ($i = \{1, 2\}$) component of $\psi_{(S,n)}$, where $S = \{A, B\}$. First of all, we notice that the change of gear can happen only at a velocity less than or equal to a_n. The tangent cone to $\mathcal{S}_{(A,n),(B,n)}$ is $\{(v, w, v, w) : v, w \in \mathbb{R}\}$ if the switching velocity is strictly less than a_n and $\{(v, w, v, w) : v, w \in \mathbb{R}, w \leq 0\}$ if the switching velocity is equal to a_n. From the switching condition, we get that the first scalar component of the covector ψ does not jump at time t. If $x_n^{(2)}(t) < a_n$, then also the second scalar component of ψ does not jump at time t, otherwise we have that $\psi^{(2)}_{(A,n)}(t-) \leqslant \psi^{(2)}_{(B,n)}(t+)$. By the Hamiltonian value condition we obtain that

$$\sup_{u\in[0,1]} \left\{ u\psi^{(2)}_{(A,n)}(t-)g_n(x^{(2)}_n(t-)) \right\} = \sup_{u\in[-1,0]} \left\{ u\psi^{(2)}_{(B,n)}(t+)h_n(x^{(2)}_{n+1}(t+)) \right\}.$$

We have some possibilities:

- $\psi^{(2)}_{(A,n)}(t-) = \psi^{(2)}_{(B,n)}(t+) = 0$. In this case we have either that $u = 0$ or that $|u| = 1$ in a neighborhood of t.

- $\psi^{(2)}_{(A,n)}(t-) = \psi^{(2)}_{(B,n)}(t+) \neq 0$. This case is not possible.

- $0 < \psi^{(2)}_{(A,n)}(t-) < \psi^{(2)}_{(B,n)}(t+)$. This implies that $\psi^{(2)}_{(A,n)}(t-) = 0$, which is a contradiction.

- $\psi^{(2)}_{(A,n)}(t-) < \psi^{(2)}_{(B,n)}(t+) < 0$. As before.

- $\psi^{(2)}_{(A,n)}(t-) < 0 < \psi^{(2)}_{(B,n)}(t+)$. In this case $u = 0$ in a neighborhood of t.

- $0 = \psi^{(2)}_{(A,n)}(t-) < \psi^{(2)}_{(B,n)}(t+)$. In this case $u = 0$ on a right neighborhood of t. For a non fictitious switching, we get $u = 1$ on a left neighborhood of t, but then we can replace this trajectory defining $u = 0$ after t and operating no location switching at t.

- $\psi^{(2)}_{(A,n)}(t-) < \psi^{(2)}_{(B,n)}(t+) = 0$. As before.

So we have proved the second claim.

Last claim can be proved in an entirely similar way. □

The conclusion again matches with intuition. Indeed after the acceleration part the driver should turn to the braking maneuver using always the lowest possible gear to decelerate. If the change happens at a velocity less then a_n, then we pass from control $u = 1$ maximum acceleration to control $u = -1$ maximum braking. If the change happens at a velocity a_n, then this means that we stopped acceleration at the maximum gear to not overpass the maximum velocity for braking: a_n. (This last case does not happen in reality but only because we chose never vanishing gaussian for the acceleration.) Notice that we have to use the additional information of zero final velocity and the non–jump condition for covectors to completely determine the optimal trajectory.

We show some simulation results. We suppose that $g_i(z) = \exp(-0.5(z - k_i)^2)$ for some $k_1 < k_2 < ... < k_n$ and $h_i(z) = \exp(z - s_i) - 1$ for some $s_1 < s_2 < ... < s_n$. We consider a car with five gears with $k_1 = 0.5$, $k_2 = 3$, $k_3 = 6$, $k_4 = 9$, $k_5 = 12$, $a_1 = 1.5$, $a_2 = 5$, $a_3 = 8$, $a_4 = 11$, $a_5 = 14$, $s_1 = 0$, $s_2 = 1$, $s_3 = 2$, $s_4 = 4.5$, $s_5 = 6$ and we suppose that it has to reach the position $(200, 0)$.

Now we describe in detail the braking strategy. If the pilot is accelerating using the fifth gear and he did not reach the final position he starts to brake using the function $h_5(z)$ until the velocity is equal to 11. When the last condition is verified he changes the gear from five to four. Likewise when he brakes using the gears 4, 3, 2 he changes gear when the velocity respectively is equal to 8, 5, 1.5. Finally when the pilot brakes using the first gear, he remains in the first gear until the velocity is equal to zero or until he did not reach the final position.

