# Numerical Methods for Porous Media and Shallow Water Flows

# &

# An Algebra of Singular Semiclassical Pseudodifferential Operators

by

Gerardo Hernández-Dueñas

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
2011

Doctoral Committee:

Professor Smadar Karni, Co-Chair
Professor Alejandro Uribe-Ahumada, Co-Chair
Professor Daniel M. Burns Jr
Assistant Professor Christiane Jablonowski
Professor Robert Krasny

To my son Gerardo, my wife Yadira, my parents Gerardo and Lilia and all my family.

# ACKNOWLEDGEMENTS

My family and God have always been my greatest moral support and I cannot imagine finishing my Ph.D. without their help. I especially thank my wife Yadira Baeza Rodríguez for being such a great spouse and for her company and support every day. I sincerely thank my parents Gerardo Hernández Trujillo and Lilia Dueñas Casillas for their care and love, and for always encouraging me to pursue a career, which is greatly appreciated. I am also very grateful to my lovely son, Gerardo Hernández Baeza, who brought a new meaning to my life. He is responsible for many wonderful memories that I have from Ann Arbor, including the experience of cutting his umbilical cord. My siblings Liliana and María, my aunts Gloria and Celia, my grandmother María, and my nephews and nieces are also very special to me.

I deeply thank my advisors, Professors Smadar Karni and Alejandro Uribe, for their guidance and support, and especially for their friendship. I have been in contact with Professor Alejandro Uribe even before my arrival to Ann Arbor and I still remember when he picked me up from the airport on my first day here. I met Professor Smadar Karni later when I took one of her courses. A question that I had at the end of a session ended up in our collaboration. I thank Professor Karni for supporting me with her grant (NSF DMS 0609766) as a Graduate Student Research Assistant and for helping me gain teaching experience as a Graduate Student Instructor for two semesters. I am very grateful for both of my advisors, and their knowledge, passion for mathematics, and professionalism will always be inspirations to me.

I greatly appreciate the participation of the doctoral committee in my thesis de-

fense. It consisted of Professors Smadar Karni, Alejandro Uribe, Daniel M. Burns, Christiane Jablonowski and Robert Krasny.

A very valuable part of my life here has been the time spent with my friends. They have made my stay here very enjoyable and my Ph.D. would not have been possible without their help. I am especially thankful for the friendship, help and support of Daniel Hernández, Emily Witt, Luis Nuñez, Chelsea Walton, José Luis González, Luis Serrano, Jose Manuel Gómez, Alfonso Tercero, Tere Tercero, Brian Jennings, Geri Izbicki, and Sue Sierra. My wife and I became parents during my Ph.D., and it was a wonderful and rewarding but also hard experience. My wife and I will always be very grateful with Daniel and Emily for their help and company before and after my son's birth. This is something I will never forget.

Professor Enrique Farias was the first person who encouraged me to take mathematics as a profession and as a lifestyle. He recruited me to participate in the Mexican Mathematics Olympiad (MMO) and I thank him for believing in me, and for his help finding financial support. I also thank Professor Martin Isaias for his friendship and for teaching me how to write rigorous mathematical proofs for the MMO competition. They really made a big change in my life and their advice and guidance helped me discover the beauty of mathematics. I also met some of my best friends there: Fernando Diaz, Pablo Diaz, Eira Delgado, Alci Cárdenas, Alicia Sánchez and Verónica Arevalo. I also thank Professors Alberto Martinez Montoya, Gerardo Lopez Aguirre, and Julio Mecina for the education received in my hometown.

Coming to the University of Michigan was not an easy task at all, and I got help from many friends. My undergraduate advisor, Professor Xavier Gomez-Mont from CIMAT (Mexican Research Center for Mathematics) was a wonderful mentor and I really thank him for encouraging me to apply to The University of Michigan, and for all the financial support provided. Professor Adolfo Sanchez supported me in the application for the CONACYT (Mexican Council for Science and Technology)

fellowship. I thank Professor Helga Fetter Nathansky for being such a great professor and for putting me in contact with Professor Alejandro Uribe through Professor Berta Gamboa. My friends Vidal Alcalá, Luis López, and Areli Vasquez helped me a lot with the CONACYT fellowship application process. Professor Ignacio Barradas and Genoveva Aguilera were a great help in getting all my documents on time from the University of Guanajuato. Professor Lawrence Nash and Susy Owen helped me a lot with my English-improving endeavors. I finally thank Professors Helga Fetter Nathansky, Xavier Gomez-Mont, Renato Iturriaga, and Adolfo Sanchez for writing letters of recommendation to support my application to The University of Michigan.

The following professors have contributed to my mathematical formation, which I graciously acknowledge: Jorge Balbás, David Barret, Daniel Burns, Charles Doering, Hendrikus Derksen, Sergey Fomin, Juha Heinonen, Lizhen Ji, Smadar Karni, Peter Miller, Jeffrey Rauch, Peter Scott, Karen Smith, Joel Smoller, Alejandro Uribe and Divakar Viswanath. I thank Professors Jorge Balbás, William Breslin, Daniel Burns, Smadar Karni, Angela Kubena, Jeffrey Rauch and Alejandro Uribe for writing letters of recommendation regarding my research and teaching for my postdoctoral applications. I also learned a lot from the informal meetings with Geri Izbicki, Dave Starinshak, Paul Ullrich, Daniel Zaide and Professors Smadar Karni and Philip Roe.

I give special thanks to all the staff members of the Department. In particular, I am very grateful to Bennet Fauber for his help and support with the software and hardware necessary for the numerical results.

I would also like to thank my sponsors, CONACYT for proving the financial support through the fellowship number CONACYT #160147, which made my stay at Michigan possible. The fellowship covered tuition, living expenses, and partial coverage for medical insurance for 5 years. I also thank PROMEP (Mexican Program for the Quality Improvement of Professor's Teaching) and the Colima State Government for providing additional financial support.

# TABLE OF CONTENTS

# LIST OF FIGURES

**Figure**

# LIST OF APPENDICES

**Appendix**

# ABSTRACT

Numerical Methods for Porous Media and Shallow Water Flows
&
An Algebra of Singular Semiclassical Pseudodifferential Operators

by

Gerardo Hernández-Dueñas

Co-Chairs:

Smadar Karni and Alejandro Uribe-Ahumada

**Numerical Methods for Hyperbolic Systems.** This work considers the Baer-Nunziato model for two-phase flows in porous media with discontinuous porosity. Numerical discretizations may fail to correctly capture the jump conditions across the so-called compaction wave, and yield incorrect solutions. This work formulates the system using the Riemann Invariants across the porosity jump, and proposes a hybrid algorithm that uses the Riemann Invariants formulation across the compaction wave, and the conservative formulation away from the compaction wave. The hybrid scheme is described and numerical results are presented.

The shallow water equations for flows in channels with arbitrary cross-section are also considered. The system forms a hyperbolic set of balance laws. Exact steady-state solutions are available and are controlled by the relation between the bottom topography and the channel geometry. In this work a Roe-type upwind

scheme for the system is developed. Considerations of conservation, near steady-state accuracy, velocity regularization and positivity near dry states are discussed. Numerical solutions are presented.

**Semiclassical Analysis.** This work is also concerned with the quantization of compact symplectic manifolds with boundary, contained in cotangent bundles. The boundary is always foliated by curves tangent to the kernel of the pull-back of the symplectic form. Assuming the fibrating-and-Bohr-Sommerfeld conditions, an algebra of semiclassical pseudodifferential operators with singular symbols that are naturally associated to the symplectic manifold is constructed. A symbolic calculus is developed, where elements in the algebra have principal symbols on the diagonal and on the flow-out given by the fibers of the foliation, with a singularity on the intersection. The singular principal symbols on the flow-out are identified with families of classical pseudodifferential operators acting on smooth half-densities on the fibers. It is shown that the algebra admits projectors under mild conditions, a generalized Szegö limit theorem is proved, and some corresponding singular propagators are studied. The action of the present singular propagators on coherent states is studied numerically, and it is observed that the center of the coherent states may split into more than one when it approaches to the boundary.

# CHAPTER I

# Introduction to Hyperbolic Balance Laws

This work is concerned with numerical methods for hyperbolic balance laws. Roughly speaking, these are equations of the form of hyperbolic conservation laws with non-zero right-hand side terms. In this introductory chapter, we provide a quick overview of solutions of hyperbolic conservation laws and some standard numerical techniques used in the field. More details can be found in [LeV92, LeV02].

We start by considering nonlinear hyperbolic conservation laws, which are equations of the form

$$W_t + F(W)_x = 0. \tag{1.1}$$

Here $W(x,t) = (W_1(x,t), W_2(x,t), \ldots, W_N(x,t))$ is the vector of conserved quantities, $N$ is the dimension of the system, $F(W)$ is the vector of flux functions, where the Jacobian $\frac{\partial F}{\partial W}$ has real eigenvalues and a complete set of eigenvectors. Except when otherwise stated, we denote vector quantities with upper case letters, and scalars with lower case letters.

Many physical phenomena can be modeled by this class of equations. The Euler equations and the shallow water equations over flat topography are examples of such systems.

## 1.1 Weak Solutions

### 1.1.1 Shock Waves and the Entropy Condition

Hyperbolic conservation laws are distinguished by the propagation of information along characteristics at finite speed. We illustrate it by considering the scalar case

$$\begin{cases} u_t + f(u)_x = 0, \\ u(x,0) = u_0(x) \end{cases} \tag{1.2}$$

where $f(u)$ is a convex (or a concave) function. The simplest example is the Burgers' equation in which $f(u) = \frac{1}{2}u^2$, and

$$u_t + uu_x = 0.$$

The solution $u$ is constant along the characteristics

$$x'(t) = u(x(t), t)$$

since

$$\frac{d}{dt}u(x(t),t) = u_t(x(t),t) + u_x(x(t),t)x'(t) = u_t + uu_x = 0.$$

Furthermore, characteristics are straight lines and if the initial data are smooth, this can be used to determine the solution $u(x,t)$ for small enough $t$ such that the characteristics do not cross: For each $(x,t)$ we can solve the equation

$$x = \xi + u_0(\xi)t$$

for $\xi$ and then

$$u(x,t) = u_0(\xi).$$

Here $(\xi, 0)$ is the point in the $x$-axis in the $(x,t)$-plane that is connected to $(x,t)$ through the characteristic passing through $(\xi, 0)$. The characteristics intersect when faster information overtakes slower information, and multi-valued solutions appear (see Figure 1.1). It can be shown that if the initial condition is smooth and $u_0'(x)$ is negative somewhere, then the characteristics intersect at time

$$T = \frac{-1}{\min u_0'(x)}.$$

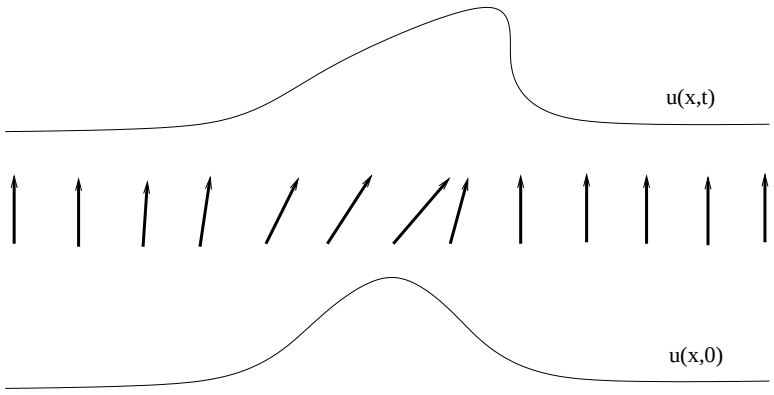



Figure 1.1: Characteristics and solution for the Burgers' equation.

For $t > T$, solutions are extended as weak solutions, which are piecewise smooth, separated by curves of discontinuity called shock waves. These weak solutions satisfy an integral version of equation (1.1), which is obtained by integrating the equation over the domain $[x_1, x_1 + \Delta x] \times [t_1, t_1 + \Delta t]$:

$$\begin{aligned} \int_{x_1}^{x_1+\Delta x} W(x, t_1 + \Delta t)dx - \int_{x_1}^{x_1+\Delta x} W(x, t_1)dx \\ + \int_{t_1}^{t_1+\Delta t} F(W(x_1 + \Delta x, t))dt - \int_{t_1}^{t_1+\Delta t} F(W(x_1, t))dt = 0 \end{aligned} \tag{1.3}$$

Notice that condition (1.3) does not require the solution to be smooth. Near shock

waves, one can integrate equation (1.3) over a small finite volume (see Figure 1.2). Assuming that the solution is approximately constant on each side of the box, one obtains

$$\Delta x \;\; (W_\ell - W_r) + \Delta t \, (F_r - F_\ell) = O(\Delta t^2),$$

where $W_\ell, W_r, F_\ell, F_r$ are the states and fluxes on left and right, respectively, and $\Delta x / \Delta t \to s$ as $\Delta x, \Delta t \to 0$.

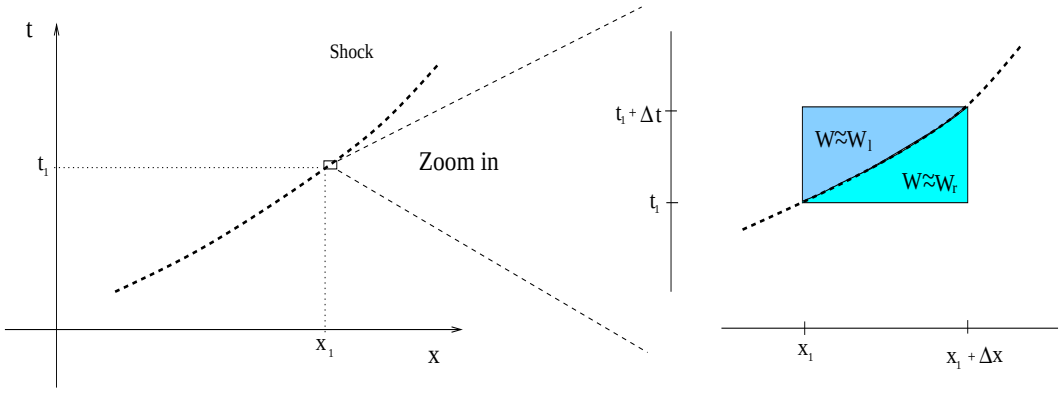



Figure 1.2: Shock waves in weak solutions.

This yields the Rankine-Hugoniot jump conditions

$$[F(W)] = s[W],$$

where $s$ is the speed of the shock, and $[\cdot] = ()_r - ()_\ell$ denotes the jump. Weak solutions are not unique, even in the scalar case. The physically relevant solutions is selected by the *entropy condition*. For convex (concave) functions $f$, the entropy condition says

$$f'(u_\ell) > s > f'(u_r),$$

which reduces to the condition $f'(u_\ell) > f'(u_r)$.

Another way to formulate the entropy condition uses a entropy function $\eta(u)$ and a corresponding entropy flux $\psi(u)$ such that $\psi'(u) = \eta'(u) f'(u)$. The function $u(x,t)$ is

the entropy solution of (1.2) if, for all convex entropy function $\eta(u)$ and corresponding entropy flux $\psi(u)$, the inequality

$$\eta(u)_t + \psi(u)_x \leq 0$$

is satisfied in the weak sense. It can be shown that this condition selects the weak solution which is the limit of the viscous equation as the viscosity tends to zero.

### 1.1.2 Rarefaction Fans

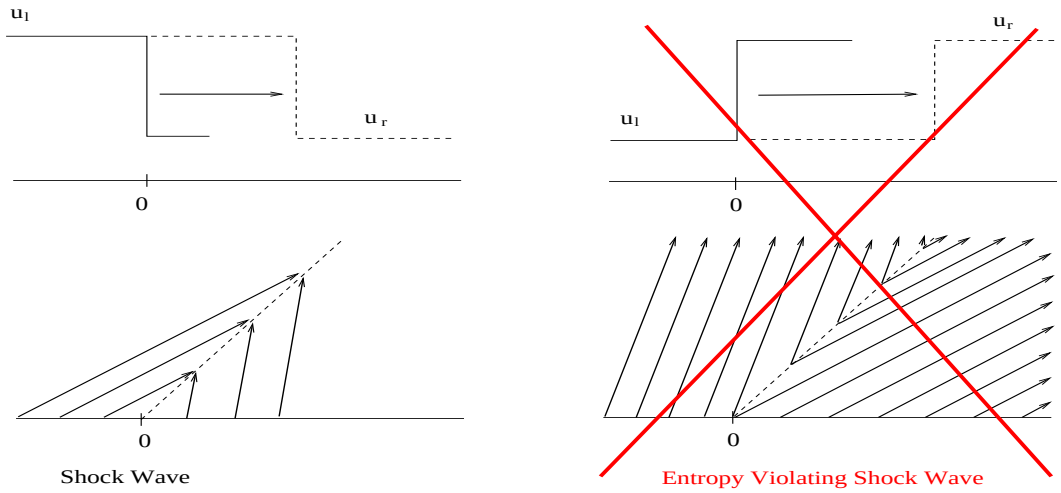



Figure 1.3: Shock wave (left) and entropy-violating shock (right).

Figure 1.3 shows a shock wave that satisfies the entropy condition (left) and a shock wave that violates it (right). The entropy-violating shock wave has characteristics emerging from the shock wave, and makes it unstable to perturbation. Either smearing out the initial profile a little, or adding some viscosity to the system, will cause the shock to be replaced by a rarefaction fan of characteristics, as in Figure 1.4.

The Riemann problem consists of piecewise constant initial conditions

$$W(x,0) = \begin{cases} W_\ell & \text{if } x < 0, \\ W_r & \text{if } x > 0. \end{cases}$$

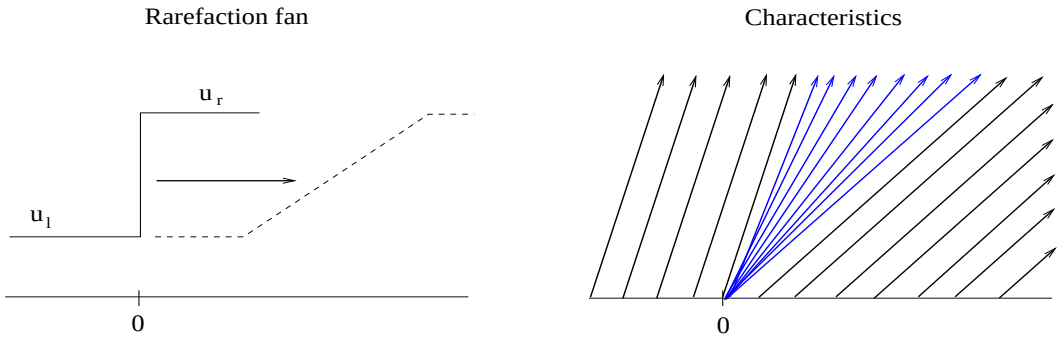



Figure 1.4: Rarefaction fan for a scalar solution.

The solution for the Riemann problem may contain shock waves or rarefaction fans. Rarefaction fans are smooth solutions of the form $W(x,t) = \omega(x/t)$, which are constant along rays between the two states that are connected through the rarefaction. Such a function $W(x,t) = \omega(x/t)$ is a solution of equation (1.1) if and only if

$$F'(\omega(\xi))\omega'(\xi) = \omega'(\xi)\ \xi,$$

where $F'(W) = \frac{\partial F}{\partial W}$ is the Jacobian matrix. This implies that

$$\omega'(\xi) = \alpha(\xi)\ r_p(\omega(\xi)),\ \xi = \lambda_p(\omega(\xi)),$$

where $r_p$ is an eigenvector of the Jacobian matrix, $\alpha(\xi)$ is a scalar, and $\lambda_p$ is the corresponding eigenvalue. Rarefaction waves are characterized by quantities that remain constant across the fan, also known as Riemann invariants. The Riemann invariants associated to the $p$-th characteristic field are found by integrating the equations

$$\frac{dW_1}{r_p^1} = \frac{dW_2}{r_p^2} = \ldots = \frac{dW_N}{r_p^N},$$

where $r_p = (r_p^1, r_p^2, \ldots, r_p^N)$. Therefore, each $p$-rarefaction in a $N$-dimensional system has $N-1$ Riemann invariants.

Two states $W_\ell$, $W_r$ are connected through a $p$-rarefaction wave provided they lie

on the same integral curve, and

$$\lambda_p(W_\ell) < \lambda_p(W_r).$$

For a nonlinear system it might happen that in one of the characteristic fields the eigenvalue $\lambda_p(u)$ is constant along the integral curves of this field, and hence

$$\nabla\lambda_p(W) \cdot r_p(W) \equiv 0 \ \ \forall\, W.$$

In this case we say that the $p$th field is *linearly degenerate*. The Euler equations have this property, and the linearly degenerate field is known as *contact discontinuity*.

If the field is genuinely nonlinear (not linearly degenerate), we can find the solution inside the rarefaction by expressing the coefficient $\alpha(\xi)$ in terms of $\lambda_p, r_p$. The equation reduces to

$$\begin{cases} \omega'(\xi) = \frac{1}{\nabla\lambda_p(\omega(\xi))\cdot r_p(\omega(\xi))} r_p(\omega(\xi)), \ \ \lambda_p(W_\ell) = \xi_1 \leq \xi \leq \xi_2 = \lambda_p(W_r) \\ \\ \omega(\xi_1) = W_\ell \end{cases}$$

## 1.2 Source Terms

We are particularly interested in conservation laws with source terms, also called balance laws, which are of the form

$$W_t + F(W)_x = S\left(x, W\right). \tag{1.4}$$

They model a wide variety of physical phenomena, where the source terms $S$ generally describe geometrical effects, and do not involve the derivatives of the solution. Examples include source terms that describe the bottom topography of a domain,

the geometry of a channel, or the effect of a non-cartesian geometry in the Euler equations. The corresponding equation (1.3) becomes

$$\int_{x_1}^{x_1+\Delta x} W(x,t_1+\Delta t)dx - \int_{x_1}^{x_1+\Delta x} W(x,t_1)dx + \\ + \int_{t_1}^{t_1+\Delta t} F(W(x_1+\Delta x,t))dt - \int_{t_1}^{t_1+\Delta t} F(W(x_1,t))dt = \int_{t_1}^{t_1+\Delta t}\int_{x_1}^{x_1+\Delta x} S(x,W)dxdt \tag{1.5}$$

If the source terms are bounded, they do not affect the jump conditions and weak solutions are obtained analogously. One can easily verify it by integrating equation (1.5) on a small domain $[x_1+\Delta x, t_1+\Delta t]$

$$\Delta x\ (W_\ell - W_r) + \Delta t\,(F_r - F_\ell) = O(\Delta t^2) + \int_{t_1}^{t_1+\Delta t}\int_{x_1}^{x_1+\Delta x} S(x,W)dxdt = O(\Delta t^2),$$

which gives the same jump conditions

$$[F] = s[W].$$

Additional complications may arise when the source terms involve derivatives of the solution, as observed in the Baer-Nunziato model studied in Chapter III.

## 1.3 Numerical Methods

### 1.3.1 Conservation

Due to the non-linearity of $F(W)$, exact solutions are rarely available. Solving those systems then relies on numerical methods. Even when exact solutions may be unattainable, much is known about the general mathematical structure of the solutions. Numerical methods may be carefully built according to the structure and

properties of the system. One such property is conservation. Godunov's method ([God59]) considers initial data that are picewise constant, solves the Riemann problem between the cells, and integrates over each cell to update the solution in the next time step. Figure 1.5 shows the schematic in the case where two waves are present. The time step $\Delta t$ is required in principle to be small enough so that neighboring Riemann problems are independent of each other. When two waves interact, they slow down and the constraint on the time step can be relaxed. This yields the CFL condition

$$\lambda_{\max}\frac{\Delta t}{\Delta x} \leq 1,$$

where $\lambda_{\max}$ is the maximum eigenvalue of the Jacobian matrix over all the cells.

Conservative methods have the general form for the update of the cell averages

$$\overline{W_j^{n+1}} = \overline{W_j^n} - \frac{\Delta t}{\Delta x}\left(F_{j+1/2} - F_{j-1/2}\right), \tag{1.6}$$

where $F_{j+1/2} = F(W_j, W_{j+1})$ is the numerical flux function. Different choices of the numerical fluxes generate different methods. The Lax-Wendroff theorem ([LW60]) ensures that if the numerical solution converges to some function as the grid is refined, the function has to be a weak solution of the conservation law.

Godunov schemes are conservative by construction. However, they require the exact solution of the Riemann problem on each cell interface. This task might be not possible to do systematically, or computationally expensive.

Instead of solving the Riemann problem exactly between cells, one can use approximate Riemann solutions. In [Roe81], P.L. Roe proposed an upwind scheme based on linearizing the Jacobian matrix at the interface between cells by a matrix $\hat{A}(W_\ell, W_r)$ satisfying the following conditions:

(i) $\hat{A}(W_\ell, W_r)\Delta W = \Delta F$, where $\Delta W = W_r - W_\ell,\ \Delta F = F_r - F_\ell$

(ii) $\hat{A}(W_\ell, W_r)$ is diagonalizable with real eigenvectors, and

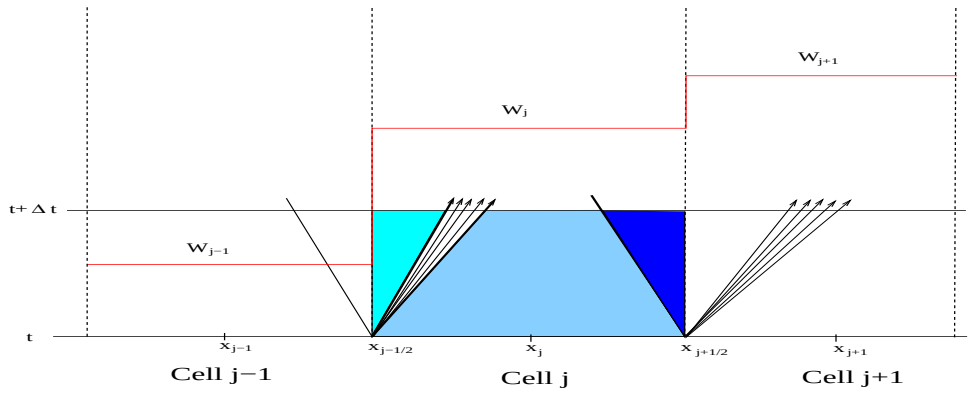



Figure 1.5: Godunov's methods.

(iii) $\hat{A}(W_\ell, W_r) \to F'(\bar{W})$ smoothly as $W_\ell, W_r \to \bar{W}$

The Roe scheme is then given by

$$W_j^{n+1} = W_j^n - \frac{\Delta t}{\Delta x}\left\{A^+_{j-\frac{1}{2}}\left(W_j^n - W_{j-1}^n\right) + A^-_{j+\frac{1}{2}}\left(W_{j+1}^n - W_j^n\right)\right\}.$$

where,

$$A^+\Delta W = \sum_{\hat{\lambda}_k>0}\alpha_k\hat{\lambda}_k\hat{r}_k\ , \qquad A^-\Delta W = \sum_{\hat{\lambda}_k<0}\alpha_k\hat{\lambda}_k\hat{r}_k, \qquad \Delta W = \sum_k \alpha_k\hat{r}_k.$$

Here, $\hat{\lambda}_k$ and $\hat{r}_k$ are the eigenvalues and eigenvectors of the local linearization of the Jacobian matrix, and $\alpha_k$ are the wave strengths associated with $\Delta W$. The matrices $A^\pm$ are given by

$$A^\pm = \frac{1}{2}R\left(\Lambda \pm |\Lambda|\right)R^{-1},$$

where $R$ is the matrix of eigenvectors, and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$. If such a linearization exists, this method can be written in the form (1.6), with

$$F_{j+1/2} = F(W_j, W_{j+1}) = \frac{1}{2}\left(F(W_j) + F(W_{j+1})\right) - \sum_k |\hat{\lambda}_k|\alpha_k\,\hat{r}_k,$$

where the eigenvalues and eigenvectors correspond to the linearization $\hat{A}(W_j, W_{j+1})$.

We have developed Roe-type upwind schemes in Chapters III and IV to compute out numerical results. The scheme is extended to a second order scheme as a limited Lax-Wendroff scheme (see [LeV92] for more details).

### 1.3.2 Numerical Treatment of Source Terms

A first numerical approach for balance laws is an operator splitting algorithm, where the problem is divided into two subproblems:

(i) Update $W_j^n$ with any numerical scheme for the equation $W_t + F(W)_x = 0$ to obtain $W_j^{n,*}$.

(ii) Update $W_j^{n,*}$ using an ODE integration to account for the source terms: $W_t = S(x, W)$ to obtain $W_j^{n+1}$.

The problems modeled by the equations above often have steady-state solutions when a delicate balance between the flux gradient and the source terms is present. One is often interested in near steady-state solutions, and splitting techniques perform poorly in those cases, as they fail to maintain such a balance. In Chapters III and IV, we use the Roe-type upwind scheme proposed in [Roe87] that projects the source terms onto the characteristics fields. This is given by:

$$W_j^{n+1} = W_j^n - \frac{\Delta t}{\Delta x}\left\{A^+_{j-\frac{1}{2}}\left(W_j^n - W_{j-1}^n\right) + A^-_{j+\frac{1}{2}}\left(W_{j+1}^n - W_j^n\right)\right\}.$$

Here,

$$A^+\Delta W = \sum_{\hat{\lambda}_k>0}(\alpha_k\hat{\lambda}_k - \beta_k)\hat{r}_k\ , \qquad A^-\Delta W = \sum_{\hat{\lambda}_k<0}(\alpha_k\hat{\lambda}_k - \beta_k)\hat{r}_k$$

where $\hat{\lambda}_k$ and $\hat{r}_k$ are the eigenvalues and eigenvectors of some local linearization of the flux Jacobian, to be specified, and $\alpha_k$ and $\beta_k$ are the wave strengths associated

with $\Delta W$ and the source:

$$\Delta W = \sum_k \alpha_k \hat{r}_k, \qquad \Delta x\, S = \sum_k \beta_k \hat{r}_k.$$

## 1.4 Description of Chapters II, III and IV

In the next three chapters, we introduce models that have additional numerical difficulties and show how the existing numerical methods may fail to correctly compute the solutions. We then propose techniques that significantly improves the numerical results, and compare them to the exact solutions.

The traffic flow model with variable coefficients is a scalar conservation law in which the flux function may depend not only on the solution, but also on the spatial variables. Chapter II shows additional complications on the structure of the solution for the Riemann problem when the coefficients are discontinuous. We construct the exact solution for the Riemann problem in the present context, and show how the entropy condition selects a unique solution.

The traffic flow model with variable coefficients is a motivation to study gas flows in porous media, which is a related model in systems. This is done in Chapter III, which considers the Baer-Nunziato model for two-phase flows in porous media, with discontinuous porosity. Computing solutions of the Riemann problem rests on capturing the jump in the solution across the porosity jump. Numerical discretizations may fail to correctly capture the jump in the solution across the porosity jump, also called compaction wave, and yield incorrect solutions. Using the Riemann invariants across the compaction wave as variables will trivially recognize and respect the correct jump conditions. We have formulated the system using the Riemann Invariants across the porosity jump, and propose a hybrid algorithm that uses the Riemann Invariants formulation across the compaction wave, and the conservative formulation away from the compaction wave. This algorithm describes what formulations we use near the

compaction wave, and it can be discretized by any numerical scheme. A reduced model is studied first, followed by a generalization of the hybrid approach to the full system. Numerical results are presented. Appendix A shows how to obtain the eigenvectors for various formulations of the equations.

Chapter IV considers the shallow water equations for flows through channels with arbitrary cross-section. The system forms a hyperbolic set of balance laws. Exact steady-state solutions are available and are controlled by the relation between the bottom topography and the channel geometry. The equations can be derived by depth-width averaging, and the properties of the system are described. We construct a Roe-type upwind scheme for the system. Considerations of conservation, near steady-state accuracy, velocity regularization and positivity near dry states are discussed. The scheme obtained here is conservative and preserves steady states of rest. We show how this property enables the scheme to accurately compute near general steady-state flows. Numerical solutions are presented, illustrating the merits of the scheme for a variety of flows and demonstrating the effect of the interplay between topography and geometry on the solution. We also show the robustness of the scheme near dry states in drainage problems. The last section studies positivity in Roe-type schemes for shallow water flows in the absence of topography and geometry.

# CHAPTER II

# The Traffic Flow

The traffic flow equations models the interaction between vehicles, drivers and the infrastructure (highways, signage and traffic control devices). See for instance [Pic09] for a model on flows on complicated domains and networks. In this chapter we assume a simple relation between the velocity of the cars and the density, which greatly simplifies the model (see discussion on [LeV92]). The aim of this chapter is to show the additional complications that the Riemann problem has even in scalar equations when the flux depends discontinuously on the spatial variables. Similar complications arise in the study of gas flows in porous media with discontinuous porosity, which is studied in Chapter III.

## 2.1 Weak Solutions in the Traffic Equation

We begin by considering the scalar problem of traffic flow:

$$\begin{cases} q_t + f(q)_x = 0 \\ q(x, t = 0) = q_0(x). \end{cases} \tag{2.1}$$

Here $q(x,t)$ is the density of the cars at time $t$ and position $x$, $f(q) = uq$ is the flux, where $u(x,t) = u_m\ (1 - q(x,t))$ is the velocity of the cars, $u_m$ being the maximum

velocity. The velocity is maximum at zero density, and zero at density $q = 1$. Figure 2.1 shows how shock waves may appear in finite time along 1D road segments.

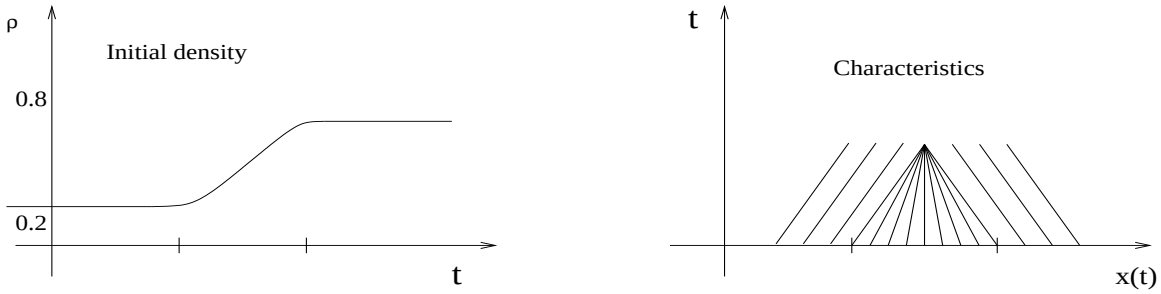



Figure 2.1: Characteristics for the traffic equation.

The solution for a Riemann problem is very simple here, and consists of just one wave. Notice that

$$[f] = [u_m(q - q^2)] = [q]\ u_m\ (1 - 2\bar{q}), \quad \text{where } \bar{q} = \frac{q_\ell + q_r}{2}.$$

If $q_\ell < q_r$, then the two states are connected through a shock wave with speed $s = u_m(1 - 2\bar{q})$, and the solution is given by

$$q(x,t) = \begin{cases} q_\ell & \text{if } x < st, \\ q_r & \text{if } x > st \end{cases}$$

Furthermore, the characteristic speed is

$$f'(q) = u_m(1 - 2q).$$

If $q_r < q_\ell$, then the two states are connected through a rarefaction, and the solution

for the Riemann problem is

$$q(x,t) = \begin{cases} q_\ell & \text{if } x/t \le f'(q_\ell) \\ \frac{-1}{2u_m}(x/t - u_m) & \text{if } f'(q_\ell) < x/t \le f'(q_r) \\ q_r & \text{if } f'(q_r) \le x/t \end{cases}$$

Figure 2.2 (left) shows the graph of $f(q)$, two states connected through a shock wave,

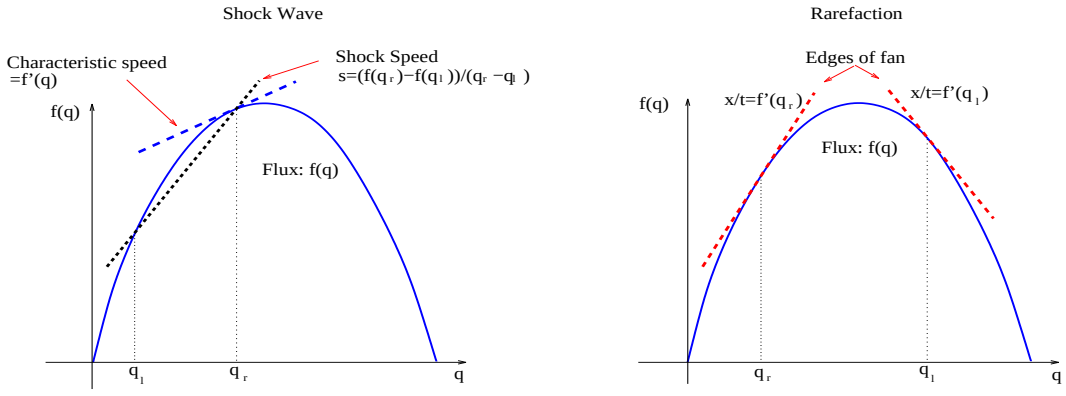



Figure 2.2: Graph of flux function with data connected through a shock wave (left) and its speed and rarefaction fan (right) with its edges.

and the speed of the shock interpreted as the slope of the line joining the two points on the graph. The sign of the slope determines the direction of the shock wave. The slope $f'(q)$ of the tangent lines are the characteristic speeds, and Figure 2.2 (right) shows the edges of a rarefaction fan as the slopes of the tangent lines at each point.

## 2.2 The Traffic Flow with Variable Coefficients

In a more realistic model, the maximum velocities may depend on the variable $x$ due to conditions of the road, speed limit, etc. We consider now the traffic equation

with variable coefficients:

$$\begin{cases} q(x,t)_t + f(q(x,t),x)_x = 0 \\ q(x,0) = q_0(x), \end{cases} \tag{2.2}$$

where now the maximum velocity $u_m = u_m(x)$ depends on $x$, $u(x,t) = u_m(x)\ (1 - q(x))$. The flux $f(q,x) = u_m(x)\ q(x,t)\ (1-q(x,t))$ now depends on both, the solution $q$ and the spatial variable $x$. Along the characteristic curves:

$$x'(t) = \frac{\partial f}{\partial q}(q(x(t),t),t) \tag{2.3}$$

we have

$$\frac{d}{dt}\left(q(x(t)t)\right) = q_x(x(t),t)\,x'(t) + q_t(x(t),t) = q_x(x(t),t)\,(x'(t) - f_q(x(t),t)) - f_x(x(t),t)$$

$$= -f_x(x(t),t)$$

Godunov's method necessitates an exact Riemann solver on each cell interface. We now illustrate how to obtain the solution for the Riemann problem in the traffic flow, which may become complicated in the presence of discontinuous flux functions. In the Riemann problem, we will also consider $u_m(x)$ piecewise constant

$$u_m(x) = \begin{cases} u_{m,\ell} & \text{if } x < 0 \\ u_{m,r} & \text{if } x > 0, \end{cases}$$

which defines two fluxes:

$$f_\ell(q) = u_{m,\ell}\ q(1-q),\ \ f_r(q) = u_{m,r}\ q(1-q).$$

The discontinuity in $u_m$ introduces a stationary wave, which we call interface. Thus,

$[f] = s[q] = 0$ across the interface, and the jump conditions reduces to

$$f_\ell(q_{m,\ell}) = f_r(q_{m,r}),$$

where $q_{m,\ell}, q_{m,r}$ are states connected through the interface. A typical solution for the Riemann problem is illustrated in Figure 2.3. The states $q_\ell, q_{m,\ell}$ $(q_{m,r}, q_r)$ are connected either through a left-going (right-going) shock wave or a left-going (right-going) rarefaction fan.

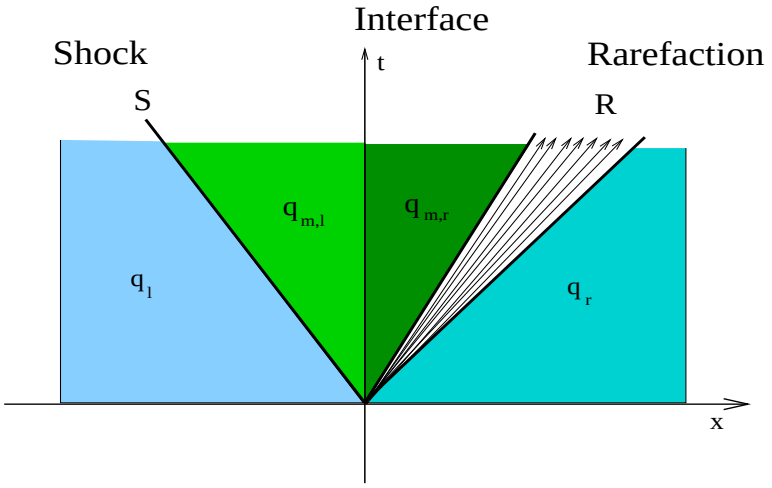



Figure 2.3: Schematic for Riemann problem in traffic equation with variable coefficients.

In Figure 2.5, two states $q_\ell, q_{m,\ell}$ in the dotted curve generate either a right-going shock wave, or a rarefaction fan with positive and negative characteristics speeds. Thus, we cannot have both states on the dotted curves. Similarly, the graph of $f_r$ cannot have both states $q_{m,r}, q_r$ on the dotted curve.