We fix the time of acceleration using the gears 1, 2, 3 to the theoretical values, prescribed by the Hybrid Maximum Principle, and the deceleration as described above. To approximate the minimum time we have used the following cost function

$$time + 2|x_{fin} - M| + 2|y_{fin} - V|$$

where x_{fin} and y_{fin} are respectively the final position and the velocity reached by the car, (M, V) the final position to be reached. The minimum of the simulation is obtained for $t_4 = 4.74$ and $t_5 = 8.5$ which is a good approximation of the theoretical minimum.

Notice that for both problems we first prescribe the supposed optimal strategy and then apply a Maximum Principle to single out the optimal trajectory. In other cases, when one has not clearly a candidate optimal strategy, one should check more strategies and relative optimal trajectories and then minimize over a finite set of possibilities. A key point is that the system avoids the so called Fuller or Zeno phenomenon if fictitious location switchings are not admitted.

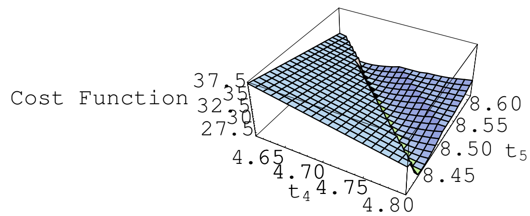


Figure 8.4: Cost as function of gear changes.

8.3 HNP for a car with gears.

In this section we consider a simple example of a hybrid system Σ not satisfying the property (H). We consider again the same model of a car, but, for simplicity, with only two gears. So $\mathcal{Q} = \{A\} \times \{1,2\}$. The hybrid system is the same as before, but we put further restrictions for controls usable after location switchings. We assume to have a $\mathcal{C}^1$ function $\zeta : \mathbb{R} \to [0,1]$ such that, if r is the switching velocity from gear 1 to 2, we can use only the controls of the form $\zeta(r)u(\cdot)$, $u \in \mathcal{U}$, after the switching time. Thus the switching set $\mathcal{S}_{(A,1),(A,2)}$ is given by the elements

$$((A,1),(z_1,z_2),(A,2),(z_1,z_2),u(\cdot),\tau)$$

such that $u(\cdot) \in \zeta(x_1^{(2)}(t_1))\mathcal{U}_{(A,2)}$. We consider the optimal control problem with initial condition $(0,0)$ and free terminal point. Moreover, instead of minimum time, we consider a zero running cost $L \equiv 0$ and the final cost

$$\varphi_{(A,1),(A,2)}((z_1,z_2),(z_3,z_4)) = z_3$$

to be maximized over all trajectories defined on $[0,T]$. We consider the family of trajectories that have a location switching at time t_1 (assuming $t_1 < T$) and using the results of the previous section, we want to determine if the control $u \equiv 1$ maximizes the cost $\varphi_{(A,1),(A,2)}$. Every $\tau \in]0,t_1[$ is a Lebesgue point for the function

$$t \mapsto f_{(A,1)}(x_1(t),1),$$

where $x_1(\cdot)$ is the trajectory corresponding to $u \equiv 1$ with $x_1(0) = 0$. Consider a needle variation based at time τ, see equation (7.2.8). By Lemma 7.2.2, the derivative $v_1(t)$ of $x_1^\varepsilon(t)$ w.r.t. ε calculated at $\varepsilon = 0$ evolves according to

$$\begin{cases} \dot{v}_1(t) = D_x f_{(A,1)}(x_1(t), 1) \cdot v_1(t), \\ v_1(\tau) = f_{(A,1)}(x_1(\tau), \omega) - f_{(A,1)}(x_1(\tau), 1), \end{cases} \tag{8.3.1}$$

while in the second location

$$\begin{cases} \dot{v}_2(t) = D_u f_{(A,2)}(x_2(t), 1) c_\omega + D_x f_{(A,2)}(x_2(t), 1) \cdot v_2(t), \\ v_2(t_1) = v_1(t_1), \end{cases}$$

with c_ω is a constant depending by ζ and ω. If we consider

$$\begin{cases} \dot{\psi}_2(t) = -\psi_2(t) D_x f_{(A,2)}(x_2(t), 1), \\ \psi_2(T) = (1, 0), \end{cases} \tag{8.3.2}$$

and

$$\begin{cases} \dot{q}_2(t) = -\psi_2(t) D_u f_{(A,2)}(x_2(t), 1) c_\omega, \\ q_2(T) = 0, \end{cases}$$

we have that $\psi_2(t) \cdot v_2(t) + q_2(t)$ is constant in the interval $[t_1, T]$. In particular it is non positive, since $\psi_2(T) \cdot v_2(T) + q_2(T) \leqslant 0$, if the control 1 maximizes the cost functional. If we write $\psi_2(t) = (\psi_2^1(t), \psi_2^2(t))$, then the system (8.3.2) becomes