Consider maximum velocities $u_{m,r} > u_{m,\ell}$ and two states $q_\ell < q_r < 1/2$. One can find the solution to the Riemann problem if we know the wave configuration, i.e., if we know the presence or lack of shock waves or rarefaction fans on either side of the interface. See Figures 2.4 and 2.5 for the following analysis. The state $q_\ell$ cannot be connected with another state $q_{m,\ell} < 1/2$ through a shock because it would have a

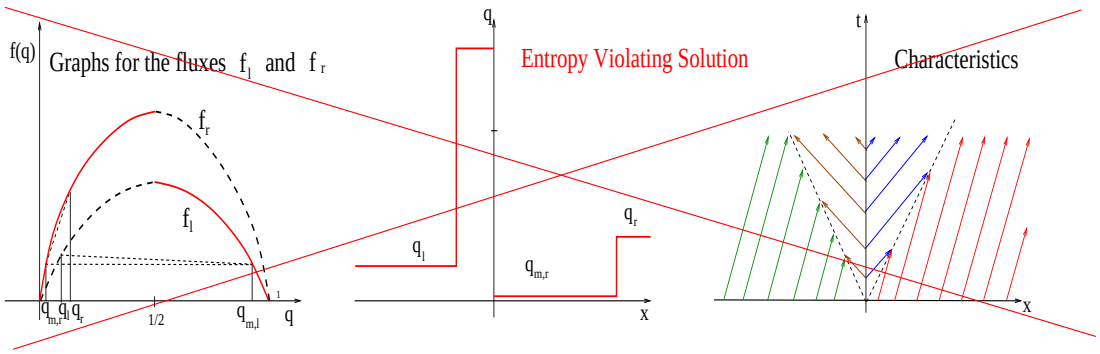



Figure 2.4: Graphs of the fluxes, weak solution and characteristics of the Riemann problem for a entropy violating solution

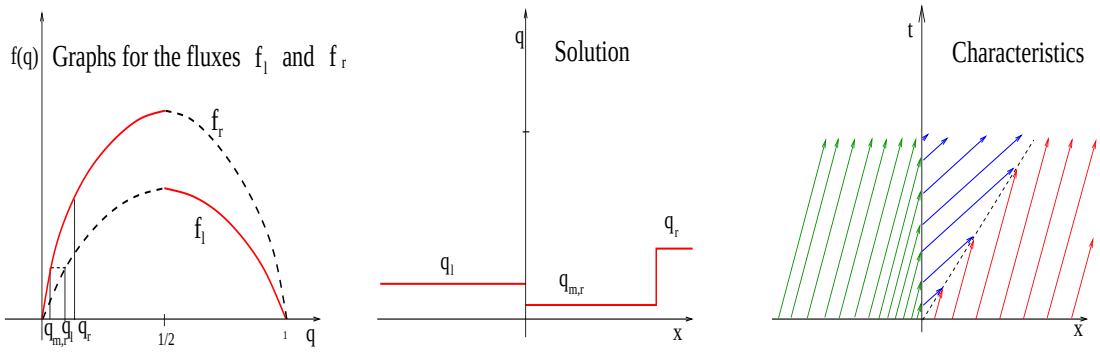



Figure 2.5: Graphs of the fluxes, weak solution and characteristics of the Riemann problem for $u_{m,r} > u_{m,\ell}, q_\ell < q_r < 1/2$.

positive speed. Likewise it cannot be connected to any $q_{m,\ell}$ through a rarefaction fan because it would have the edges to the right of the interface ($u_{m,\ell}(1-2q_\ell) > 0$). If $q_\ell$ is connected to a state $q_{m,\ell} > 1/2$, the characteristics would cross the interface from right to left. Since $q_r < 1/2$ the solution would violate the entropy conditions, as shown if Figure 2.4. Therefore, the only wave configuration that satisfies the entropy conditions is $q_\ell$ is connected directly to a state on the right of the interface. We find $q_{m,r} < 1/2$ such that $f_\ell(q_\ell) = f_r(q_{m,r})$. The states $q_{m,r}, q_r$ are connected through a shock wave. The physically relevant solution is then shown in Figure 2.5.

One can show that this weak solution is stable to perturbations by smearing out

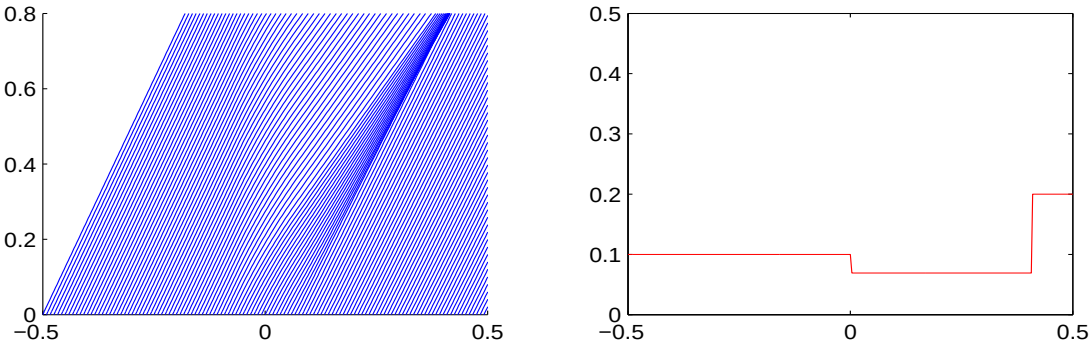



Figure 2.6: Characteristic curves to the smeared profile (left) and weak solution (right) of Figure 2.5.

the initial profile:

$$q_0(x,\epsilon) = \begin{cases} q_\ell, x \le -\epsilon \\ q_\ell + \frac{q_r - q_\ell}{2\epsilon}(x+\epsilon), -\epsilon < x \le \epsilon \\ q_r, x > \epsilon \end{cases}$$

and similarly for $u_m(x)$. Figure 2.6 (left) shows the characteristic curves in equation (2.3) for initial data $q_\ell = 0.1, q_r = 0.2, u_{m,\ell} = .5, u_{m,r} = .7, \epsilon = 0.07$. We can observe that the characteristic curves in the smeared profile are converging to the wave configuration in Figure 2.5. Figure 2.6 (right) shows the weak solution to the Riemann problem.

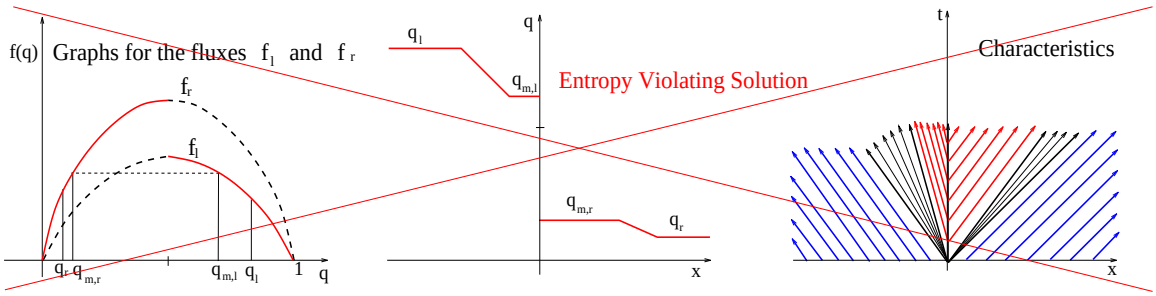



Figure 2.7: Schematic, wrong weak solution and characteristics of Riemann problem for $u_{m,r} > u_{m,\ell}$, $q_\ell > 1/2, q_r < 1/2$.

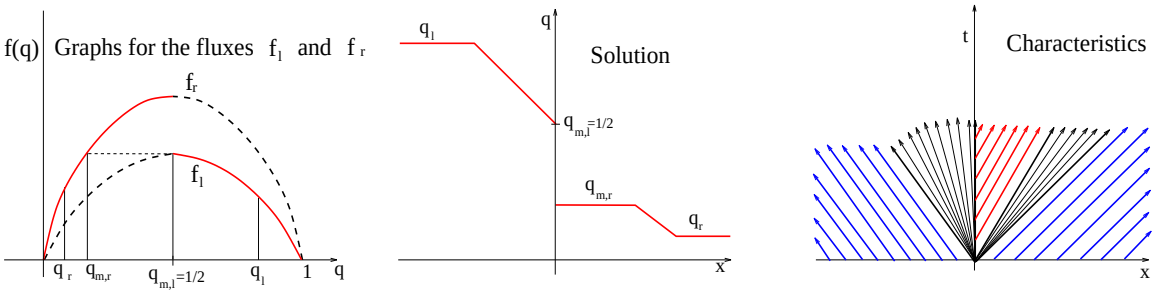



Figure 2.8: Schematic, physically relevant weak solution and characteristics of Riemann problem for $u_{m,r} > u_{m,\ell}$, $q_\ell > 1/2, q_r < 1/2$.

Now consider states $q_\ell > 1/2$, $q_r < 1/2$, and $u_{m,r} > u_{m,\ell}$. The state $q_\ell$ can only be connected to states $q_{m,\ell} \geq 1/2$, otherwise it would develop a fan crossing the interface. Similarly, $q_r$ can be connected only to states $q_{m,r} \leq 1/2$. Figure 2.7 shows one such a solution, and we observe that the characteristics are emerging from the interface, which violates the entropy condition. However, the boundary case $q_{m,\ell} = 1/2$ has a rarefaction that ends at the interface. Figure 2.8 shows the graph of each flux, the weak solution and schematic. We observe that the characteristics are converging to a rarefaction that ends at the interface, a rarefaction on the right and characteristics coming out of the interface on the right. We can show that the weak solution is stable to perturbations. Figure 2.9 shows the characteristic curves for the smeared solution with data $q_\ell = 0.8, q_r = 0.1, u_{m,\ell} = .5, u_{m,r} = .7, \epsilon = 0.03$.

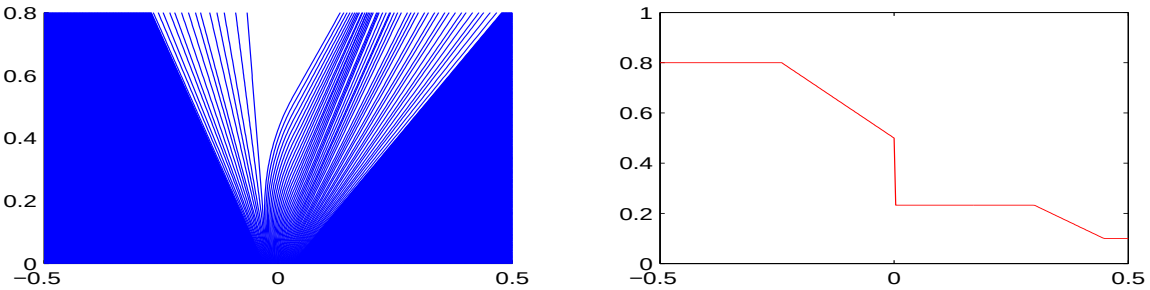



Figure 2.9: Characteristic curves to the smeared profile (left) and weak solution (right) of Figure 2.8.

## 2.3 Conclusions

The solution to the Riemann problem in the traffic flow model with constant coefficients can be easily computed and the configuration consists of only one wave. When the flux depends discontinuously on the spatial variable, the structure of the configuration consists of an additional stationary wave and one needs to determine the correct configuration for given left and right states when we solve the Riemann problem. Similar complications arise in the study of gas flows in porous media with discontinuous porosity, which is studied in the next chapter.

# CHAPTER III

# Flows in Porous Media

The generic formation of shock waves in solutions of nonlinear hyperbolic conservation laws necessitates the reliance on conservative formulations in numerical computations. Isolated waves are characterized by flow invariances which simplify the structure of the underlying flow but are not readily seen by the conservative formulations. Multimaterial flows are one such example ([AK01, Kar94, Kar96]). Naive discretizations based on conservative flow formulations fail to respect the fact that the pressure and velocity are constant across the material front, and are often plagued by pressure and other oscillations across propagating material interfaces. Recognizing and respecting material front data is automatically accomplished by writing the evolution equations directly in terms of the pressure and velocity variables, a property which is trivially inherited by *any* consistent discretization of the pressure (and velocity) evolution equation. The pressure and the velocity are the Riemann invariants across material fronts, and formulating the flow equations in terms of these variables brings out the underlying simple structure of the flow, as these variables simply 'do nothing' across the front and in doing nothing, provide the correct statement of the jump conditions. Another example is near steady-state computations of shallow water flows. Here, the mass flow rate and total energy are constant in the smooth steady-state limit. It has been recognized that incorporating these variables into the

design of the numerical scheme makes it possible for the numerical scheme to recognize and respect certain steady-state flows, which result in better accuracy for near steady-state flows ([Rus05, NXS07a]).

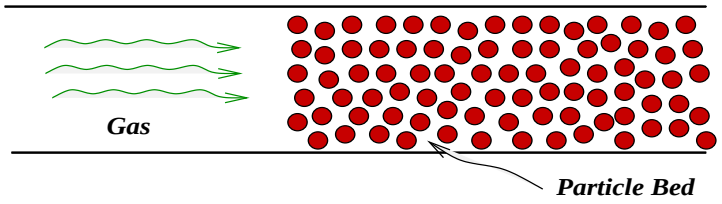



Figure 3.1: A schematic (left) and typical weak solution (right) for the Riemann problem.

In this part of the thesis, we consider the Baer-Nunziato model [BN86] for two-phase flow in porous media, with piecewise constant porosity. A schematic is given in Figure 3.1. Solutions for the Riemann problem consist of two single-phase compressible Euler sub-systems and an additional wave called a compaction wave, which carries changes in porosity and provides the means through which the two phases are coupled. Computing solutions of the Riemann problem rests on capturing the jump conditions across the porosity jump. A recent numerical study [Low05], showed that numerical methods may have difficulties to compute it correctly, and suggested that this failure is due to the inability of the conservative formulation to preserve constant gas entropy across the compaction wave front, resulting in an incorrect jump across the wave front and yielding an incorrect solution. The gas entropy $\eta$ is one of the Riemann invariants across the porosity jump. The evolution equation for the gas entropy is given by

$$\frac{\partial \eta}{\partial t} + u \frac{\partial \eta}{\partial x} = 0,$$

and has the attractive property that if the gas entropy is constant in the data, $\eta_x = 0$, it will remain constant in the solution $\eta_t = 0$, a property which is inherited by *any* numerical discretization. Although not in conservation form, this equation is *exact* across a compaction wave front, and provides a way of imposing the jump

conditions across it. We follow the rationale of the above examples and formulate the Baer-Nunziato model in terms of the Riemann invariants across the compaction wave. We examine the merit of a hybrid algorithm that uses the Riemann invariants formulation across the compaction wave, and the conservative formulation away from the compaction wave.

## 3.1 The Equations

### 3.1.1 The Baer-Nunziato Model

The Baer-Nunziato (BN) model was originally proposed in [BN86] to describe the deflagration-to-detonation transition in reactive granular material. Neglecting the terms due to combustion processes, drag and heat transfer, and focusing on the hydrodynamic part of the system, the system is given by

$$\begin{array}{lcll}
(\phi_g\rho_g)_t & + & (\phi_g\rho_g u_g)_x & = 0 \\
(\phi_g\rho_g u_g)_t & + & (\phi_g\rho_g u_g^2 + \phi_g p_g)_x & = p_g(\phi_g)_x \\
(\phi_g E_g)_t & + & (u_g(\phi_g E_g + \phi_g p_g))_x & = p_g u_s(\phi_g)_x \\
(\phi_s\rho_s)_t & + & (\phi_s\rho_s u_s)_x & = 0 \\
(\phi_s\rho_s u_s)_t & + & (\phi_s\rho_s u_s^2 + \phi_s p_s)_x & = p_g(\phi_s)_x \\
(\phi_s E_s)_t & + & (u_s(\phi_s E_s + \phi_s p_s))_x & = p_g u_s(\phi_s)_x \\
(\phi_s)_t & + & u_s(\phi_s)_x & = 0
\end{array} \tag{3.1}$$

Here the subscript $(\,)_{g,s}$ denote gas and solid phases respectively, $\rho,\ u,\ p$ and $E$ denote the density, velocity, pressure and energy of the respective phases, both assumed ideal fluids and satisfy the Equation of State

$$E = \frac{1}{2}\rho u^2 + \frac{p}{\gamma - 1}, \tag{3.2}$$

where $\gamma$ is the ratio of specific heats, and $\phi$ is the porosity, satisfying

$$\phi_g + \phi_s = 1 \ . \tag{3.3}$$

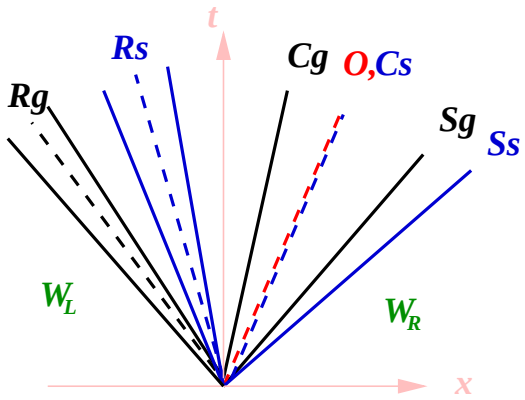



Figure 3.2: Typical solution for the Riemann problem for the BN system (3.1) .

The coupling between the phases occurs through changes in porosity. The system is inherently nonconservative due to momentum and energy exchange between the phases. The underlying conservation is revealed upon addition of the momenta (energy) equations of the individual phases. The presence of the non-conservative terms has major consequences both theoretically and computationally.

The eigenvalues and eigenvectors for this system are given by

$$\Lambda = diag\left(u_g - c_g, u_g, u_g + c_g, u_s - c_s, u_s, u_s + c_s, u_s\right), \tag{3.4}$$

$$R^C = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & \rho_g c_g^2 \\ u_g - c_g & u_g & u_g + c_g & 0 & 0 & 0 & \rho_g c_g^2 u_s \\ h_g - u_g c_g & \frac{1}{2}u_g^2 & h_g + u_g c_g & 0 & 0 & 0 & \rho_g(v_g^2 h_g - u_g c_g^2 v_g) - (v_g^2 - c_g^2)E_g \\ 0 & 0 & 0 & 1 & 1 & 1 & (v_g^2 - c_g^2)\left(\rho_s + \frac{p_g - p_s}{c_s^2}\right) \\ 0 & 0 & 0 & u_s - c_s & u_s & u_s + c_s & \left(v_g^2 - c_g^2\right)\left(\rho_s + \frac{p_g - p_s}{c_s^2}\right) u_s \\ 0 & 0 & 0 & h_s - u_s c_s & \frac{1}{2}u_s^2 & h_s + u_s c_s & (v_g^2 - c_g^2)\left(E_s + \frac{p_g - p_s}{c_s^2} h_s\right) \\ 0 & 0 & 0 & 0 & 0 & 0 & v_g^2 - c_g^2 \end{pmatrix}. \tag{3.5}$$

Here $v_g = u_g - u_s$, $h = \frac{1}{2}u^2 + \frac{c^2}{\gamma - 1}$ is the specific enthalpy and $c = \sqrt{\frac{\gamma p}{\rho}}$ the speed of sound. We observe the three familiar waves in each phase, and an additional so-called compaction wave. A solution for a typical Riemann problem is illustrated in Figure 3.2, where we have used $(\cdot)_{g,s}$ to denote the gas and solid phases respectively, and $O$ to denote the compaction wave. We note that in this model, porosity changes are carried with the solid phase, resulting in $u_s$ being a double eigenvalue, corresponding to a solid contact discontinuity *and* a compaction wave. The Riemann invariants across a compaction wave are [AW04b, BN86, Low05, EB92]

$$[u_s] = 0\ , [\eta_g] = 0\ ,\ [\eta_s] = 0\ ,\ [Q] = 0\ ,\ [P] = 0\ ,\ [H] = 0.$$

Here $\eta = \frac{p}{\rho^\gamma}$ is the entropy, $Q = \phi_g \rho_g v_g$ is the gas mass flux, $H = \frac{1}{2}v_g^2 + \frac{c_g^2}{\gamma_g - 1}$ is the gas enthalpy and $P = \phi_g p_g + \phi_s p_s + \phi_g \rho_g v_g^2$ is the sum of phase momenta fluxes as observed *in the frame of reference of the compaction wave* moving with speed $u_s$.

We note that the compaction wave is a linearly degenerate field, $\lambda = u_s$ is constant across the wave, and the solution simply translates with the constant speed $u_s$. In this case, the Hugoniot curve and the integral curve agree, and the Riemann invariants also represent conserved quantities across a porosity jump. See [EB92, LeV02] for

more details.

The system is only conditionally hyperbolic, and may lose hyperbolicity if $v_g^2 = c_g^2$. This corresponds, for example, to a flow in which the gas rarefaction straddles the compaction wave. In this case, the set of eigenvectors is no longer linearly independent. For certain data, the solution for the Riemann problem may be nonunique, and continuous dependence on coefficients may be used to select the physically relevant solution (for further discussion, see [AW04b, BN86, Low05, EB92, SWK06]).

### 3.1.2 Reduced Flow Model

If the solid phase is assumed stationary, $u_s = 0$, and incompressible, $\delta\rho_s = 0$, the porosity $\phi$ becomes a function of $x$ alone and system (3.1) reduces to

$$\begin{aligned} (\phi\rho)_t &+ (\phi\rho u)_x = 0 \\ (\phi\rho u)_t &+ (\phi\rho u^2 + \phi p)_x = p\phi_x \\ (\phi E)_t &+ (u(\phi E + \phi p))_x = 0, \end{aligned} \tag{3.6}$$

where we have removed the subscripts $()_g$ to avoid notation clutter. Note that the role of $\phi = \phi(x)$ here is that of a variable coefficient. The reduced system (3.6) is in fact the Euler equations through a channel with variable geometry. Here the porosity can be viewed as a variable cross sectional area. We consider gas flow through porous media with piece-wise constant porosity $\phi(x)$, as illustrated in Figure 3.3. This model

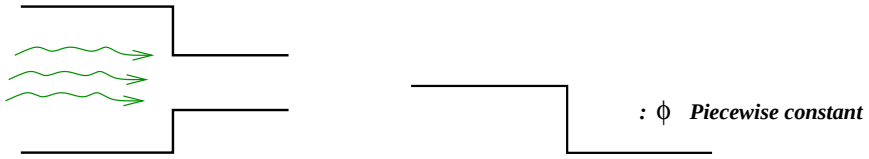



Figure 3.3: Gas flow through a channel with discontinuous area variation.

is studied in [AW04a], and was considered as a simplification of the BN system (3.1) in [Low05].

System (3.6) can be rewritten in quasilinear form using the primitive variables

$$W_t^P + A(W^P)W_x^P = 0,$$

where

$$W^P = \begin{pmatrix} \rho \\ u \\ p \\ \phi \end{pmatrix}, \quad A(W^P) = \begin{pmatrix} u & \rho & 0 & \frac{\rho u}{\phi} \\ 0 & u & \frac{1}{\rho} & 0 \\ 0 & \gamma p & u & \frac{\gamma p u}{\phi} \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.7}$$

The eigenvectors and eigenvalues of the system are

$$R^P = \begin{pmatrix} 1 & 1 & 1 & -\rho u^2 \\ -c/\rho & 0 & c/\rho & uc^2 \\ c^2 & 0 & c^2 & -\rho u^2 c^2 \\ 0 & 0 & 0 & \phi(u^2 - c^2) \end{pmatrix} \qquad \Lambda = \text{diag}\left(u - c,\ u,\ u + c,\ 0\right). \tag{3.8}$$

The first three eigenfields are analogous to single-phase Euler flow. The last eigenvec-

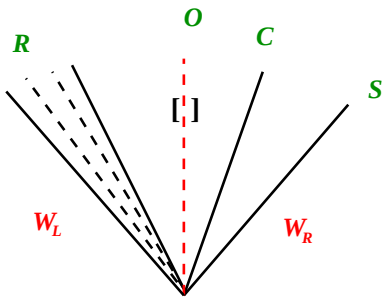



Figure 3.4: Schematic of the Riemann problem of the reduced system (3.6).

tor corresponds to the stationary porosity interface, across which the jump conditions are [AW04a, BN86, Low05, EB92]

$$[\phi\rho u] = 0,\ [\eta] = 0,\ [h] = 0. \tag{3.9}$$

We note that while the compaction wave is linearly degenerate, it does not behave like a contact discontinuity. More specifically, it can be seen from (3.9) that if $[p] = 0$ as is the case across a contact discuntinuity, then (3.9b) implies $[\rho] = 0$, (3.9c) then implies $[u] = 0$ and (3.9a) then implies that $[\phi] = 0$. It follows that if $[\phi] \neq 0$, as is the case across the porosity jump, $p,\ \rho$ and $u$ are generally discontinuous (see discussion in [Low05]).

Figure 3.4 shows a typical solution of a Riemann problem for the system (3.6) consisting of a rarefaction (R), a shock (S), a contact wave (C) and a compaction wave (O). Exact solutions to Riemann data may be obtained by embedding the interface jump conditions (3.9) in a rootfinding iteration of two single-phase Riemann solutions. Alternatively, one can start from a given state, say $W_R$, and build up a solution of a pre-determined wave structure using known relationships that hold across the single-phase wavefronts and porosity jump. We note that for certain data, the solution for the Riemann problem is not unique, and some consideration of continuous dependence on coefficients may be used to select the relevant solution (see [AW04a] for more details). In the numerical examples in this paper, we have used both approaches to generate exact solutions. When using an iterative solver for given left/right states, we have used the rootfinding algorithm proposed in [SWK06].

## 3.2 Numerical Scheme

The hybrid algorithm that we propose is formulated at the differential equations level, and may be discretized by one's favorite choice of scheme. The conservative formulation (3.1) has the form

$$W_t + F(W)_x = S,$$

where $F(W)$ is the flux function and $S$ denotes the nonconservative products. The examples in the next two sections were obtained using a wave propagation Roe-type upwind scheme [Roe81]

$$W_j^{n+1} = W_j^n - \frac{\Delta t}{\Delta x}\left\{A^+_{j-\frac{1}{2}}\left(W_j^n - W_{j-1}^n\right) + A^-_{j+\frac{1}{2}}\left(W_{j+1}^n - W_j^n\right)\right\} \tag{3.10}$$

with

$$A^+\Delta W = \sum_k \alpha_k \hat{\lambda}_k^+ \hat{r}_k \ , \ \hat{\lambda}_k^+ = \max\left(0, \ \hat{\lambda}_k\right)$$

$$A^-\Delta W = \sum_k \alpha_k \hat{\lambda}_k^- \hat{r}_k \ , \ \hat{\lambda}_k^- = \min\left(0, \ \hat{\lambda}_k\right)$$

where $\hat{r}_k$ and $\hat{\lambda}_k$ are the eigenvectors and eigenvalues of some linearization of the Jacobian matrix $A = F'(W)$, and $\alpha_k$ are the wave strengths determined by $\Delta W = \sum_k \alpha_k \hat{r}_k$ .

The Roe linearization is used where applicable, other variables are linearized by simple arithmetic averages

$$\begin{aligned}
&\hat{\rho} = \sqrt{\rho_L \rho_R} && \\
&\hat{u} = \frac{\sqrt{\rho_L}u_L + \sqrt{\rho_R}u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} && \hat{h} = \frac{\sqrt{\rho_L}h_L + \sqrt{\rho_R}h_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} \\
&\overline{c^2} = (\gamma - 1)\left(\hat{h} - \frac{1}{2}\hat{u}^2\right) && \hat{\eta} = \frac{\eta_L + \eta_R}{2}
\end{aligned}$$

The conservative formulation of the model is solved in a split-step algorithm, the upwind scheme (3.10) is applied to the Jacobian matrix $A = F'(W^C)$, the source terms in the momentum and energy equations are approximated by

$$\begin{aligned}
p\,\frac{\partial \phi}{\partial x} &\approx \frac{1}{2}\left(p_{j-\frac{1}{2}}\frac{\phi_{j+1}-\phi_{j-1}}{\Delta x} + p_{j+\frac{1}{2}}\frac{\phi_{j+1}-\phi_j}{\Delta x}\right) \\
up\,\frac{\partial \phi}{\partial x} &\approx \frac{1}{2}\left(u_{j-\frac{1}{2}}p_{j-\frac{1}{2}}\frac{\phi_j-\phi_{j-1}}{\Delta x} + u_{j+\frac{1}{2}}p_{j+\frac{1}{2}}\frac{\phi_{j+1}-\phi_j}{\Delta x}\right),
\end{aligned} \tag{3.11}$$

where $(\cdot)_{j\pm\frac{1}{2}} = \frac{1}{2}\{(\cdot)_j + (\cdot)_{j\pm1}\}$. The compaction wave equation in (3.1) is approximated by the upwind scheme. The nonconservative formulations based on the Riemann invariants are approximated by (3.10), with the matrix $A$ denoting the coefficient matrix in the quasilinear form of the equations.

## 3.3 Numerics - Reduced System

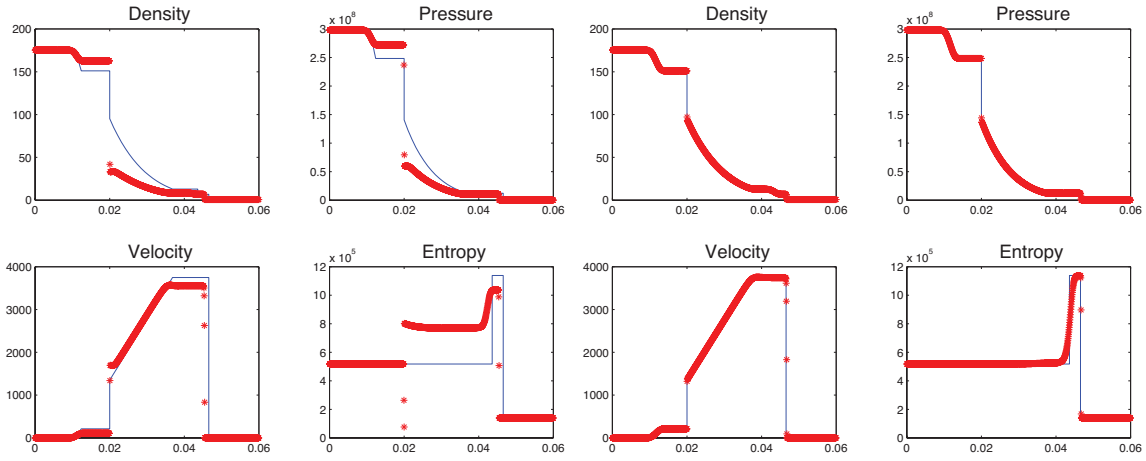



Figure 3.5: Computed (red) and exact (blue solid line) solution for initial data (3.12) by the conservative (left) and hybrid (right) formulations.

In the following numerical examples, the initial data are given in terms of $W = (\rho,\ u,\ p,\ \phi)$. Consider the Riemann problem

$$\begin{aligned} W_L &= (1.6934 \times 10^2,\ 0,\ 2.96 \times 10^8,\ 0.25) \\ W_R &= (7.6278 \times 10^{-1},\ 0,\ 1.00 \times 10^5,\ 1.00) \end{aligned} \tag{3.12}$$

corresponding to a rarefaction wave that straddles the porosity interface, followed by a right moving contact discontinuity and a shock. This problem was considered in [Low05], here $\gamma = 1.23$, $CFL = 0.8$ and the grid has 2000 points. The numerical

solution based on the conservative formulation and the exact solution are shown in Figure 3.5 (left). The computation is in noticeable error, the solution appears to jump incorrectly across the porosity change yielding an incorrect solution. As can be seen in the figure and was put forward in [Low05] by way of explanation, the gas entropy fails to remain constant across the porosity jump, resulting in incorrect jump across the wave front and yielding incorrect solution. In the next section, we propose a more direct way to impose the jump conditions across the porosity jump by formulating the equations in terms of the entropy.

### 3.3.1 A Hybrid Formulation

By differentiating $\eta = p/\rho^\gamma$ and using the primitive variable formulation (3.7) it is straightforward to derive the entropy evolution equation

$$\eta_t + u\eta_x = 0. \tag{3.13}$$

Equation (3.13) has the attractive property that when applied to data in which $\eta$ is constant, $\eta_x = 0$, $\eta$ will trivially remain constant in the solution, $\eta_t = 0$. This property is easily inherited by *any* consistent discretization of (3.13), which makes (3.13) a suitable equation to use across the porosity jump where $[\eta] = 0$. We propose that across the porosity jump, the momentum equation be replaced by the entropy evolution equation (3.13). Away from the porosity jump, the flow reduces to two uncoupled single-phase Euler flows and generically is expected to develop shock waves. In these regions, we revert back to the conservative formulation. We therefore propose the following hybrid formulation:

(i) Away from the porosity jump solve

$$\begin{aligned} (\phi\rho)_t + (\phi\rho u)_x &= 0 \\ (\phi\rho u)_t + (\phi\rho u^2 + \phi p)_x &= p\phi_x \\ (\phi E)_t + (u(\phi E + \phi p))_x &= 0 \end{aligned} \tag{3.14}$$

(ii) Across the porosity jump solve

$$\begin{aligned} (\phi\rho)_t + (\phi\rho u)_x &= 0 \\ \eta_t + u\eta_x &= 0 \\ (\phi E)_t + (u(\phi E + \phi p))_x &= 0 \end{aligned} \tag{3.15}$$

We note that the energy flux may be written as $u(\phi E + \phi p) = \phi\rho uh$. It is straightforward to see that if the data correspond to a porosity wave, hence satisfy (3.9), all spatial derivatives in (3.15) vanish. Consequently, the entropy, enthalpy and mass flow rate will remain constant in the analytic solution, as well as in any discrete approximation based on (3.15). Thus, data corresponding to compaction waves are recognized and respected by this formulation. This hybrid technique necessitates to track the compaction wave front when we switch between the conservative and RI formulation. We explain how we do it at the end of this chapter.

We also note that the conservative formulation is only used *away* from the porosity jump, hence the source term on its right hand side in fact vanishes and the system effectively reduces to two uncoupled standard Euler sub-systems.

The eigenvectors, eigenvalues and wave strengths for the conservative system are given by

$$R^C = \begin{pmatrix} 1 & 1 & 1 \\ u-c & u & u+c \\ h-uc & \frac{1}{2}u^2 & h+uc \end{pmatrix} \quad \begin{aligned} \alpha_1^C &= \frac{\Delta\tilde{p} - \tilde{\rho}c\ \Delta u}{2c^2} \\ \alpha_2^C &= \frac{c^2\ \Delta\tilde{\rho} - \Delta\tilde{p}}{c^2} \\ \alpha_3^C &= \frac{\Delta\tilde{p} + \tilde{\rho}c\ \Delta u}{2c^2} \end{aligned} \quad \Lambda = diag(u-c\ ,\ u\ ,\ u+c)$$

where we have used $(\tilde{\cdot}) = \phi(\cdot)$ to denote the respective quantities scaled by $\phi$, and $\Delta(\cdot) = ()_R - ()_L$ to denote the difference between left and right quantities.

The eigenvectors of the nonconservative system (3.15) are derived in Appendix A. Decomposing $A\Delta W^{NC} = \sum_k \alpha_k \hat{\lambda}_k \hat{r}_k = \sum_k z_k \hat{r}_k$ in terms of the Riemann invariants enables the scheme to recognize and respect porosity interface data. After some algebra we obtain the so-called f-wave strengths (see [BLMR02])

$$R^{NC} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -\frac{\gamma\eta}{\tilde{\rho}} & 0 \\ h-uc & \frac{1}{2}u^2 & h+uc \end{pmatrix}$$

$$\begin{aligned} z_1 = \alpha_1\lambda_1 &= \frac{1}{2}\Delta\left(\tilde{\rho}u\right) + \frac{\tilde{p}\left(c + (\gamma-1)u\right)}{2(\gamma-1)\ \eta c^2}\Delta\eta - \frac{\tilde{\rho}}{2c}\Delta h \\ z_2 = \alpha_2\lambda_2 &= -\frac{\tilde{p}}{\eta c^2}\ u\Delta\eta \\ z_3 = \alpha_3\lambda_3 &= \frac{1}{2}\Delta\left(\tilde{\rho}u\right) - \frac{\tilde{p}\left(c - (\gamma-1)u\right)}{2(\gamma-1)\ \eta c^2}\Delta\eta + \frac{\tilde{\rho}}{2c}\Delta h \end{aligned}$$

### 3.3.2 Numerical Results

The next example corresponds to porosity interface data, extracted from the test in Figure 3.5.

$$\begin{aligned} W_L &= (1.5113 \times 10^2,\ 2.1231 \times 10^2,\ 2.4836 \times 10^8,\ 1.00) \\ W_R &= (9.5199 \times 10^1,\ 1.3482 \times 10^3,\ 1.4067 \times 10^8,\ 0.25) \end{aligned} \tag{3.16}$$

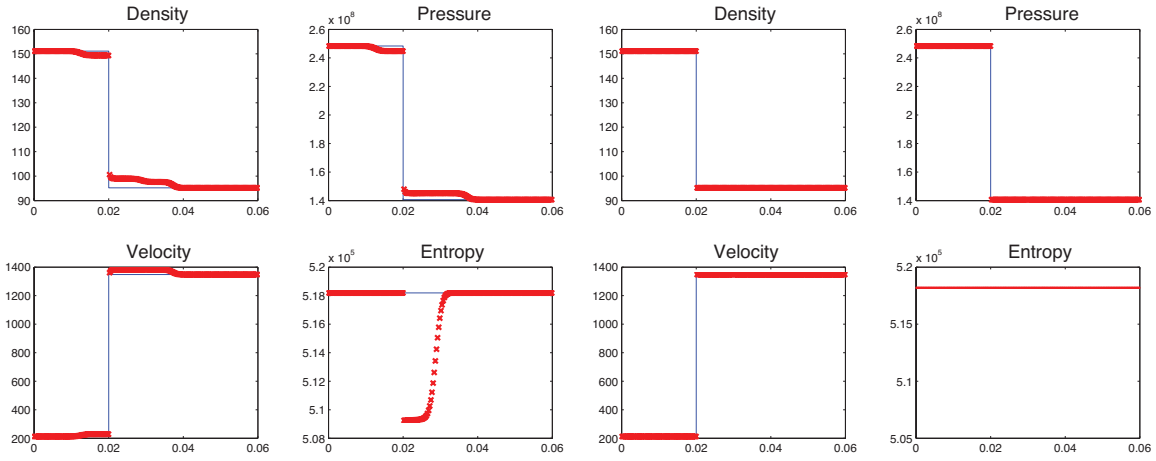



Figure 3.6: Computed (red) and exact (blue solid line) solutions corresponding to interface data (3.16): conservative (left) and hybrid (right) formulations.

Numerical and exact solution are shown in Figure 3.6, for the conservative (left) and hybrid (right) formulations. We note that the conservative formulation has difficulties keeping the entropy constant across the porosity jump, leading to erroneous waves structure. Using the entropy equation across the porosity jump and the conservative formulation everywhere else makes it possible to recognize and respect the interface data and produces a clean and error free solution.

We recompute the Riemann problem (3.12), this time with the hybrid formulation. Results are shown in Figure 3.5 (right) and are in excellent agreement with the exact solution, also shown. The hybrid formulation clearly recognizes and respects interface data, and yields the correct jump in the solution.

Figure 3.7 shows the computed solution by the hybrid scheme for two more Riemann problems. On the left, the solution for the initial data

$$\begin{aligned} W_L &= (1.0555, -1.0651, 1.5, 1.00) \\ W_R &= (1.0000, -1.0000, 1.0, 1.25) \end{aligned} \tag{3.17}$$

corresponding to a left going rarefaction and a right going shock and on the right the solution for the initial data

$$\begin{aligned} W_L &= (0.6894, -1.6941, 1.5, 1.00) \\ W_R &= (1.0000, -0.5000, 1.0, 1.25) \end{aligned} \tag{3.18}$$

producing a left and right moving rarefactions. In both examples, $\gamma = 1.4$, the CFL number is 0.8 and the grid has 400 points. Again, the jump in the solution across the

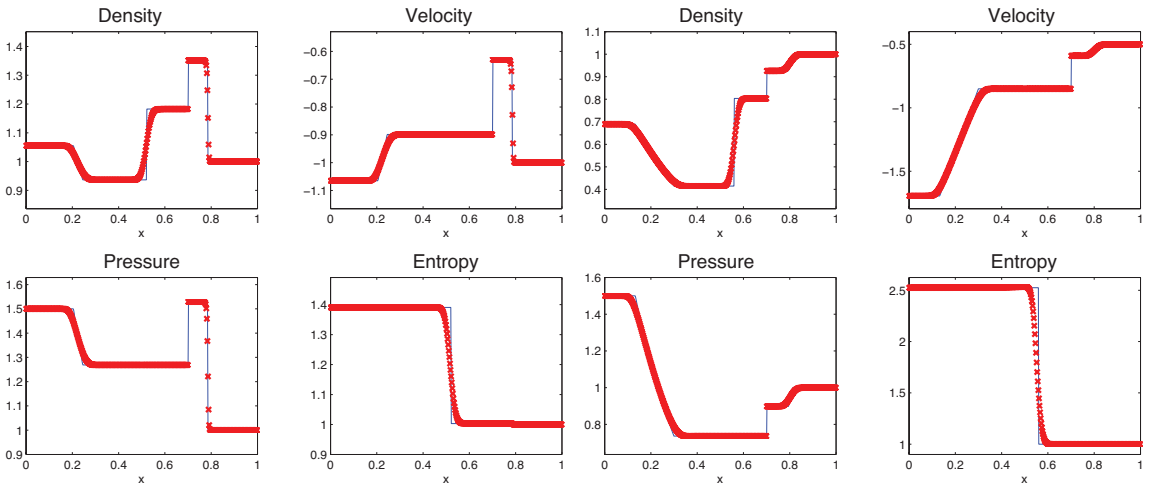



Figure 3.7: Computed (red) and exact (blue solid line) solutions corresponding to initial data (3.17) (left) and (3.18) (right).

interface is captured very well, and the computed solutions are in excellent agreement with the exact solutions, also shown.