$$\begin{cases} \dot{\psi}_2^1(t) = 0, \\ \dot{\psi}_2^2(t) = -\psi_2^1(t) - \psi_2^2(t) g_2'(x_2^{(2)}(t)), \\ \psi_2^1(T) = 1, \\ \psi_2^2(T) = 0, \end{cases}$$

and so, we have that

$$\psi_2^1(t) = 1,$$
$$\psi_2^2(t) = \frac{x_2^{(2)}(T) - x_2^{(2)}(t)}{g_2(x_2^{(2)}(t))}$$

and

$$q_2(t) = c_\omega \int_T^t x_2^{(2)}(s) ds - c_\omega x_2^{(2)}(T)(t - T).$$

We evaluate ψ_2 and q_2 at the point t_1:

$$\psi_2^1(t_1) = 1,$$
$$\psi_2^2(t_1) = \frac{x_2^{(2)}(T) - x_2^{(2)}(t_1)}{g_2(x_2^{(2)}(t_1))}$$
$$q_2(t_1) = c_\omega \int_T^{t_1} x_2^{(2)}(s)ds - c_\omega x_2^{(2)}(T)(t_1 - T). \tag{8.3.3}$$

It is clear that the covector does not jump, see [46, 59, 110]. In the first location $(A, 1)$, the adjoint covector is given by

$$\begin{cases} \dot{\psi}_1(t) = -\psi_1(t) \cdot D_x f_{(A,1)}(x_1(t), 1), \\ \psi_1^1(t_1) = 1, \\ \psi_1^2(t_1) = \frac{x_2^{(2)}(T) - x_2^{(2)}(t_1)}{g_2(x_2^{(2)}(t_1))}, \end{cases}$$

and so, the solution to this system is

$$\psi_1^1(t) = 1,$$
$$\psi_1^2(t) = \frac{x_2^{(2)}(T) - x_2^{(2)}(t_1)}{g_2(x_2^{(2)}(t_1))} \cdot \frac{g_1(x_1^{(2)}(t_1))}{g_1(x_1^{(2)}(t))} - \frac{1}{g_1(x_1^{(2)}(t))} \left[x_1^{(2)}(t) - x_1^{(2)}(t_1) \right].$$

In the first location the scalar product $v_1(t) \cdot \psi_1(t)$ is constant, so we have that $v_1(\tau) \cdot \psi_1(\tau) + q_2(t_1) \leqslant 0$. This implies that

$$(\omega - 1) \left[\frac{g_1(x_1^{(2)}(t_1))}{g_2(x_2^{(2)}(t_1))} (x_2^{(2)}(T) - x_2^{(2)}(t_1)) - (x_1^{(2)}(\tau) - x_2^{(2)}(t_1)) \right] +$$
$$+ c_\omega \left[\int_T^{t_1} x_2^{(2)}(s)ds + (T - t_1) x_2^{(2)}(T) \right] \leqslant 0.$$

Since t_1 is such that $g_1(x_1^{(2)}(t_1)) = g_2(x_2^{(2)}(t_1))$, then

$$(\omega - 1) \left[x_2^{(2)}(T) - x_1^{(2)}(\tau) \right] + c_\omega \left[\int_T^{t_1} x_2^{(2)}(s)ds + (T - t_1) x_2^{(2)}(T) \right] \leqslant 0. \tag{8.3.4}$$

Now, the two terms in the square brackets are both strictly positive. If we know that $c_\omega \leqslant 0$ for every $\omega \in [-1, 1]$, then the control $u(t) = 1$ satisfies the HNP.

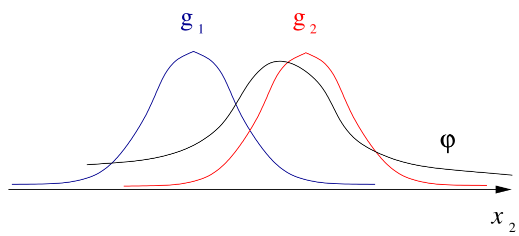


Figure 8.5: $\zeta'(x_2(t_1)) > 0$

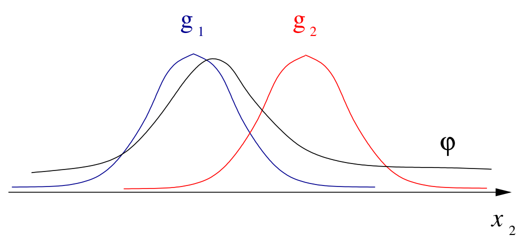


Figure 8.6: $\zeta'(x_2(t_1)) < 0$

By a simple computation, $c_\omega = \zeta'(x_2^{(2)}(t_1))v_1^2(t_1)$, where v_1^2 is the second component of v_1. By (8.3.1) we have that $v_1^2(t_1) < 0$ for all $\omega \neq 1$, and $v_1^2(t_1) = 0$ if $\omega = 1$. Therefore three cases are possible:

1. $\zeta'(x_2^{(2)}(t_1)) = 0$. In this case $c_\omega = 0$ for all $\omega \in [-1, 1]$ and so (8.3.4) becomes
$$(\omega - 1)\left[x_2^{(2)}(T) - x_1^{(2)}(\tau)\right] \leqslant 0$$
which is always true. Thus the control $u \equiv 1$ satisfies HNP.

2. $\zeta'(x_2^{(2)}(t_1)) > 0$. So ζ is an increasing function at $x_2^{(2)}(t_1)$, see Figure 8.5. In this case $c_\omega \leqslant 0$ for every ω and so $u \equiv 1$ is a control satisfying HNP.

3. $\zeta'(x_2^{(2)}(t_1)) < 0$. The function ζ is decreasing at $x_2^{(2)}(t_1)$, as in Figure 8.6. Thus $c_\omega > 0$ $\forall \omega \in [-1, 1[$ and so the control $u \equiv 1$ does not satisfy HNP. In this case it is not optimal. In fact, it is better to use a control $u \not\equiv 1$ or to choose a switching time less than t_1 in order to have a greater acceleration in the second location.

Thus in the third case a non–optimal trajectory, satisfying the classical HMP can be excluded by applying HNP. The results are confirmed by simulations.

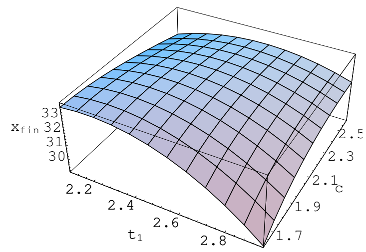


Figure 8.7: Final position as function of gear change and maximum of ζ.

8.3.1 Simulations.

We report the simulation results using the following function

$$\zeta(r) = e^{(-0.5(r-c))},$$

where r is the switching velocity from gear 1 to 2.

The following Figure 8.7 shows the maximal final position reached by the car in function of the time of switching from gear 1 to 2 and in function of the centre of the gaussian. The figure shows that the maximum of the cost is obtained for the switching time equal to 2.2 and for $c = 1.8$ according to HNP, in fact $\zeta'(x_2^{(2)}(t_1)) = 1.75452$.

Bibliography

[1] A. A. Agrachev, R. V. Gamkrelidze, A. V. Sarychev, *Local invariants of smooth control systems*, Acta Appl. Math. **14** (1989), no. 3, pp. 191–237.

[2] A. A. Agrachev, Yu. L. Sachkov, Control Theory from the Geometric Viewpoint, Springer–Verlag Berlin Heidelberg 2004.

[3] A. A. Agrachev, G. Stefani, P. Zezza, *An invariant second variation in optimal control*, Internat. J. Control **71** (1998), no. 5, pp. 689–715.

[4] J. H. Albert, *Genericity of simple eigenvalues for elliptic PDE's*, Proc. Amer. Math. Soc. **48** (1975), pp. 413–418.

[5] G. Alberti, L. Ambrosio, P. Cannarsa, *On the singularities of convex functions*, Manuscripta Math., **76** (1992), pp. 421–435.

[6] D. Amadori, *Initial–boundary value problems for systems of conservation laws*, NoDEA, **4** (1997), pp. 1–42.

[7] D. Amadori, R. M. Colombo, *Continuous dependence for 2x2 conservation laws with boundary*, J. Differential Equations, **138** (1997), pp. 229–266.

[8] F. Ancona, A. Marson, *On the attainable set for scalar nonlinear conservation laws with boundary control*, SIAM J. Control Optim. **36** (1998), pp. 290–312.

[9] F. Ancona, A. Marson, *Scalar non–linear conservation laws with integrable boundary data*, Nonlinear Analysis, **35** (1999), pp. 687–710.

[10] P. J. Antsaklis, A. Nerode eds., *Special issue on hybrid systems*, IEEE Trans. Automat. Control **43** (1998), no. 4.

[11] P. J. Antsaklis, X. XU, *A dynamic programming approach for optimal control of switched systems*, Proceedings of the 39th IEEE Conference on Decision and Control, Sidney, (2000), pp. 1822–1827.

[12] S. A. Avdonin, S. A. Ivanov, Families of exponentials, Cambridge University Press, 1995.

[13] A. Aw, M. Rascle, *Resurrection of "Second Order" Models of traffic flow?*, SIAM J. App. Math., **60** (2000), pp. 916–938.