## 3.4 Numerics - BN System

We now generalize the computational framework proposed in Section 3.3 to the full Baer-Nunziato system (3.1). We propose the following strategy:

(i) Away from the compaction wave, solve for the conservative variables

$$W^C = (\phi_g\rho_g,\ \phi_g\rho_g u_g,\ \phi_g E_g,\ \phi_s\rho_s,\ \phi_s\rho_s u_s,\ \phi_s E_s,\ \phi_s)\ ; \tag{3.19}$$

(ii) Across the compaction wave, solve the nonconservative system in terms of the Riemann invariants

$$W^{RI} = (u_s,\ \eta_g,\ \eta_s,\ Q,\ P,\ H,\ \phi_s)\ . \tag{3.20}$$

The eigenvalues of the system are given in (3.4), the eigenvectors for the conservative formulation are given in (3.5). The corresponding wave strengths are

$$\begin{aligned}
\alpha_1^C &= \frac{\Delta\tilde{p}_g - \tilde{\rho}_g c_g \Delta u_g}{2c_g^2} - \frac{p_g(c_g + (\gamma_g - 1)v_g)}{2c_g^2(v_g - c_g)}\Delta\phi_s \\
\alpha_2^C &= \frac{-\Delta\tilde{p}_g + c_g^2\Delta\tilde{\rho}_g}{c_g^2} + \frac{(\gamma_g - 1)p_g}{c_g^2}\Delta\phi_s \\
\alpha_3^C &= \frac{\Delta\tilde{p}_g + \tilde{\rho}_g c_g \Delta u_g}{2c_g^2} + \frac{p_g(c_g - (\gamma_g - 1)v_g)}{2c_g^2(v_g + c_g)}\Delta\phi_s \\
\alpha_4^C &= \frac{\Delta\tilde{p}_s - \tilde{\rho}_s c_s \Delta u_s}{2c_s^2} - \frac{p_g}{2c_s^2}\Delta\phi_s \\
\alpha_5^C &= \frac{-\Delta\tilde{p}_s + c_s^2\Delta\tilde{\rho}_s}{c_s^2} - \frac{(\gamma_s - 1)p_s}{c_s^2}\Delta\phi_s \\
\alpha_6^C &= \frac{\Delta\tilde{p}_s + \tilde{\rho}_s c_s \Delta u_s}{2c_s^2} - \frac{p_g}{2c_s^2}\Delta\phi_s \\
\alpha_7^C &= \frac{1}{c_g^2 - v_g^2}\Delta\phi_s
\end{aligned}$$

where we have used $\tilde{(\cdot)}$ to denote the respective quantities scaled by the corresponding porosities $\phi_{g,s}$. The last eigenfield, corresponding to the compaction wave, does not play a significant role in the present context since it is proposed to use this formulation only away from the compaction wave front, where the porosity does not vary and $\alpha_7$

vanishes. In that case, we also note that the rest of the wave strengths reduces to the standard expressions for the Euler system. The eigenstructure in terms of the Riemann invariants $W^{RI}$ is given by

$$R^{RI} = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & -v_g \dfrac{\tilde{p}_g}{\eta_g\, c_g^2} & 1 & \tilde{\rho}_g & 0 & 0 & \tilde{\rho}_g \\ v_g - c_g & -v_g^2 \dfrac{\tilde{p}_g}{\eta_g c_g^2} & v_g + c_g & 2\tilde{\rho}_g v_g + c_s \tilde{\rho}_s & 0 & 0 & 2\tilde{\rho}_g v_g - c_s \tilde{\rho}_s \\ -\dfrac{c_g}{\tilde{\rho}_g} & \dfrac{1}{\gamma_g - 1} \dfrac{\tilde{p}_g}{\eta_g\, \tilde{\rho}_g} & \dfrac{c_g}{\tilde{\rho}_g} & v_g & 0 & 0 & v_g \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \tag{3.21}$$

corresponding to the wave speed ordering

$$\Lambda = diag\left(u_g - c_g, u_g, u_g + c_g, u_s - c_s, u_s, u_s, u_s + c_s\right).$$

We note that the eigenfield corresponding to the double eigenvalue $u_s$ is spanned by two linearly independent eigenvectors, one corresponding to the contact discontinuity carrying only changes in solid entropy $\eta_s$, and one corresponding to the compaction wave, carrying only changes in the porosity $\phi_s$.

The wave strengths expressed in terms of $W^{RI}$ are given by

$$\alpha_1 = -\frac{\tilde{\rho}_g\left(v_g - c_g\right)}{2c_g}\Delta u_s + \frac{\tilde{p}_g\left(c_g + (\gamma_g - 1)v_g\right)}{2(\gamma_g - 1)\eta_g c_g^2}\Delta\eta_g + \frac{1}{2}\Delta Q - \frac{\tilde{\rho}_g}{2c_g}\Delta H,$$

$$\alpha_3 = \frac{\tilde{\rho}_g(v_g + c_g)}{2c_g}\Delta u_s - \frac{\tilde{p}_g\left(c_g - (\gamma_g - 1)v_g\right)}{2(\gamma_g - 1)\eta_g c_g^2}\Delta\eta_g + \frac{1}{2}\Delta Q + \frac{\tilde{\rho}_g}{2c_g}\Delta H,$$

$$\alpha_2 = \Delta\eta_g, \quad \alpha_4 = -\frac{1}{2}\Delta u_s + \frac{M}{2\tilde{\rho}_s c_s}, \quad \alpha_5 = \Delta\eta_s, \quad \alpha_6 = \Delta\phi_s, \quad \alpha_7 = -\frac{1}{2}\Delta u_s - \frac{M}{2\tilde{\rho}_s c_s},$$

where $M = \dfrac{\tilde{p}_g}{(\gamma_g - 1)\eta_g}\Delta\eta_g - v_g\Delta Q + \Delta P - \tilde{\rho}_g\Delta H$.

### 3.4.1 Numerical Examples

In all the following examples, initial data are given in terms of the primitive variables $W = (\rho_g, u_g, p_g, \rho_s, u_s, p_s, \phi_s)$.

#### Isolated Propagating Porosity Interface

The next example

$$\begin{aligned} W_L &= (1.0000, 2.0000, 0.5000, 2.0, 0.3,\ 5.000, 0.8) \\ W_R &= (0.2304, 2.4082, 0.0640, 3.0, 0.3, 13.0547, 0.3) \end{aligned} \tag{3.22}$$

corresponds to a *moving* compaction wave/solid contact, propagating with speed $u_s$. The CFL number is 0.8, the grid size is 400 points, and $\gamma_g = \gamma_s = 1.4$. It is not difficult to show that for an isolated moving compaction wave, the RI-formulation is in fact *exact*: it reduces to *linear advection* in the solid entropy and the porosity with advection speed $u_s$, and keeps the other variables constant. At the discrete level, this is reflected in the fact that for isolated interface data, all the wave strengths vanish except for $\alpha_5$ and $\alpha_6$, which reduce to $\alpha_5 = \Delta\eta_s$ and $\alpha_6 = \Delta\phi_s$. Figures 3.8 and 3.9 show the results by both the conservative and the hybrid formulation. The results clearly illustrate that the hybrid formulation recognizes and respects moving interface data. The Riemann invariants for the latter computation are shown in Figure 3.10 and confirm the ability of the formulation to treat correctly moving compaction waves.

We next consider Riemann problems corresponding to some degenerate wave configurations, here degeneracy in the sense that certain waves may be missing in the solution or that certain wave speeds may coincide. We examine the merit of the Riemann invariants based algorithm in these flow computations.

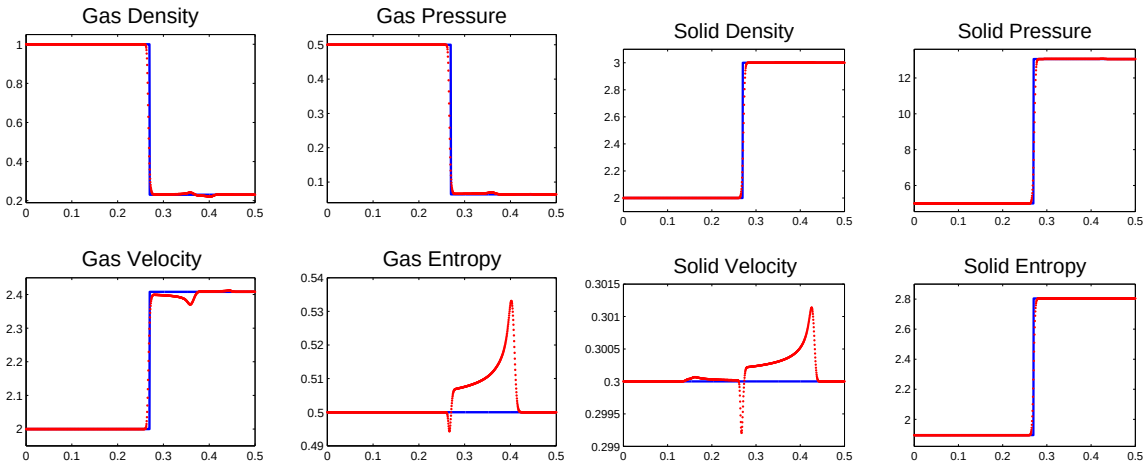



Figure 3.8: Computed (red dots) and exact (blue solid line) solutions for initial data (3.22) by the conservative formulation.

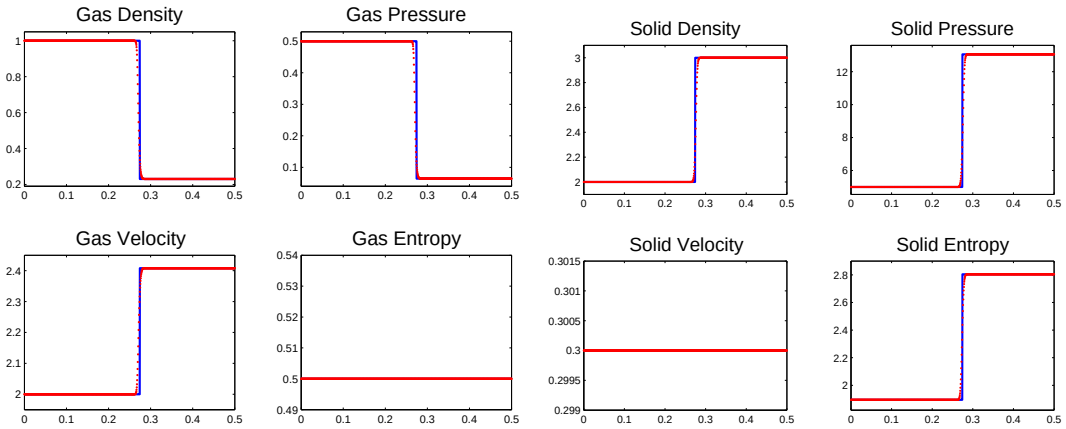



Figure 3.9: Computed (red dots) and exact (blue solid line) solutions for initial data (3.22) by the hybrid formulation.

### Compaction Wave attached to a Gas Rarefaction

This degenerate case is a borderline case where the system loses strict hyperbolicity due to coinciding eigenvectors. The compaction wave is attached to the edge of the rarefaction fan in the gas, as illustrated in Figure 3.11. The initial data are given by

$$\begin{aligned} W_L &= (0.2800, -3.3761, 0.1051, 2.0, -1.0, 1.9990, 0.5) \\ W_R &= (0.4666, -2.6668, 0.2148, 0.5, -1.0, 8.3989, 0.1) \end{aligned} \tag{3.23}$$

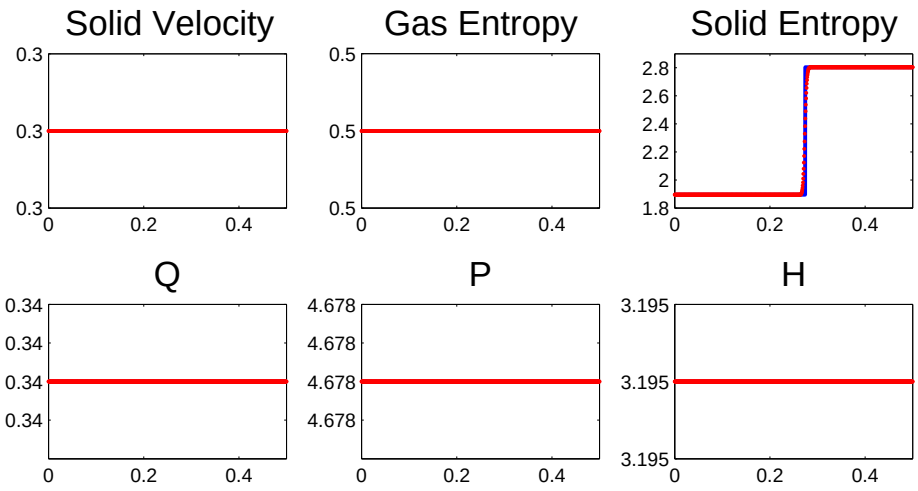



Figure 3.10: Computed (red) and exact (blue solid line) Riemann invariants for initial data (3.22) by hybrid formulation.

following [AW04b]. Results by the conservative formulation are shown in Figure 3.12, and exhibit errors near the compaction wave front. The solution shows noticeable improvement in resolving the compaction wave front in the hybrid formulation results, shown in Figure 3.13. Finally, as this example does not involve any shock waves, we have also computed solutions *entirely* by the nonconservative Riemann invariants formulation, whose eigenvectors are given by (3.21). The results are shown in Figure 3.14 and show further improvements over the hybrid formulation, indicating that some of the remaining errors in Figure 3.13 are in fact due to the conservative part of the hybridization.

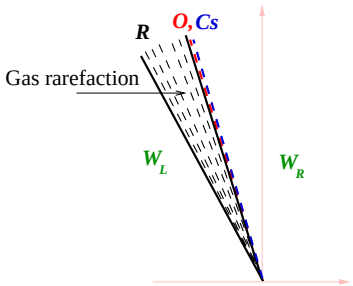



Figure 3.11: Schematic for a compaction wave right at the edge of a gas rarefaction fan.

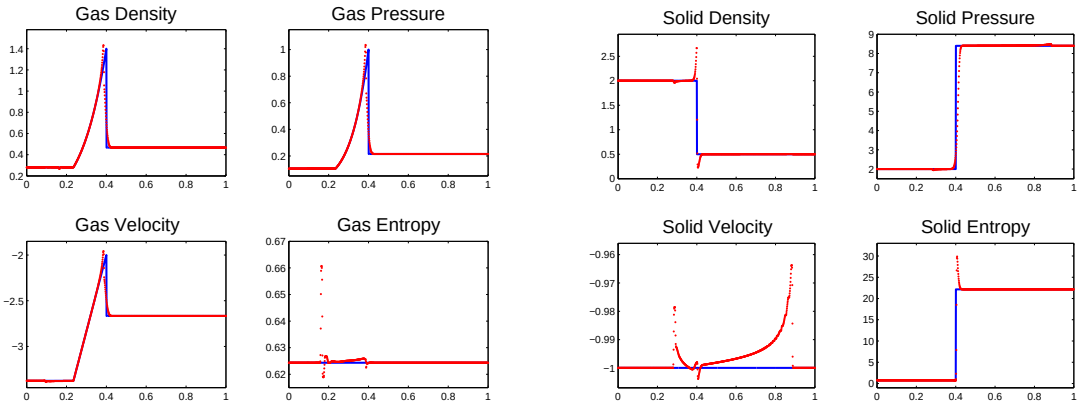



Figure 3.12: Numerical results for data (3.23) by conservative formulation.

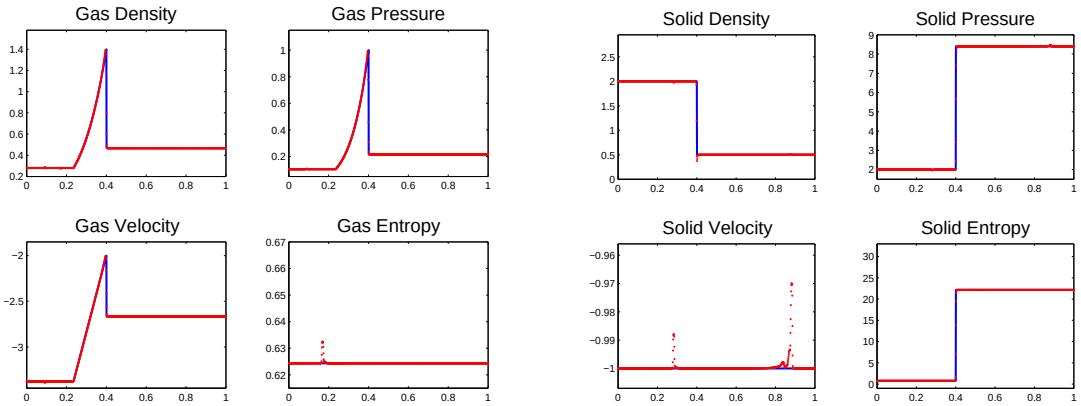



Figure 3.13: Numerical results for data (3.23) by hybrid formulation.

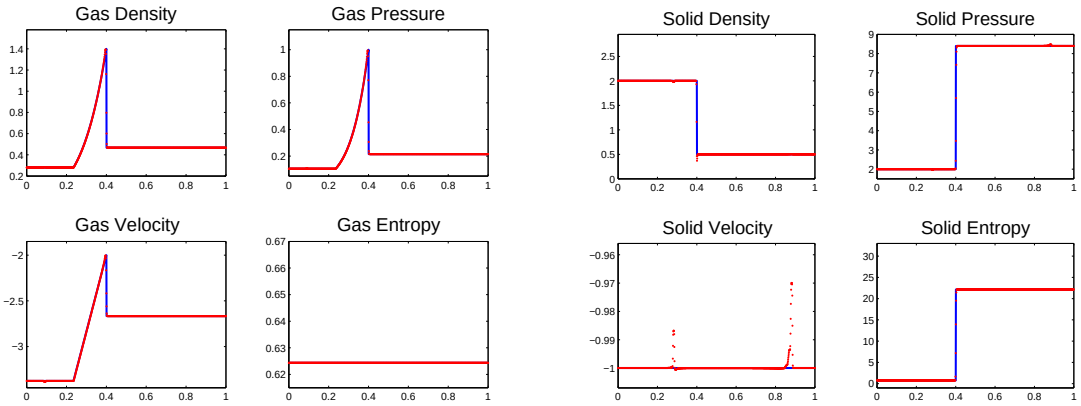



Figure 3.14: Numerical results for data (3.23) by the Riemann invariants formulation.

### Stationary solid phase

The following initial data

$$\begin{aligned} W_L &= (5.92, -0.74, 6.680, 0.550,\ 0.00,\ 0.3504, 0.3) \\ W_R &= (2.02,\ 0.86, 1.870, 1.264, -0.115, 1.1234, 0.7) \end{aligned} \tag{3.24}$$

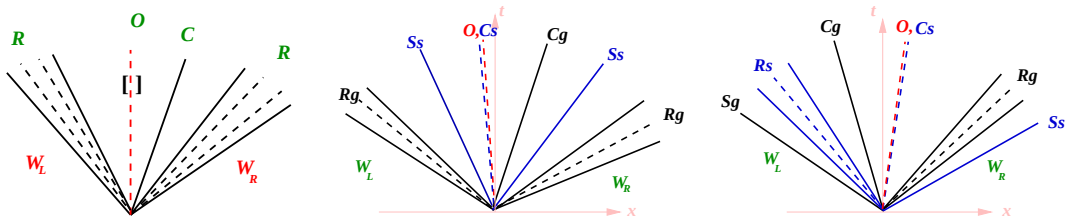



Figure 3.15: Schematic of solution corresponding to initial data (3.24) (left) , (3.25) (center) and (3.26) (right).

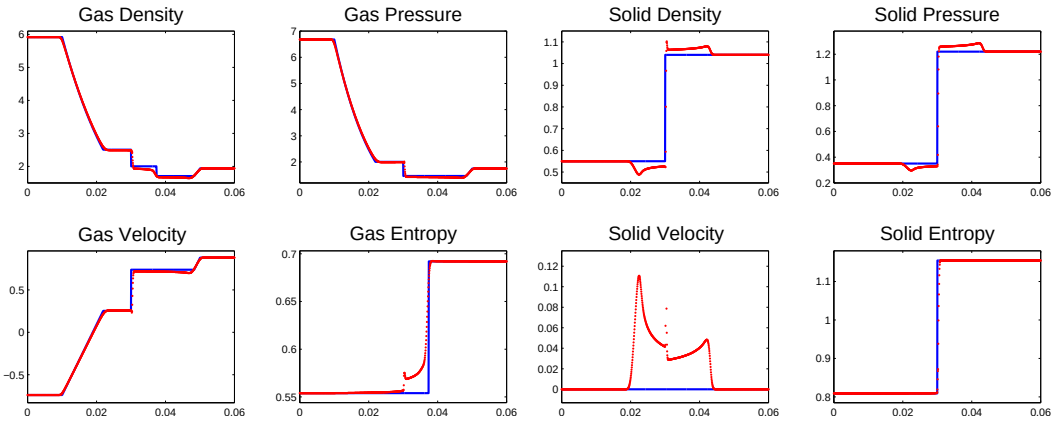



Figure 3.16: Computed (red) and exact (blue slid line) solutions corresponding to initial data (3.24) by the conservative formulation.

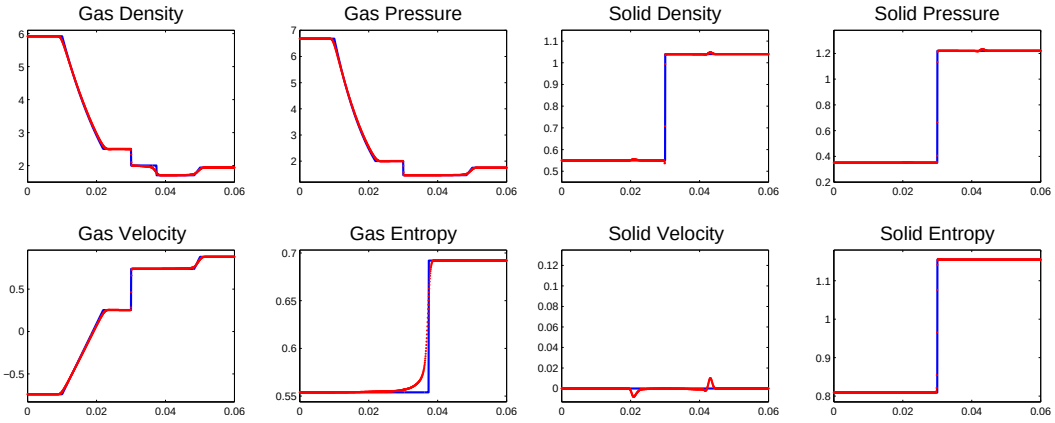



Figure 3.17: Computed (red) and exact (blue solid line) solutions corresponding to initial data (3.24) by the hybrid formulation.

is taken from [Low05] and correspond to the Riemann solution in which the gas phase moves through a *stationary* solid phase, see schematic in Figure 3.15 (left).

Figures 3.16 and 3.17 show the computed and exact solutions by the conservative and hybrid formulation respectively. The challenge in this problem is to compute

correctly the jump across the porosity interface. An additional and unrelated source of computational difficulty is the absence of certain waves from the solution, namely the acoustic waves in the solid phase. We note errors reminiscent of start-up errors, propagating at speeds corresponding to 'missing' waves, which can be observed in both sets of results. We can see that the hybrid strategy recognizes the Riemann invariant across the porosity interface, and produces a clean jump in the solution across the stationary interface at $x = 0.03$.

The following initial data

$$\begin{aligned} W_L &= (5.71, -0.75, 6.36, 0.553, -0.0553, 0.4527, 0.3) \\ W_R &= (1.94,\ 0.88, 1.75, 1.040,\ 0.00, 1.2200, 0.7) \end{aligned} \tag{3.25}$$

is taken from [Low05] and correspond to the Riemann solution schematic in Figure 3.15 (center). Figures 3.18 and 3.19 show the computed and exact solutions by the conservative formulation and the hybrid formulation respectively for data (3.25). In this case the interface is moving slightly to the left. The hybrid strategy produces clean solutions with the correct jumps at the interface.

**Full Wave Configuration**

The solution for the Riemann problem for the initial data (see [SWK06])

$$\begin{aligned} W_L &= (0.2, 0, 0.3, 1, 0, 1, 0.8) \\ W_R &= (1, 0, 1, 1, 0, 1.01, 0.3) \end{aligned} \tag{3.26}$$

is depicted by the schematic in Figure 3.15 (right). Figures 3.20 and 3.21 show the exact and computed solutions for the conservative and hybrid formulations respectively.

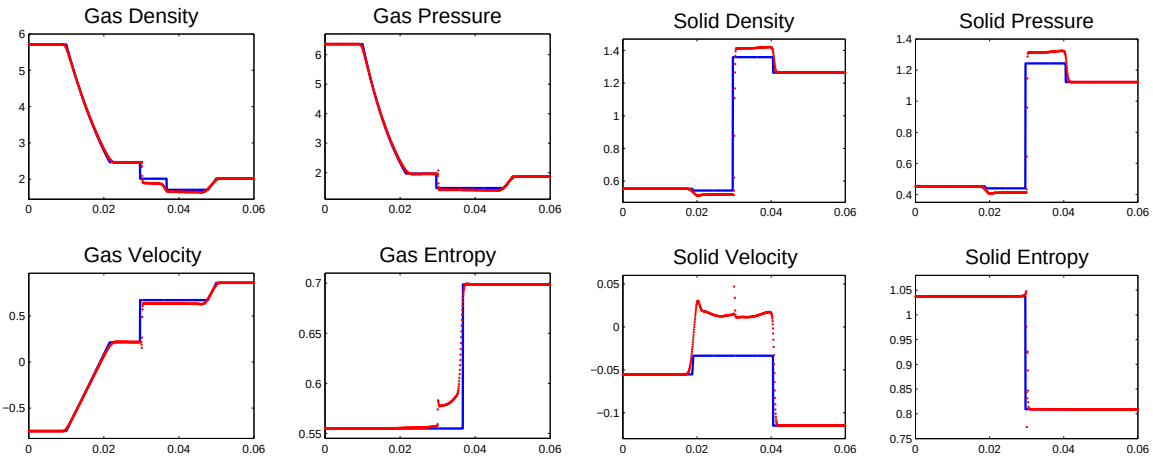



Figure 3.18: Computed (red dots) and exact (blue solid line) solutions corresponding to initial data (3.25) by the conservative formulation.

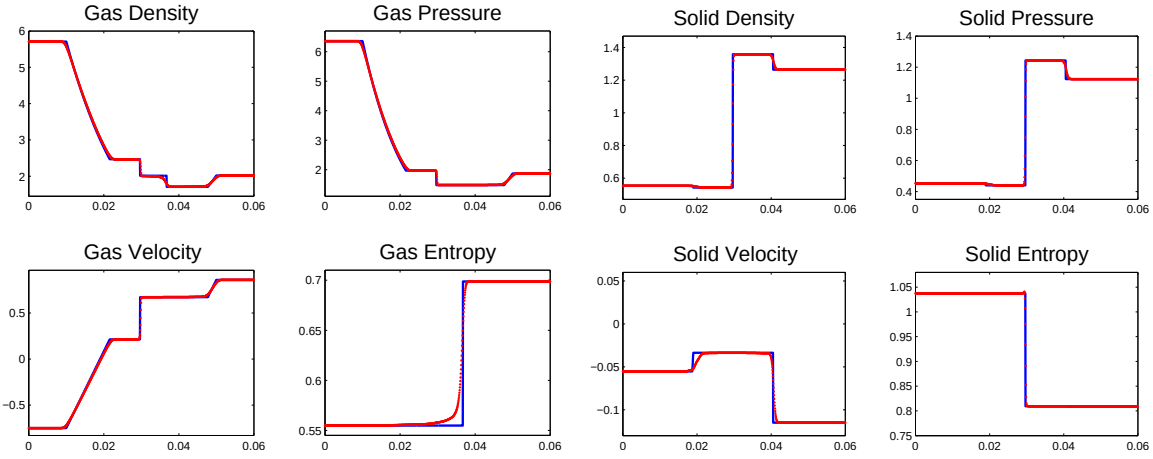



Figure 3.19: Computed (red dots) and exact (blue solid line) solutions corresponding to initial data (3.25) by the hybrid formulation.

### Shock Refraction at a Porosity Interface

In this example, a shock wave in the gas phase propagates to the right and hits a porosity interface, see schematic in Figure 3.22. Initial data are given by

$$\begin{aligned} W_L &= (2.9330, 0.4136, 2.5000, 0.5476, 0.00, 0.3280, 0.3000) \\ W_M &= (2.5154, 0.2480, 2.0155, 0.5476, 0.00, 0.3280, 0.3000) \\ W_R &= (2.0462, 0.7114, 1.5096, 1.0400, 0.00, 1.2200, 0.7000) \end{aligned} \tag{3.27}$$

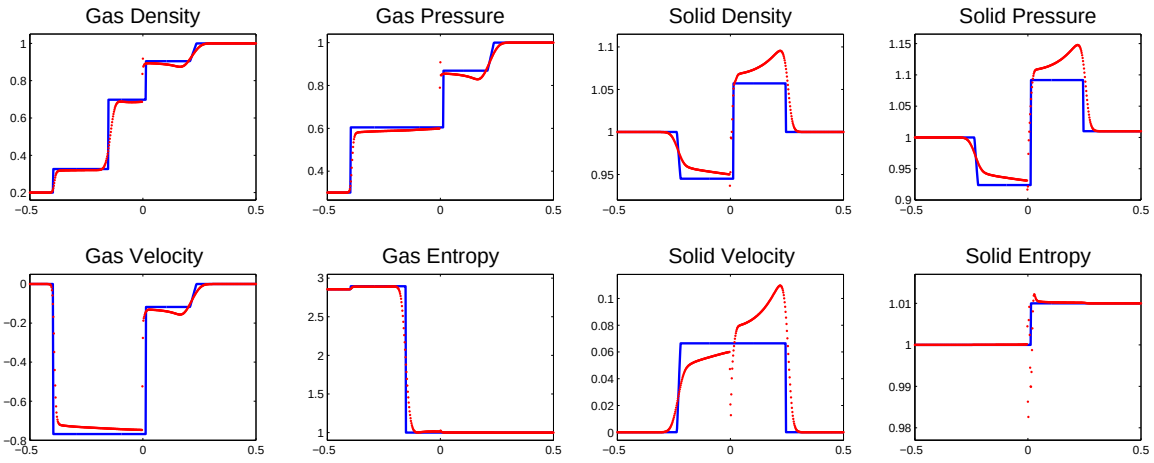



Figure 3.20: Computed (red dots) and exact (blue solid line) solutions corresponding to initial data (3.26) by the conservative formulation.

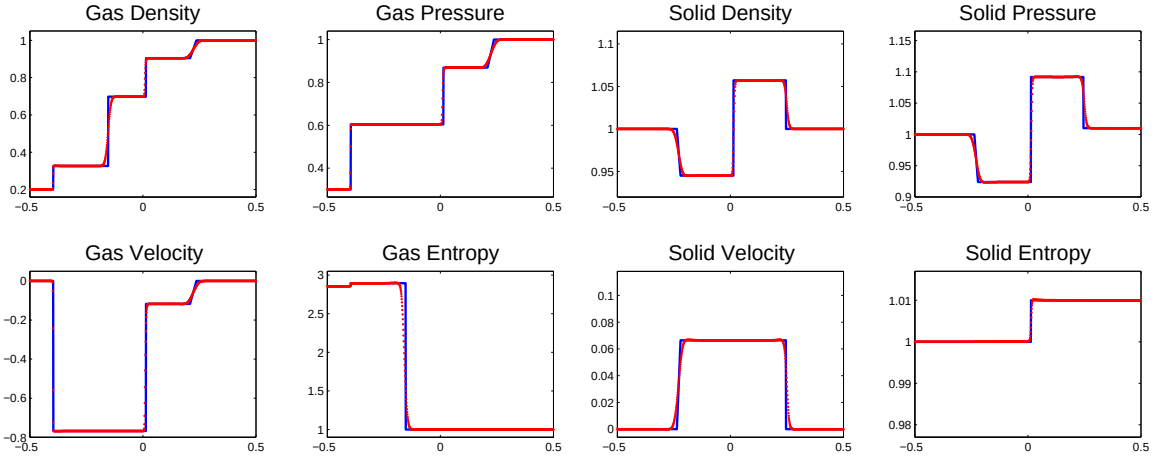



Figure 3.21: Computed (red dots) and exact (blue solid line) solutions corresponding to initial data (3.26) by the hybrid formulation.

here $W_M$ is the pre-shock state to the left of the porosity interface. Figures 3.23 and 3.24 show the computed solution by the conservative and hybrid formulations at a later time, after the shock has refracted at the porosity interface. Here, the conservative formulation is solved by the unsplit scheme with source upwinding proposed in [BLMR02], in which the source terms are projected onto the eigenvectors of the system. This solution is in visible error in computing the jump across the porosity interface, which results in inaccuracies in the other waves in the solution. The hybrid

formulation produces a cleaner more accurate approximation.

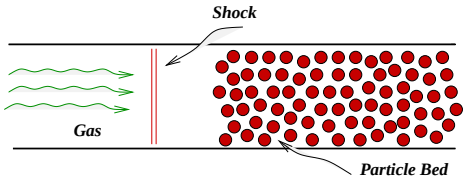



Figure 3.22: A shock hitting a porosity jump.

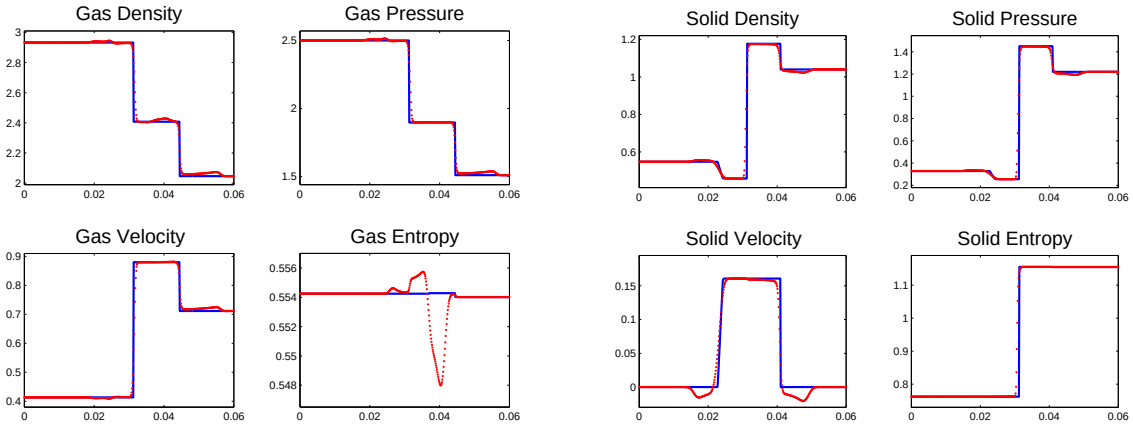



Figure 3.23: A shock refracting at a porosity interface. Conservative method.

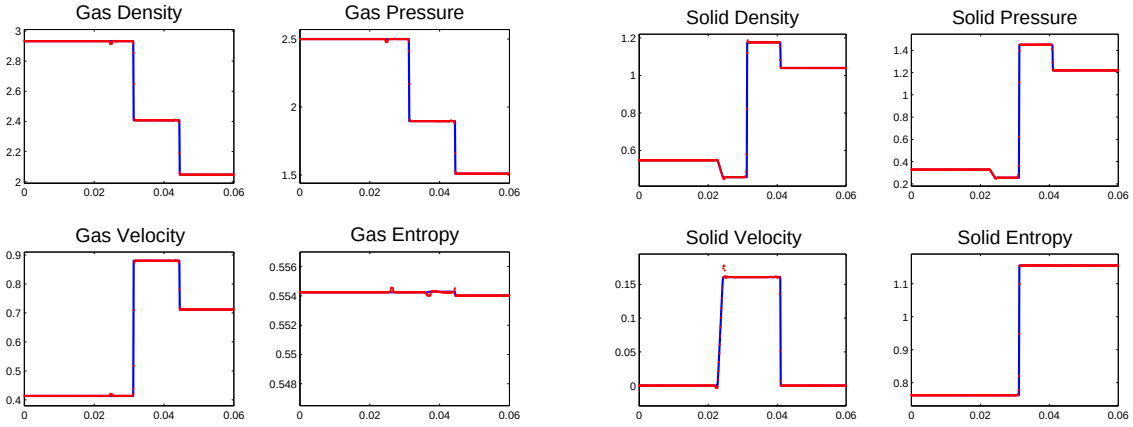



Figure 3.24: A shock refracting at a porosity interface. Hybrid method.

### Coinciding Shocks in Gas/Solid Phases

The following data

$$\begin{aligned} W_L &= (0.5806, 1.5833, 1.375, 0.2068, 1.4166, 0.0416, 0.1) \\ W_R &= (0.4890, -0.70138, 0.986, 2.2263, 0.9366, 6.0, 0.2) \end{aligned} \tag{3.28}$$

is taken from [AW04b] and correspond to the eigenstructure on Figure 3.25. Here, both phases have left moving shock waves, which are moving at the same speed. The gas phase has a right moving shock which falls within a right moving rarefaction in the solid phase. Here $\gamma_g = \gamma_s = 1.4$. Results by the hybrid formulation are shown in Figure 3.26 and are in good agreement with the exact solution.

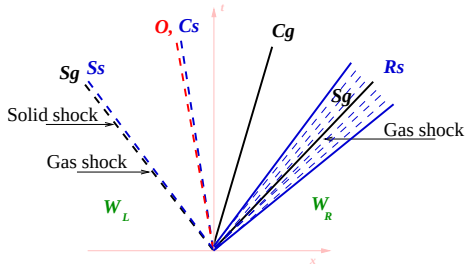



Figure 3.25: Schematic of solution corresponding to initial data (3.28), coinciding shock and rarefactions.

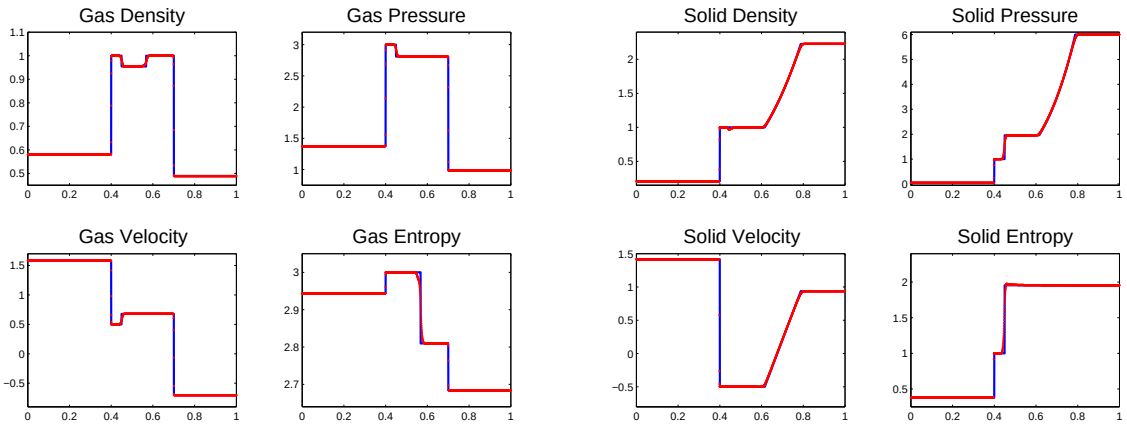



Figure 3.26: Computed ans exact solution corresponding to data (3.28). Hybrid formulation.

### Shock Wave in gas near Compaction Wave

We expect the hybrid formulation not to perform well in wave configurations where the shock in the gas phase is moving with the same or close speed to the compaction wave. With this added degeneracy of an additional wave moving at the same speed as the compaction wave, the Riemann invariants are no longer constant across the combined wave front and the advantage of the formulation is lost. Furthermore, reverting to the nonconservative Riemann invariants solver may result in more noticeable conservation errors in the vicinity of the shock. The next data taken following [AW04b] produces this type of solution (see schematic in Figure 3.27 )

The initial data are

$$\begin{aligned} W_L &= (6.3311, -0.7890, 1.3244, 2.1917, -0.9950, 3.00, 0.5) \\ W_R &= (0.4141, -0.6741, 0.0291, 0.6333, -1.1421, 2.5011, 0.1) \end{aligned} \tag{3.29}$$

Here $\gamma_g = \gamma_s = 1.4$. Numerical results by the hybrid formulation are shown in Figure 3.28. Considering the fact that the hybrid formulation is ill-suited for this type of flow, the results are actually surprisingly good.

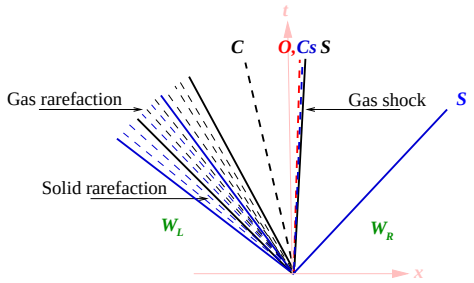



Figure 3.27: Schematic for data in (3.29), shock wave in gas near a compaction wave.

We conclude this section by making the following remarks:

*Remark I:*

In order to switch between the conservative and RI formulation, we need to track the compaction wave front. There are various ways to do that. In the computations

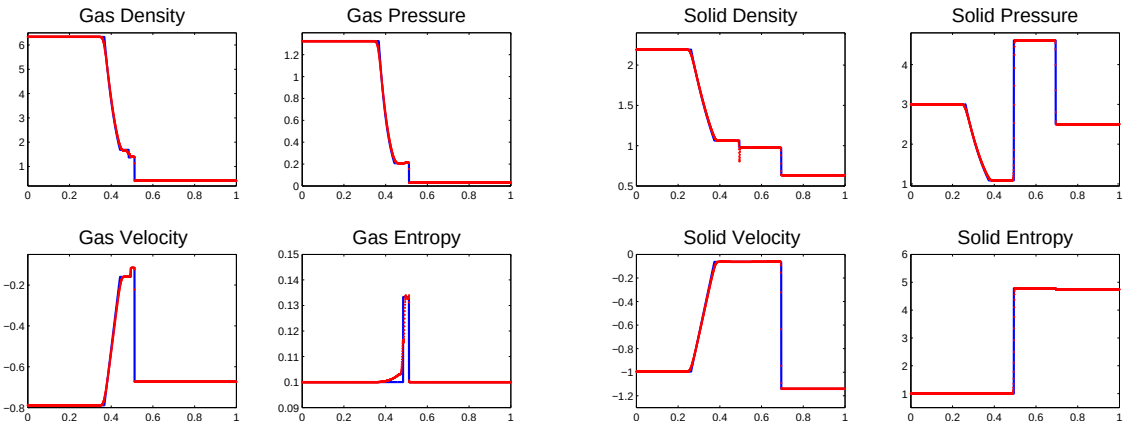



Figure 3.28: A shock wave in gas near a compaction wave. Numerical results for data (3.29) by hybrid method.

presented in this paper, we have used gradient of the porosity $\phi_s$ to locate the interface. More specifically,

$$\left| \frac{\Delta \phi_s}{(\phi_s)_R - (\phi_s)_L} \right| < \epsilon,$$

for $\epsilon$ small, typically $\epsilon = 10^{-2} - 10^{-3}$.

*Remark II:*

Converting between the conservative and primitive variables in the case of multimaterial flow involves closed form explicit formulas (see [AK01]). In the present case, recovering the conservative set of variables from the Riemann invariants involves rootfinding of $\rho_g$ in the equation

$$\frac{Q^2}{2\phi_g^2}\frac{1}{\rho_g^2} + \frac{\gamma_g}{\gamma_g - 1}\eta_g \rho_g^{\gamma_g - 1} - H = 0. \tag{3.30}$$

The equation may have more than one root, or no root at all. The latter may occur when intermediate states are generated when a large initial jump resolves itself into waves. By differentiating (3.30) with respect to $\rho_g$, it is possible to obtain a condition for the existence of the root, and identify data for which there is no root. In such cases, one may choose *not* to convert to the Riemann invariants and stick with the conservative formulation, thus avoiding altogether the need for rootfinding.

Alternatively, since the initial data itself always permits going back and forth between conservative variables and Riemann invariants, one may repeat the calculation with a smaller time step $\Delta t$, typically half of the original $\Delta t$ to allow the flow to resolve itself less abruptly.

## 3.5 Conclusions

State-of-the-art numerical schemes may fail to correctly compute the correct jump in the solution across the compaction wave, affecting the rest of the solution and producing completely erroneous numerical results. We propose a hybrid technique that uses the Riemann invariants across the compaction wave and reverts back the conservative formulation away from the porosity jump. This hybrid algorithm can be discretized by any numerical scheme. The new hybrid formulation recognizes and respects interface data, computes the correct jump in the solution across the compaction wave and produces better clean results in the solution.

In [KHD09a] we explain the hybrid algorithm and show some results. In [KHD10] we study and give details of the hybrid scheme both for the reduced model and for the full system. We show the numerical results in a variety of different challenging Riemann problems with missing waves or coinciding wave speeds.

# CHAPTER IV

# Shallow Water Flows in Channels

The shallow water equations model a variety of atmospheric and geophysical flows. They may be derived from the Euler equations by cross sectional averaging, and describe flows that are nearly horizontal. They form a set of nonlinear hyperbolic conservation laws with geometric source terms representing the topography and geometry constraining the flow. Delicate balance between the flux gradient and the geometric source terms give rise to a range of interesting flows including a variety of non-trivial equilibrium solutions. This part of the thesis is concerned with shallow water flows through channels of variable cross sectional area, where the interplay between the bottom topography and the contraction of the channel affects and controls the resulting solution.

Recent years have seen a rapidly growing interest in development of numerical methods for shallow water systems in various numerical frameworks [ABB$^+$04, BK09, Geo08, Jin01, KL02, KP07, LeV98, NPPN06, NXS07b, PS01, Roe87, Rus01], see also the recent book [Bou04] and references cited therein. Most relevant for the present work are papers involving shallow water flows in variable geometry, including [KHD09b, VC99, GNVC00] where an upwind scheme for the single layer shallow water is derived and generalized to rectangular channel flows, and [CGRGV$^+$04] where the Q-scheme [CMP01] is used to solve the two-layer shallow water system, the scheme in

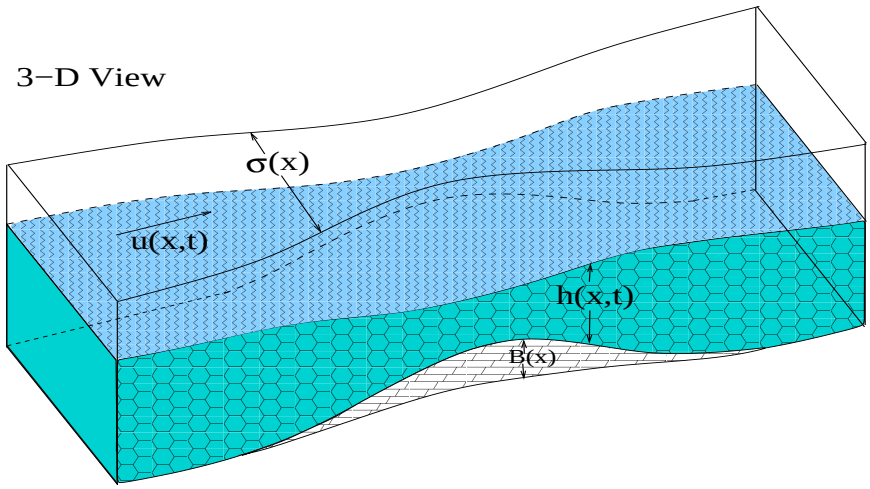



Figure 4.1: Schematic for the shallow water flow in a channel with vertical walls.

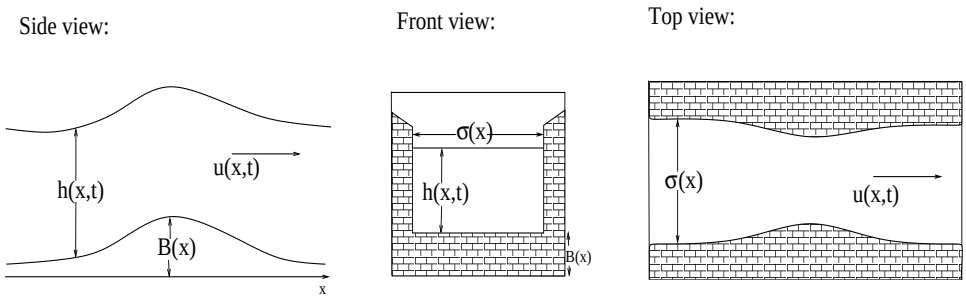



Figure 4.2: Side, front and top view of the flow.

[VS03, ČŽVS04] based on central WENO reconstruction and [BK09] using a central-upwind scheme [KP07].

Figure 4.1 shows a 3D view of a channel with vertical walls that we use to introduce notation (in the next section). Figure 4.2 shows the side (left), front (middle) and top (right) view of the flow in the channel.

## 4.1 The Model

We are interested in studying flows in more general channels, where the cross sections are not necessarily vertical walls. The geometry of the channel may become very complicated, as seen in Figure 4.3. While the model is formulated for channels

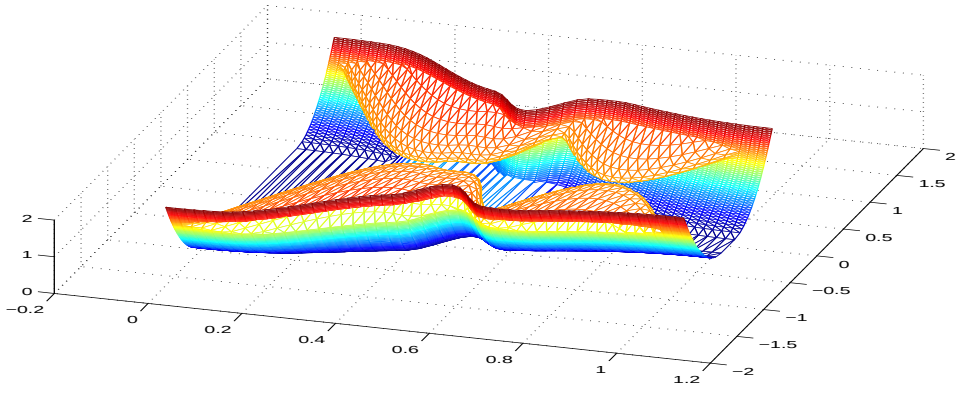



Figure 4.3: Channels with non-vertical walls.

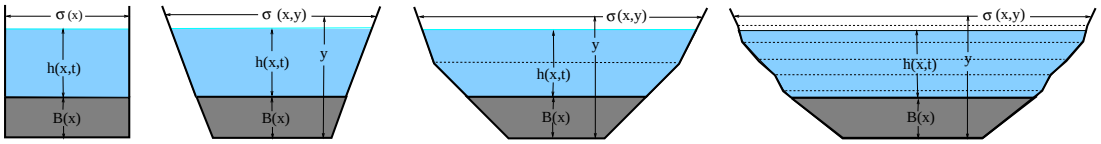



Figure 4.4: Schematic of channel cross section

with a general cross section, we approximate the channel walls by piecewise linear segments, and the cross section by piecewise trapezoids, see Figure 4.4. The work in rectangular channels appears in [KHD09b], and has been extended to channels with arbitrary cross-section in [HDK10].

The shallow water equations for flow through channels with variable cross section is given by

$$\begin{pmatrix} A \\ Au \end{pmatrix}_t + \begin{pmatrix} Au \\ Au^2 + I_1 \end{pmatrix}_x = \begin{pmatrix} 0 \\ I_2 - g\sigma_B(x)hB_x \end{pmatrix} \tag{4.1}$$

where $h$ denotes the depth of the layer, $u$ the velocity, $B(x)$ the bottom topography, $\sigma(x, y)$ the channel geometry that depends on both the horizontal $(x)$ and vertical direction $(y)$, $A = \int_B^{B+h} \sigma(x, y)dy$ is the cross-sectional wet area, $Au = Q$ is the flow rate or discharge, $\sigma_B(x) = \sigma(x, B(x))$ the bottom channel width, and $g$ the

gravitational constant. We further use $w = h + B$ to denote the total surface height,

$$I_1 = g \int_B^w (w - y)\sigma(x, y)dy = Ap,$$

where $p$ denotes the cross-sectional average of the hydrostatic pressure, and (See Figure 4.4)

$$I_2 = g \int_B^w (w - y)\sigma_x(x, y)\, dy.$$

Written in quasilinear form, the system is given by

$$\begin{pmatrix} A \\ Au \end{pmatrix}_t + \begin{pmatrix} 0 & 1 \\ c^2 - u^2 & 2u \end{pmatrix} \begin{pmatrix} A \\ Au \end{pmatrix}_x = \begin{pmatrix} 0 \\ c^2 \left(hI_3 - \sigma_B B_x\right) \end{pmatrix} \tag{4.2}$$

where $I_3 = \frac{1}{h} \int_B^w \sigma_x(x, y)\, dy$ is the averaged width variation, and the "speed of sound" is given by $c^2 = gA/\sigma_T$, where $\sigma_T(x) = \sigma(x, h + B)$ is the width of the channel at the top surface. Notice that $c^2$ reduces to the familiar expression $c^2 = gh$ for rectangular channels. The system is hyperbolic, with eigenvectors and eigenvalues

$$R = \begin{pmatrix} 1 & 1 \\ u - c & u + c \end{pmatrix} \qquad \Lambda = \begin{pmatrix} u - c & 0 \\ 0 & u + c \end{pmatrix}, \tag{4.3}$$

and is characterized by the nondimensional Froude number $F$, where $F^2 = \dfrac{u^2}{c^2}$. The flow is described as subcritical for $F^2 < 1$ and supercritical for $F^2 > 1$.

Strict hyperbolicity is lost for $h = 0$, when eigenvectors coincide, representing a so-called "dry state".

### 4.1.1 Derivation of The Model

Consider the Euler equations in three space dimensions:

$$\begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \end{pmatrix}_t + \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \end{pmatrix}_x + \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho vw \end{pmatrix}_y + \begin{pmatrix} \rho w \\ \rho uw \\ \rho vw \\ \rho w^2 + p \end{pmatrix}_z = \vec{0}$$

We will use the fundamental theorem of calculus

$$\int_a^b \partial f dr = \partial \int_a^b f dr - \left[ f\Big|_b \partial b - f\Big|_a \partial a \right]. \tag{4.4}$$

To Integrate the mass conservation equation over the wet area, we write

$$\begin{aligned}
\int_B^\omega \int_0^{\sigma(x,y)} \tfrac{\partial \rho}{\partial t} dz dy &= \tfrac{\partial}{\partial t}\left( \int_B^\omega \int_0^{\sigma(x,y)} \rho dz dy \right) - \tfrac{\partial h}{\partial t} \int_0^{\sigma(x,y)} \rho dz\Big|_{y=\omega}, \\
\int_B^\omega \int_0^{\sigma(x,y)} \tfrac{\partial}{\partial x}(\rho u) dz dy &= \tfrac{\partial}{\partial x}\left( \int_B^\omega \int_0^{\sigma(x,y)} \rho u dz dy \right) - \int_0^{\sigma(x,y)} \rho u dz\Big|_{y=\omega} \tfrac{\partial(\omega)}{\partial x} \\
&\quad + \int_0^{\sigma(x,y)} \rho u dz\Big|_{y=B} \tfrac{\partial B}{\partial x} - \int_B^\omega \rho u\Big|_{z=\sigma(x,y)} \tfrac{\partial \sigma(x,y)}{\partial x} dy, \\
\int_B^\omega \int_0^{\sigma(x,y)} \tfrac{\partial}{\partial y}(\rho v) dz dy &= \int_0^{\sigma(x,y)} \rho v dz\Big|_{y=\omega} - \int_0^{\sigma(x,y)} \rho v dz\Big|_{y=B} - \int_B^\omega \tfrac{\partial \sigma(x,y)}{\partial y} \rho v\Big|_{z=\sigma(x,y)} dy, \\
\int_B^\omega \int_0^{\sigma(x,y)} \tfrac{\partial}{\partial z}(\rho w) dz dy &= \int_B^\omega \left[ \rho w\Big|_{z=\sigma(x,y)} - \rho w\Big|_{z=0} \right] dy.
\end{aligned}$$

Adding the terms together yields

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\int_B^{\omega} \rho dz dy\right) + \frac{\partial}{\partial x}\left(\int_B^{\omega}\int_0^{\sigma(x,y)} \rho u dz dy\right) \\
&- \int_0^{\sigma(x,y)} \rho\left[\frac{\partial h}{\partial t} + u\frac{\partial(\omega)}{\partial x} - v\right]\Big|_{y=\omega} dz + \int_0^{\sigma(x,y)} \rho\left[u\frac{\partial B}{\partial x} - v\right]\Big|_{y=B} dz \\
&+ \int_B^{\omega}\left[-\rho u\Big|_{z=\sigma(x,y)}\frac{\partial\sigma(x,y)}{\partial x} - \rho v\Big|_{z=\sigma(x,y)}\frac{\partial\sigma(x,y)}{\partial y} + \rho w\Big|_{z=\sigma(x,y)} - \rho w\Big|_{z=0}\right] dy = 0.
\end{aligned} \tag{4.5}$$

Similarly, we integrate the $x$-momentum equation

$$\begin{aligned}
\int_B^{\omega}\int_0^{\sigma(x,y)} \tfrac{\partial}{\partial t}(\rho u) dz dy &= \tfrac{\partial}{\partial t}\left(\int_B^{\omega}\int_0^{\sigma(x,y)} \rho u dz dy\right) - \int_0^{\sigma(x,y)} \rho u dz\Big|_{z=\omega}\tfrac{\partial\omega}{\partial t}, \\
\int_B^{\omega}\int_0^{\sigma(x,y)} \tfrac{\partial}{\partial x}\left(\rho u^2 + p\right) dz dy &= \tfrac{\partial}{\partial x}\left(\int_B^{\omega}\int_0^{\sigma(x,y)} \left(\rho u^2 + p\right) dz dy\right) \\
-\int_0^{\sigma(x,y)} \left(\rho u^2 + p\right) dz\Big|_{z=\omega}\tfrac{\partial\omega}{\partial x} &+ \int_0^{\sigma(x,y)} \left(\rho u^2 + p\right) dz\Big|_{y=B}\tfrac{\partial B}{\partial x} \\
&\quad - \int_B^{\omega} \left(\rho u^2 + p\right)\Big|_{z=\sigma(x,y)}\tfrac{\partial\sigma(x,y)}{\partial x} dy, \\
\int_B^{\omega}\int_0^{\sigma(x,y)} \tfrac{\partial}{\partial y}(\rho u v) dz dy &= \int_0^{\sigma(x,y)} \rho u v dz\Big|_{y=\omega} - \int_0^{\sigma(x,y)} \rho u v dz\Big|_{y=B} \\
&- \int_B^{\omega} \tfrac{\sigma(x,y)}{\partial y}\rho u v\Big|_{z=\sigma(x,y)} dy, \\
\int_B^{\omega}\int_0^{\sigma(x,y)} \tfrac{\partial}{\partial z}(\rho u w) dz dy &= \int_B^{\omega}\left[\rho u w\Big|_{z=\sigma(x,y)} - \rho u w_{z=0}\right] dy.
\end{aligned}$$

Adding the terms together gives

$$\begin{aligned}
&\tfrac{\partial}{\partial t}\left(\int_B^{\omega}\int_0^{\sigma(x,y)} \rho u dz dy\right) + \tfrac{\partial}{\partial x}\left(\int_B^{\omega}\int_0^{\sigma(x,y)} \left(\rho u^2 + p\right) dz dy\right) \\
&- \int_0^{\sigma(x,y)}\left[\rho u\tfrac{\partial h}{\partial t} + (\rho u^2 + p)\tfrac{\partial(\omega)}{\partial x} - \rho u v\right]\Big|_{y=\omega} dz + \int_0^{\sigma(x,y)}\left[(\rho u^2 + p)\tfrac{\partial B}{\partial x} - \rho u v\right]\Big|_{y=B} dz \\
&+ \int_B^{\omega}\left[-(\rho u^2 + p)\Big|_{z=\sigma(x,y)}\tfrac{\partial\sigma(x,y)}{\partial x} - \rho u v\Big|_{z=\sigma(x,y)}\tfrac{\partial\sigma(x,y)}{\partial y} + \rho u w\Big|_{z=\sigma(x,y)} - \rho u w\Big|_{z=0}\right] dy = 0.
\end{aligned} \tag{4.6}$$

We impose the following physical assumptions:

(i) The pressure is given by hydrostatic balance:

$p = p(y) = \rho g(h + B - y)$, where $y$ is the height and $g$ is the gravitational constant.

(ii) $\rho = 1$. Constant density.

(iii) The flow is nearly horizontal: We replace $\overline{u^2} = \bar{u}^2$,

and the following geometric assumptions:

(iv) The top surface is a streamline: $\left[\frac{\partial h}{\partial t} + u\frac{\partial(\omega)}{\partial x} - v\right]\Big|_{y=h+B} = 0.$

(v) The flow is tangential to the bottom and channel walls, i.e., normal velocity vanishes:

$\left[u\frac{\partial B}{\partial x} - v\right]\Big|_{y=B} = 0$, $[u\frac{\partial \sigma(x,y)}{\partial x} + v\frac{\partial \sigma(x,y)}{\partial y} - w]\Big|_{z=\sigma(x,y)} = 0$ and $w\Big|_{z=0} = w\Big|_{z=\sigma(x,y)} = 0.$

The assumptions above, together with equations (4.5) and (4.6) gives the shallow water equation (4.1).

### 4.1.2 Steady-State Solutions

**Proposition IV.1.** *For smooth flows, one may express the governing equations as*

$$\begin{array}{rcrcl} A_t & + & Q_x & = & 0 \\ Q_t & + & u\,Q_x + A\,E_x & = & 0 \end{array}, \tag{4.7}$$

*where $Q = Au$ is the discharge, and $E = \frac{1}{2}u^2 + g(h + B)$ is the energy.*

*Remark* IV.2. The quantities $Q$ and $E$ are known as the equilibrium variables.

*Proof.* Using the relation

$$A_x = \int_B^\omega \sigma_x(x, y)dy + \sigma_T\omega_x - \sigma_B B_x,$$

where $\omega = h + B$ is the total height and $\sigma_T = \sigma(x, h + B)$ is the width at the top surface, we get (see equation (4.2))

$$\left(c^2 - u^2\right) + 2uQ_x - c^2 \left( \int_B^\omega \sigma_x(x,y) - \sigma_B B_x \right) = (c^2 - u^2)A_x + 2uQ_x - \frac{gA}{\sigma_T}\left(A_x - \sigma_T \omega_x\right)$$

$$= -u^2 A_x + 2uQ_x + gA\omega_x = uQ_x + u\left(uA_x + u_x A\right) - u^2 A_x + gA(h+B)_x$$

$$= uQ_x + AE_x,$$

as desired. □

As a result, smooth steady-state solutions are characterized by two constants, the flow rate $Q$, and the energy $E$

$$Q \equiv Au = Const\ , \qquad E \equiv \frac{1}{2}u^2 + g(h+B) = Const,$$

of which it is easy to recognize the steady state of rest

$$u = 0\ , \qquad h + B = Const\ .$$

Exact smooth solutions can be found by rootfinding

$$\frac{1}{2}\frac{Q^2}{A^2} + g(h+B) - E = 0\ ,\ A = A(h). \tag{4.8}$$

In the straight channel case ($\sigma = const.$), smooth steady solutions satisfy (here $()' = \frac{d}{dx}()$)

$$(F^2 - 1)h' = B'.$$

At the crest $B' = 0$, and the solution is either critical ($F^2 = 1$) or symmetric ($h' = 0$).

If the channel is rectangular with variable cross sectional width ($\sigma = \sigma(x)$), then

$$(F^2 - 1)h' = B' - \frac{h\sigma'}{\sigma}F^2 \ . \tag{4.9}$$

If the crest ($B' = 0$) and the throat ($\sigma' = 0$) occur at the same point, the right hand side of (4.9) vanishes there and the flow is either symmetric or reaches criticality at that point. Otherwise, criticality occurs where

$$B' = \frac{h}{\sigma}F^2\sigma'$$

which is somewhere between the crest and the throat. For general channels ($\sigma = \sigma(x, y)$), smooth steady-state solutions satisfy

$$(F^2 - 1)h' = \left(1 - \frac{(\sigma_T - \sigma_B)\ F^2}{\sigma_T}\right) B' - \frac{hI_3}{\sigma_T}F^2$$

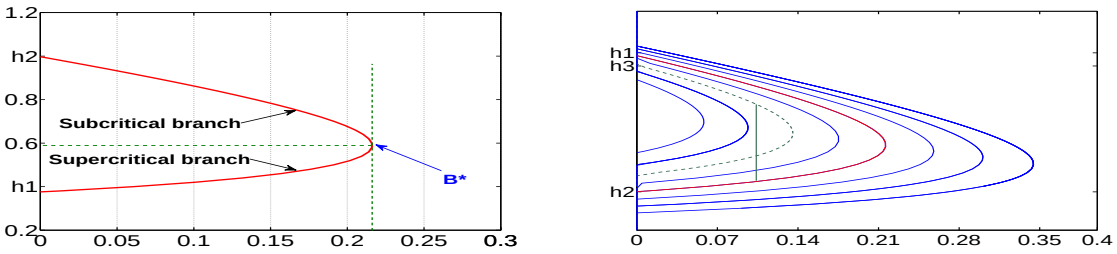



Figure 4.5: Smooth (left) and discontinuous (right) steady-state solutions, $h$ vs. $B$

In unsteady flow calculations, one typically imposes the discharge $Q$ at inflow and the depth of the layer $h$ at outflow. Depending on boundary conditions, steady states may be smooth or not smooth, may remain subcritical or become supercritical. The structure of the steady-state solution may be used to better understand the role of the boundary conditions in unsteady flows. We illustrate this in Figure 4.5. For simplicity, we consider the straight channel case, $\sigma = 1$, for given $Q$ and $E$. At a

given elevation $B(x)$, $h$ can be found from equation (4.8) by rootfinding. It is easy to compute $B(h)$ and reverse their roles to plot $h(B)$, see Figure 4.5. We observe that only values of $B$ below some $B^*$ can be supported for given $Q$ and $E$. Within that range, for any given value of $B$, there are two possible values of $h$, corresponding to subcritical flow (top branch) and to supercritical flow (bottom branch). At $B^*$, one has $B'(h) = 0$, which can be easily shown to imply $F^2 = u^2/(gh) = 1$, that is the flow is critical at $B^*$.

Consider a flow from left to right over a bump in $B$ that vanishes near the domain boundaries. A solution that starts off as subcritical at inflow, accelerates as the flow runs over the bump, and its Froude number increases. The solution moves to the right along the top subcritical branch of the curve, until it reaches the crest at some $B_{max} < B^*$, beyond which the flow starts decelerating, its Froude number decreases, and the solution moves back along the top subcritical branch, to meet the boundary condition at outflow. For the case $B_{max} = B^*$, the solution moves along the subcritical branch all the way to $B^*$, becomes critical and depending on the boundary conditions, it either returns to the subcritical branch or turns around to the supercritical branch. If it continues along the supercritical branch, its Froude number continues to increase, to meet the boundary condition at outflow. This flow accelerates smoothly from sub- to supercritical flow (similar to Laval nozzle flow in converging-diverging channels). In reference to Figure 4.5, $h_1$ and $h_2$ are the only boundary conditions at outflow that produce smooth solutions: $h_1$ produces a symmetric subcritical flow, and $h_2$ an asymmetric transcritical flow. To adjust to any other boundary condition at outflow the flow must form a discontinuity. Figure 4.5 (right) shows several curves of $h$ vs. $B$ for the same $Q$ but different values of $E$. Each one of those curves corresponds to a different smooth steady solution. A flow that starts off as subcritical at inflow along the red curve, and needs to adjust to $h_3$ at outflow, becomes critical as it reaches $B_{max}$ then supercritical along the bottom red branch. It then jumps from the red curve to

the top (subcritical) branch of the dashed green curve, a curve that corresponds to a (lower) value of $E$, and continues smoothly along this branch to meet the outflow boundary condition. The jump between curves occurs at the point where the shock jump conditions are satisfied. Symmetric and asymmetric transcritical solutions are illustrated in Figure 4.6

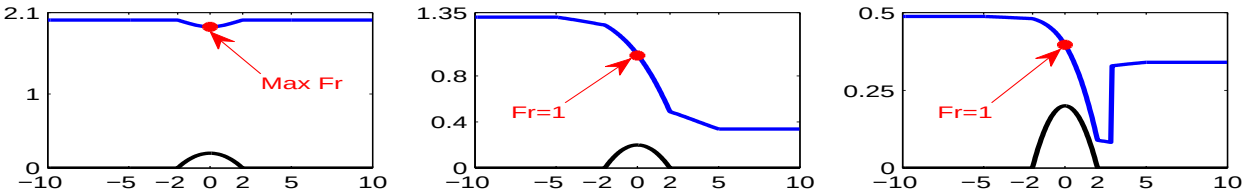



Figure 4.6: Steady flow in a channel: Symmetric subcritical (left); Asymmetric transcritical (middle); and Asymmetric transcritical with a shock (right)

### 4.1.3 The Entropy Function

**Proposition IV.3.** *The system is endowed with an entropy function*

$$\mathcal{E} = AE - I_1$$

*satisfying an entropy inequality*

$$\frac{\partial \mathcal{E}}{\partial t} + \frac{\partial}{\partial x}(QE) \leq 0$$

*Proof.* We consider the shallow water equations with small friction term

$$\begin{aligned} &A_t + Q_x = 0 \\ &Q_t + (Au^2 + I_1)_x = I_2 - g\sigma_B(x)hB_x - \epsilon u, \end{aligned}$$

for some positive $\epsilon << 1$ that represents a friction coefficient. Then

$$uQ_t + (c^2 - u^2)uA_x + 2u^2Q_x - c^2u(hI_3 - \sigma_B B_x) = -\epsilon u^2 \leq 0$$

Using the fact that $A_t = -Q_x$, one can write the inequality above as

$$\left(\frac{1}{2}Au^2\right)_t + \frac{3}{2}u^2Q_x - u^3A_x + uc^2\left(A_x - hI_3 + \sigma_B B_x\right)$$

Using that

$$A_x = \int_B^{\omega} \sigma_x(x,y)dy + \sigma_T\omega_x - \sigma_B B_x,$$

where $\omega = h + B$ is the total height and $\sigma_T = \sigma(x, h+B)$ is the width at the top surface, we get

$$\left(\frac{1}{2}Au^2\right)_t + \frac{1}{2}u^3A_x + \frac{3}{2}Au^2u_x + gQ\omega_x \leq 0$$

One can rewrite the inequality above as

$$\left(A\left(\frac{1}{2}u^2 + gB\right)\right)_t - ghQ_x + (QE)_x \leq 0$$

Notice that

$$\frac{\partial}{\partial t}I_1 = gA\ \omega_t,\ A_t = \sigma_T\ \omega_t,\ (Agh)_t = -ghQ_x - c^2Q_x.$$

This implies that

$$(AE - I_1)_t + (QE)_x \leq 0,$$

which concludes the proof. □

## 4.2 Numerical Method

We write system (4.2) as

$$W_t + A(W)W_x = S(W)$$

and use a Roe-type upwind scheme [Roe81], with upwinding of the geometric source terms as proposed in [Roe87]. The scheme has the general form

$$W_j^{n+1} = W_j^n - \frac{\Delta t}{\Delta x}\left\{A^+_{j-\frac{1}{2}}\left(W_j^n - W_{j-1}^n\right) + A^-_{j+\frac{1}{2}}\left(W_{j+1}^n - W_j^n\right)\right\}. \tag{4.10}$$

Here,

$$A^+\Delta W = \sum_{\hat{\lambda}_k>0}(\alpha_k\hat{\lambda}_k - \beta_k)\hat{r}_k\ , \qquad A^-\Delta W = \sum_{\hat{\lambda}_k<0}(\alpha_k\hat{\lambda}_k - \beta_k)\hat{r}_k \tag{4.11}$$

where $\hat{\lambda}_k$ and $\hat{r}_k$ are the eigenvalues and eigenvectors of some local linearization of the flux jacobian, to be specified, and $\alpha_k$ and $\beta_k$ are the wave strengths associated with $\Delta W$ and the source

$$\Delta W = \sum_k \alpha_k\hat{r}_k, \qquad \Delta x\, S = \sum_k \beta_k\hat{r}_k \tag{4.12}$$

given by

$$\begin{aligned}
\alpha_1 &= \frac{(\widehat{u}+\widehat{c})\Delta A - \Delta Q}{2\widehat{c}}\ , \qquad \beta_1 = \frac{\widehat{c}^2\left(\widehat{\sigma_B}\Delta B - \Delta x\ \widehat{hI_3}\right) + \widehat{G}}{2\widehat{c}} \\
\alpha_2 &= -\frac{(\widehat{u}-\widehat{c})\Delta A - \Delta Q}{2\widehat{c}}\ , \quad \beta_2 = -\frac{\widehat{c}^2\left(\widehat{\sigma_B}\Delta B - \Delta x\ \widehat{hI_3}\right) + \widehat{G}}{2\widehat{c}}.
\end{aligned} \tag{4.13}$$

where we have used $\bar{(\,)} = \frac{()_L+()_R}{2}$ to denote arithmetic averages, and $\widehat{(\,)}$ to denote other linearized quantities as defined below

$$\begin{aligned} \widehat{A} &= \tfrac{1}{2}\left[\int_{B_L}^{w_L} + \int_{B_R}^{w_R}\right] \bar{\sigma}(y)dy,\ \bar{\sigma}(y) = \tfrac{1}{2}\big(\sigma_L(y) + \sigma_R(y)\big), \\ \widehat{u} &= \frac{\sqrt{A_L}u_L + \sqrt{A_R}u_R}{\sqrt{A_L}+\sqrt{A_R}},\ \widehat{c}^2 = \tfrac{g\widehat{A}}{\widehat{\sigma_T}}, \end{aligned} \tag{4.14}$$

$\widehat{\sigma_T}$ and $\widehat{\sigma_B}$ are the linearized widths at the top/bottom surface

$$\widehat{\sigma_T}\ \Delta(h+B) = \int_{w_L}^{w_R} \bar{\sigma}(y)dy\ ,\ \widehat{\sigma_B}\ \Delta B = \int_{B_L}^{B_R} \bar{\sigma}(y)dy, \tag{4.15}$$

and

$$\Delta x\ \widehat{hI_3} = \frac{1}{2}\left[\int_{B_L}^{w_L} + \int_{B_R}^{w_R}\right] \Delta\sigma(y)dy\ ,\ \widehat{G} = g\int_{w_L}^{w_R} \overline{(w-y)\sigma(x,y)}dy, \tag{4.16}$$

The above linearization is conservative, and respects steady state of rest (see Section 4.2.1 for details), and in the case of vertical walls $\sigma(x,y) = \sigma(x)$, reduces to

$$\widehat{A} = \bar{\sigma}\ \bar{h},\ \ \widehat{\sigma_T} = \widehat{\sigma_B} = \bar{\sigma} \qquad \widehat{c}^2 = g\bar{h},\ \ \widehat{u} = \frac{\sqrt{A_L}u_L + \sqrt{A_R}u_R}{\sqrt{A_L}+\sqrt{A_R}},$$

$$\Delta x\ \widehat{hI_3} = \bar{h}\ \Delta\sigma, \qquad \widehat{G} = \tfrac{g}{4}\Delta\sigma\ (\Delta(h+B))^2$$

### 4.2.1 Derivation of the Linearization

#### 4.2.1.1 Conservation

We consider system (4.1), and relate the conservative and quasilinear forms. At the differential level we have

$$\begin{pmatrix} Au \\ Au^2 + I_1 \end{pmatrix}_x = \begin{pmatrix} 0 & 1 \\ c^2 - u^2 & 2u \end{pmatrix} \begin{pmatrix} A \\ Au \end{pmatrix}_x + \begin{pmatrix} 0 \\ c^2\left(-hI_3 + \sigma_B B_x\right) + \underline{I_2 - g\sigma_B h B_x} \end{pmatrix}, \tag{4.17}$$

where the geometric terms on the right hand side arise from careful application of the Fundamental Theorem of Calculus (FTC) to $I_1$. We further note that the underlined geometric term in (4.17) cancels out with an identical term in the geometric source in (4.1), and while it appears in the derivation of the method, it 'washes out' and ends up not playing a role in the method. We focus on the second component of this vector equation, and seek a discrete analogue. We use the following discrete version of the FTC

$$\begin{aligned} \Delta \int_{a(x)}^{b(x)} f(y,x)dy \equiv & \int_{a_R}^{b_R} f(y,x_R)dy - \int_{a_L}^{b_L} f(y,x_L)dy \\ = & \frac{1}{2}\left(\int_{a_L}^{b_L} + \int_{a_R}^{b_R}\right)\Delta f(y)dy + \int_{b_L}^{b_R} \bar{f}(y)dy - \int_{a_L}^{a_R} \bar{f}(y)dy, \end{aligned} \tag{4.18}$$

where we have used here and in what follows $\Delta(\ ) = (\ )_R - (\ )_L$, and $\bar{(\ )} = \big((\ )_R + (\ )_L\big)/2$.

The discrete version of (4.17) requires the flux difference $\Delta\Big(Au^2 + I_1\Big)$. We begin by seeking a linearization of $\widehat{u}$ for which

$$\Delta(Au^2) = 2\widehat{u}\Delta(Au) - \widehat{u}^2\Delta A$$

which is satisfied by the familiar expression

$$\widehat{u} = \frac{\sqrt{A_L}u_L + \sqrt{A_R}u_R}{\sqrt{A_L} + \sqrt{A_R}}.$$

We next apply the discrete FTC (4.18) to express $\Delta I_1$

$$\begin{aligned}
\Delta I_1 &= g\Delta \int_B^w (w-y)\sigma(x,y)dy = \frac{g}{2}\left[\int_{B_L}^{w_L} + \int_{B_R}^{w_R}\right] \Delta\Big((w-y)\sigma(x,y)\Big)dy \\
&+ g\int_{w_L}^{w_R} \overline{(w-y)\sigma(x,y)}dy - g\int_{B_L}^{B_R} \overline{(w-y)\sigma(x,y)}dy \\
&= \frac{g}{2}\left[\int_{B_L}^{w_L} + \int_{B_R}^{w_R}\right] \left(\bar{\sigma}(y)\Delta w + (\bar{w}-y)\Delta\sigma(y)\right)dy \\
&+ g\int_{w_L}^{w_R} \overline{(w-y)\sigma(x,y)}dy - g\int_{B_L}^{B_R} \overline{(w-y)\sigma(x,y)}dy \\
&= g\widehat{A}\Delta w + \Delta x\ \widehat{I_2} + \widehat{G} - g\widehat{\sigma_B h}\ \Delta B,
\end{aligned} \tag{4.19}$$

where we define

$$\widehat{A} := \tfrac{1}{2}\left[\textstyle\int_{B_L}^{w_L} + \int_{B_R}^{w_R}\right] \bar{\sigma}(y)dy,\ \Delta x\ \widehat{I_2} := \tfrac{g}{2}\left[\textstyle\int_{B_L}^{w_L} + \int_{B_R}^{w_R}\right] (\bar{w}-y)\Delta\sigma(y)dy,$$

$$\widehat{G} := g\textstyle\int_{w_L}^{w_R} \overline{(w-y)\sigma(x,y)}dy,\ g\widehat{\sigma_B h}\Delta B := g\int_{B_L}^{B_R} \overline{(w-y)\sigma(x,y)}dy.$$

In order to express $\Delta w$ in (4.19) in terms of the conserved variables, we apply the discrete FTC (4.18) to the wet area $A(x) = \int_B^w \sigma(x,y)dy$ and obtain

$$\begin{aligned}
\Delta A &= \frac{1}{2}\left[\int_{B_L}^{w_L} + \int_{B_R}^{w_R}\right] \Delta\sigma(y)dy + \int_{w_L}^{w_R} \bar{\sigma}(y)dy - \int_{B_L}^{B_R} \bar{\sigma}(y)dy \\
&= \Delta x\ \widehat{hI_3} + \widehat{\sigma_T}\Delta(h+B) - \widehat{\sigma_B}\Delta B\ ,
\end{aligned} \tag{4.20}$$

where we define

$$\Delta x\ \widehat{hI_3} := \frac{1}{2}\left[\int_{B_L}^{w_L} + \int_{B_R}^{w_R}\right]\Delta\sigma(y)dy,\ \widehat{\sigma_T}\ \Delta(h+B) := \int_{w_L}^{w_R}\bar{\sigma}(y)dy\ ,$$

$$\text{and } \widehat{\sigma_B}\ \Delta B := \int_{B_L}^{B_R}\bar{\sigma}(y)dy.$$

Rearranging (4.20) yields

$$\Delta(h+B) = \frac{1}{\widehat{\sigma_T}}\left\{\Delta A + \widehat{\sigma_B}\Delta B - \Delta x\ \widehat{hI_3}\right\}. \tag{4.21}$$

and $\Delta I_1$ becomes

$$\Delta I_1 = \frac{g\widehat{A}}{\widehat{\sigma_T}}\left[\Delta A + \widehat{\sigma_B}\Delta B - \Delta x\ \widehat{hI_3}\right] + \widehat{G} + \Delta x\ \widehat{I_2} - g\widehat{\sigma_B h}\ \Delta B.$$

This suggest to define

$$\widehat{c}^2 = \frac{g\widehat{A}}{\widehat{\sigma_T}}.$$

yielding

$$\Delta I_1 = \widehat{c}^2\Delta A + \widehat{c}^2\left[-\Delta x\ \widehat{hI_3} + \widehat{\sigma_B}\Delta B\right] + \widehat{G} + \Delta x\ \widehat{I_2} - g\widehat{\sigma_B h}\ \Delta B$$

The discrete version of (4.17) then becomes

$$\begin{aligned}\Delta\begin{pmatrix} Au \\ Au^2 + I_1\end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ \widehat{c}^2 - \widehat{u}^2 & 2\widehat{u}\end{pmatrix}\begin{pmatrix}\Delta A \\ \Delta(Au)\end{pmatrix} \\ &+ \begin{pmatrix} 0 \\ \widehat{c}^2\left(-\Delta x\ \widehat{hI_3} + \widehat{\sigma_B}\ \Delta B\right) + \widehat{G} + \Delta x\ \widehat{I_2} - g\widehat{\sigma_B h}\ \Delta B\end{pmatrix},\end{aligned} \tag{4.22}$$

and the last two terms cancel out with identical terms in the numerical approximation of the geometric source, in the same way that they do at the differential equation level in (4.1).

It is easy to verify that

$$\Delta \begin{pmatrix} Au \\ Au^2 + I_1 \end{pmatrix} - \Delta x \begin{pmatrix} 0 \\ \Delta x\ \widehat{I_2} - g\widehat{\sigma_B h}\ \Delta B \end{pmatrix} = \sum_k \Big(\hat{\lambda}_k \alpha_k - \beta_k\Big)\hat{r}_k$$

where

$$\alpha_1 = \frac{(\widehat{u} + \widehat{c})\Delta A - \Delta\big(Au\big)}{2\widehat{c}}\ , \qquad \beta_1 = \frac{\widehat{c}^2\left(\widehat{\sigma_B}\Delta B - \Delta x\ \widehat{hI_3}\right) + \widehat{G}}{2\widehat{c}}$$

$$\alpha_2 = -\frac{(\widehat{u} - \widehat{c})\Delta A - \Delta\big(Au\big)}{2\widehat{c}}\ , \quad \beta_2 = -\frac{\widehat{c}^2\left(\widehat{\sigma_B}\Delta B - \Delta x\ \widehat{hI_3}\right) + \widehat{G}}{2\widehat{c}}.$$

#### 4.2.1.2 Respecting Steady State of Rest

Consider the total fluctuation in the first wave family

$$\alpha_1\hat{\lambda}_1 - \beta_1 = (\widehat{u} - \widehat{c})\frac{(\hat{u} + \hat{c})\Delta A - \Delta\big(Au\big)}{2\widehat{c}} - \frac{\widehat{c}^2\left(\widehat{\sigma_B}\Delta B - \Delta x\widehat{hI_3}\right) + \widehat{G}}{2\hat{c}}$$

which, for steady state of rest, $u = 0,\ \Delta(h + B) = 0$, reduces to

$$\alpha_1\hat{\lambda}_1 - \beta_1 = -\frac{\widehat{c}}{2}\left(\Delta A + \widehat{\sigma_B}\Delta B - \Delta x\widehat{hI_3}\right) = -(\alpha_2\hat{\lambda}_2 - \beta_2)$$

and observe that (4.21) implies the total fluctuation vanishes, which insures that steady states of rest are recognized and respected.

### A Comment about More General Steady States

The above version of the scheme respects steady state of rest. It is generally not easy to design a scheme that respects *all* steady states, even if smooth, and often necessitates nontrivial rootfinding (see [CPMP07, NPPN06, NXS07b]). We would like to make the following observations.

As seen in equation (4.7), one may express the governing equations in terms of the equilibrium variables $Q$ and $E$ as follows (for smooth flows)

$$\begin{array}{rcrcl} (A)_t & + & Q_x & = & 0 \\ (Au)_t & + & u\,Q_x + A\,E_x & = & 0 \end{array} .$$

This formulation trivially respects *all* smooth steady states, and does not require resorting to rootfinding. Of course, system (4.7) is not in conservation form, but for smooth flows, computed solutions are conservative to the order of the numerical approximation, which can be as high as one wishes. Where (4.7) falls short is in handling discontinuous flows.

In [BLMR02], a method was proposed for conservation laws with spatially varying flux functions. The method uses the so-called f-waves, and is suitable for computations of near steady-state flows in that the entire residual

$$\Delta F - \Delta x \hat{S}$$

is decomposed onto the characteristic fields, for some linearization of the source $\hat{S}$. If a source linearization can be found so that the steady state is recognized on the discrete level, the residual is identically zero and so are its projections onto the characteristic fields.