[14] P. Baiti, A. Bressan, *The semigroup generated by a Temple class system with large data*, Differ. Integral Equ. **10**, 1997, pp. 401–418.

[15] M. Bardi, I. Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations, Birkhäuser Boston, Boston, 1997.

[16] C. Bardos, A. Y. Le Roux, J. C. Nedelec, *First Order Quasilinear Equations with Boundary Conditions*, Commun. Partial Differential Equations, **4** (1979), pp. 1017–1034.

[17] C. A. Berenstein, R. Gay, Complex Variables. An Introduction, Springer–Verlag, New York, 1991.

[18] R. M. Bianchini, *Good needle-like variations*. Differential geometry and control (Boulder, CO, 1997), 91–101, Proc. Sympos. Pure Math., **64**, Amer. Math. Soc., Providence, RI, 1999.

[19] V. G. Boltyanskii, *Sufficient conditions for optimality and the justification of the dynamic programming method*, SIAM J. Control, **4** (1966), pp. 326–361.

[20] U. Boscain, B. Piccoli, *Geometric control approach to synthesis theory. Control theory and its applications*, Rend. Sem. Mat. Univ. Politec. Torino, **56** (1998), pp. 53–68.

[21] U. Boscain, B. Piccoli, Optimal Syntheses for Control Systems on 2–D Manifolds, Springer SMAI, Vol. **43**.

[22] M. S. Branicky, V. S. Borkar, S. K. Mitter, *A unified framework for hybrid control: model and optimal control theory*, IEEE Trans. Automatic Control **43** (1998), pp. 31–45.

[23] D. Bresch, J. Simon, *On the normal variations of a domain*, ESAIM: COCV **3** (1998), pp. 251–261.

[24] A. Bressan, *A high order test for optimality of bang-bang controls*, SIAM J. Control Optim. **23** (1985), no. 1, pp. 38–48.

[25] A. Bressan, Lecture Notes on the Mathematical Theory of Control, SISSA, Trieste, Italy, 1994.

[26] A. Bressan, Hyperbolic Systems of Conservation Laws - The One-dimensional Cauchy Problem, Oxford Univ. Press, 2000.

[27] A. Bressan, G. Crasta, B. Piccoli, *Well Posedness of the Cauchy Problem for $n \times n$*, Amer. Math. Soc. Memoir, **146** (2000).

[28] A. Bressan, A. Marson, *A Variational Calculus for Discontinuous Solutions of Systems of Conservation Laws*, Comm. Part. Diff. Equat., **20** (1995), pp. 1491–1552.

[29] A. Bressan, B. Piccoli, *A generic classification of time-optimal planar stabilizing feedbacks*, SIAM J. Control Optim. **36** (1998), pp. 12–32.

[30] H. Brezis, Analyse fonctionnelle: Théorie et applications, Masson, 1987.

[31] P. Brunovský, *Existence of regular syntheses for general problems*, J. Differential Equations, **38** (1980), pp. 317–343.

[32] P. Cannarsa, A. Mennucci, C. Sinestrari, *Regularity results for solutions of a class of Hamilton–Jacobi equations*, Arch. Ration. Mech. Anal., **140** (1997), pp. 197–223.

[33] C. Cassandras, D. L. Pepyne, Y. Wardi, *Optimal control of a class of hybrid systems*, IEEE Trans. Automat. Control **46** (2001), pp. 398–415.

[34] Y. Chitour, J.–M. Coron, M. Garavello, *Obstructions for steady–state controllability of some linear partial differential equations*, Preprint Université Paris–sud.

[35] Y. Chitour, B. Piccoli, *Traffic circles and timing of traffic lights for car flow*, Preprint.

[36] F. H. Clarke, Optimization and Nonsmooth Analysis, Canad. Math. Soc. Ser. Monogr. Adv. Texts, Wiley, New York, 1983.

[37] F. H. Clarke, R. B. Vinter, *Optimal Multiprocesses*, SIAM J. Control Optim. **27** (1989), pp. 1072–1091.

[38] F. H. Clarke, R. B. Vinter, *Applications of the Theory of Optimal Multiprocesses*, SIAM J. Control Optim. **27** (1989), pp. 1048–1071.

[39] G. M. Coclite, M. Garavello, B. Piccoli, *Traffic Flow on a Road Network*, submitted to SIAM J. Math. Anal.

[40] R. M. Colombo, *Hyperbolic Phase Transitions in Traffic Flow*, SIAM J. Appl. Math., **63** (2002), pp. 708–721.