For rectangular channels, we write the fluctuations in terms of the equilibrium variables, $\Delta Q$ and $\Delta E$. Using repeatedly the identity $\Delta(AB) = \bar{A}\Delta B + \bar{B}\Delta A$ where

$(\bar{\cdot})$ indicates arithmetic average, we obtain

$$\begin{aligned}
\Delta\left(\sigma h u^2 + \tfrac{g}{2}\sigma h^2\right) &\equiv \bar{u}\Delta Q + \overline{\sigma h u}\Delta u + \tfrac{g}{2}\bar{\sigma}\Delta h^2 + \tfrac{g}{2}\,\overline{h^2}\,\Delta\sigma \\
&= \bar{u}\Delta Q + \bar{Q}\Delta u + g\bar{\sigma}\,\bar{h}\Delta h + \tfrac{g}{2}\,\overline{h^2}\,\Delta\sigma \\
&= \bar{u}\Delta Q + \left(\bar{Q} - \bar{\sigma}\,\bar{h}\,\bar{u}\right)\Delta u + \bar{\sigma}\,\bar{h}\,\bar{u}\Delta u + g\bar{\sigma}\,\bar{h}\Delta h + \tfrac{g}{2}\,\overline{h^2}\,\Delta\sigma \\
&= \bar{u}\Delta Q + \bar{\sigma}\,\bar{h}\,\Delta\left(\tfrac{1}{2}u^2 + gh\right) + \tfrac{g}{2}\,\overline{h^2}\,\Delta\sigma + \left(\bar{Q} - \bar{\sigma}\,\bar{h}\,\bar{u}\right)\Delta u\ .
\end{aligned}$$

Using

$$\Delta x\big(\widehat{S}\big)^{(2)} = -g\widehat{\sigma h}\Delta B + \frac{1}{2}g\widehat{h^2}\Delta\sigma := -g\bar{\sigma}\bar{h}\Delta B + \frac{1}{2}\,g\overline{h^2}\Delta\sigma$$

we obtain the discrete identity

$$\begin{aligned}
\left(\Delta F - \Delta x\hat{S}\right)^{(2)} &= \Delta\left(\sigma h u^2 + \tfrac{g}{2}\sigma h^2\right) + g\bar{\sigma}\bar{h}\Delta B - \tfrac{1}{2}\,g\overline{h^2}\Delta\sigma \\
&= \bar{u}\Delta Q + \bar{\sigma}\,\bar{h}\,\Delta E + \underbrace{(\bar{Q} - \bar{\sigma}\,\bar{h}\,\bar{u})\Delta u}_{\text{Conservation Correction Term}}\ .
\end{aligned}$$

We make the following comments:

(i) It is easy to see that $\Delta F - \Delta x\hat{S} \equiv 0$ for steady state of rest.

(ii) For more general smooth steady states, $\left(\Delta F - \Delta x\hat{S}\right)^{(2)} = (\bar{Q} - \bar{\sigma}\,\bar{h}\,\bar{u})\Delta u \neq 0$. We observe that

$$\left(\bar{Q} - \bar{\sigma}\,\bar{h}\,\bar{u}\right)\Delta u = \left\{\frac{\Delta u\Delta\left(\sigma h\right) + \Delta h\Delta\left(\sigma u\right) + \Delta\sigma\Delta\left(hu\right)}{8}\right\}\Delta u = O(\Delta x)^3$$

is small for smooth flow, so the residual that is being decomposed is very small, which may explain good behaviour of the method for general (smooth) steady states.

(iii) The wave strengths expressed in terms of $\Delta Q$ and $\Delta E$ are given by

$$Z_1 = \tfrac{1}{2}\Delta Q - \tfrac{\bar{\sigma}}{2g}\ \bar{c}\ \Delta E - \tfrac{\bar{Q}-\bar{\sigma}\ \bar{h}\ \bar{u}}{2\bar{c}}\Delta u$$
$$Z_2 = \tfrac{1}{2}\Delta Q + \tfrac{\bar{\sigma}}{2g}\ \bar{c}\ \Delta E + \tfrac{\bar{Q}-\bar{\sigma}\ \bar{h}\ \bar{u}}{2\bar{c}}\Delta u. \tag{4.23}$$

We have used both versions of the upwind scheme (4.10)-(4.16) and (4.23) in the computations of the next section. In general, we have found them to give very similar results.

**Entropy Fix**

It is known that Roe-type schemes require an entropy fix [LeV92]. We have implemented an entropy fix following [HH83], as discussed in [LeV92]. It is our experience that implementing an entropy fix is crucial for computations of drainage problems, where the flow develops centered rarefactions in regions of very thin layers (see Section 4.3).

Straight channels with flat topography corresponds to the regular shallow water equations with no source terms. In the present context, a numerical scheme is called positivity preserving if the numerical solution given by the scheme has positive layer's depth in all the cells provided that the depth is positive in the initial data. Following [Dub99], we developed a positivity preserving numerical scheme in Appendix 4.4 using positive Roe matrices.

**Velocity Regularization and Positivity**

When $h << 1$ is very small, for example in drainage problems, recovering the velocity $u$ in the standard way $u = Q/A$ becomes inaccurate and may cause instabilities.

This is often remedied by regularizing the velocity, for example

$$u = \frac{2Q}{A + max(A, \epsilon)}. \tag{4.24}$$

Typically, $\epsilon = O(10^{-5})$. Other formulas may be used [BK09, KP07]. For drainage problems, we have also used

$$u = sign(Au) \left\{ max \left( 2 \left( E_{StSt} - g(h + B) \right), 0 \right) \right\}^{1/2} \quad \text{for } h < \epsilon \tag{4.25}$$

which replaces $u$ in very thin layers by a value consistent with the steady-state solution towards which the solution is converging (e.g. $E_{StSt}$ is the final steady-state energy for the drainage problem). This formula often gives very smooth and clean convergence (see numerical results in Section 4.3). The current version of the scheme is not positive, but has proven to be extremely robust in maintaining positivity, for example in drainage problems (see Section 4.3).

## 4.3 Numerical Results

The numerical scheme is formulated in terms of integrals over general channel cross sections. In this section, we present results for various shallow flows through channels with various geometries, including rectangular, trapezoidal and general (multi-trapezoidal) cross sectional areas. Unless otherwise stated, the examples use $g = 9.81$, a grid of 200 points, and a CFL number of 0.9.

### 4.3.1 Rectangular Channels

The examples in this subsection involve channels with rectangular variable (in $x$) cross sections.

### Small Perturbation to Steady State

In the first example, the initial data is a small perturbation to steady state of rest. By design, the scheme (4.10)-(4.16) preserves steady state of rest, and the propagation of small perturbations thereof is computed very accurately. Computed solutions are shown in Figure 4.7, for centered and off-centered channel contractions. Once the small perturbation leaves the computational domain, the unperturbed steady state is recovered.

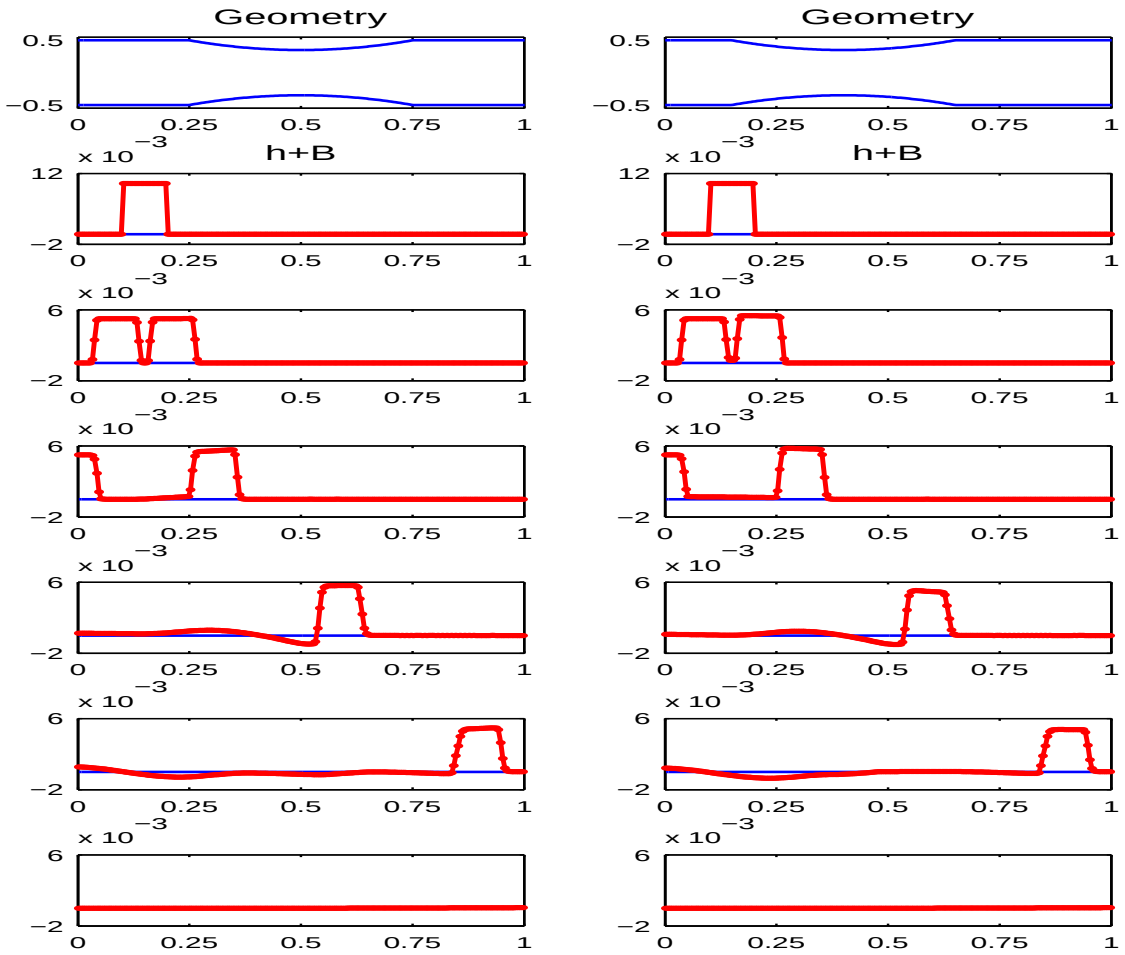



Figure 4.7: Propagation of small perturbation to steady state of rest through a contracting rectangular channel, $\epsilon = 10^{-2}$: Centered contraction (left) and off centered contraction (right) T=0,0.02,0.05, 0.15,0.25 and 0.5.

A comparison with results by the central-upwind scheme [BK09] are shown in

Figure 4.8 for $\epsilon = 10^{-5}$, on a grid of 200 points. Results are similar, with the upwind scheme better able to maintain a sharp profile of the perturbation.

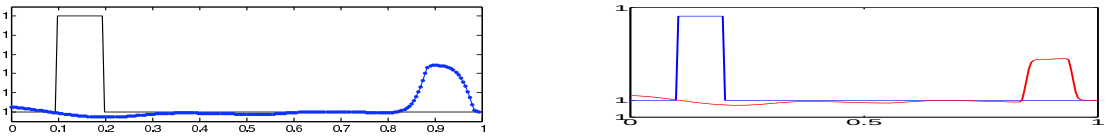


Figure 4.8: Propagation of small disturbance to steady state of rest through a contracting rectangular channel, $\epsilon = 10^{-5}$. Total water height, $w = h + B$, at $t = 0.25$ (dots) over initial conditions (solid line): central-upwind [BK09] (left) and upwind (right) schemes.

The propagation of a small perturbation to a non-rest steady state is shown in Figure 4.9. Again, once the perturbation leaves the computational domain, the unperturbed steady state is recovered, indicated by $Q$ and $E$ going back to their constant unperturbed levels.

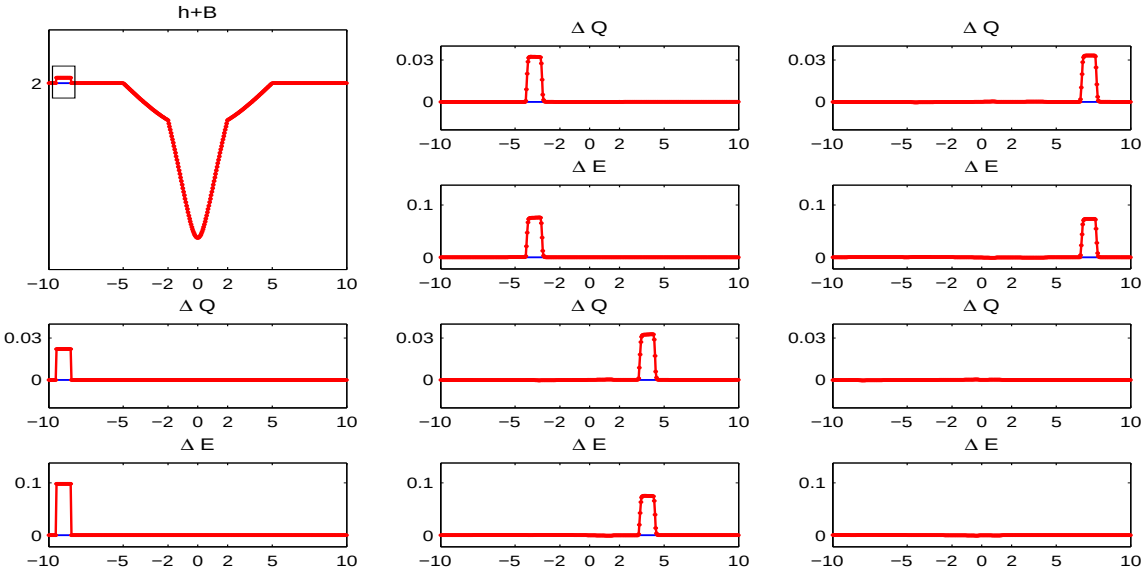



Figure 4.9: Propagation of small perturbation to non rest steady state through a contracting rectangular channel, $\epsilon = 10^{-2}$, centered contraction. The total height for the initial perturbation (top left) and the equilibrium variables for T = 0,0.8,1.9,2.4 and 4 are shown.

**Convergence to Steady State**

The next set of examples illustrates the long time convergence of transient solutions to a steady state. The channel has vertical walls, with a parabolic contraction. In all cases, the flow discharge $Q$ was imposed at inflow, and the depth of the layer $h$ was imposed at (subcritical) outflow. Figure 4.10 (top) shows a schematic of the geometry (not drawn to scale) and includes a straight channel, a channel with centered contraction and a channel with off-centered contraction. Below the geometry, Figure 4.10 shows the topography $B$ and water level $h + B$ at steady state for (i) subcritical flow (top middle), (ii) smooth transcritical flow (bottom middle) and (iii) transcritical flow with a jump (bottom). The following boundary conditions were specified for (i) straight channel (ii) centered and (iii) off-centered contraction respectively. The bottom topography in all the examples is $B(x) = \max\{(0.05(4 - x^2), 0)\}$. The geometry is given by a parabolic contraction extending from $x_0$ to $x_1$, as specified in the following table. Computed solutions are in excellent agreement with exact solutions, also shown.

| | $Q_{in}$ | $h_{out}$ | $\sigma_{min}$ | $x_0$ | $x_1$ | |
|---|---|---|---|---|---|---|
| | | | 1.0 | $-10$ | 10 | : straight |
| subcritical flow: | 4.42 | 2.0 | 0.9 | $-5$ | 5 | : centered |
| | | | 0.9 | $-3$ | 9.5 | : off-centered |
| | | 0.4058 | 1.0 | $-10$ | 10 | : straight |
| smooth transcritical flow: | 1.53 | 0.3384 | 0.7 | $-5$ | 5 | : centered |
| | | 0.3356 | 0.6 | $-3$ | 9.5 | : off-centered |
| | | 0.34 | 1.0 | $-10$ | 10 | : straight |
| transc. flow with jump: | 0.18 | 0.34 | 0.66 | $-5$ | 5 | : centered |
| | | 0.39 | 0.4 | $-3$ | 9.5 | : off-centered |

Figure 4.11 shows a comparison between the present upwind scheme and the central scheme of [BK09]. The examples were taken from [BK09]. See Figure 9 (bottom), Figure 10 (top) and Figure 11 (bottom) therein. By design, both schemes respect steady state of rest, but neither is able to preserve general steady states. The closeness of $Q$ and $E$ to being constants is a good measure of how well the schemes do in approximating general (smooth) steady states. It is striking to note that the present upwind scheme converges to $Q$ and $E$ with relative errors consistently 2-4 orders of magnitude better than the central-upwind scheme.

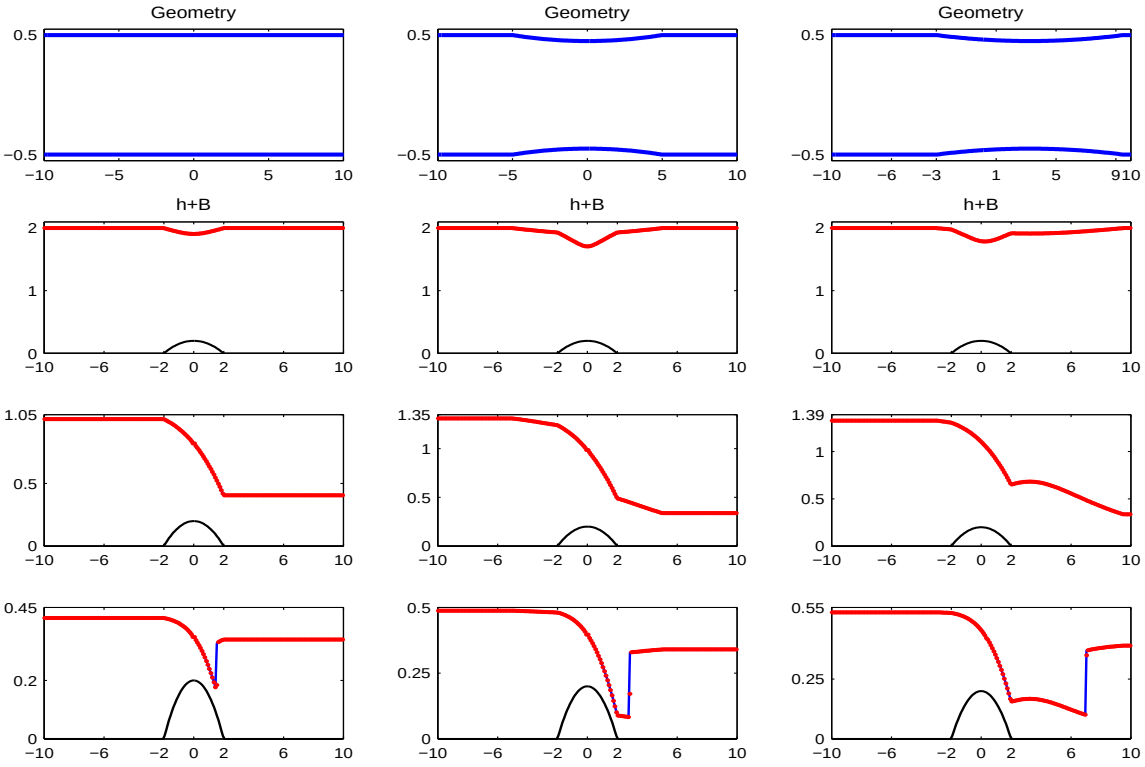



Figure 4.10: Numerical (symbol) and exact (solid line) water level in steady-state solutions: Geometry (top), subcritical flow (top middle), smooth transcritical flow (bottom middle) and transcritical flow with a jump (bottom); Rectangular channel with straight walls (left), centered contraction (middle) and off-center contraction (right).

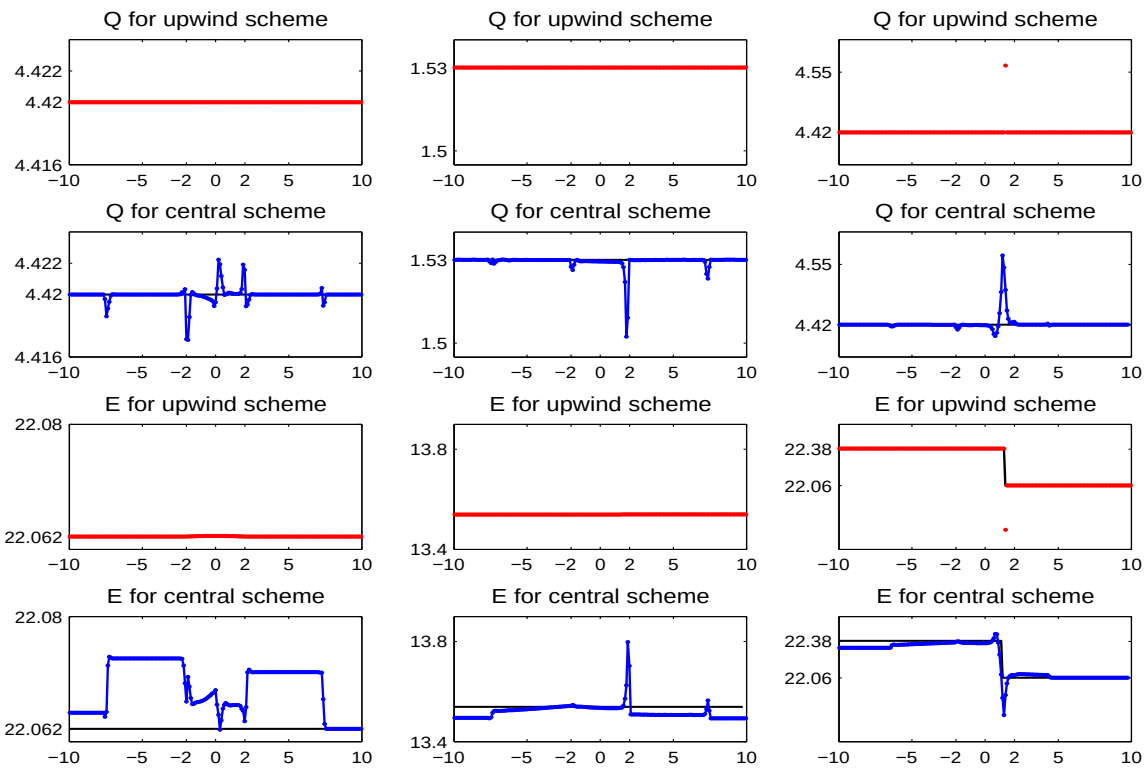



Figure 4.11: Comparison between the upwind and the central schemes for convergence to steady states. For channels with vertical walls, equilibrium variables $Q$ and $E$ are shown for a subcritical (left), smooth transcritical (middle), and discontinuous transcritical (right) flow.

### Reservoir Drainage after Dam Break

In the next example, a reservoir is being drained through a contracting channel. The water is initially at rest $u = 0$, leveled at $h + B = 0.8$. The water drains through the right boundary, the left boundary is assumed a line of symmetry of the domain and wall boundary conditions are applied, trapping the water to the left of the bump. Computed solution is shown in Figure 4.12 for various intermediate times, and the reservoir has essentially drained by $T = 15$. The equilibrium variables $Q$ and $E$ are also shown for the solution at the final time.

Figure 4.13 compares the computed solution using the velocity regularization (4.24) and (4.25) respectively. It can be observed that the regularization (4.24) re-

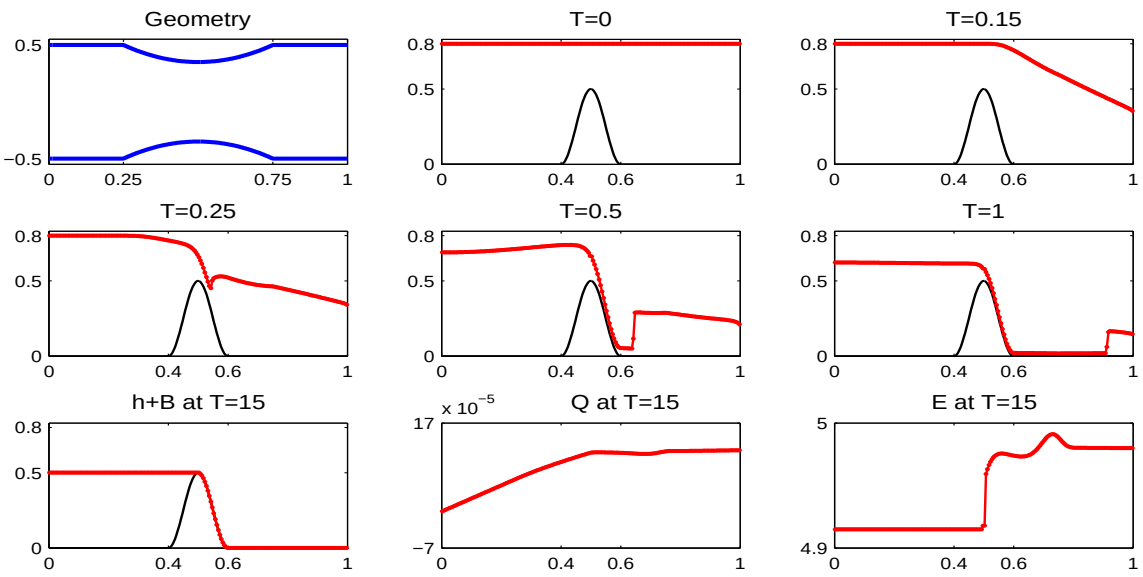



Figure 4.12: Reservoir drainage after dam break. The equilibrium variables at T=15 are also shown.

sults in a more noisy drained solution, while regularization (4.25), which makes use of the steady-state energy $E_{StSt}$ in the trough converges to a cleaner and generally more accurate solution.

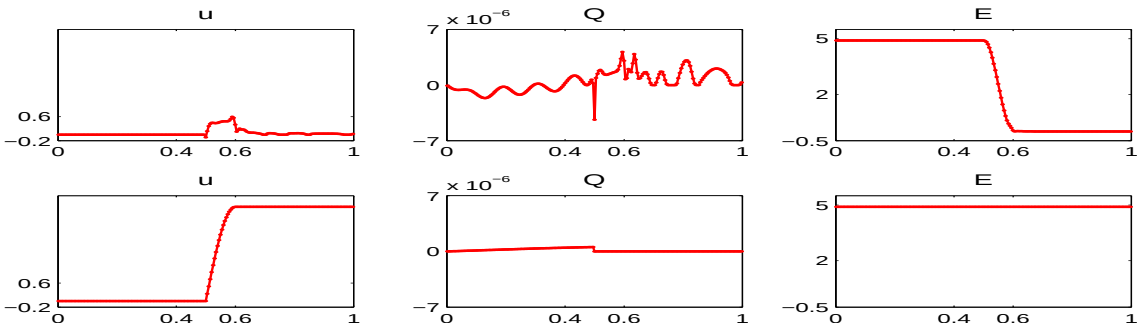



Figure 4.13: Reservoir drainage after dam break. Comparison at $T = 50$ of the computed velocity and equilibrium variables using regularization (4.24) (top) and regularization (4.25) (bottom).

Figure 4.14 shows reservoir drainage through a contracting channel, this time over a double bump topography. The water now gets trapped in two troughs. Despite the

fact that the scheme is not positive, the computed solution remains positive and we are able to integrate this solution for very long time until drainage is reached. The equilibrium variables $Q$ and $E$ corresponding to the final time are also shown.

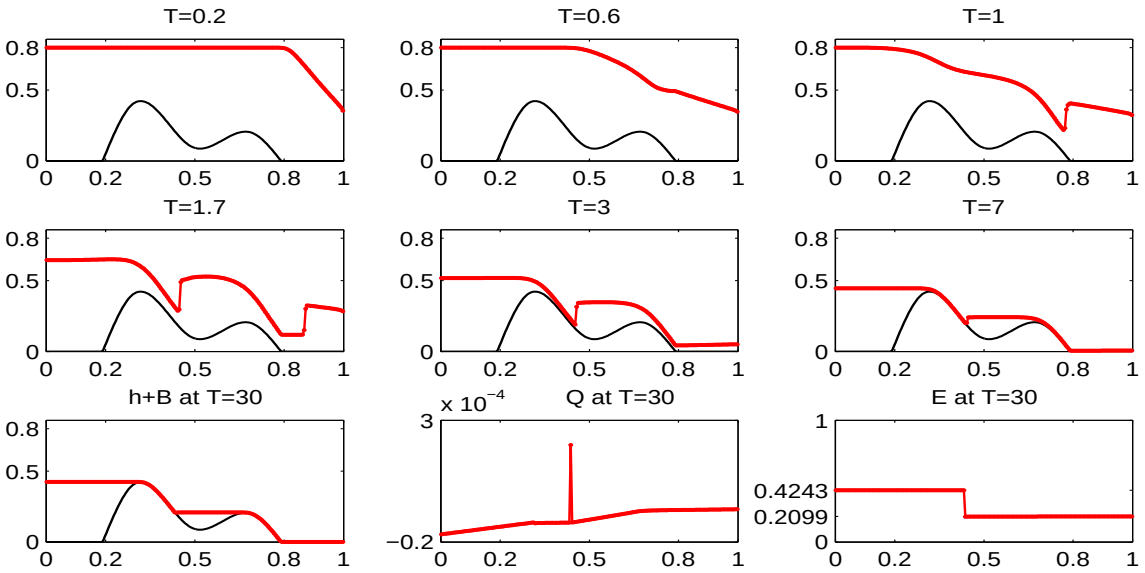



Figure 4.14: Reservoir drainage after dam break. The equilibrium variables are also shown for T=30.

### 4.3.2 More general channels

The following tests involve channels of general cross section described by $\sigma(x, y)$. We present examples for channels consisting of one, two or several trapezoids. Exact solutions are also computed and are compared to computations. We use $\chi_{[a,b]}(x)$ to denote the characteristic function of the interval $[a, b]$.

#### Propagation of Small Perturbation to Non-Rest Steady State

In this test, the topography is a cosine bump $B(x) = \chi_{[0.4,0.6]}(x)\frac{1}{4}(cos(\pi(x - 1/2)/0.1) + 1)$. The channel has a trapezoidal cross section with variable (in x) wall inclination $\sigma(x, y) = \sigma_B(x) + m(x)y$, with $m(x) = 2 + \chi_{[0.4,0.8]}(x)\frac{1}{4}(cos(\pi(x -$

.6)/0.2) + 1), and $\sigma_B(x) = \min(1, 0.7 + 4.8(x - 0.4)^2)$. In this example, the steady-state flow is subcritical, with $Q = 4$ and $h_{out} = 1.4$. The size of the perturbation is $\epsilon = 2 * 10^{-3}$. The initial disturbance to the interface, as well as the relative errors for the equilibrium variables are shown in Figure 4.15. We observe that the unperturbed steady state is recovered very accurately.

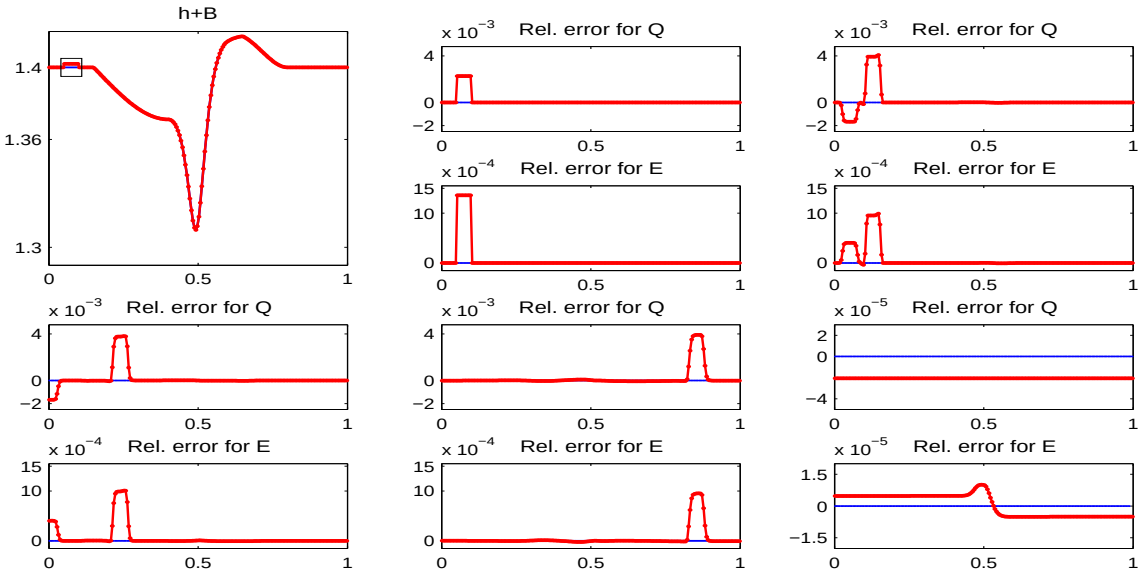



Figure 4.15: Small perturbation to steady state of non-rest for a trapezoidal channel, $\epsilon = 2 * 10^{-3}$, non-centered contraction. The total height for the initial perturbation (top left) and the relative errors for the equilibrium variables for T = 0,0.014,0.04,0.19, and 20 are shown.

### Convergence to Steady States for Trapezoidal Channel

We next study the convergence of transient solutions to steady state. The topography and geometry are the same as in the previous test. For subcritical flow, $Q = 4.42$ and $h_{out} = 1.47$. For smooth transcritical flow, $Q = 8.4992$ and $h_{out} = 1.0388$. For discontinuous transcritical flow, $Q = 1.1104$ and $h_{out} = 0.7195$. Computed and exact solutions are shown in Figure 4.16, with very good agreement.

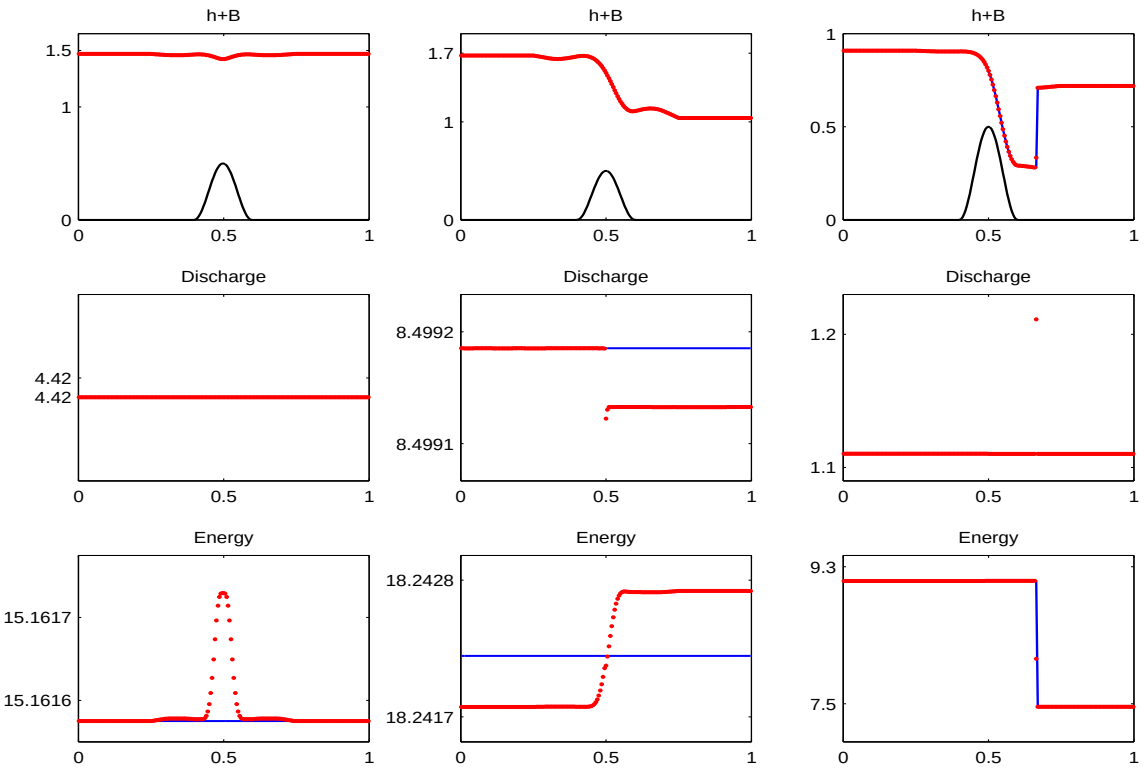



Figure 4.16: Numerical (symbol) and exact (solid line) solutions in steady-state flows: Water level (top), discharge (middle) and energy (bottom); Subcritical flow (left), smooth transcritical flow (middle) and transcritical flow with a jump (right)

### Convergence to Steady States for Piecewise Trapezoidal Channel

In this test, each cross section of the channel consists of two trapezoids, with variable (in x) wall inclination. The bottom trapezoid, with height $y = 1.2$, is the same as in the previous example, and the wall of the top trapezoid has twice the slope of the bottom one. Convergence of transient solutions to steady state are shown in Figure 4.17 for subcritical (left), smooth transcritical (middle) and discontinuos (right) flows. Agreement between computed and exact solutions is excellent.

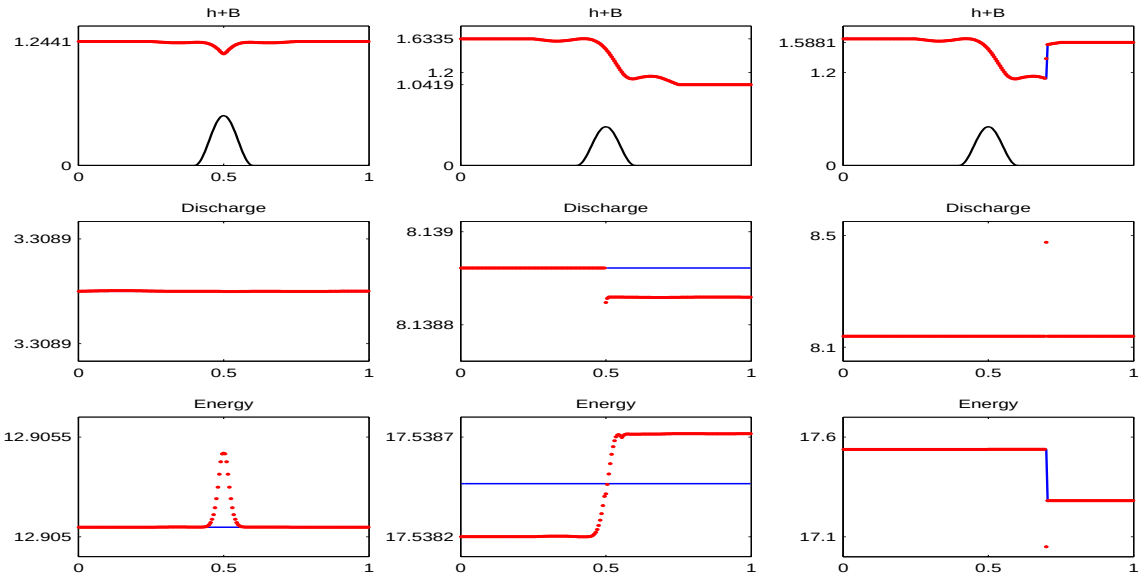



Figure 4.17: Convergence to steady state for a trapezoidal channel (two trapezoids). Total height (top), discharge (middle), and energy (bottom) are shown for subcritical (left), smooth transcritical (center) and discontinuous transcritical (right) flows.

### Convergence to Steady State for General Channels

The last example concerns flow in a channel whose geometry is given by

$$\begin{aligned}\sigma(x,y) =& 1+\frac{3}{4}cos(\pi x)-\frac{1}{4}\chi_{[0.4,0.6]}(x)\left(cos(\pi(x-1/2)/0.1)+1\right)\\ &+\sqrt{y}\left(1-\frac{1}{4}\chi_{[0.1,0.7]}(x)\left(cos\left(\pi(x-0.4)/0.3\right)+1\right)\right)\\ &-2\chi_{[0,1]}\left(\frac{(x-0.3)^2+(y-1.4)^2}{r_1^2}\right)cos\left(\sqrt{\frac{(x-0.3)^2+(y-1.4)^2}{r_1^2}}\frac{\pi}{2}\right)\\ &-1.6\chi_{[0,1]}\left(\frac{(x-0.75)^2+(y-1.4)^2}{r_2^2}\right)cos\left(\sqrt{\frac{(x-0.75)^2+(y-1.4)^2}{r_1^2}}\frac{\pi}{2}\right),\end{aligned}$$

where $r_1 = 0.28$, $r_2 = 0.2$. The topography is a 3-bump spline with nodes $(x, y)$=$(0.2, 0)$, $(0.3, 0.6)$, $(0.4, 0.4)$, $(0.5, 0.5)$, $(0.6, 0.2)$, $(0.7, 0.3)$ and $(0.8, 0)$, shown on top left of Figure 4.18. In this example $Q = 2.0583$ and $h_{out} = 1.5$. The cross section is approximated by 50 trapezoids. The results in Figure 4.18 show excellent agreement

between the computed and exact steady-state solutions. The relative errors for the equilibrium variables $Q$ and $E$ are of orders $10^{-6}$ and $10^{-4}$ respectively. Figure 4.19 shows a 3D view of the flow.

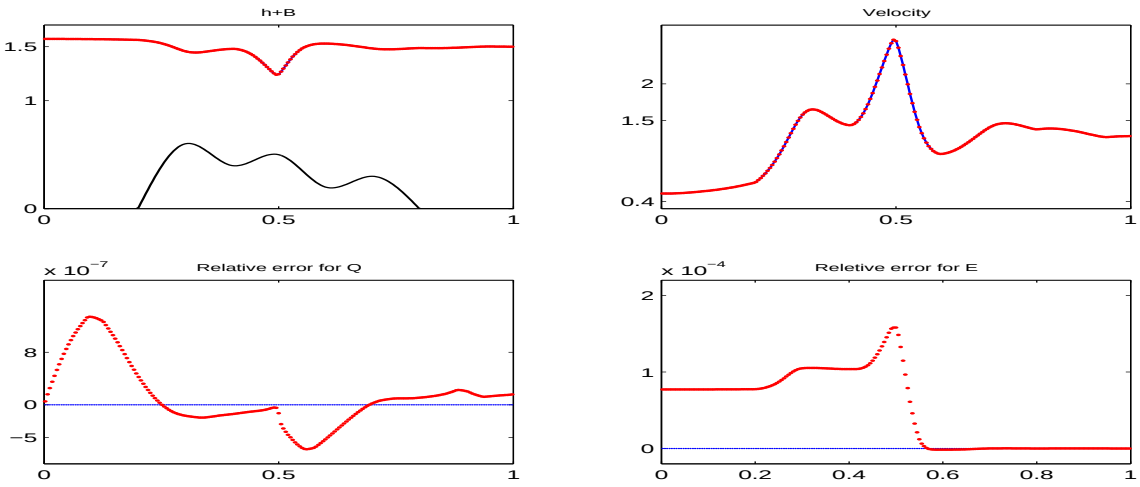



Figure 4.18: Convergence to a subcritical flow. Exact and numerical solutions are plotted with excellent agreement. The top surface, topography (top left), velocity (top right), and relative errors for the equilibrium variables (bottom) are shown.

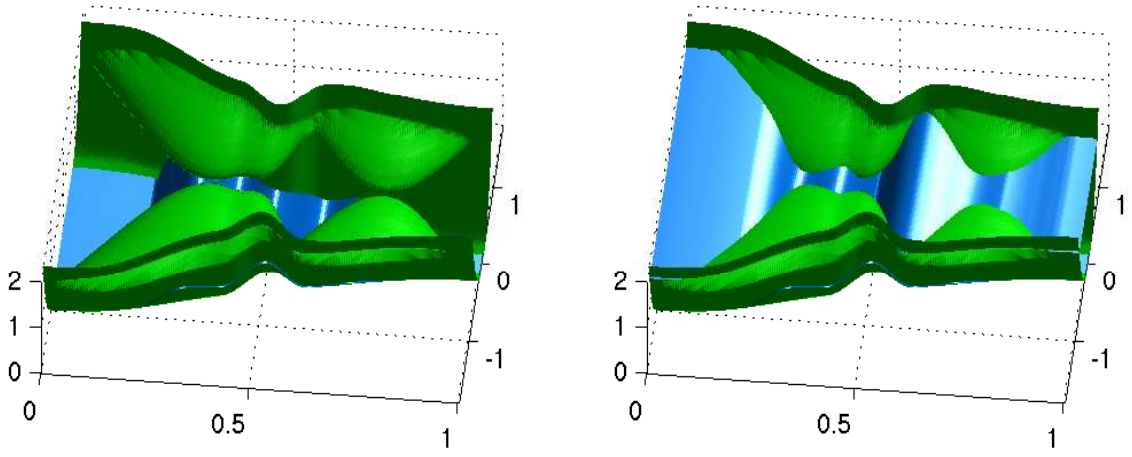


Figure 4.19: 3D view of the channel (left) and the channels with the subcritical flow (right) given in Figure 4.18.

## 4.4 Positivity Preserving Roe Scheme in the Absence of Source Terms

In the present context, a numerical scheme is called positivity preserving if the numerical solution given by the scheme has positive layer's depth in all the cells provided that the depth is positive in the initial data. The schemes presented in this chapter are not formally positivity preserving schemes, but they have shown to be very robust near dry states. In this section we will derive a positivity preserving Roe scheme for the regular shallow water flows (no topography and no geometry).

Let us consider the regular shallow water equations

$$\begin{pmatrix} h \\ hu \end{pmatrix}_t + \begin{pmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \end{pmatrix}_x = 0.$$

### 4.4.1 Positivity in Lagrangian Coordinates

Following [Dub99], we express the equations above in Lagrangian coordinates

$$(x,t) \to (m,t), \text{ where } m = \int_{x_0}^{x} h(\tilde{x},t)d\tilde{x} = \tilde{m}(x),$$

where $m = \tilde{m}(x)$ is a function of $x$, and doesn't depend on $t$. In the new coordinates,

$$\tilde{h}(m,t) = h(\tilde{m}^{-1}(m),t), \ \tilde{u}(m,t) = u(\tilde{m}^{-1}(m),t).$$

Define $\tilde{v} = \frac{1}{h}$. In this coordinates, the shallow water equations become

$$\begin{pmatrix} \tilde{v} \\ \tilde{u} \end{pmatrix}_t + \begin{pmatrix} -\tilde{u} \\ \frac{g}{2}\tilde{v}^{-2} \end{pmatrix}_m = 0$$

The shallow water equations share the same weak solutions in both coordinate systems. However, near dry states, $\tilde{v} = 1/\tilde{h}$ is not small and positivity is easier to achieve.

The flux, coefficient matrix, and eigensystem are given by

$$\tilde{F} = \begin{pmatrix} -\tilde{u} \\ \frac{g}{2}\tilde{v}^{-2} \end{pmatrix}, \tilde{A} = \begin{pmatrix} 0 & -1 \\ -g\tilde{v}^{-3} & 0 \end{pmatrix}, \Lambda = \begin{pmatrix} -C & 0 \\ 0 & C \end{pmatrix}, R = \begin{pmatrix} 1 & 1 \\ C & -C \end{pmatrix}$$

Here $C = ch$ is the speed of sound in the Lagrangian coordinates. Dropping the $\tilde{}$, the wave strengths are given by

$$\alpha_1 = \frac{1}{2}\Delta v + \frac{\Delta u}{2\bar{C}},\ \alpha_2 = \frac{1}{2}\Delta v - \frac{\Delta u}{2\bar{C}},$$

where $\bar{C}$ is the linearization for $C$, to be specified. For conservation, we require

$$\frac{g}{2}\Delta v^{-2} = -\bar{C}^2 \Delta v,$$

which is satisfied for

$$\bar{C} = C_{Roe} = \sqrt{h_L h_R}\ \sqrt{g\bar{h}},\ \bar{h} = \frac{h_L + h_R}{2}.$$

The approach in [Dub99] suggests to modify the coefficient matrix to one of the form

$$\begin{pmatrix} 0 & -1 \\ \chi & \omega \end{pmatrix},$$

where $\chi$ approximates $-gv^{-3}$, and $\omega$ is small. The modified matrix needs to ensure

conservation, i.e., it need to satisfy

$$\begin{pmatrix} 0 & -1 \\ \chi & \omega \end{pmatrix} \begin{pmatrix} \Delta\tilde{v} \\ \Delta\tilde{u} \end{pmatrix} = \Delta\tilde{F},$$

which holds if and only if

$$\frac{g}{2}\Delta v^{-2} = \chi\Delta v + \omega\Delta u.$$

It can be shown that this relation is satisfied by

$$\omega = -B\frac{g}{2}\Delta v^{-2}\ \Delta u,\ \chi = -g\left(1 + B\left(\Delta u\right)^2\right)\ \overline{v^{-1}}\ h_L\ h_R = -\left(1 + B\left(\Delta u\right)^2\right) C_{Roe}^2,$$

where $B$ is a constant to be determined. The spectral information in Lagrangian coordinates becomes:

$$\Lambda^L = \begin{pmatrix} \frac{\omega-\sqrt{\omega^2-4\chi}}{2} & 0 \\ 0 & \frac{\omega+\sqrt{\omega^2-4\chi}}{2} \end{pmatrix}, \qquad \alpha_1^L = \frac{2\chi}{\sqrt{\omega^2-4\chi}\ (\omega-\sqrt{\omega^2-4\chi})}\Delta v + \frac{\Delta u}{\sqrt{\omega^2-4\chi}},$$

$$R^L = \begin{pmatrix} 1 & 1 \\ -\frac{2\chi}{\omega+\sqrt{\omega^2-4\chi}} & -\frac{2\chi}{\omega-\sqrt{\omega^2-4\chi}} \end{pmatrix}, \quad \alpha_2^L = \frac{\omega\sqrt{\omega^2-4\chi}}{\sqrt{\omega^2-4\chi}\ (\omega-\sqrt{\omega^2-4\chi})}\Delta v - \frac{\Delta u}{\sqrt{\omega^2-4\chi}}. \tag{4.26}$$

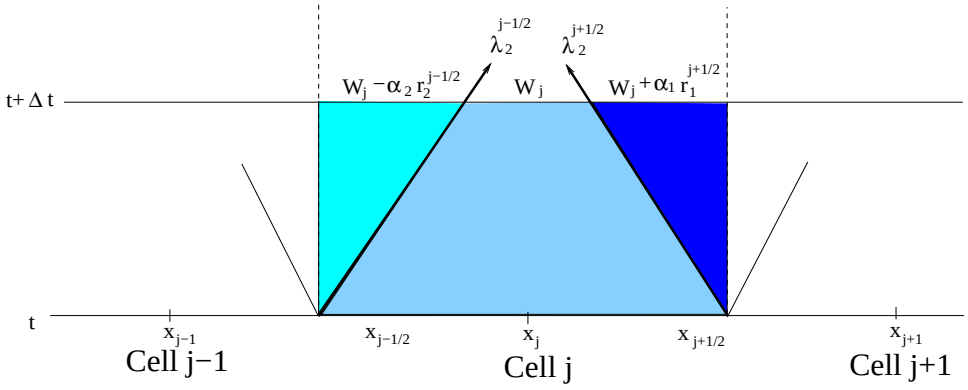



Figure 4.20: Schematic for Roe schemes.

In the absence of source terms, the Roe scheme can be expressed as integration

over the next time step of the states given by an approximate Riemann solver (see Figure 4.20). Positivity can be achieved then if the intermediate state $v_L + \alpha_1$ is positive. This is equivalent to have

$$\sqrt{\omega^2 - 4\chi}\left(\sqrt{\omega^2 - 4\chi} - \omega\right) v_L - 2\chi\Delta v + \left(\sqrt{\omega^2 - 4\chi} - \omega\right)\ \Delta u > 0.$$

Since $\omega^2 - 4\chi = 4C_{Roe}^2\left[1 + B\left(\Delta u\right)^2 + \frac{1}{4}B^2\left(\Delta u\right)^2\left(\Delta v\right)^2 C_{Roe}^2\right]$, we get

$$\left(\sqrt{\omega^2 - 4\chi} + \omega\right)\left(\left(\sqrt{\omega^2 - 4\chi} - \omega\right)\bar{v} + \Delta u - \omega v_L\right) > 0$$

$$\iff \sqrt{\omega^2 - 4\chi}\ \bar{v} + \frac{\omega}{2}\ \Delta v\ \Delta u > 0$$

$$\iff \left(\frac{\sqrt{\omega^2 - 4\chi} - \omega}{2}\right)\ v_L + \left(\frac{\sqrt{\omega^2 - 4\chi} + \omega}{2}\right)\ v_R + \Delta u > 0$$

It is enough to have

$$\sqrt{\omega^2 - 4\chi}v_{min} + \Delta u > 0,\ \text{where } v_{\min} = \min(v_L, v_R),$$

which is satisfied if

$$B \geq \frac{1}{4C_{Roe}^2 v_{min}^2}.$$

Summary:

**Proposition IV.4.** *The Roe scheme in Lagrangian coordinates given by equation* (4.26) *is positivity preserving if the linearization is chosen with*

$$B \geq \frac{1}{4C_{Roe}^2 v_{min}^2}.$$

### 4.4.2 Positivity in Eulerian Coordinates

One now needs to carry the positivity over the Eulerian coordinates. The middle state in Lagrangian coordinates is given by

$$v_s = v_L + \alpha_1 = \frac{v_L + v_R}{2} + \frac{\omega\ \Delta v}{2\sqrt{\omega^2 - 4\chi}} + \frac{\Delta u}{\sqrt{\omega^2 - 4\chi}}$$

$$u_s = u_L - \frac{2\chi}{\omega + \sqrt{\omega^2 - 4\chi}}\ \alpha_1$$

$$= \frac{u_L + u_R}{2} - \frac{\omega}{2\sqrt{\omega^2 - 4\chi}}\Delta u - \frac{\chi \Delta v}{\sqrt{\omega^2 - 4\chi}}$$

So, we want to impose these states in the Eulerian coordinates:

$$h_s = \frac{1}{v_s},\ Q_s = \frac{u_s}{v_s}.$$

The task now is to find the eigenvectors and eigenvalues that produces these intermediate states. We now refer to $\alpha_{1,2}$ the wave strengths in the Eulerian coordinates. So $\alpha_1$ is given by $\alpha_1 = h_s - h_L$. One can show that

$$\alpha_1 = -\frac{h_L}{v_s}\ \left[\frac{\sqrt{\omega^2 - 4\chi} + \omega}{2\sqrt{\omega^2 - 4\chi}}\Delta v + \frac{\Delta u}{\sqrt{\omega^2 - 4\chi}}\right].$$

After some tedious computations, one can show that

$$Q_s - Q_L = \left(u_L + \frac{\omega - \sqrt{\omega^2 - 4\chi}}{2} v_L\right)\alpha_1,$$

which give the first eigenvector and eigenvalue:

$$r_1^E = \begin{pmatrix} 1 \\ u_L + \frac{\omega - \sqrt{\omega^2 - 4\chi}}{2} v_L \end{pmatrix}, \alpha_1 = -\frac{h_L}{v_s}\ \left[\frac{\sqrt{\omega^2 - 4\chi} + \omega}{2\sqrt{\omega^2 - 4\chi}}\Delta v + \frac{\Delta u}{\sqrt{\omega^2 - 4\chi}}\right].$$

Similarly, one can show that

$$\alpha_2 = h_R - h_s = \frac{h_R}{v_s}\left[-\frac{\sqrt{\omega^2-4\chi}-\omega}{2\sqrt{\omega^2-4\chi}}\Delta v + \frac{\Delta u}{\sqrt{\omega^2-4\chi}}\right],$$

and

$$Q_r - Q_s = \left(u_R + \frac{\sqrt{\omega^2-4\chi}+\omega}{2}v_R\right)\alpha_2$$

The complete information of the eigensystem is in the following

**Theorem IV.5.** *Consider the Roe scheme with eigensystem*

$$\Lambda^E = \begin{pmatrix} u_L + \frac{\omega-\sqrt{\omega^2-4\chi}}{2}\ v_l & 0 \\ 0 & u_R + \frac{\omega+\sqrt{\omega^2-4\chi}}{2}\ v_R \end{pmatrix}, \quad \alpha_1^E = \frac{-h_L}{v_s}\left[\frac{\sqrt{\omega^2-4\chi}+\omega}{2\sqrt{\omega^2-4\chi}}\Delta v + \frac{\Delta u}{\sqrt{\omega^2-4\chi}}\right],$$

$$R^E = \begin{pmatrix} 1 & 1 \\ u_L + \frac{\omega-\sqrt{\omega^2-4\chi}}{2}v_L & u_R + \frac{\sqrt{\omega^2-4\chi}+\omega}{2}v_R \end{pmatrix}, \quad \alpha_2^E = \frac{h_R}{v_s}\left[-\frac{\sqrt{\omega^2-4\chi}-\omega}{2\sqrt{\omega^2-4\chi}}\Delta v + \frac{\Delta u}{\sqrt{\omega^2-4\chi}}\right],$$

*where*

$$\omega = -B\frac{g}{2}\Delta v^{-2}\ \Delta u,\ \chi = -g\left(1 + B\left(\Delta u\right)^2\right)\ \overline{v^{-1}}\ h_L\ h_R.$$

*The scheme is positivity preserving provided that*

$$B \geq \frac{1}{4C_{Roe}^2 v_{min}^2},\ \textit{where } C_{Roe} = \sqrt{h_L h_R}\ \sqrt{g\bar{h}},\ \textit{and } v_{\min} = \min(v_L, v_R).$$

Figure 4.21 shows the approximated Riemann solver ($h_L = 0.01, u_L = -1$) and ($h_R = 0.01, u_R = 1$). We notice that the standard approximated Riemann solver (dotted line) has an intermediate state that is negative. Although after integrating we may end up with a positive state in the next time step, one cannot guarantee that this will always happen. On the contrary, one can check that the linearization given in Theorem IV.5 (solid line), the intermediate state is guaranteed to be positive, producing a positivity preserving Roe scheme.

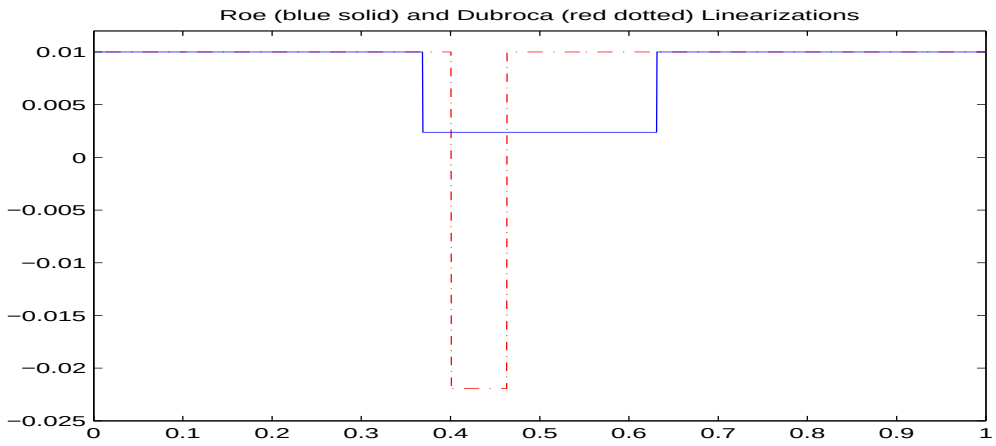



Figure 4.21: Intermediate states using the Standard Roe linearization (dotted line), and the linearization in Theorem IV.5 (solid line).

## 4.5 Conclusions

We considered shallow water flows in channels or arbitrary cross sections. Delicate balance between the flux gradient and the source terms give rise to equilibrium solutions. Naive discretizations may fail to correctly compute near steady-state solutions, and one may observe large errors in perturbations and convergence to steady-state flows.

We developed a conservative numerical scheme that recognizes and respect steady states of rest. We show that this property enables the scheme to correctly capture near general steady states. The bottom topography and the contraction of the channel affects and controls the resulting solution. Testing the scheme in a variety of flows, we show the merits of the scheme comparing to exact solutions and other schemes. The robustness of the scheme near dry states is shown in drainage problems.

## 4.6 Future work

Our numerical method for shallow water flows has proven to be very robust near dry states. We are interested on positivity preserving Roe-type schemes, or central schemes in channels with arbitrary cross sections.

Hyperbolicity in *two-layer* shallow water flows is conditional and obtaining good numerical methods is challenging, even when the channel has vertical walls. We would like to generalize the present Roe-type scheme to two-layer shallow water flows with arbitrary cross sections. Other problems we are interested in include shallow water flows with moving topography, multi-layer flows, and flows in a pipes.

# CHAPTER V

# Introduction to Singular Semiclassical Pseudodifferential Operators. Background.

In semiclassical analysis, the main goal is to study the relationships between the mathematics of classical and quantum mechanics. Traditionally one goes from the quantum setting to the classical setting in this field by letting Planck's constant tend to zero. Classical objects are symplectic manifolds $(X, \omega)$ (phase spaces), together with the set of $C^\infty$ real-valued functions in $(X, \omega)$ (classical observables), endowed with the Poisson bracket $\{\cdot, \cdot\}$. The corresponding quantum objects consist of Hilbert spaces with an inner product $(\mathcal{H}, \langle \cdot, \cdot \rangle)$, together with the set of self-adjoint operators (quantum observables), and the commutator operation. In general we want to assign to each classical observable a quantum equivalent, respecting the Poisson bracket and the commutator, at least asymptotically. This process, referred to as quantization, is important as it is studied in analysis as well as in physics. The main example of this is the so-called Weyl quantization, where $X = T^*(\mathbb{R}^n)$ and $\mathcal{H} = L^2(\mathbb{R}^n)$. This quantization comes from the representation theory of the Heisenberg group, and the next section exposes its basic properties. We refer the reader to [EZ03, Mar02] for more details.

## 5.1 A quick overview of semiclassical analysis.

The space of Schwartz functions on $\mathbb{R}^k$, $\mathcal{S}(\mathbb{R}^k)$, is the set of rapidly decreasing functions

$$\mathcal{S}(\mathbb{R}^k) = \left\{ a \in C^\infty(\mathbb{R}^k) \big| \sup_{\mathbb{R}^k} \left| x^\alpha \partial^\beta a \right| \leq \infty \text{ for all multi-index } \alpha, \beta \right\}.$$

The semiclassical Fourier transform is a mathematical operation that decomposes a signal into its constituent frequencies. More precisely, for any $\varphi \in \mathcal{S}(\mathbb{R}^k)$, the semiclassical Fourier transform of $\varphi$, denoted by $\mathcal{F}_\hbar(\varphi)$ or $\hat{\varphi}_\hbar$, is given by

$$\mathcal{F}_\hbar(\varphi)(p) := \frac{1}{(2\pi\hbar)^{n/2}} \int e^{-ix\cdot p/\hbar} u(x) dx,$$

where the parameter $\hbar$ is known as the Planck's constant, and $p$ as the momentum variable.

**Proposition** (Properties of the Fourier transform)

*(1) The mapping $\mathcal{F}_\hbar : \mathcal{S}(\mathbb{R}^k) \to \mathcal{S}(\mathbb{R}^k)$ is an isomorphism, and the Fourier inversion formula is given by*

$$\mathcal{F}_\hbar^{-1}(\psi)(x) = \frac{1}{(2\pi\hbar)^{n/2}} \int e^{ix\cdot p/\hbar} \psi(p) dp.$$

*Therefore,*

$$\varphi(x) = \frac{1}{(2\pi\hbar)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot p/\hbar} \hat{\varphi}_\hbar(p) dp$$

*(2) In addition, for any multi-index $\alpha$,*

$$(\hbar D)^\alpha = \frac{\hbar^{|\alpha|}}{i^{|\alpha|}} \partial^\alpha,$$

*and* $\varphi \in \mathcal{S}(\mathbb{R}^k)$, *we have*

$$(\hbar D_p)^\alpha \left(\mathcal{F}_\hbar \varphi\right) = \mathcal{F}_\hbar((-x)^\alpha \varphi),$$

*and*

$$\mathcal{F}_\hbar \left((\hbar D_x)^\alpha \varphi\right) = p^\alpha \mathcal{F}_\hbar \varphi \tag{5.1}$$

This shows that for any Schwartz function $u \in \mathcal{S}(\mathbb{R}^n)$, and any multi-index $\alpha$,

$$(\hbar D_x)^\alpha u(x) = \frac{1}{(2\pi\hbar)^n} \int e^{i(x-y)p/\hbar} p^\alpha u(y) dy dp.$$

In the integral above we have first integrated with respect to $y$, followed by the integration in $p$. The Fourier transform in equation (5.1) relates the operator $(\hbar D)^\alpha$ with the function in phase space $a : T^*(\mathbb{R}^n) \to \mathbb{R}$ given by $a(x,p) = p^\alpha$. The following section explains this relation in detail.

### 5.1.1 Weyl quantization

Functions $a : T^*\mathbb{R}^n \to \mathbb{C}$ on phase space are known as *classical observables*. If $a \in \mathcal{S}(\mathbb{R}^{2n})$ is a Schwartz function, we define the operator $a^W(x, \hbar D)$ acting on functions $u \in \mathcal{S}(\mathbb{R}^n)$, as follows:

$$a^W(x,\hbar D)(u)(x) = \frac{1}{(2\pi\hbar)^n} \int e^{i(x-y)\cdot p/\hbar} a\left(\frac{x+y}{2}, p\right) u(y) dy dp. \tag{5.2}$$

For any $a \in \mathcal{S}(\mathbb{R}^{2n})$, $a^W(x,\hbar D) : \mathcal{S}' \to \mathcal{S}$ is continuous, and its adjoint operator satisfies (see [EZ03] Theorem 4.1)

$$\left(a^W(x,\hbar D)\right)^* = \left(\bar{a}(x,\hbar D)\right)^W.$$

**Proposition** *The composition of two Weyl quantizations* $a^W(x,\hbar D) \circ b^W(x,\hbar D)$

*(where $a, b \in \mathcal{S}$) is again the quantization of another symbol, denoted by $a\#b$ (Moyal product), given by*

$$a\#b(x,p) =$$

$$= \frac{1}{(\pi\hbar)^{2n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} e^{-2i\hbar^{-1}(\langle \eta_1, y_2\rangle - \langle y_1, \eta_2\rangle)} a(x+y_1, p+\eta_1) b(x+y_2, p+\eta_2) dy d\eta_1 dy_2 d\eta_2,$$

*where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{R}^n$.*

Applying the stationary phase theorem to $a\#b$, we obtain the expansion

$$a\#b = ab + \frac{\hbar}{2i}\{a,b\} + O(\hbar^2), \tag{5.3}$$

where

$$\{f,g\} = \langle \partial_p f, \partial_x g\rangle - \langle \partial_x f, \partial_p g\rangle \text{ is the Poisson bracket.}$$

In particular, this implies

**Proposition**

$$\left[a^W, b^W\right] = \frac{\hbar}{i}\{a,b\}^W + O(\hbar^2). \tag{5.4}$$

The symbol $a\#b$ depends on $\hbar$ and has the asymptotic expansion given in equation (5.3). The Weyl quantization can be extended to a wider class of symbols, described below.

**Definition V.1.** One defines the symbol classes $S^k_\delta$ as:

$$S^k_\delta := \left\{ a(x,p,\hbar) \in C^\infty \middle| |\partial^\alpha a| \leq C_\alpha \hbar^{-\delta|\alpha|-k} m \text{ for all multiindices } \alpha \right\},$$

where $m : \mathbb{R}^{2n} \to (0,\infty)$ is an order function, i.e., $\partial^\alpha m = O(m)$ for any multi-index $\alpha$.

If $a \in S_\delta(m) := S^0_\delta(m)$, then (Theorem 4.12 in [EZ03], see also[Mar02])

$$a^W(x, \hbar D) : \mathcal{S} \to \mathcal{S},$$

and if $a \in S = S_0(1)$, then

$$a^W(x, \hbar D) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$$

is a bounded operator, and

$$\left|\left|a^W(x, \hbar D)\right|\right|_{L^2 \to L^2} \leq C \sum_{|\alpha| \leq 2n+1} \sup_{\mathbb{R}^n} |\partial^\alpha a|$$

Thus $L^2(\mathbb{R}^n)$ is a quantization of the phase space $T^*(\mathbb{R}^n)$ since we can assign operators in $L^2(\mathbb{R}^n)$ (quantum observables) to symbols in $T^*(\mathbb{R}^n)$ (classical observables) such the asymptotic expansion in equation (5.4) holds.

Typical Weyl quantizations of classical observables include the following examples:

(i) If $a(x, p) = p^\alpha$, where $\alpha$ is a multi-index, then

$$a^W(x, \hbar D) = (\hbar D)^\alpha, \text{ where } D = \frac{1}{i}\frac{\partial}{\partial x}.$$

(ii) If $a(x, p) = V(x)$, then

$$a^W(x, \hbar D) = M_V, \text{ the multiplication operator by } V(x)$$

(iii) If $a(x, p) = \langle x, p \rangle$, then

$$a^W(x, \hbar D) = \frac{\langle \hbar D, M_x \rangle + \langle M_x, \hbar D \rangle}{2}$$

### 5.1.2 Other quantizations. Semiclassical Pseudodifferential Operators.

**Definition V.2.** A quantization for symbols $a(x, y, p)$ that depend on both variables, $x$ and $y$, is defined as

$$Op_\hbar(a)u(x;\hbar) = \frac{1}{(2\pi h)^n}\int e^{i(x-y)p/\hbar}a(x,y,p)u(y)dydp.$$

**Definition V.3.** For any $a \in S(\langle p\rangle^m)$ and $t \in [0,1]$ we have $a((1-t)x+ty,p) \in S(\langle p\rangle^m)$. We define

$$Op_\hbar^t(a) := Op_\hbar(a((1-t)x+ty,p)).$$

The Weyl quantization is recovered when $t = 1/2$. A single operator can be expressed in the different quantizations using different amplitudes. The following proposition shows that all the different amplitudes share the same principal term in the expansion on $\hbar$, called *principal symbol*.

**Proposition** *Let $a = a(x,y,p) \in S(\langle p\rangle^m)$ and $t \in [0,1]$. Then there exists a unique $a_t(x,p) \in S(\langle p\rangle^m)$ such that*

$$Op_\hbar(a) = Op_\hbar^t(a_t).$$

*where $a_t$ is given by the oscillatory integral*

$$a_t(x,p) = \frac{1}{(2\pi\hbar)^n}\int e^{i(p'-p)\theta/\hbar}a(x+t\theta, x-(1-t)\theta, p')dp'd\theta,$$

*and satisfies*

$$a_t(x,p) \sim \sum_{\alpha\in\mathbb{Z}^n}\frac{(-1)^{|\alpha|}\hbar^{|\alpha|}}{i^{|\alpha|}\alpha!}\partial_p^\alpha\partial_\theta^\alpha a(x+t\theta, x-(1-t)\theta, p)\Big|_{\theta=0}$$

$$\sim a(x,x,p) + O(\hbar).$$

This notion of quantization can be extended to manifolds. For any $C^\infty$ manifold $M$, there is a canonical symplectic structure for $X = T^*M$, where in local coordinates, $\omega = dx \wedge dp$. The symplectic manifold $(T^*M, \omega)$ can be quantized to $(L^2(M), \langle \cdot, \cdot \rangle)$ by *semiclassical pseudodifferential operators* (sc-ΨDO). Locally, the Schwartz kernel of those operators can be written as an oscillatory integral as in equation (5.2). The oscillatory integral depends on the coordinate system chosen. However, it can be shown that the principal symbol is globally defined and independent of the coordinates.

The following chart summarizes the classical and quantum objects in the standard case $X = T^*M$:

| | Classical Mechanics: | Quantum Mechanics: |
|---|---|---|
| Phase Space | $X = (T^*M, \omega)$ | $\mathcal{H} = (L^2(M), \langle \cdot, \cdot \rangle)$ |
| Observables | $f : T^*M \longrightarrow \mathbb{R}$ | $A : \mathcal{H} \longrightarrow \mathcal{H}$ self-adjoint |
| Equations of Motion | Hamilton Equations: $\frac{dx_j}{dt} = \frac{\partial H}{\partial p_j}$, $\frac{dp_j}{dt} = -\frac{\partial H}{\partial x_j}$ | Schrödinger Equation: $i\hbar \frac{\partial \psi}{\partial t} = \widehat{H}\psi$ |
| | Poisson bracket: $\{f, g\} = \sum_{i=1}^{n} (\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i})$ | Quantum Parenthesis: $\{A, B\}_Q = \frac{i}{\hbar}[A, B]$ |

### 5.1.3 Semiclassical Lagrangian States

We refer the reader to [PU95] for more details on Semiclassical Lagrangian states. Given a manifold $M$, and a Lagrangian submanifold $\Lambda \subset T^*M$ of the cotangent bundle $T^*M$ (not necessarily conic), we associate semiclassical families of functions to $\Lambda$. Locally, these families can be written as

$$\psi_\hbar(x) = \frac{1}{(2\pi\hbar)^{\frac{n+2N}{4}}} \int e^{i\hbar^{-1}\phi(x,\theta)} a(x,\theta) d\theta,$$

where $\theta \in \mathbb{R}^N$ is $N$-dimensional, $a(x, \theta)$ is the amplitude, and belongs to suitable symbol classes, and $\phi(x, \theta)$ is a phase function that parametrizes $\Lambda$, i.e., $d_{x,\theta}\phi(x, \theta) \neq 0$

for all $(x,\theta) \in M \times \mathbb{R}^N$, and in the support of $a(x,\theta)$,

$$d_\theta \phi(x,\theta) = 0 \text{ implies } (x, d_x\phi(x,\theta)) \in \Lambda.$$

These objects are defined globally in [GS10]. Under mild conditions on $\Lambda$, we can construct a symbolic calculus where each semiclassical state has an invariantly defined principal symbol. In [PU95], semiclassical states are related to Lagrangian distributions studied in [Hör85]. This relations enables us to define a symbolic calculus for the semiclassical states. We now give the details on the relation, and leave details of the Lagrangian distributions in Appendix B.

Given a $C^\infty$ manifold $M$, the pre-quantum circle bundle of $T^*M$ can be identified with the following submanifold of $T^*(M \times S^1)$:

$$Z = \{(x,\theta;\xi,\kappa) \in T^*(M \times S^1) \; ; \; \kappa = 1\},$$

with the obvious circle action. The connection form, $\alpha$, is the pull-back to $Z$ of the canonical one form of $T^*(M \times S^1)$.

**Definition V.4.** ([PU95]) A Lagrangian submanifold $\Lambda \subset T^*M$ will be called *admissible* iff there exists a conic Lagrangian submanifold,

$$\tilde{\Lambda} \subset T^*(M \times S^1) \cap \{\kappa > 0\}$$

such that

$$\Lambda = \left(\tilde{\Lambda} \cap Z\right)/S^1. \tag{5.5}$$

We call such a $\tilde{\Lambda}$ a *homogenization* of $\Lambda$. Notice that $\Lambda$ is not required to be conic.

It is not hard to see that $\Lambda$ is admissible if and only if it satisfies the following

Bohr-Sommerfeld condition: There exists $\varphi : \Lambda \to S^1$ such that

$$\iota^* \eta = d \log \varphi,$$

where $\iota : \Lambda \to T^*M$ is the inclusion and $\eta$ the canonical one form of $T^*M$. Given such a $\varphi$, a homogenization of $\Lambda$ can be defined by:

$$\tilde{\Lambda} = \{e^{i\theta} = \varphi(\lambda),\ \kappa > 0\}. \tag{5.6}$$

**Definition V.5.** Let $M$ be a $C^\infty$ manifold and consider a family of smooth functions $\{\psi_\hbar\}$. The $\hbar$ *-transform* of the family $\psi_\hbar$ is the following distribution (if the series converges weakly) in $M \times S^1$:

$$\Psi(x, \theta) = \sum_{m=0}^{\infty} \psi_{1/m}(x) e^{im\theta}.$$

The main point of the previous two definitions is the following lemma.

**Lemma** ([PU95]) *The $\hbar$-transform of a semiclassical state associated to an admissible Lagrangian is a Lagrangian distribution associated to a homogenization of the Lagrangian submanifold.*

This enables us to define a symbolic calculus for the class of semiclassical states, and semiclassical Fourier integral operators. Roughly speaking, semiclassical Fourier integral operators are operators whose Schwartz kernel are semiclassical states.

We now introduce the semiclassical notion of wavefront set, known as the *frequency set*.

**Definition V.6.** Given a family of semiclassical states $\psi_\hbar$, a point $(x_0, p_0)$ is not in the frequency set, $FS(\psi_\hbar)$, if and only if there exists a cut-off function $\rho$ with $\rho(x_0) \neq 0$ and a neighborhood $V$ of $p_0$ such that for each $k$, $\alpha$, there exists a $C > 0$

such that for each $p \in V$,

$$|D^{\alpha}\mathcal{F}_{\hbar}(\rho\psi_{\hbar})(p)| \leq C\hbar^{k},$$

where $\hbar = 1/N$.

**Proposition** *Let $\pi : \tilde{\Lambda} \to \Lambda$ be the projection in equation (7.5). Given a family semiclassical family $\psi_{\hbar}$ and its homegenization $\Psi$, we have*

$$FS(\psi_{\hbar}) = \pi\left(WF(\Psi) \cap Z\right).$$

The next section introduces notation, and provides the main results of this part of the thesis.

## 5.2 Main Results. Quantizations of Symplectic Manifolds with Boundary.

The goal of this part of the thesis can be easily motivated using a simple example. Let us consider $M = S^1$, and let $Q : T^*(S^1) \to \mathbb{R}$ be a smooth function in phase space. The function $Q(\theta, p)$ is $2\pi$-periodic in $\theta$, and can be expressed as

$$Q(\theta, p) = \sum_{k=-\infty}^{\infty} f_k(p)e^{ik\theta}.$$

One of the "Szegö limit theorems" that is proved in the classic text by Grenander and Szegö, [GS58], is the following:

**Theorem:** *For each positive integer $N$ form the $(N+1) \times (N+1)$ matrix, $T_N$, whose $(i, j)$ entry is*

$$(T_N)_{ij} = f_{i-j}\Big(\frac{i+j}{2N}\Big), \quad 0 \leq i,\, j \leq N. \tag{5.7}$$

*Then, for any integer $m \geq 0$*

$$\mathrm{Tr}(T_N^m) \sim \frac{N}{2\pi} \int_0^1 \int_0^{2\pi} Q(\theta, p)^m \, d\theta \, dp.$$

*as $N \to \infty$.*

We begin by recasting the previous theorem in semiclassical terms. The standard basis of $L^2(S^1)$ is

$$\left\{ \frac{1}{\sqrt{2\pi}} e^{ik\theta} \right\}_{k=-\infty}^{\infty}.$$

The elements in the basis above are the eigenfunctions of the semiclassical operator $\widehat{P}(\hbar) = \hbar D_\theta$, where $D_\theta = \frac{1}{i}\frac{\partial}{\partial \theta}$, with eigenvalues $\hbar k$. In particular,

$$SpecP(\hbar) = \hbar \mathbb{Z}.$$

Let $\widehat{Q}(\hbar)$ be the Weyl quantization of $Q$, which we think of as an operator on $L^2(S^1)$, $S^1 = \mathbb{R}/2\pi\mathbb{Z}$, and let

$$\Pi_N : L^2(S^1) \to \mathcal{H}_N := \mathrm{span}\{e^{ikx}, 0 \leq k \leq N\}$$

be the orthogonal projection, where $N = 1/\hbar$. Then a calculation shows that the matrix (5.7) is the matrix of the "cut" operator

$$\Pi_N \, \widehat{Q}_N \, \Pi_N$$

in the exponential basis, where $\widehat{Q}_N = \widehat{Q}(1/N)$. Note that $\mathcal{H}_N$ is the span of the eigenvectors of $\hat{P} = \hbar D_x$ with eigenvalues in $[0, 1]$, and so the region in phase space, $X = T^*S^1$,

$$X_c := \{(x, p) \in T^*S^1 \; ; \; p \in [0, 1]\},$$

is the classical counterpart to $\mathcal{H}_N$. The result in the previous theorem can be restated as

$$\frac{\mathrm{Tr}\left(\left(\Pi_N \widehat{Q}_N \Pi_N\right)^m\right)}{N} \sim \frac{1}{2\pi} \int\limits_{X_c} Q(x,p)^m dxdp,$$

and suggests that $(\Pi_N \widehat{Q}_N \Pi_N)^m$ behaves as if it had for symbol the restriction of $Q^m$ to $X_c$. The projector in the eigenfunctions with eigenvalues in $[0,1]$ is then related to the manifold with boundary $X_c = P^{-1}[0,1] \subset T^*(S^1)$. One would like to think of $\Pi \widehat{Q} \Pi$ as a quantum observable of some quantization of $X_c$. In this part of the thesis we will make the discussion above precise, in a general setting.

The Weyl quantization of $a(x,p) = 1$ in $L^2(\mathbb{R}^n)$ is the identity operator. One would like to think of the projector above as the quantization of the characteristic function $\chi_{X_c}$ of $X_c$. Of course, the characteristic function is discontinuous, and the standard semiclassical theory is for smooth amplitudes only. The stationary phase theorem does not directly apply to discontinuous amplitudes, and the symbolic calculus does not make sense. In this thesis we will also develop a symbolic calculus for operators such as $\Pi \widehat{Q} \Pi$.

There has been interest in quantizing symplectic manifolds with boundary in a broader setting. For instance, in the physics literature, Bojowald and Strobl [BS00, BS03] considered for the classical setting a symplectic submanifold of $T^*M$ with boundary, which is obtained by a symplectic cut that will be described below. Furthermore, they proposed as a corresponding Hilbert space the image of an associated projector. We will assign an algebra of semiclassical operators to symplectic manifold with boundary, as explain below.

Let $M$ be a smooth compact manifold, and $X_c \subset X := T^*M$ a closed compact domain with a smooth boundary. We assume $X_c$ is "fibrating and Bohr- Sommerfeld", by which we mean the following. The boundary $\partial X_c$ is always foliated by curves tangent to the kernel of the pull-back of the symplectic form. The fibrating condition

is that there exists a manifold $S$ and a submersion $\pi : \partial X_c \to S$ whose fibers are the leaves of the null-foliation. This is satisfied iff $X_c$ has a globally defining function whose Hamilton flow on $\partial X_c$ is periodic with a common minimal period. The Bohr-Sommerfeld condition is that all leaves, $\gamma$, of $\pi$ satisfy

$$\int_\gamma \alpha \in 2\pi\mathbb{Z},$$

where $\alpha$ is the tautological one-form in $T^*M$. We will show that under these conditions there exist spaces $J^{\ell,m}$ of pseudodifferential operators with singular symbols naturally associated with $X_c$. (Here $(\ell, m)$ is a bi-degree, to be explained later). The frequency sets of their Schwartz kernels are contained in the union of Lagrangian submanifolds of $T^*(M \times M)$

$$\Delta'_c \cup \mathcal{F}\partial X'_c,$$

where

$$\Delta_c = \{(\overline{x}, \overline{x}) \in X_c \times X_c\}$$

and $\mathcal{F}\partial X_c$ is the flow-out

$$\mathcal{F}\partial X_c = \{(\overline{x}, \overline{y}) \in \partial X_c \times \partial X_c \; ; \; \overline{x}, \; \overline{y} \text{ in the same leaf } \}.$$

Intuitively speaking, the diagonal part, $\Delta_c$, is expected to be a part of any pseudodifferential operator calculus associated with $X_c$. The flow-out part, $\mathcal{F}\partial X_c$, is there because the fibrating-and-Bohr-Sommerfeld conditions imply that there should be a significant part of Hilbert space associated with the symplectic reduction of the boundary of $X_c$. The two symbols, one on $\Delta_c$ and one on $\mathcal{F}\partial X_c$, have a compatibility condition that comes about most naturally in our setting.

In particular, the space $J^{-1/2,1/2}$ is closed under composition. The Bohr-Sommerfeld condition is needed for the existence of a global symbolic calculus, and goes along

with having to restricting Planck's constant to take the values $\hbar = 1/N$, $N = 1, 2, \ldots$. The fibrating condition is needed in order for $\mathcal{F}\partial X_c$ to be a closed submanifold of $X \times X$.

The Schwartz kernels are semiclassical analogues of the oscillatory integrals with singular symbols of Melrose-Uhlmann and Guillemin-Uhlmann, [MU79] and [GU81], associated to a pair of intersecting conic Lagrangian submanifolds. See also [Jos98], where a precise calculus for a more generalized class is discussed, and [HV01] where a connection with a class of Legendre distributions is explained. In the conic case, the realization that if the Lagrangians are the diagonal and a flow-out one obtains a symbol calculus is due to Antoniano and Uhlmann, [AU85]. This is possible because

$$\Delta_c \circ \mathcal{F}\partial X_c,\ \mathcal{F}\partial X_c \circ \Delta_c,\ \mathcal{F}\partial X_c \circ \mathcal{F}\partial X_c \subset \mathcal{F}\partial X_c\ ; \quad \Delta_c \circ \Delta_c \subset \Delta_c,$$

so composing two operators with wave-front set in $\Delta_c \cup \mathcal{F}\partial X_c$ produces an operator with wave-front set contained in the same union. We believe that the present semiclassical setting provides a very natural expression for the Antoniano-Uhlmann algebra.

In the Zoll case, there have been numerous papers aimed at refining the general result (5.7), mostly in terms of the reminder estimate. For instance, in [GO97a], the second term in the Szegö formula for Zoll manifolds is proved. Other references in this direction are [Gio03, GO97b, LS96, LRS98, Wid73].

Here we are not focusing on remainder estimates, but on the fact that there is an operator algebra with a symbolic calculus that provides a quantization of $X_c$ and, among other things, a broader (symbolic) setting for Szegö limit theorems. Furthermore, the existence of the operator algebra allows us to go farther in the analysis of the operators $\Pi\widehat{Q}\Pi$ mentioned above, with respect to previous work.

The issues we raise here are also connected with work of Guillemin and Lerman,

[GL02], on the relationship between symplectic cutting and quantization, in the homogeneous (non-semiclassical) category. They consider the case when $X_c = \phi^{-1}(-\infty, 0]$ where $\phi : X \to \mathbb{R}$ is the moment map for a homogeneous action of the circle group on $S^1$. In this setting one can form the symplectic cut

$$Y = X_c/\sim,$$

Their work centers on the algebra of operators $\{\Pi\widehat{Q}\Pi \ ; \ [\Pi, \widehat{Q}] = 0\}$ where $\widehat{Q}$ ranges over (non-semiclassical) $\Psi$DOs on $M$ and $\Pi$ is a spectral projector, as above, associated to a quantization of the circle action by a Fourier integral operator. Roughly speaking they show that such an algebra can be considered a quantization of the symplectic manifold $Y$. We now state one of our main results.