[41] J.–M. Coron, *On the controllability of the 2–D incompressible Navier–Stokes equations with the Navier slip boundary conditions*, ESAIM COCV **1** (1996), pp. 35–75.

[42] J.–M. Coron, *Some open problems in control theory*, Proceeding of Symposia in Pure Mathematics **64** (1999), pp. 149–162.

[43] R. Courant, D. Hilbert, Methods of Mathematical Physics, Volume I, Interscience Publishers, New York, 1953.

[44] C. Daganzo, *Requiem for second-order fluid approximation to traffic flow*, Transportation Res., Part B, **29**, (1995), pp. 277–286.

[45] C. Dafermos, Hyperbolic Conservation Laws in Continuum Physics, Springer–Verlag, 1999.

[46] C. D'Apice, R. Manzo, M. Garavello, B. Piccoli, *Hybrid optimal control: case study of a car with gears*, Internat. J. Control **76**, (2003), pp. 1272–1284.

[47] B. De Schutter, *Optimal control of a class of linear hybrid systems with saturation*, SIAM J. Control Optim. **39** (2000), pp. 835–851.

[48] J. I. Diaz, *Obstruction and some approximate controllability results for the Burgers equation and related problems*, in Control of partial differential equations and applications (Laredo, 1994), pp. 63–76, Lecture Notes in Pure and Appl. Math., **174**, Dekker, New York, 1996.

[49] M. D. Di Benedetto ed., *Hybrid systems in control*, Internat. J. Robust Nonlinear Control **11** (2001), no. 5.

[50] M. D. Di Benedetto, A. Sangiovanni–Vincentelli eds., Hybrid System Computation and Control, Lecture Notes in Computer Science **2034**, Springer–Verlag, Berlin Heidelberg, 2001.

[51] L. C. Evans, R. F. Gariepy, Measure Theory and Fine Properties of Functions, Stud. Adv. Math., CRC, Boca Raton, FL, 1992.

[52] H. O. Fattorini, *Control in finite time of differential equations in Banach space*, Comm. Pure Appl. Math. **19** (1966), pp. 17–34.

[53] W. H. Fleming, R. W. Rishel, Deterministic and Stochastic Optimal Control, Springer–Verlag, New York, 1975.

[54] G. B. Folland, Real Analysis: Modern Techniques and their Applications, J. Wiley and sons, 1984.

[55] A. V. Fursikov, Optimal Control of Distributed Systems. Theory and Applications, Translations of Mathematical Monographs. American Mathematical Society, 2000.

[56] A. V. Fursikov, O. Yu. Imanuvilov, *On controllability of certain systems simulating a fluid flow*, in Flow Control, IMA vol. in Math. and its Appl., M. D. Gunzburger ed., Springer–Verlag, New York, **68** (1994), pp. 149–184.

[57] A. V. Fursikov, O. Yu. Imanuvilov, *Local exact controllability of the Navier–Stokes equations*, C. R. Acad. Sci. Paris, **323** (1996), pp. 275–280.

[58] M. Garavello, *Verification Theorems for Hamilton–Jacobi–Bellman equations*, SIAM J. Control Optim. **42** (2003), no. 5, pp. 1623–1642.

[59] M. Garavello, B. Piccoli, *Hybrid Necessary Principle*, SIAM J. Control Optim., to appear.

[60] J. M. Greenberg, *Extension and Amplifications of a Traffic Model of Aw and Rascle*, SIAM J. Appl. Math., **62** (2001), pp. 729-745.

[61] I. M. Guelfand, G. E. Chilov, Les Distributions, Tome 2, Espaces fondamentaux, Dunod, Paris, 1964.

[62] R. S. Hamilton, *The inverse function theorem of Nash and Moser*, Bull. Amer. Math. Soc. (N.S.) **7** (1982), pp. 65–222.

[63] P. Hartman, Ordinary Differential Equations, S. H. Hartman, Baltimore, 1973.

[64] S. Hedlund, A. Rantzer, *Optimal control of hybrid systems*, Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, 1999, pp. 3972–3977.

[65] H. Holden, N. H. Risebro, *A Mathematical Model of Traffic Flow on a Network of Unidirectional Roads*, SIAM J. Math. Anal., **26** (1995), pp. 999–1017.

[66] T. Horsin, *On the controllability of the Burgers equation*, ESAIM COCV **3** (1998), pp. 83–95.

[67] V. Ivrii, *Microlocal analysis and precise spectral asymptotics*, SMM, Springer, 1991.

[68] V. Jurdjevic, *Geometric control theory*, Cambridge University Press, 1997.

[69] R. E. Kalman, P. L. Falb, M. A. Arbib, Topics in mathematical system theory, International Series in Pure and Applied Mathematics, McGraw–Hill Book Company, New York, 1969.