**Theorem V.7.** *There exist vector spaces of semiclassical pseudodifferential operators with singular symbols, $J^{\ell,m}(M \times M; \Delta, \mathcal{F}\partial X_c)$, where $\hbar$ is restricted to the sequence $1/N, N = 1, 2, \ldots$, such that:*

1. *The frequency set of the Schwartz kernel of any operator in the algebra is contained in the union $\Delta \cup \mathcal{F}\partial X_c$.*

2. *Let $\Sigma = \Delta \cap \mathcal{F}\partial X_c$. Then $J^{\ell,m}$ are microlocally Lagrangian states of order $\ell + m$ on $\Delta \setminus \Sigma$ and $\ell$ on $\mathcal{F}\partial X_c \setminus \Sigma$, and therefore there are symbol maps:*

$$\sigma_0 : J \to |\wedge|^{1/2}(\Delta \setminus \Sigma), \quad \sigma_1 : J \to |\wedge|^{1/2}(\mathcal{F}\partial X_c \setminus \Sigma)$$

   *(where $|\wedge|^{1/2}$ denotes the space of half-densities, and we will ignore Maslov factors). There is in fact a symbolic calculus, that will be described below.*

3. *$J^{\ell,m} \circ J^{\tilde{\ell},\tilde{m}} \subset J^{\ell+\tilde{\ell}+1/2, m+\tilde{m}-1/2}$, in particular $J^{-1/2,1/2}$ is an algebra.*

*4. Assume that $X_c$ is compact. Then every*

$$\widehat{A} \in J^{-1/2,1/2}(M \times M; \Delta, \mathcal{F}\partial X_c)$$

*with microsupport contained in $X_c$ is smoothing and*

$$\mathrm{Tr}(\widehat{A}) = (2\pi)^{-n}\ \hbar^{-n} \int\limits_{X_c} \sigma_0(\widehat{A})\, \frac{\omega^n}{n!} + O(\hbar^{-n+1} \log(1/\hbar))$$

*where $\omega$ is the symplectic form of $T^*M$ and $n$ is the dimension of $M$.*

The manifold $X_c$ has an associated operator algebra, $\mathcal{A}_{X_c}$, which consists of elements in the algebra $J\,(M \times M; \Delta, \partial X_c)$ which are microlocally of order $\mathrm{O}(\hbar^\infty)$ in the complement $T^*M \setminus X_c$.

For any $A \in J^{\ell,m}\,(M \times M; \Delta, \mathcal{F}\partial X_c)$, the symbol $\sigma_1(A)$ in the flow-out $\mathcal{F}\partial X_c$ can be identified with a family of classical pseudodifferential operators of order $m' = m - 1/2$ acting on smooth half-densities on the fibers $F_s$ of the foliation (see Proposition VII.17). The two symbols satisfy a compatibility condition in the intersection $\Sigma = \Delta \cap \mathcal{F}\partial X_c$ (see Theorem VII.18). Under this identification, the results become more natural, as described in the following

**Theorem V.8.** *Let $F_s$ be a fiber $\pi : \partial X_c \to S$ above $s \in S$. Let $\sigma_1(A)_s$ and $\sigma_1(B)_s$ be the corresponding classical pseudodifferential operators acting on $F_s$. Then*

*(i) For any $A \in J^{\ell,m}$, $B \in J^{\tilde{\ell},\tilde{m}}$, $A \circ B \in J^{\ell+\tilde{\ell}+1/2, m+\tilde{m}-1/2}$, and*

$$\sigma_0(A \circ B) = \sigma_0(A)\sigma_0(B), \ \textit{and}$$

$$\sigma_1(A \circ B)_s = \sigma_1(A)_s \circ \sigma_1(B)_s$$

*(ii) The adjoint $A^*$ is also in $J^{\ell,m}$, and*

$$\sigma_0(A^*) = \overline{\sigma_0(A)},\ \sigma_1(A^*)_s = (\sigma_1(A)_s)^*$$

*(iii) If $A \in \mathcal{A}_{X_c}^{\ell,m}$ and $m' > 0$*

$$\mathrm{Tr}(A) = (2\pi)^{-n}\hbar^{-\ell-m-n} \int\limits_{X_c} \sigma_0(x,x,p,-p)dxdp + O(\hbar^{-\ell-m-n+1}).$$

*(iv) If $A \in \mathcal{A}_{X_c}^{\ell,m}$ and $m' \leq -4$*

$$\mathrm{Tr}(A) = (2\pi)^{-n+1/2}\hbar^{-n-\ell+1/2} \int\limits_{S} \mathrm{Tr}\left(\sigma_1(A)_s\right) ds + O(\hbar^{-n-\ell+3/2}).$$

In Section 7.4, we also show that under the "Bohr-Sommerfeld and Zoll" conditions, there exists a defining function $P : T^*M \to \mathbb{R}$ of $\partial X_c = P^{-1}(0)$ such that zero is a regular value of $P$, $X_c = \{P \geq 0\}$, and the Hamilton flow $\Phi_t^P$ is $2\pi$-periodic in a neighborhood of $\partial X_c$. From now on we fix this defining function. Several results can be expressed in a very natural way in terms of the Hamilton flow of $P$.

**Theorem V.9.** *Assume in addition that $\partial X_c$ is of contact type. Then there exist orthogonal projections $\Pi \in J^{-1/2,1/2}$. Let $\widehat{Q}(\hbar)$ be a zeroth-order semiclassical pseudodifferential operator with compact microsupport. Then $\Pi\widehat{Q}\Pi$ is in the class $J^{-1/2,1/2}(M \times M\ ;\ \Delta, \mathcal{F}\partial X_c)$. The symbols, ignoring Maslov factors, are as follows:*

$$\sigma_0\left(\Pi\widehat{Q}\Pi\right)(\overline{x},\overline{x}) = \chi_{X_c}(\overline{x})Q(\overline{x})\ \sqrt{dx \wedge dp},$$

$$\sigma_1\left(\Pi\widehat{Q}\Pi\right)_s = \Pi_{F_s} M_{|_{Q_{F_s}}} \Pi_{F_s}$$

*where $\overline{x} = (x,p) \in T^*M$, $F_s$ is a fiber above $s \in S$, $Q_{|_{F_s}}$ is the restriction of $Q$ to $F_s$, $M_{Q_{|_{F_s}}}$ is the operator "multiplication by $Q_{|_{F_s}}$", and $\Pi_{F_s}$ is the Szegö projector in the*

*orbit $F_s$, i.e., for $u : F_s \to \mathbb{C}$ smooth, $u(\Phi_s^P \bar{y}) = \sum_j u_j(\bar{y}) \frac{e^{ijs}}{\sqrt{2\pi}}$,*

$$[\Pi_{F_s} u](\bar{x}) = \sum_{j \geq 0} \frac{u_r(\overline{x})}{\sqrt{2\pi}}.$$

As an immediate corollary we obtain the following Szegö limit theorem:

**Corollary V.10.** *Assume that $X_c$ is compact and $\hbar = 1/N$. Then for any integer $m \geq 0$*

$$\mathrm{Tr}(\Pi \widehat{Q} \Pi)^m = (2\pi)^{-n} \, N^n \int\limits_{X_c} Q^m \, \frac{\omega^n}{n!} + O(N^{n-1} \log(N)).$$

The propagator $e^{-it\hbar^{-1}\widehat{Q}}$ is well known to be a Fourier integral operator. When $\widehat{Q}$ is replaced by $\Pi_N \widehat{Q}_N \Pi_N$, the corresponding propagator becomes "singular". Suppose that the Hamilton flow of the principal symbol of $\widehat{Q}$ preserves the region $X_c$, i.e., for each $\sigma \in \partial X_c$, $\Xi_Q(\sigma) \in T_\sigma \partial X_c$. This is equivalent to the condition that $Q$ is constant along the orbits of $P$ in the boundary $\partial X_c$. We will also assume that the Poisson bracket $\{P, Q\}$ vanishes to second order at $X_c$, which is equivalent to the commutativity of the Hamilton flows $\Phi_s^Q$ and $\Phi_s^Q$ at the boundary $\partial X_c$. Our next main result describes the propagator in this case:

The classes $J^{\ell,m}$ can be similarly defined for intersecting pairs of submanifolds, i.e., submanifolds that intersect cleanly in $T^*M$. We can now state the following:

**Theorem V.11.** *Suppose $\widehat{Q}$ is a zeroth-order semiclassical pseudodifferential operator with principal symbol $Q$ being constant along the orbits $\partial X_c$ generated by $\Phi_t^P$. Assume that the Poisson bracket $\{Q, P\}$ vanishes to second order on $\partial X_c$, and sub$\widehat{Q}(\hbar) = 0$. Then for $\hbar = \frac{1}{N}$,*

$$\Pi e^{-it\hbar^{-1}\Pi \widehat{Q} \Pi} \in J^{-1/2,1/2}\left(M \times M; \Delta_c(t), \mathcal{F}\partial X_c(t)\right),$$

*where*

$$\Delta_c(t) = \Delta(t) \cap X_c \times X_c,\ \Delta(t) = \left\{ (\overline{x}, \overline{y}) \,\middle|\, \overline{x}, \overline{y} \in T^*(M \times M), \overline{x} = \Phi_t^Q(\overline{y}) \right\}$$

$$\mathcal{F}\partial X_c(t) = \left\{ (\overline{x}, \overline{y}) \,\middle|\, \overline{x}, \overline{y} \in \partial X_c,\ \exists s \in \mathbb{R} \text{ such that } \overline{x} = \Phi_s^P \Phi_t^Q(\overline{y}) \right\}$$

**Corollary V.12.** *Egorov-type theorem. Under the conditions above, and for any* $\widehat{A} \in$ *sc-*$\Psi$*DO of order zero, we have*

$$e^{it\hbar^{-1}\Pi\widehat{Q}\Pi}\Pi\widehat{A}\Pi e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi} \in J^{-1/2,1/2}\left(M \times M, \Delta, \mathcal{F}\partial X_c\right),$$

*with the following principal symbols:*

$$\sigma_0(\overline{x}, \overline{x}) = \chi_{X_c}(\overline{x})\, a(\Phi_t^Q(\overline{x})),\ \textit{and}$$

$$\sigma_1\left(e^{it\hbar^{-1}\Pi\widehat{Q}\Pi}\Pi\widehat{A}\Pi e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi}\right)_s = \Pi_{F_s} M_{(a\circ\Phi_t^Q)_{|_{F_s}}} \Pi_{F_s}.$$

**Remarks:**

(1) The result on the trace (part (4) of the Theorem V.7) extends to the spaces $J^{\ell,m}$ provided $m \geq -1/2$. If $m$ is sufficiently negative, the leading contribution to the trace comes from the asymptotic singularity of the kernel of the operator at $\Sigma$.

(2) The symbol calculus for $\sigma_0$ is just the usual pseudodifferential calculus. The symbol calculus for $\sigma_1$ is more complicated (and non-commutative). In fact $\sigma_1$ comes with an extension to a distribution on $\mathcal{F}\partial X_c$ conormal to $\Sigma$, and there exists a formula for the smooth part of the $\sigma_1$ of the composition in terms of the corresponding extensions of the factors.

(3) The restriction that $\hbar = 1/N$ is necessary to have a well-defined global symbol on the flow-out $\mathcal{F}\partial X_c$, see [GS77] chapter VII, §0. Locally elements of the $J^{\ell,m}$ are

given by semiclassical oscillatory integrals with $\hbar \to 0$ continuously.

(4) The error estimate $O(\hbar \log(1/\hbar))$ is sharp for the class $J^{-1/2,1/2}$, but in Corollary V.10 the error should be $O(\hbar^2)$, see Remark VIII.4.

(5) The idea that the image of $\Pi$ quantizes $X_c$ and that the operators $\Pi\widehat{Q}\Pi$ can be considered as associated observables appeared first in the physics literature, see [BS00] §II E and [BS03] (where a connection with symplectic cutting is also made). The operators $\Pi\widehat{Q}\Pi$ do not form an algebra, however, while the class $J^{-1/2,1/2}$ does.

(6) Our work also implies that the results of V. Guillemin and E. Lerman in [GL02] also hold in the semiclassical case.

### 5.2.1 A Numerical Analysis of Propagation of Coherent States

Let $\widehat{Q}$ be a zeroth order semiclassical pseudodifferential operator with symbol $Q$. It is well-known that, if $\psi_{(x_0,p_0)}$ is a coherent state with center at $(x_0, p_0)$, then $e^{-it\hbar^{-1}\widehat{Q}}(\psi_{(x_0,p_0)})$ is a coherent state (appropriately "squeezed") with center at $\Phi_t^Q(x_0, p_0)$, where $\Phi^Q$ is the Hamilton flow of $Q$. If the flow $\Phi^Q$ preserves $X_c$ and the center $(x_0, p_0)$ is in the interior of $X_c$, then the same conclusion holds for the propagation $e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi}(\psi_{(x_0,p_0)})$ of the coherent state by $\Pi\widehat{Q}\Pi$, as the trajectory of the center will remain away from the boundary $\partial X_c$ and everything is as if we were in the boundaryless case.

In Chapter XI we present results of a numerical calculation of $e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi}(\psi_{(x_0,p_0)})$ in an example where the Hamilton flow of $Q$ does not preserve $X_c$, that is, trajectories of $\Phi^Q$ cross the boundary $\partial X_c$.

We consider the Harmonic oscillator $\widehat{P} = \frac{1}{2}(x^2 - \hbar^2\partial_x^2)$ in $\mathbb{R}^1$, and the corresponding projector $\Pi$ onto the span of its eigenfunctions of with eigenvalues less than or equal to one. We take $Q = x^2 - p^2$, and $\widehat{Q}$ the obvious quantization of $Q$.

We now take a coherent state centered inside the interior of $X_c$, and numerically

compute its propagation under $\Pi\widehat{Q}\Pi$. Figure 5.1 consists of plots of the Husimi density in the $z$ variable of the initial projected coherent state, its propagation at a time when the center of the coherent state approximately hits the boundary, and after the center hits the boundary. We observe that after the collision time the coherent state splits into two localized states with centers on inward trajectories and with the same initial energy.

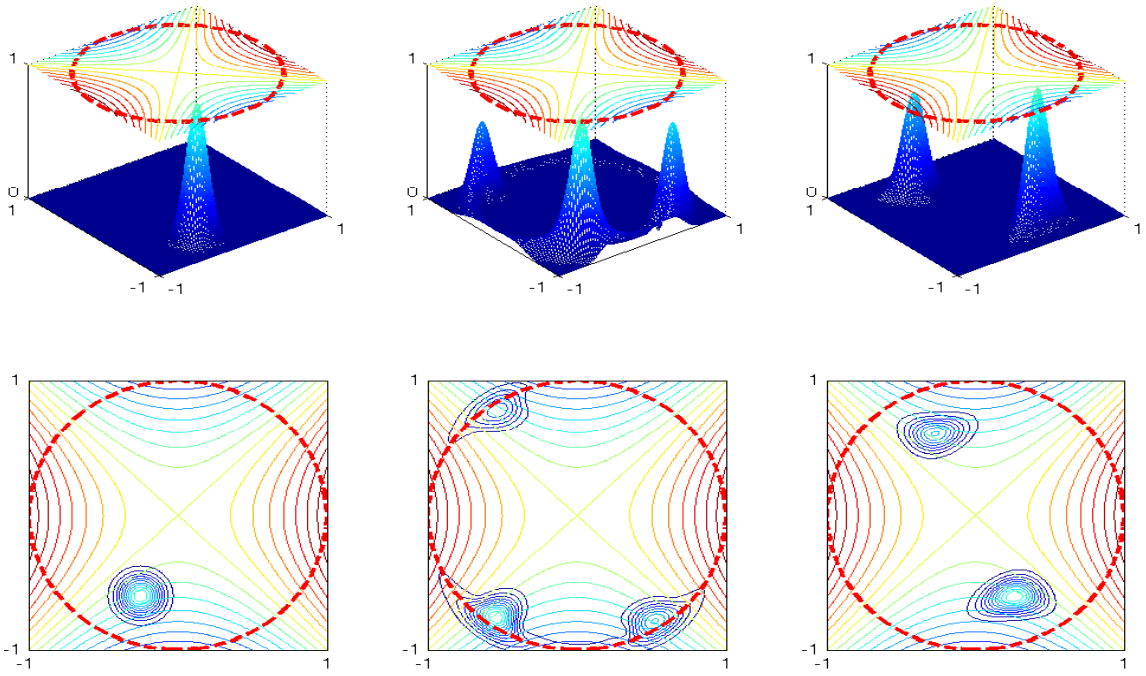



Figure 5.1:
A coherent state is propagated by $e^{-itN\Pi_N\hat{Q}\Pi_N}$ ($Q = x^2 - p^2$) at time $t = 0$ (left), $t = .25$ (middle) and $t = 0.5$ (right). The contour plots of the Husimi densities at each time and of $Q$ are also given (bottom).

The splitting happens immediately after the center collides with the boundary; thus one can speak of infinite-propagation speed along the boundary. Note that the evolution is time-reversible, so that in some cases the opposite phenomenon will occur, namely, two localized states with same classical energy will hit the boundary at the same time and combine into one.

## 5.3 Conclusions

This chapter outlines the basic definitions and results in the theory of semiclassical analysis, specifically in the theory of semiclassical pseudodifferential operators. We illustrated how the principal symbol stays invariant and described how to construct a symbolic calculus for the class of pseudodifferential operators on manifolds by using technique developed by Paul and Uribe in [PU95]. We motivated the goals of this part of the thesis and describe the main theoretical and numerical results. The next section describes the order in which the results are proved.

## 5.4 Description of Chapters VI, VII, VIII, IX, X and XI

In Chapters VI the spaces $J^{\ell,m}$ are defined in a model case by oscillatory integrals with amplitudes having expansions in $\hbar$ with coefficients that are classical symbols. In Chapter VII, the spaces are defined globally on manifolds. We develop a symbolic calculus, where elements in the algebra have principal symbols on the diagonal and on the flow-out given by the fibers of the foliation, with a singularity on the intersection. We identify the singular principal symbols on the flow-out with a family of classical pseudodifferential operators acting on smooth half-densities on the fibers, and expose a symbolic compatibility condition at the intersection. The theorem on the trace is proved in Chapter VIII. In Chapter IX, we prove that the algebra admits projectors under mild conditions and prove a generalized Szegö limit theorem. In Chapter X, we study "singular" propagators given by certain elements in the algebra, and an Egorov-type Theorem is proved. In Chapter XI, the action of the present singular propagators on coherent states is studied numerically for cases not covered in Chapter X. We observe that the center of the coherent states split into more than one when they approach to the boundary. Appendix B provides details of the theory of Lagrangian distributions and their symbols.

# CHAPTER VI

# The model case

In this chapter we define the symbol classes of Theorem V.7 in a model case. In general, the operator classes will be defined as images of the model case by suitable semiclassical Fourier integral operators.

## 6.1 Definitions

We begin by discussing the microlocal model case: $M = \mathbb{R}^n$ and $X_c^n = \{p_1 \geq 0\}$, where $(x_1, \ldots, x_n, p_1, \ldots, p_n)$ are canonical coordinates in $T^*\mathbb{R}^n$. Let $\mathcal{F}\partial X_c^n$ be the flow-out of $\partial X_c^n = \{p_1 = 0\}$, which is

$$\mathcal{F}\partial X_c^n = \left\{(((x_1, x'), p_1 = 0, p'); ((y_1, y' = x'), p_1 = 0, p')) \,\middle|\, (x,p), (y,p) \in T^*(\mathbb{R}^n)\right\}.$$

We will use this case to define operator classes denoted $J^{\ell,m}(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n, \mathcal{F}\partial X_c^n)$, where $\Delta^n \subset T^*\mathbb{R}^n \times T^*\mathbb{R}^n$ is the diagonal.

Roughly speaking, elements in $J^{\ell,m}(\mathbb{R}^n \times \mathbb{R}^n, \Delta^n, \mathcal{F}\partial X_c^n)$ will be defined as oscillatory integrals with amplitudes as follows:

1. We denote by $\mathcal{A}^{\ell,m}$ the class of all smooth functions $a(s, x, y, p, \sigma, \hbar)$ with com-

pact support in $s, x, y, p$ such that, as $\hbar \to 0$

$$a(s, x, y, p, \sigma, \hbar) \sim \sum_{j=-\ell}^{\infty} \hbar^j a_j(s, x, y, p, \sigma)$$

(in a sense that will be explained below) where, for each $j$, $a_j(s, x, y, p, \sigma)$ is a polyhomogeneous classical symbol in $\sigma$ of degree $m$:

$$a_j(s, x, y, p, \sigma) \sim \sum_{r=m}^{-\infty} a_{j,r}(s, x, y, p, \sigma),$$

and for all $\lambda > 0$,

$$a_{j,r}(s, x, y, p, \lambda\sigma) = \lambda^r a_{j,r}(s, x, y, p, \sigma).$$

2. The operators in $J^{\ell,m}(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n, \mathcal{F}\partial X_c^n)$ are those whose Schwartz kernels are of the form

$$A(x, y, \hbar) = \frac{1}{(2\pi\hbar)^n} \int e^{\frac{i}{\hbar}\left[(x_1 - y_1 - s)p_1 + (x' - y')p'\right] + is\sigma} \, a(s, x, y, p, \sigma, \hbar) \, ds \, dp \, d\sigma, \tag{6.1}$$

where we have split the variables: $x = (x_1, x')$, $y = (y_1, y')$ $(x', y' \in \mathbb{R}^{n-1})$.

We now give the details.

**Definition VI.1.** Let $(x, y)$ and $s$ be the standard coordinates on $\mathbb{R}^{2n}$ and $\mathbb{R}$ respectively, and let $p$, $\sigma$ be the dual coordinates to $x$, $s$. We denote by $z = (x, y, s)$ coordinates in $\mathbb{R}^{2n} \times \mathbb{R}$. Define $\mathcal{A}^{\ell,m}$ to be the space of smooth families $a(s, x, y, p, \sigma, \hbar)$ compactly supported in $z$, $p$ such that

$$\left| (\partial/\partial z)^\alpha \, (\partial/\partial p)^\beta \, (\partial/\partial\sigma)^\gamma \, a \right| \leq C\hbar^{-\ell}(1 + |\sigma|)^{m-|\gamma|}, \tag{6.2}$$

for some constant $C = C(\alpha, \beta, \gamma)$ and $\hbar \in (0, h_0]$ for some fixed $h_0 > 0$.

Let $a_j(s,x,y,p,\sigma) \in S^m$ be a sequence of classical symbols in the $\sigma$ variable, of order $m$, and compactly supported in $(s,x,y,p)$, i.e., $a_j$ is smooth,

$$\left|(\partial/\partial z)^\alpha (\partial/\partial p)^\beta (\partial/\partial\sigma)^\gamma a_j(s,x,y,p,\sigma)\right| \leq C(\alpha,\beta,\gamma)(1+|\sigma|)^{m-|\gamma|},$$

for some constant $C = C(\alpha,\beta,\gamma)$, and there exists a sequence $a_{j,r}$ of smooth functions in $\sigma \neq 0$, homogeneous of degree $r$ in $\sigma$, such that $a_j \sim \sum_{r=m}^{-\infty} a_{j,r}$ in the standard sense. Given $a \in \mathcal{A}^{\ell,m}$, we will say that $a \sim \sum_{j=-\ell}^{\infty} \hbar^j a_j$ iff for all integers $K \geq 0$

$$a - \sum_{j=-\ell}^{-\ell+K} \hbar^j a_j \in \mathcal{A}^{\ell-K-1,m}.$$

**Definition VI.2.** Denote by $\mathcal{A}^{\ell,m}_{\text{classical}} \subset \mathcal{A}^{\ell,m}$ the set of all $a(s,x,y,p,\sigma,\hbar)$ that satisfy $a \sim \sum_{j=-\ell}^{\infty} \hbar^j a_j$, as above, and let $\Sigma^n = \Delta^n \cap \mathcal{F}\partial X^n_c$. Define $J^{\ell,m}(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n, \mathcal{F}\partial X^n_c)$ to be the set of kernels of the form

$$A(x,y,\hbar) + F_1(x,y,\hbar) + F_2(x,y,\hbar)$$

where $A(x,y,\hbar)$ is as in (6.1) with $a \in \mathcal{A}^{\ell',m'}_{\text{classical}}$ where

$$\ell' = \ell + 1/2, \quad m' = m - 1/2,$$

and $F_1$, $F_2$ are the kernels of semiclassical Fourier integral operator in sc-$I^{\ell+m}(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n \setminus \Sigma^n)$ and sc$I^{\ell}(\mathbb{R}^n \times \mathbb{R}^n; \mathcal{F}\partial X^n_c)$, respectively.

*Remark* VI.3. As we will see, it is necessary that $F_1$ and $F_2$ appear in the definition in order for the classes $J$ to be closed under composition.

We have yet to give sense to the formula in equation (6.1). Denote by $D = 1 + D_s^2$ where $D_s = \frac{1}{i}\frac{\partial}{\partial s}$. Notice that $De^{is\sigma} = (1+\sigma^2)e^{is\sigma}$. For $a \in A^{\ell',m'}_{\text{classical}}$, we define

$A(x, y, \hbar)$ as

$$A(x, y, \hbar) = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}(x-y)p+is\sigma} \frac{1}{(1+\sigma^2)^k} D^k \left( e^{-i\hbar^{-1}sp_1} a(s, x, y, p, \sigma, \hbar) \right) dsdpd\sigma. \tag{6.3}$$

The integral above is absolutely convergent for $k >> 0$ large enough since $D^k \left( e^{-i\hbar^{-1}sp_1} a \right)$ is $O(\sigma^m)$. It can be easily checked that the definition above does not depend on $k$ using integration by parts.

Equivalently, one can define the integral in (6.1) as an iterated integral, where one first integrates over the variables $(s, p)$ with respect to which the amplitude is compactly supported. The resulting function of $\sigma$ is rapidly decreasing and therefore integrable. Both of these interpretations are useful in proofs.

We now state the first property of the kernels we have just defined:

**Proposition VI.4.** *For any $A \in J^{\ell,m}(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n, \mathcal{F}\partial X_c^n)$, the frequency set of $A$ is contained in the union of the Lagrangians $\Delta^n$ and $\mathcal{F}\partial X_c^n$:*

$$FS(A) \subset \Delta^{n\prime} \cup \mathcal{F}\partial X_c^{n\prime}.$$

*Moreover, away from the intersection $\Sigma = \Delta^n \cap \mathcal{F}\partial X_c^n$, $A$ is microlocally in the space sc-$I^{\ell+m}(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n \setminus \Sigma^n)$ and sc-$I^{\ell}(\mathbb{R}^n \times \mathbb{R}^n; \mathcal{F}\partial X_c^n \setminus \Sigma^n)$, where sc-$I$ denotes the spaces of semiclassical Fourier integral operators.*

*Remark* VI.5. This is a consequence of Theorem VII.14, proved in Section 7.3.

## 6.2 The symbol maps

By the second part of Propositon VI.4, one has two symbol maps:

$$J^{\ell,m}\left(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n, \mathcal{F}\partial X_c^n\right) \begin{array}{ll} \sigma_0 \nearrow & |\wedge|^{1/2}(\Delta^n \setminus \Sigma) \\ \\ \sigma_1 \searrow & |\wedge|^{1/2}(\mathcal{F}\partial X_c^n \setminus \Sigma). \end{array} \tag{6.4}$$

It is easy to see that, for $A$ as above, they are given by the following formulae:

$$\begin{aligned} \sigma_0(A) &:= 2\pi\ a_{-\ell',m'}(s,x,y,p,\sigma)\sqrt{dxdp}\Big|_{y=x,s=0,p_1=\sigma} \quad \text{and} \\ \sigma_1(A) &:= \sqrt{2\pi}\ \int a_{-\ell'}(s,x,y,p,\sigma)e^{is\sigma}d\sigma\sqrt{dxdy_1dp'}\Big|_{y'=x',p_1=0,s=x_1-y_1}. \end{aligned} \tag{6.5}$$

Notice that $\sigma_0$ has a singularity as $\sigma = p_1$ converges to zero, that is, as the point where $\sigma_0$ is evaluated tends to the intersection, $\Sigma$. The same is true of $\sigma_1$, and the leading singularities of $\sigma_0$ and $\sigma_1$ are Fourier transforms of each other. This is exactly as in §5 of [GU81] , and (appropriately understood) will be true in the manifold case as well.

**Proposition VI.6.** *One has the following exact sequence:*

$$0 \to J^{\ell,m-1} \oplus J^{\ell-1,m} \to J^{\ell,m} \xrightarrow{\sigma_0} R^m(\Delta^n \setminus \Sigma) \to 0.$$

Here $R^m(\Delta^n \setminus \Sigma)$ is, roughly speaking, the space of smooth functions on $\Delta^n \setminus \Sigma$ that have singularities of degree $m$ at $\Sigma$ in the normal directions (for details we refer to [GU81], §5 and §6).

The classes of operators with kernels in the $J^{\ell,m}$ are closed under composition in the following sense:

**Proposition VI.7.** *The composition of properly supported operators with kernels*

*in* $J^{\ell,m}$ *and* $J^{\tilde{\ell},\tilde{m}}$*, respectively, is an operator with kernel in* $J^{\ell+\tilde{\ell}+1/2,m+\tilde{m}-1/2}$*. In particular* $J^{-1/2,1/2}$ *is an algebra:*

$$J^{-1/2,1/2} \circ J^{-1/2,1/2} \subset J^{-1/2,1/2},$$

*and for any* $u \in J^{\ell,m}$*,* $v \in J^{\tilde{\ell},\tilde{m}}$ *, we obtain (ignoring Maslov factors and half-densities):*

$$\sigma_0(u \circ v)(\bar{x},\bar{x}) = \sigma_0(u)(\bar{x},\bar{x})\sigma_0(\bar{x},\bar{x}),\ \sigma_1(\bar{x},\bar{y}) = \frac{1}{\sqrt{2\pi}} \int \sigma_1(u)(\bar{x},\bar{z}(t))\sigma_1(\bar{z}(t),\bar{y})dt,$$

*where* $\bar{z}(t)$ *is the bicharacteristic curve joining* $\bar{x}$ *and* $\bar{y}$*.*

Note that, by the symbol calculus for semiclassical FIOs ([GS10]) there are formulae for the symbols $\sigma_0$, $\sigma_1$, of the composition. (Microlocally near $\Delta \setminus \Sigma$ the operators are pseudodifferential and the symbol calculus for $\sigma_0$ is the usual one). We will have more to say about the symbol calculus in the next section.

*Proof.* (Our proof follows the lines of the proof of Theorem 0.1 in [AU85].) It is not hard to show that the classes $J$ are closed under composition by operators $F_1$, $F_2$ as in definition VI.2. Therefore we start with kernels as in (6.1). It is also not hard to show that it suffices to prove the theorem in the case when the amplitudes of these kernels are independent of $\hbar$, in which case $\ell = \tilde{\ell} = -1/2$. Let us therefore consider

$$u(x,y) = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}((x-y)p - sp_1) + is\sigma} a(s,x,y,p,\sigma) ds dp d\sigma,$$

$$v(y,z) = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}((y-z)\tilde{p} - \tilde{s}\tilde{p}_1) + i\tilde{s}\tilde{\sigma}} \tilde{a}(\tilde{s},y,z,\tilde{p},\tilde{\sigma}) d\tilde{s} d\tilde{p} d\tilde{\sigma},$$

where $a$ and $\tilde{a}$ are classical symbols in $\sigma$ and $\tilde{\sigma}$ of orders $m' = m - 1/2$ and $\tilde{m}' = \tilde{m} - 1/2$, respectively. For $\chi_0$ a function that vanishes near the origin, and 1 outside

a compact support, we have

$$u(x,y) = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}((x-y)p - sp_1) + is\sigma} a(s,x,y,p,\sigma)\chi_0(\sigma) ds dp d\sigma + \tilde{u},$$

$$v(y,z) = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}((y-z)\tilde{p} - \tilde{s}\tilde{p}_1) + i\tilde{s}\tilde{\sigma}} \tilde{a}(\tilde{s},y,z,\tilde{p},\tilde{\sigma})\chi_0(\tilde{\sigma}) d\tilde{s} d\tilde{p} d\tilde{\sigma} + \tilde{v},$$

where $\tilde{u}$, $\tilde{v} \in \text{sc-}I^{\ell}\left(\mathbb{R}^n \times \mathbb{R}^n; \mathcal{F}\partial X_c^n\right)$. We can therefore assume that $a$, $\tilde{a}$ are zero near $\sigma = 0$, $\tilde{\sigma} = 0$ respectively. The composition has Schwartz kernel:

$$\omega(x,z) = \frac{1}{(2\pi\hbar)^n} \int e^{is\sigma + i\tilde{s}\tilde{\sigma}} D ds d\tilde{s} d\sigma d\tilde{\sigma} dp,$$

with

$$D = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}\phi} \left(a \cdot \tilde{a}\right) \left(s,\tilde{s},x,y,z,p,\tilde{p},\sigma,\tilde{\sigma}\right) dy d\tilde{p},$$

where

$$\phi = (x-y)p + (y-z)\tilde{p} - \tilde{s}\tilde{p}_1 - sp_1.$$

The critical points for this phase for each $x, z, p$ fixed are: $\tilde{p} = p$, $y' = z'$, $y_1 = z_1 - \tilde{s}$. So, by the stationary phase theorem, we obtain:

$$D \sim e^{i\hbar^{-1}\phi(y=(z_1-\tilde{s},z'),\tilde{p}=p)} a\Big(s,x,y=(z_1-\tilde{s},z'),p,\sigma\Big)\ \tilde{a}\Big(\tilde{s},y=(z_1-\tilde{s},z'),z,p,\tilde{\sigma}\Big)$$

as $\hbar \to 0$. Stationary phase in fact gives us a complete asymptotic expansion in increasing powers of $\hbar$, with coefficients derivatives of $a$ and $\tilde{a}$ evaluated at the critical points. It suffices to consider the contribution to $\omega$ of the leading term in the expansion of $D$ (the other terms are treated in the same manner):

$$\omega_0(x,z) := \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}[(x'-z')p' + (x_1-z_1-s-\tilde{s})p_1] + is\sigma + i\tilde{s}\tilde{\sigma}}$$
$$a(s,x,(z_1-\tilde{s},z'),p,\sigma)\,\tilde{a}(\tilde{s},(z_1-\tilde{s},z'),z,p,\sigma)\,ds d\tilde{s} dp d\sigma d\tilde{\sigma}.$$

Making the change of variables $t = s + \tilde{s}$, we can write

$$\omega_0(x,z) = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}[(x_1-z_1-t)p_1+(x'-z')p']+it\tilde{\sigma}} b(t,x,z,p,\tilde{\sigma}) dt dp d\tilde{\sigma},$$

where

$$b(t,x,z,p,\tilde{\sigma}) = \int e^{is(\sigma-\tilde{\sigma})} a(s,x,(z_1-t+s,z'),p,\sigma)\, \tilde{a}(t-s,(z_1-t+s,z'),z,p,\tilde{\sigma})\, ds d\sigma.$$

Next, we split $\omega_0$ in three parts, as follows. Let $\chi_k = \chi_k(\sigma,\tilde{\sigma})$, $k = 1,2$ be smooth, classical symbols of degree zero such that

$$\chi_1 = \begin{cases} 1 & \text{for } |\sigma| \leq \frac{1}{2}\epsilon|\tilde{\sigma}| \\ 0 & \text{for } |\sigma| \geq \epsilon|\tilde{\sigma}| \end{cases} \quad \text{and} \quad \chi_2 = \begin{cases} 1 & \text{for } |\tilde{\sigma}| \leq \frac{1}{2}\epsilon|\sigma| \\ 0 & \text{for } |\tilde{\sigma}| \geq \epsilon|\sigma| \end{cases}$$

for $\epsilon << 1$. We let

$$\Upsilon_k(x,z) = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}[(x_1-z_1-t)p_1+(x'-z')p']+it\tilde{\sigma}} b(t,x,z,p,\tilde{\sigma})\, \chi_k\, dt dp d\tilde{\sigma} \quad k = 1,2,$$

and

$$\Upsilon_0 = \omega_0 - \Upsilon_1 - \Upsilon_2.$$

We will show that $\Upsilon_k$, $k = 1,2$ is a semiclassical state in the flow out while $\Upsilon_0$ is an integral of the form (6.1). Let us rewrite, for $k = 1,2$,

$$\Upsilon_k = \frac{1}{(2\pi\hbar)^n} \int e^{\hbar^{-1}(x'-z')p'} \left( \int e^{i\hbar^{-1}(x_1-z_1-t)p_1} C_k(x,z,t,p)\, dp_1\, dt \right) dp'$$

where

$$C_k(x,z,t,p) := \int e^{it\tilde{\sigma}+is(\sigma-\tilde{\sigma})} \chi_k a\, \tilde{a}\, ds d\sigma d\tilde{\sigma},$$

interpreted as an iterated integral (first with respect to $s$). On the support of $\chi_1$ one

has $|\sigma| \leq \epsilon|\tilde{\sigma}|$ which implies $|\sigma - \tilde{\sigma}| \geq |\tilde{\sigma}| - |\sigma| \geq (1-\epsilon)|\tilde{\sigma}|$, and therefore

$$\frac{1}{|\sigma - \tilde{\sigma}|^N} \leq \frac{1}{(1-\epsilon)^N} \frac{1}{|\tilde{\sigma}|^N} \leq \left(\frac{\epsilon}{1-\epsilon}\right)^N \frac{1}{|\sigma|^N}$$

for all $N > 0$. Since $a \cdot \tilde{a}$ vanish near $\tilde{\sigma} = 0 = \sigma$ and $\chi_1$ vanishes in a conic neighborhood of the diagonal, we can integrate by parts repeatedly to obtain

$$C_1(x, z, t, p) = \int e^{it\tilde{\sigma} + is(\sigma - \tilde{\sigma})} \frac{\chi_1}{(-i)^N (\sigma - \tilde{\sigma})^N} D_s^N(a\tilde{a}) ds d\sigma d\tilde{\sigma}.$$

Since $D_s^N(a\tilde{a})$ is of the same order in $\sigma, \tilde{\sigma}$ and $N$ is arbitrary, $C_1$ is Schwartz in the variable $t$. Applying once again the method of stationary phase this implies that

$$\int e^{i\hbar^{-1}(x_1 - z_1 - t)p_1} C_1(x, z, t, p)\, dp_1\, dt$$

is a semiclassical symbol and therefore $\Upsilon_1$ is as desired. The proof for $\Upsilon_2$ is analogous. To show that $\Upsilon_0$ is as desired we only need to show that

$$\begin{aligned} b_0(t, x, z, p, \tilde{\sigma}) = \int & e^{is(\sigma - \tilde{\sigma})} a(s, x, (z_1 - t + s, z'), p, \sigma) \\ & \tilde{a}(t - s, (z_1 - t + s, z'), z, p, \tilde{\sigma})\ (1 - \chi_1(\sigma, \tilde{\sigma}) - \chi_2(\sigma, \tilde{\sigma}))\ ds d\sigma. \end{aligned}$$

is a classical symbol in $\tilde{\sigma}$ of order $m' + \tilde{m}'$. Making the change of variables $\tau = \frac{\sigma}{\tilde{\sigma}}$, we obtain:

$$\begin{aligned} b_0 = \tilde{\sigma} \int & e^{is\tilde{(}\tau - 1)} a(s, x, (z_1 - t + s, z'), p, \tilde{\sigma}\tau) \\ & \tilde{a}(t - s, (z_1 - t + s, z'), z, p, \tilde{\sigma})(1 - \chi_1(\tilde{\sigma}\tau, \tilde{\sigma}) - \chi_2(\tilde{\sigma}\tau, \tilde{\sigma})) ds d\tau. \end{aligned}$$

Notice that $\tau$ is bounded in the support of the amplitude, and so using stationary phase in $\tilde{\sigma}$, we get an expansion for $b_0$ in terms of $\tilde{\sigma}$, where the leading term is

$$b \sim 2\pi a_{m'}(0, x, (z_1 - t, z'), p, \tilde{\sigma})\ \tilde{a}_{\tilde{m}'}(t, (z_1 - t, z'), z, p, \tilde{\sigma}),$$

and we conclude $b$ is a classical symbol in $\tilde{\sigma}$ of degree $m' + \tilde{m}'$. The symbol of $\omega(x, z)$ in the diagonal is easily computed as:

$$\begin{aligned}\sigma_0(\omega)(x, z = x) &= 2\pi b_{m'+\tilde{m}'}(t = 0, x, z = x, p, p_1 = \tilde{\sigma}) \\ &= (2\pi)^2 a_{m'}(0, x, z = x, p, \tilde{\sigma} = p_1)\tilde{a}_{\tilde{m}'}(t = 0, z = x, z = x, p, \tilde{\sigma} = p_1) \\ &= \sigma_0(u)(x, z = x)\sigma_0(v)(z = x, x).\end{aligned}$$

The symbol on the flow-out is the sum of the principal symbols of each summand at $(t = z_1 - x_1, x' = z', z_1 - t + s = x_1 + s)$, which after the cancellation of $\chi_1, \chi_2$, reduces to

$$\begin{aligned}&\sqrt{2\pi}\int e^{it\tilde{\sigma}}e^{is(\sigma-\tilde{\sigma})}a(s, x, (x_1 + s, x'), (0, p'), \sigma) \\ &\qquad\qquad \tilde{a}(t - s, (x - 1 + s, x'),(z_1, x'), (0, p'), \tilde{\sigma})dsd\sigma d\tilde{\sigma} \\ &= \frac{1}{\sqrt{2\pi}}\int \left(\sqrt{2\pi}\int e^{is\sigma}a(s, x, (x_1 + s, x'), (0, p'), \sigma)d\sigma\right) \\ &\qquad\qquad \left(\sqrt{2\pi}\int e^{i(t-s)\tilde{\sigma}}\tilde{a}(t - s, (x_1 + s, x'), \Big(z_1, x'), (0, p'), \tilde{\sigma})d\tilde{\sigma}\Big)ds \\ &= \frac{1}{\sqrt{2\pi}}\int \sigma_1(u)(x, (x_1 + s, x'), (0, p'))\sigma_1(v)((x_1 + s, x'),(z_1, x'), (0, p'))ds.\end{aligned}$$

□

## 6.3 Conclusions

We provided an oscillatory integral form for elements in the class $J^{\ll,m}$, where the amplitudes have asymptotic expansions in $\hbar$ and each coefficient is a polyhomogeneous classical symbol in one of the variables. The symbol maps were described in the model case, observing that the two principal symbols have singularities on the intersection of the diagonal and flow-out. In the last part of this chapter we proved that the class forms an algebra and describe the principal symbols of the composition of elements in the algebra.