[70] A. Klar, Kinetic and macroscopic traffic flow models, lecture notes for the XX School of Computational Mathematics “Computational aspects in kinetic models”, Piano di Sorrento (Italy), September 22–28, 2002.

[71] A. Klar, R. D. Kühne, R. Wegener, *Mathematical models for vehicular traffic*, Surveys Math. Indust. **6** (1996), no. 3, pp. 215–239.

[72] A. Klar, R. Wegener, *Enskog–like kinetic models for vehicular traffic*, J. Statist. Phys. **87** (1997), no. 1-2, pp. 91–114.

[73] C. Klingenberg, N. H. Risebro, *Convex Conservation Laws with Discontinuous Coefficients. Existence, uniqueness and asymptotic behavior*, Comm. Partial Differential Equations, **20** (1995), pp. 1959–1990.

[74] A. J. Krener, *The high order maximal principle and its application to singular extremals*, SIAM J. Control Optimization **15** (1977), no. 2, pp. 256–293.

[75] J. C. Lagarias, A. M. Odlyzko, *Effective versions of the Chebotarev density theorem*, Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975), pp. 409–464. Academic Press, London, 1977.

[76] B. Ya Levin, Lectures on entire functions, Translations of mathematical monographs, **150**, American mathematical society, Rhode Island, USA, 1996.

[77] M. J. Lighthill, G. B. Whitham, *On kinetic waves. II. Theory of Traffic Flows on Long Crowded Roads*, Proc. Roy. Soc. London Ser. A, **229** (1955), pp. 317–345.

[78] J.–L. Lions, E. Zuazua, *Approximate controllability of a hydro-elastic coupled system*, ESAIM: COCV **1** (1996), pp. 1–15.

[79] J. Lygeros, G. J. Pappas, S. Sastry, *An introduction to hybrid systems modeling, analysis and control*, Preprints of the First Nonlinear Control Network Pedagogical School (1999), Athens, Greece, pp. 307–329.

[80] M. Malisoff, *On the Bellman equation for control problems with exit times and unbounded cost functionals*, in Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, AZ, 1999, pp. 23–28.

[81] M. Malisoff, H. J. Sussmann, *Further results on the Bellman equation for optimal control problems with exit times and nonnegative instantaneous costs*, in Proceedings of the 39th Annual IEEE Conference on Decision and Control (Sidney, Australia), IEEE Control Systems Society, Piscataway, NJ, 2000.

[82] S. Micu, E. Zuazua, *On the lack of null controllability of the heat equation on the half line*, Transactions of the A.M.S. **353** (2001), pp. 1635–1659.

[83] S. Micu, E. Zuazua, *On the lack of null controllability of the heat equation on the half space*, Portugaliae Mathematica **58** (2001), pp. 1–24.

[84] S. Micu, E. Zuazua, *On the controllability of a fractional order parabolic equation*, preprint, Universidad Autónoma de Madrid, 2004.

[85] P. M. Morse, H. Feshbach, Methods of Theoretical Physics, Part I, McGraw–Hill Book Company, Tokyo, 1953.

[86] A. S. Morse, C. C. Pantelides, S. Sastry, J. M. Schumacher eds., *Special issue on hybrid systems*, Automatica J. IFAC **35** (1999).

[87] J. H. Ortega, E. Zuazua, *Generic simplicity of the spectrum and stabilization for a plate equation*, SIAM J. Control Optim. **39** (2000), pp. 1585–1614.

[88] J. H. Ortega, E. Zuazua, *Generic simplicity of the eigenvalues of the Stokes system in two space dimensions*, Adv. Differential Equations **6** (2001), pp. 987–1023.

[89] J. H. Ortega, E. Zuazua, *On a constrained approxiamte controllability problem for the heat equation*, J. of Optm. Theory and App. **1** (2001), pp. 29–64.

[90] H. J. Payne, *Models of freeway traffic and control*, Simulation Council (1971).

[91] N. Petit, P. Rouchon, *Dynamics and solutions to some control problems for water–tank systems*, IEEE Trans. Automat. Control **47** (2002), pp. 594–609.

[92] B. Piccoli, *Classification of generic singularities for the planar time-optimal synthesis*, SIAM J. Control Optim., **34** (1996), pp. 1914–1946.

[93] B. Piccoli, *Hybrid systems and Optimal control*, Proceedings of the 37th IEEE Conference on Decision and Control, Tampa, 1998, pp. 13–18.

[94] B. Piccoli, *Infinite time regular synthesis*, ESAIM Control Optim. Calc. Var., **3** (1998), pp. 381–405.

[95] B. Piccoli, *Necessary conditions for hybrid optimization*, Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, 1999, pp. 410–415.

[96] B. Piccoli, H. J. Sussmann, *Regular synthesis and sufficient conditions for optimality*, SIAM J. Control Optim. **39** (2000), pp. 359–410.

[97] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishenko, The Mathematical Theory of Optimal Processes, John Wiley, New York, 1962.

[98] M. H. Protter, H. F. Weinberger, *Maximum Principles in Differential Equations*, Prentice–Hall, New Jersey, 1967.

[99] P. I. Richards, *Shock Waves on the Highway*, Oper. Res., **4** (1956), pp. 42–51.

[100] P. Riedinger, F. Kratz, C. Iungand, C. Zanne, *Linear quadratic optimization for hybrid systems*, Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, 1999, pp. 2466–2470.

[101] P. Rouchon, *Calcul symbolique et génération de trajectoires pour certaines EDP avec contrôle frontière*, lecture at the Collège de France, Décembre 2003.

[102] D. L. Russell, *Controllability and stabilizability theory for linear partial differential equations. Recent progress and open questions.* SIAM Rev. **20**, (1978), pp. 639–739.

[103] H. Schättler, *Regularity properties of optimal trajectories: recently developed techniques*, in Nonlinear controllability and optimal control, pp. 351–381, Monogr. Textbooks Pure Appl. Math. **133**, Dekker, New York, 1990.

[104] D. Serre, *Solutions à variations bornées pour certains systèmes hyperboliques de lois de conservation*, J. Differential Equations **68**, (1987), pp. 137–168.

[105] M. A. Shubin, Pseudodifferential operators and spectral theory, Springer Series in Soviet Math., Springer-Verlag, (1987).

[106] J. Simon, Differenciacion de problemas de contorno respecto del dominio, Universidad de Sevilla (1991) 122 pages.

[107] G. Stefani, *Higher order variations: how can they be defined in order to have good properties?* Nonsmooth analysis and geometric methods in deterministic optimal control (Minneapolis, MN, 1993), 227–237, IMA Vol. Math. Appl., **78**, Springer, New York, 1996.

[108] H. J. Sussmann, *Multidifferential calculus: chain rule, open mapping and transversal intersection theorems*. In Optimal Control: Theory, Algorithms and Applications, W. W. Hager and P. M. Pardalos Editors, Kluwer Academic Publishers, 1998, pp. 436–487.

[109] H. J. Sussmann, *A Maximum Principle for hybrid optimal control problems*, Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, 1999.

[110] H. J. Sussmann, *A nonsmooth hybrid maximum principle*. Stability and stabilization of nonlinear systems (Ghent, 1999), 325–354, Lecture Notes in Control and Inform. Sci., **246**, Springer, London, 1999.

[111] H. J. Sussmann, *New theories of set–valued differentials and new versions of the Maximum Principle of optimal control theory*, in A. Isidori, F. Lamnabhi–Lagarrigue and W. Respondek (eds.) *Nonlinear control in the year 2000*, Springer–Verlag (2000), pp. 487–526.

[112] H. J. Sussmann, *Set-valued differentials and the Hybrid Maximum Principle*, Proceedings of the 39th IEEE Conference on Decision and Control, Sidney, 2000.

[113] M. E. Taylor, Partial differerential equations I, Applied mathematical sciences, **115**, Springer–Verlag, New York, 1996.

[114] M. E. Taylor, Partial Differential Equations III, Nonlinear Equations, Applied Mathematical Sciences **117**, Springer–Verlag, 1996.

[115] B. Temple, *Global Solutions of the Cauchy Problem for a Class of* 2×2 *Nonstrictly Hyperbolic Conservation Laws*, Advances in Applied Mathematics, **3** (1982), pp. 335–375.

[116] J. D. Towers, *A Difference Scheme for Conservation Laws with a Discontinuous Flux – the Nonconvex Case*, SIAM J. Numer. Anal., **39** (2001), pp. 1197–1218.

[117] K. Uhlenbeck, *Generic properties of eigenfunctions*, Amer. J. Math. **98** (1976), pp. 1059–1078.

[118] R. Vinter, Optimal Control, Birkhäuser Boston, Boston, 2000.

[119] G. B. Whitham, Linear and Nonlinear Waves, Pure and Applied Math., Wiley–Interscience, New York, 1974.

[120] L. Zajicek, *On the points of multiplicity of monotone operators*, Comment. Math. Univ. Carolin., **19** (1978), pp. 179–189.

[121] M. I. Zelikin, V. F. Borisov, Theory of Chattering Control with Applications to Astronautics, Robotics, Economics, and Engineering, Birkhäuser Boston, Boston, 1994.