# CHAPTER VII

# The manifold case

## 7.1 The spaces $J^{\ell,m}$ on manifolds

In this section we extend the definition of the spaces $J^{\ell,m}$ to the manifold case. Let $M$ be a $C^\infty$ manifold of dimension $n$, and $X_c \subset T^*M$ be a compact bounded domain with smooth boundary contained in the cotangent bundle. The boundary $\partial X_c$ is then foliated by curves tangent to the kernel of the pull-back of the symplectic form. In addition, we assume that the fibrating and Bohr-Sommerfeld conditions are satisfied, i.e., the leaves of the null foliation are the fibers of a submersion $\pi : \partial X_c \to S$, and for each leaf $\gamma \subset \partial X_c$,

$$\int_\gamma \alpha \in 2\pi\mathbb{Z}, \tag{7.1}$$

where $\alpha$ is the tautological one-form in $T^*M$.

**Definition VII.1.** Given $X_c$ as above, define the flow-out

$$\mathcal{F}\partial X_c := \left\{(\overline{x},\overline{y}) \;\middle|\; \overline{x},\overline{y} \in \partial X_c,\ \overline{x},\overline{y} \text{ are in the same leaf}\right\} \subset T^*\left(M \times M\right). \tag{7.2}$$

The condition that $\partial X_c$ be fibrating easily implies that $\mathcal{F}\partial X$ is a closed submanifold of $T^*\left(M \times M\right)$.

**Lemma VII.2.** *There exists an open neighborhood $\partial X_c \subset U \subset T^*M$ of $\partial X_c$, and a map*

$$P : U \to \mathbb{R} \tag{7.3}$$

*such that zero is a regular value of $P$, $\partial X_c = P^{-1}(0)$, the orbits of the Hamilton flow generated by $P$ are $2\pi$ periodic in the boundary $\partial X_c$ and coincide with the leaves of the foliation.*

*Remark* VII.3. From now on we will fix a defining function of $X_c$, $P$, with the properties of this Lemma.

*Proof.* There exists $U$ and $F : U \to \mathbb{R}$ such that $\partial X_c = F^{-1}(0)$, where zero is a regular value. The Hamiltonian vector field $\Xi_F$ is tangent to $F^{-1}(0) = \partial X_c$, and therefore the trajectories of the Hamiltonian $F$ coincide set-theoretically with the leaves of the foliation. In particular, they are periodic. For each $\overline{x} \in \partial X_c$, let $T(\overline{x})$ denote the minimal period of the trajectory through $\overline{x}$. The fibrating condition can be seen to imply that the function $T$ is smooth. Notice that the fibers of the foliation are circles. For $\overline{x} \in X_c, s = \pi(\overline{x})$, we take an small neighborhood $U$ os s in $S$, and a local trivialization

$$\pi^{-1}(U) \cong U \times S^1,$$

and we put coordinates $x_1, \ldots, x_{2n-2}, \theta$ on $U \times S^1$. The restriction of the vector field $\Xi_F$ of $F$ to $\pi^{-1}(U)$ in the coordinates above is of the form

$$\Xi_F = G(x_1, \ldots, x_{2n-2}, \theta)\partial_\theta,$$

where the function $G(x_1, \ldots, x_{2n-2}, \theta)$ never vanishes since zero is a regular value of $F$. The evolution in $\theta$ is then given by

$$\theta'(t) = G(x_1, \ldots, x_{2n-2}, \theta(t)),$$

which implies

$$t = \int_0^{\theta(t)} \frac{d\theta}{G(x_1, \ldots, x_{2n-2}, \theta)}.$$

The period of the trajectories is then a smooth functions given by

$$T(x_1, \ldots, x_{2n-2}) = \int_0^{2\pi} \frac{d\theta}{G(x_1, \ldots, x_{2n-2}, \theta)}.$$

Extend this function to a smooth function $T : U \to \mathbb{R}^+$. The defining function that satisfies the conclusions of the lemma is then

$$P = \frac{T}{2\pi}\ F$$

□

*Remark* VII.4. Notice that the boundary $\partial X_c$ of $X_c$ may not be connected. In those cases, the flow-out $\mathcal{F}\partial X_c$ consist of the union of symplectic manifolds that do not intersect with each other.

Let $\Delta \subset T^*(M \times M)$ be the diagonal in $T^*(M \times M)$.

**Lemma VII.5.** *The diagonal and the flow-out $(\Delta, \mathcal{F}\partial X_c)$ intersect cleanly.*

*Proof.* Let $(\overline{x}, \overline{x}) \in \Delta \cap \mathcal{F}\partial X_c$. It is easy to show that

$$T_{(\overline{x},\overline{x})}\mathcal{F}\partial X_c = \left\{(\delta\overline{x} + r\Xi_P(\overline{x}), \delta\overline{x}) \middle| \delta\overline{x} \in T_{\overline{x}}\partial X_c, r \in \mathbb{R}\right\},$$

and so

$$T_{(\overline{x},\overline{x})}\Delta \cap T_{(\overline{x},\overline{x})}\mathcal{F}\partial X_c = \left\{(\delta\overline{x}, \delta\overline{x}) \middle| \delta\overline{x} \in T_{\overline{x}}\partial X_c\right\} = T_{(\overline{x},\overline{x})}\left(\Delta \cap \mathcal{F}\partial X_c\right)$$

□

By an analogue to proposition 2.1 in [GU81], there exists a locally finite covering $\{U_i\}_i$ of $\Delta \cap \mathcal{F}\partial X_c$, $U_i \subset T^*(M \times M)$ open and contractible, where each $U_i$ intersects only one connected component of $\mathcal{F}\partial X_c$, and for each $U_i$ a canonical transformation

$$\chi_i : U_i \longrightarrow T^*\mathbb{R}^{2n}$$

mapping $U_i \cap \Delta$ to $\Delta^n$, and $U_i \cap \mathcal{F}\partial X_c$ to $\mathcal{F}\partial X_c^n$, where $\Delta^n, \mathcal{F}\partial X_c^n$ are the diagonal and flow-out in the model case, respectively.

**Definition VII.6.** Let $\Sigma = \Delta \cap \mathcal{F}\partial X_c$ be the intersection of the diagonal and the flow-out. The space $J^{\ell,m}(M \times M; \Delta, \mathcal{F}\partial X_c)$ is the set of families of functions $A(x, y, \hbar)$ which can be written in the form

$$A = A_0 + A_1 + \sum \omega_i$$

where:

1. $A_0 \in \mathrm{sc}\text{-}I^{\ell+m}(M \times M; \Delta \setminus \Sigma)$ and $A_1 \in \mathrm{sc}\text{-}I^{\ell}(M \times M; \mathcal{F}\partial X_c)$, and
2. $\omega_i$ is microlocally supported in $U_i$ and is of the form

$$\omega_i = F_i(v_i),$$

where $v_i \in J^{\ell,m}(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n, \mathcal{F}\partial X_c^n)$ and $F_i$ is a semiclassical Fourier integral operator associated to $\chi_i^{-1}$.

*Remark* VII.7. As in [GU81], one can show that the definition does not depend on the choice of the semiclassical FIOs $F_i$.

**Proposition VII.8.** *In the general case the conclusion of Proposition VI.4 still holds, namely: Operators with kernel $A \in J^{\ell,m}(M \times M;\ \Delta, \mathcal{F}\partial X_c)$ have frequency set con-*

*tained in the union*

$$FS(A) \subset \Delta' \cup \mathcal{F}\partial X_c'.$$

*Away from the intersection* $\Sigma$ *they are microlocally semiclassical Fourier integral operators* $sc{-}I^{\ell+m}(M \times M; \Delta \setminus \Sigma)$ *and* $sc{-}I^{\ell}(M \times M, \mathcal{F}\partial X_c \setminus \Sigma)$ *respectively. Furthermore, one has well-defined symbol maps (ignoring Maslov factors)*

$$J^{\ell,m}(M \times M;\ \Delta, \mathcal{F}\partial X_c) \begin{array}{l} \sigma_0 \nearrow \quad |\wedge|^{1/2}(\Delta \setminus \Sigma) \\ \\ \sigma_1 \searrow \quad |\wedge|^{1/2}(\mathcal{F}\partial X_c \setminus \Sigma). \end{array} \tag{7.4}$$

The proof of the existence of the symbol calculus is the subject of the following section.

Our construction, in particular, gives a way to associating to $X_c$ an algebra of operators, which can be thought of as a quantization of $X_c$:

**Definition VII.9.** We will denote by $\mathcal{A}_{X_c}$ the space of elements in $J(M \times M; \Delta, \partial X_c)$ that are microlocally of order $\mathrm{O}(\hbar^\infty)$ in the complement of $X_c$.

It is not hard to see that $\mathcal{A}_{X_c}$ is indeed closed under composition. (Elements in this algebra correspond to amplitudes that are of order $-\infty$ in $\sigma$ as $\sigma \to -\infty$.)

## 7.2 Symbolic Calculus

Here we discuss how Proposition VII.8 can be proved (for $\hbar$ tending to zero along certain sequences) using the methods from [PU95]. We begin by recalling the ideas and results from *op. cit.* that we will need.

The pre-quantum circle bundle of $T^*M$ can be identified with the following submanifold of $T^*(M \times S^1)$:

$$Z = \{(x, \theta; \xi, \kappa) \in T^*(M \times S^1)\ ;\ \kappa = 1\},$$

with the obvious circle action. The connection form, $\alpha$, is the pull-back to $Z$ of the canonical one form of $T^*(M \times S^1)$.

**Definition VII.10.** ([PU95]) A Lagrangian submanifold $\Lambda \subset T^*M$ will be called *admissible* iff there exists a conic Lagrangian submanifold,

$$\tilde{\Lambda} \subset T^*(M \times S^1) \cap \{\kappa > 0\}$$

such that

$$\Lambda = \left(\tilde{\Lambda} \cap Z\right)/S^1. \tag{7.5}$$

We call such a $\tilde{\Lambda}$ a *homogenization* of $\Lambda$.

It is not hard to see that $\Lambda$ is admissible if and only if it satisfies the following Bohr-Sommerfeld condition: There exists $\phi : \Lambda \to S^1$ such that

$$\iota^*\eta = d \log \phi,$$

where $\iota : \Lambda \to T^*M$ is the inclusion and $\eta$ the canonical one form of $T^*M$. Given such a $\phi$, a homogenization of $\Lambda$ can be defined by:

$$\tilde{\Lambda} = \{e^{i\theta} = \phi(\lambda),\ \kappa > 0\}. \tag{7.6}$$

**Definition VII.11.** Let $M$ be a $C^\infty$ manifold and consider a family of smooth functions $\{\psi_\hbar\}$. The $\hbar$ *transform* of the family $\psi_\hbar$ is the following distribution (if the series converges weakly) in $M \times S^1$:

$$\Psi(x, \theta) = \sum_{m=0}^{\infty} \psi_{1/m}(x) e^{im\theta}.$$

The main point of the previous two definitions is the following

**Lemma VII.12.** *([PU95]) The $\hbar$ transform of a semiclassical state associated to an admissible Lagrangian is a Lagrangian distribution associated to a homogenization of the Lagrangian submanifold.*

We claim that the previous lemma generalizes to our spaces $J^{\ell,m}$ of semiclassical pseudodifferential operators with singular symbols, so that the $\hbar$ transform of their kernels are Guillemin-Uhlmann operators (i. e. those defined in [GU81]) on $M \times S^1$.

Thus we now consider $\Delta$ to be the diagonal in $T^*(M \times M)$ and $\mathcal{F}\partial X_c$ as in (7.2). Clearly a homogenization for $\Delta$ is the diagonal $\widetilde{\Delta}$ of $T^*(M \times S^1)^+ = T^*(M \times S^1) \cap \{\kappa > 0\}$. Let $P$ be a globally defining function for $\partial X_c$ with periodic Hamilton flow on it, as in (7.3). Then a homogenization for $\mathcal{F}\partial X_c$ is the flow-out of the homogenization of $P$, which is the function $\tilde{P} : T^*(U \times S^1)^+ \to \mathbb{R}$ defined as

$$\widetilde{P}(x, e^{i\theta}; \xi, \kappa) = \kappa P(x, \xi/\kappa),$$

where $U$ is the neighborhood of $\partial X_c$ described in Lemma VII.2. Specifically,

$$\widetilde{\mathcal{F}\partial X_c} = \left\{ \left( \Phi_s^{\widetilde{P}}(x, e^{i\theta}, \xi, \kappa) \ ; \ (x, e^{i\theta}, \xi, \kappa) \right) \mid s \in \mathbb{R}, \kappa > 0, \ P(x, \xi/\kappa) = 0 \right\} \qquad (7.7)$$

where $\Phi_s^{\widetilde{P}}$ is the Hamilton flow generated by the equations (here $p = \xi/\kappa$)

$$\dot{x} = \tfrac{\partial P}{\partial p}(x, p), \qquad \dot{\theta} = -p \ \tfrac{\partial P}{\partial p}(x, p)$$

$$\dot{\xi} = -\kappa \tfrac{\partial P}{\partial x}(x, p), \qquad \dot{\kappa} = 0.$$

Notice that $\widetilde{\mathcal{F}\partial X_c}$ is closed if the flow $\Phi_t^P$ is $2\pi$ periodic, which happens if the Bohr-Sommerfeld condition

$$\int_\gamma p dx \in 2\pi\mathbb{Z}$$

is satisfied for orbits $\gamma \subset \partial X_c$ of $P$. The function $\phi : \mathcal{F}\partial X_c \to S^1$ associated to this homogenization is

$$\phi(\overline{x}, \overline{y}) = e^{-i\int_{\overline{x}}^{\overline{y}} p dx},$$

where $\int_{\overline{x}}^{\overline{y}} p dx$ is the action from $\overline{x}$ to $\overline{y}$, and the integral is taken on the curve in the leaf joining $\overline{x}$ to $\overline{y}$.

## 7.3 The existence of the symbolic calculus.

Let $M$ be a $C^\infty$ manifold. The semiclassical states that will be defined in this chapter will have their frequency set contained in the union of two Lagrangian submanifolds of the cotangent bundle $T^*M$, not necessarily conic. This is the equivalent notion of *intersecting pairs* in [GU81] and [MU79].

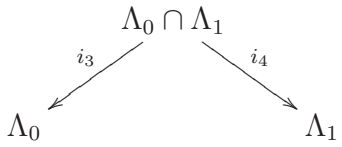
**Proposition VII.13.** *Let $(\Lambda_0, \Lambda_1)$ be two admissible Lagrangian submanifolds of $T^*M$. Then the two manifolds intersect cleanly in a codimension $k$ submanifold, if and only if there exist conic horizontal lifts $\widetilde{\Lambda_0}, \widetilde{\Lambda_1} \subset T^*(M \times S^1) \setminus 0$ such that $\left(\widetilde{\Lambda_0}, \widetilde{\Lambda_1}\right)$ form an intersecting pair in $M \times S^1$, i.e., they are conic, and intersect cleanly in a codimension $k$ submanifold.*

*Proof.* Suppose $(\Lambda_0, \Lambda_1)$ is a semiclassical intersecting pair. Since the two Lagrangian submanifolds are admissible, for $j = 1, 2$ there exist functions $\phi_j : \Lambda_j \mapsto S^1$ such that $i_j^* \eta = dlog\phi_j$ where $i_j : \Lambda_j \hookrightarrow T^*M$ are the inclusions maps. Notice that the $\phi_j$ is unique up to multiplication by constants. In that case,

$$\widetilde{\Lambda_j} = \left\{(x, e^{i\theta}; p, \kappa) \mid \kappa > 0, (x, p) \in \Lambda_j \text{ and } e^{i\theta} = \phi_j(x, p)\right\}$$

are closed conic Lagrangian submanifolds od $T^*(M \times S^1)^+$ satisfying (7.5). Consider

now the inclusions




Since $dlog\phi_j = i_j^*\eta$, then $d\left(i_3^* log\phi_1\right) = i_3^* dlog\phi_1 = i_4^* dlog\phi_2 = d\left(i_4^* log\phi_2\right)$. So, $\phi_1$ is a scalar (non-zero) multiple of $\phi_2$ on each connected component of $\Lambda_0 \cap \Lambda_1$. WLOG, we can choose the constants so that $\phi_1 = \phi_2$ on the intersection $\Lambda_0 \cap \Lambda_1$. The intersection of the horizontal lifts is then:

$$\widetilde{\Lambda_0} \cap \widetilde{\Lambda_1} = \left\{(x, e^{i\theta}; \kappa p, \kappa) \,|\, \kappa > 0 \ , \ (x,p) \in \Lambda_0 \cap \Lambda_1 \ , \ e^{i\theta} = \phi_1(x,p) = \phi_2(x,p)\right\} \quad (7.8)$$

Let us define $\Sigma = \Lambda_0 \cap \Lambda_1$, and $\widetilde{\Sigma} = \widetilde{\Lambda}_0 \cap \widetilde{\Lambda}_1$. Since $\phi_1 = \phi_2$ on the intersection of the two admissible Lagrangians, $\widetilde{\Sigma}$ is a $C^\infty$ submanifold of $T^*\left(M \times S^1\right) \backslash 0$ of codimension $k$ if and only if $\Sigma$ is a submanifold of $T^*M$ of codimension $k$.

Now consider

$$f : T^*\left(M \times S^1\right)^+ \longrightarrow T^*M \times \mathbb{R}^+ \quad (7.9)$$

$$(x, \theta \ ; \ \xi, \kappa) \longmapsto (x, \tfrac{\xi}{\kappa}; \kappa)$$

Clearly, $f\big|_{\widetilde{\Lambda_j}} : \widetilde{\Lambda_j} \to \Lambda_j \times \mathbb{R}^+$ is a diffeomorphism. Also, $f\left(\widetilde{\Lambda_0} \cap \widetilde{\Lambda_1}\right) = (\Lambda_0 \cap \Lambda_1) \times \mathbb{R}^+$. Then, in case $\widetilde{\Lambda_0} \cap \widetilde{\Lambda_1}$ and $\Lambda_0 \cap \Lambda_1$ are both manifolds, and $\widetilde{\omega} = (\omega; \phi_1(\omega), \kappa) \in \widetilde{\Lambda_0} \cap \widetilde{\Lambda_1}$, $\omega \in \Lambda_0 \cap \Lambda_1$, we have

$$T_{\widetilde{\omega}} \widetilde{\Lambda_j} = df_{\widetilde{\omega}}^{-1}\left(T_\omega \Lambda_j \times \mathbb{R}\right),$$

and

$$T_{\widetilde{\omega}}\left(\widetilde{\Lambda} \cap \widetilde{\Lambda_1}\right) = df_{\widetilde{\omega}}^{-1}\left(T_\omega\left(\Lambda_0 \cap \Lambda_1\right) \times \mathbb{R}\right)$$

Therefore, $\widetilde{\Lambda_0}$ and $\widetilde{\Lambda_1}$ intersects cleanly if and only if $\Lambda_0$ and $\Lambda_1$ intersects cleanly, as desired. $\square$

Notice that $\Delta$ and $\mathcal{F}\partial X_c$ do intersect cleanly. As a result, their homogenizations $\widetilde{\Delta}$ and $\widetilde{\mathcal{F}\partial X_c}$ are conic and intersect cleanly too, forming an intersecting pair, in the sense of [GU81, MU79]. The relationship between the semiclassical objects defined above and the operators described in [GU81] is as follows:

**Theorem VII.14.** *Let $\Delta, \mathcal{F}\partial X_c$ be as above, and $\widetilde{\Delta}, \widetilde{\mathcal{F}\partial X_c}$ their homogenization, respectively. Then $A \in J^{\ell,m}\left(M \times M \ ; \ \Delta, \mathcal{F}\partial X_c\right)$ if and only if its $\hbar$-Transform belongs to $I^{\ell,m}\left(M \times S^1 \times M \times S^1 \ ; \ \widetilde{\Delta}, \widetilde{\mathcal{F}\partial X_c}\right)$.*

*Proof.* We sketch the ideas of the proof in the model case. Let $A(x, y, \hbar)$ be as in equation (6.1). Let us assume first that the amplitude $\frac{1}{(2\pi\hbar)^n}a$ does not depend on $\hbar$ so that

$$A(x, y, \hbar) = \int e^{i\hbar^{-1}[(x_1 - y_1 - s)p_1 - (x' - y')p'] + is\sigma} a(s, x, y, p, \sigma) ds dp d\sigma,$$

where $a(s, x, y, p, \sigma)$ is a classical symbol in $\sigma$ of order $m' = m - 1/2$, and compactly supported in $s, x, y, p$. The $\hbar$-transform of $A$ is then

$$A(x, \theta, y, \alpha) = \sum_{k=1}^{\infty} \int e^{ik[(x_1 - y_1 - s)p_1 + (x' - y')p' + (\theta - \alpha)] + is\sigma} a(s, x, y, p, \sigma) ds dp d\sigma$$

$$= \int \frac{e^{i[(x_1 - y_1 - s)p_1 + (x' - y')p' + (\theta - \alpha)]}}{1 - e^{i[(x_1 - y_1 - s)p_1 + (x' - y')p' + (\theta - \alpha)]}} e^{is\sigma} a(s, x, y, p, \sigma) ds dp d\sigma.$$

This distribution is the push-forward under the projection $(s, x, y, \theta, \alpha, p) \to (x, y, \theta, \alpha)$ of the product of the distributions

$$\Gamma(s, x, y, \theta, \alpha, p) = \frac{e^{i[(x_1 - y_1 - s)p_1 + (x' - y')p' + (\theta - \alpha)]}}{1 - e^{i[(x_1 - y_1 - s)p_1 + (x' - y')p' + (\theta - \alpha)]}}$$

and

$$\Upsilon(s, x, y, \theta, \alpha, p) = \int e^{is\sigma} a(s, x, y, p, \sigma) d\sigma.$$

$\Gamma$ is a distribution in space conormal to the hypersurface $(x_1-y_1-s)p_1+(x'-y')p' = -(\theta-\alpha)$, while $\Upsilon$ is a distribution conormal to $s=0$. Therefore, $A(x,\theta,y,\alpha)$ is a Guillemin-Uhlmann distribution associated to the pair $\left(\widetilde{\Delta^n}, \widetilde{\mathcal{F}\partial X_c^n}\right)$. The general case (i.e. when $a$ also depends on $\hbar$) is an asymptotic sum of derivatives and integrals (with respect to the $\theta$ variables) of the previous case. The converse is also true, and the proof is analogue to that in [PU95] for semiclassical states. $\square$

This proposition together with Proposition 2.4 of [PU95], relating the frequency set of an $\hbar$-dependent vector and the wave-front set of its $\hbar$ transform, implies part (1) of Theorem V.7.

The symbols of operators in $J^{\ell,m}(\Delta, \mathcal{F}\partial X_c)$ are the reduction of the symbols of its $\hbar$-Transform, in the following sense. Let $\overline{x},\overline{y} \in \mathcal{F}\partial X_c \setminus \Sigma$, $\overline{x} = \Phi_s^P(\overline{y})$ for some $s \in \mathbb{R}$, and take

$$(\widetilde{x},\widetilde{y}) = \left(\Phi_s^{\widetilde{P}}(\overline{y}, e^{i\theta=0}=1, \kappa=1), (\overline{y}, e^{i\theta=0}=1, \kappa=1)\right) \in \widetilde{\mathcal{F}\partial X_c^n}.$$

Using (7.7), we obtain an isomorphism between $T_{(\widetilde{x},\widetilde{y})}\widetilde{\mathcal{F}\partial X_c}$ and $T_{(\overline{x},\overline{y})}\mathcal{F}\partial X_c \times \mathbb{R}\times\mathbb{R}$, which leads to

$$\left|T_{(\widetilde{x},\widetilde{y})}\widetilde{\mathcal{F}\partial X_c}\right|^{1/2} \cong \left|T_{(\overline{x},\overline{y})}\mathcal{F}\partial X_c \times \mathbb{R}\times\mathbb{R}\right|^{1/2} \cong \left|T_{(\overline{x},\overline{y})}\mathcal{F}\partial X_c\right|^{1/2} \tag{7.10}$$

Therefore, every half-density in $T_{\widetilde{x},\widetilde{y}}\widetilde{\mathcal{F}\partial X_c}$ will define a half-density in $T_{(\overline{x},\overline{y})}\mathcal{F}\partial X_c$. Let $\widetilde{\Sigma} = \widetilde{\mathcal{F}\partial X_c} \cap \widetilde{\Delta}$. For each family $A \in J^{\ell,m}$, denote its $\hbar$-Transform by $\widetilde{A}$. There is a well defined symbol map

$$\widetilde{\sigma_1} = \sigma_1\left(\widetilde{A}\right)_{\widetilde{\mathcal{F}\partial X_c}-\widetilde{\Sigma}} \in C^\infty\left(\widetilde{\mathcal{F}\partial X_c} \setminus \widetilde{\Sigma}, \widetilde{\Omega_1} \bigotimes \widetilde{L_1}\right),$$

where $\widetilde{\Omega}_j$ is the bundle of half-densities on $\widetilde{\mathcal{F}\partial X_c}$, and $\widetilde{L}_i$ is the corresponding Maslov bundle. Ignoring Maslov factors, we can define the symbols on $\mathcal{F}\partial X_c \setminus \Sigma$ by the

identification (7.10), i.e., restricting the symbol of the $\hbar$-Transform to $\theta = 0$, $\kappa = 1$ ($\sigma_1 = \widetilde{\sigma_1}\big|_{\theta=0,\kappa=1}$). The construction of the principal symbol on the diagonal is similar.

Let $\Sigma^n = \Delta^n \cap \mathcal{F}\partial X_c^n$, $\widetilde{\Sigma}^n = \widetilde{\Delta^n} \cap \widetilde{\mathcal{F}\partial X_c^n}$. In the model case, we know by [GU81] that $A(x, \theta, y, \alpha)$ is microlocally in

$$I^{\ell+m}\left(\mathbb{R}^n \times S^1 \times \mathbb{R}^n \times S^1 \,;\, \widetilde{\Delta^n} \setminus \widetilde{\Sigma}^n\right), \text{ and } I^{\ell}\left(\mathbb{R}^n \times S^1 \times \mathbb{R}^n \times S^1 \,;\, \widetilde{\mathcal{F}\partial X_c^n} \setminus \widetilde{\Sigma}^n\right).$$

Therefore $A_\hbar$ is microlocally in sc-$I^{\ell+m}\left(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n \setminus \Sigma^n\right)$ and sc-$I^{\ell}\left(\mathbb{R}^n \times \mathbb{R}^n; \mathcal{F}\partial X_c^n \setminus \Sigma^n\right)$ in the sense that the $\hbar$-transform is microlocally in their corresponding spaces. We have proved the following

**Proposition VII.15.**

$$\cap_\ell J^{\ell,m} = sc\text{-}I^{\infty}\left(\mathbb{R}^n \times \mathbb{R}^n; \Delta^n\right)$$

*and*

$$\cap_m J^{\ell,m} = sc\text{-}I^{\ell}\left(\mathbb{R}^n \times \mathbb{R}^n; \mathcal{F}\partial X_c^n\right).$$

## 7.4 A symbolic compatibility condition

Recall that the foliation of $\partial X_c$ is fibrating, i. e., there exists a $C^\infty$ Hausdorff manifold $S$ and a smooth fiber map

$$\pi : \partial X_c \to S, \tag{7.11}$$

whose fibers are the connected leaves of the foliation defined Section 7.1. Elements of $\mathcal{F}\partial X_c$ are pairs of points in $\partial X_c$ that lie in the same fiber of $\pi$.

Generalizing a construction in [GS79], given $s \in S$ let $\mathfrak{SO}_s$ be the $*-$algebra of all pseudodifferential operators, acting on the space of half-densities $C^\infty\left(|F_s|^{1/2}\right)$, where $F_s$ is the fiber of $\pi : \partial X_c \to S$ above $s$. Let $\mathfrak{SO}$ be the sheaf of $*-$algebras on $S$

whose stalk at $s$ is $\mathfrak{SD}_s$. We will say that a section $k$ of $\mathfrak{SD}$ is smooth if the Schwartz kernel of the operator $k(s)$ depends smoothly on $s \in S$. The Schwartz kernel theorem, applied fiber-wise to the fibers of $\pi$, together with the natural symplectic structure of $S$ yield the following:

**Proposition VII.16.** *The vector space of smooth sections of the sheaf $\mathfrak{SD}$ is naturally isomorphic to the space of half-densities on $\mathcal{F}\partial X_c \setminus \Sigma$ that extend to $\mathcal{F}\partial X_c$ as a conormal distribution to $\Sigma$*

*Proof.* Let $\gamma = (\gamma_1, \gamma_2) \in \mathcal{F}\partial X_c$, $s = \pi(\gamma_1) = \pi(\gamma_2) \in S$. One then has the following fiber product diagram:

$$\begin{array}{ccc} T_\gamma \mathcal{F}\partial X_c & \xrightarrow{d\pi_1} & T_{\gamma_1}\partial X_c \\ {\scriptstyle d\pi_2}\downarrow & & \downarrow{\scriptstyle d\pi} \\ T_{\gamma_2}\partial X_c & \xrightarrow{d\pi} & T_s S \end{array}$$

and so we get the exact sequence

$$\begin{array}{ccccccc} 0 & \longrightarrow & T_\gamma \mathcal{F}\partial X_c & \longrightarrow & T_{\gamma_1}\partial X_c \oplus T_{\gamma_2}\partial X_c & \longrightarrow & T_s S \longrightarrow 0 \\ & & v & \longmapsto & (d\pi_1 v, d\pi_2 v) & & \\ & & & & (v_1, v_2) & \longmapsto & d\pi(v_2) - d\pi(v_1) \end{array}$$

This gives a natural identification

$$\mathbb{C} \cong |T_\gamma \mathcal{F}\partial X_c|^{1/2} \otimes |T_{\gamma_1}\partial X_c \oplus T_{\gamma_2}\partial X_c|^{-1/2} \otimes |T_s S|^{1/2} .$$

as follows. For each half-densities $\mathfrak{d}_1 \in \left|T_\gamma \mathcal{F}\partial X_c\right|^{1/2}$, $\mathfrak{d}_2 \in \left|T_{\gamma_1}\partial X_c \oplus T_{\gamma_2}\partial X_c\right|^{-1/2}$, $\mathfrak{d}_3 \in \left|T_s S\right|^{1/2}$ and an ordered basis $\beta_1 \in T_\gamma \mathcal{F}\partial X_c$ , take $\beta_2 = L_1\beta_1$. We can extend $\beta_2$ to a basis $\left\{\beta_2, \tilde{\beta}_2\right\}$ of $T_{\gamma_1}\partial X_c \oplus T_{\gamma_2}\partial X_c$ and define $\beta_3 = L_2\tilde{\beta}_2$, which is a basis in $T_s S$, since $\beta_2$ is in the kernel of $L_2$.

Let us define $\lambda = \mathfrak{d}_1(\beta_1)\mathfrak{d}_2(\{\beta_2, \beta_3\})\mathfrak{d}_3(\beta_4)$. We will show that this number does not depend on the choices made. Suppose we take another basis $\tilde{\tilde{\beta}}_2$ such that $L_2\tilde{\tilde{\beta}}_2 = \beta_3$,

then $\tilde{\tilde{\beta}}_2 - \beta_2 \in ker(L_2) = Im(L_1)$, so $\tilde{\tilde{\beta}}_2 - \beta_2 = A\beta_1$ for some matrix $A$. Then

$$\begin{pmatrix} \beta_2 \\ \tilde{\tilde{\beta}}_2 \end{pmatrix} = \begin{pmatrix} I & 0 \\ A & 0 \end{pmatrix} \begin{pmatrix} \beta_2 \\ \tilde{\beta}_2 \end{pmatrix},$$

where $I$ is the identity matrix. Since this matrix has determinant 1, $\lambda$ doesn't change. If we change $\beta_1$ and $\beta_3$ by $g_1\beta_1$ and $g_3\beta_3$, $\lambda$ is given by

$$\lambda = |g_1|^{1/2}\mathfrak{d}_1(\beta_1)|g_1|^{-1/2}|g_3|^{-1/2}\mathfrak{d}_2(\beta_2)|g_3|^{1/2}\mathfrak{d}_3(\beta_3) = \mathfrak{d}_1(\beta_1)\mathfrak{d}_2(\beta_2)\mathfrak{d}_3(\beta_3),$$

as desired.

Since $S$ is a symplectic manifold, there is a canonical half-density on $T_sS$, and we get an identification

$$\left|T_\gamma \mathcal{F}\partial X_c\right|^{1/2} \cong \left|T_{\gamma_1}\partial X_c \oplus T_{\gamma_2}\partial X_c\right|^{1/2} \cong \left|T_{\gamma_1}\partial X_c\right|^{1/2} \otimes \left|T_{\gamma_2}\partial X_c\right|^{1/2}.$$

Finally, given a half density in $T_{\gamma_k}\partial X_c$, $k = 1, 2$, we need to get a half-density in $T_sF_s$, where $F_s$ is the fiber above $s$. We have the following exact sequence:

$$0 \longrightarrow \ker\ d\pi_{\gamma_k} \longrightarrow T_{\gamma_k}\partial X_c \longrightarrow T_sS \longrightarrow 0$$

which, by the same process as above, gives an identification

$$\left|T_sF_s\right|^{1/2} = \left|\ker\ d\pi_{\gamma_k}\right|^{1/2} \cong \left|T_{\gamma_k}\partial X_c\right|^{1/2}$$

Hence

$$\left|T_\gamma \mathcal{F}\partial X_c\right|^{1/2} \cong \left|T_{\gamma_1}F_s\right|^{1/2} \otimes \left|T_{\gamma_2}F_s\right|^{1/2} \cong \left|T_{(\gamma_1,\gamma_2)}(F_s \times F_s)\right|^{1/2}$$

This gives a smooth isomorphism between two line bundles over $\mathcal{F}\partial X_c$: $\left|T\mathcal{F}\partial X_c\right|^{1/2}$ and the line bundle $\Upsilon \to \mathcal{F}\partial X_c$ whose fiber over $(\gamma_1, \gamma_2)$ is $\left|T_{(\gamma_1,\gamma_2)}(F_s \times F_s)\right|^{1/2}$, where

$\pi(\gamma_1) = s = \pi(\gamma_2)$. Clearly a section of the sheaf $\mathfrak{SD}$ is a distributional section of $\Upsilon$ conormal to $\Sigma$, and by the previous isomorphism this is equivalent to a distributional section of $\left|T\mathcal{F}\partial X_c\right|^{1/2}$ conormal to $\Sigma$. $\square$

The previous isomorphism yields an algebra structure on the space of distributional half densities on $\mathcal{F}\partial X_c$ which are conormal to $\Sigma$. Analogously as in [GS79] (Proposition 2.7), one can see that the algebraic structure on this space is given, away from $\Sigma$, by the composition of half densities regarded as symbols of Fourier integral operators associated to $\mathcal{F}\partial X_c$.

Let us now take $A \in J^{\ell,m}$. The symbol $\sigma_1(A)$ in $\mathcal{F}\partial X_c \setminus \Sigma$ blows-up as the point where $\sigma_1$ is evaluated tends to the intersection. In fact, in [AU85] it was shown that this symbol has a natural extension to a distribution conormal to $\Sigma$. Using the same identification above, this determines the kernel of a Pseudodifferential operator on the fiber above each point of $S$. We have proved:

**Proposition VII.17.** *For $A \in J^{\ell,m}(M \times M, \Delta, \mathcal{F}\partial X_c)$, the symbol $\sigma_1(A)$ can be identified with a global section of the sheaf $\mathfrak{SD}$, that is, with a family of classical pseudodifferential operators of order $m' = m - 1/2$ acting on fibers of $\pi : \partial X_c \to S$ (orbits in the flow-out).*

For each $s \in S$ and $F_s$ the corresponding fiber above $s$, let us denote this operator by $\sigma_1(A)_s$:

$$C^\infty(F_s) \xrightarrow{\sigma_1(A)_s} C^\infty(F_s)$$

As a result, there is a well-defined symbol

$$\sigma\left(\sigma(A)_s\right) : T^*F_s \setminus 0 \to \mathbb{C}.$$

As we will now see, this symbol is related to the symbol $\sigma_0(A)$ of $A$ on the diagonal. (This is the compatibility of the symbols of $A$ announced earlier.)

Let $P$ be a defining function of $\partial X_c$ with a $2\pi$-periodic flow on $\partial X_c$. We have the following diffeomorphism

$$\begin{aligned} \partial X_c \times S^1 &\longrightarrow \mathcal{F}\partial X_c \\ (\overline{x}, t) &\longmapsto \left(\Phi_t^P \overline{x}, \overline{x}\right) \end{aligned}$$

from which one can see that, for any $\gamma = (\overline{x}, \overline{x}) \in \Sigma = \Delta \cap \mathcal{F}\partial X_c$, there is a natural isomorphism

$$N_\Sigma^{\mathcal{F}\partial X_c} := T_\gamma \mathcal{F}\partial X_c / T_\gamma \Sigma \cong T_{\overline{x}} F_s$$

($N_\Sigma^{\mathcal{F}\partial X_c}$ is the normal space to $\Sigma$ in $\mathcal{F}\partial X_c$ at $\gamma$). Therefore, for each $\overline{x} \in \partial X_c$, $T_{\overline{x}}^* F_s$ is isomorphic to the dual space $\left(N_\Sigma^{\mathcal{F}\partial X_c}\right)^*$.

Now let $N_\Sigma^\Delta := T_\gamma \Delta / T_\gamma \Sigma$ be the normal space to $\Sigma$ in $\Delta$ at $\gamma$. By [GU81], $N_\Sigma^{\mathcal{F}\partial X_c}$ and $N_\Sigma^\Delta$ are supplementary Lagrangian subspaces of the two-dimensional symplectic vector space $W = (T_\gamma \Sigma)^\perp / T_\gamma \Sigma$. Therefore, $N_\Sigma^{\mathcal{F}\partial X_c}$ and $N_\Sigma^\Delta$ are in duality with each other (they are canonically paired by the symplectic form of $W$). In the end we obtain a natural isomorphism

$$T_{\overline{x}}^* F_s \cong \left(N_\Sigma^{\mathcal{F}\partial X_c}\right)^* \cong N_\Sigma^\Delta.$$

The symbol of $A$ on the diagonal belongs to a class $S^{m'}(\Omega_0; \Delta, \Sigma)$ of smooth functions on $\Delta \setminus \Sigma$ which blow up at a prescribed rate at $\Sigma$ (see [GU81] for more details). Every element in this class determines a smooth function on $N_\Sigma^\Delta \setminus \{0\}$. The *compatibility condition* alluded to in the Introduction is the following:

**Theorem VII.18.** *The symbol of $A \in J^{\ell,m}(M \times M, \Delta, \mathcal{F}\partial X_c)$ on the flow-out, identified with a family $\{\sigma_1(A)_s\}_{s \in S}$ of pseudodifferential operators along the fibers $F_s$, satisfies that for each $\overline{x} \in F_s$*

$$\begin{aligned} &\sigma(\sigma_1(A)_s)_{\overline{x}}(\tau) = \lim_{\substack{\overline{y} \to \overline{x} \\ \overline{y} \in X_c \setminus \partial X_c}} \frac{\sigma_0(A)(\overline{y}, \overline{y})}{P^{m'}(\overline{y})} \\ &\textit{and} \\ &\sigma(\sigma_1(A)_s)_{\overline{x}}(-\tau) = \lim_{\substack{\overline{y} \to \overline{x} \\ \overline{y} \in T^*M \setminus X_c}} \frac{\sigma_0(A)(\overline{y}, \overline{y})}{P^{m'}(\overline{y})}, \end{aligned} \tag{7.12}$$

*where $\tau \in T^*_{\overline{x}}F_s$ is the dual of the Hamilton field of $P$ at $\overline{x}$ regarded as an element in $T_{\overline{x}}F_s$, and the two limits are taken from the interior of $X_c$ and the exterior of $X_c$, respectively. Moreover, this condition is intrinsic, i.e., it does not depend on the choice of $P$.*

*Proof.* The proof reduces to the model case, where it is immediate. One can also verify directly that changing $P$ by a multiplicative factor does not alter the relationships (7.12). Let us take $P_\varphi = \varphi P$, where $\varphi$ is a nowhere vanishing smooth function. In $\partial X_c$, the flow generated by $P_\varphi$ is still periodic, with a (possibly) different period on each orbit. In addition,

$$\Xi_{P_\varphi}\big|_{\partial X_c} = \varphi \Xi_P\big|_{\partial X_c} + P\Xi_\varphi\big|_{\partial X_c} = \varphi \Xi_P\big|_{\partial X_c}.$$

Under the new Hamiltonian, both sides of equation (7.12) are divided by $\varphi^{m'}$ in each fiber. □

The symbols in the flow-out become more natural under the present setting, as can be seen in the following symbolic version of Proposition VI.7.

**Proposition VII.19.** *The composition of properly supported operators with kernels in $J^{\ell,m}$ and $J^{\tilde{\ell},\tilde{m}}$, respectively, is an operator with kernel in $J^{\ell+\tilde{\ell}+1/2,m+\tilde{m}-1/2}$. For any $A \in J^{\ell,m}$, $B \in J^{\tilde{\ell},\tilde{m}}$, we obtain the usual symbol on the diagonal:*

$$\forall \overline{x} \in X_c \setminus \partial X_c \quad \sigma_0(A \circ B)(\bar{x},\bar{x}) = \sigma_0(A)(\bar{x},\bar{x})\sigma_0(B)(\bar{x},\bar{x}),$$

*and for any fiber $F_s$ above $s \in S$,*

$$\sigma_1(A \circ B)_s = \sigma_1(A)_s \circ \sigma_1(B)_s \ .$$

## 7.5 The adjoint

The class $J^{\ell,m}$ is closed under the operation of taking adjoints, and information about the symbol of the adjoint is given in the following

**Proposition VII.20.** *Let $A \in J^{\ell,m}(M \times M; \Delta, \mathcal{F}\partial X_c)$, then the adjoint $A^*$ belongs again to $J^{\ell,m}(M \times M; \Delta, \mathcal{F}\partial X_c)$, and*

$$\sigma_0(A^*)(\overline{x}, \overline{x}) = \overline{\sigma_0(A)(\overline{x}, \overline{x})} \quad \textit{for } (\overline{x}, \overline{x}) \in \Delta \setminus \Sigma, \textit{ and}$$

$$\sigma_1(A^*)_s = (\sigma_1(A)_s)^*, \quad \textit{for each } s \in S.$$

*Proof.* The first statement is as in the usual theory of $\hbar$-pseudodifferential operators. It is enough to prove the rest in the model case. Take $A$ with Schwartz kernel

$$K(x, y, h) = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}((x-y)\cdot p - sp_1) + is\sigma} a(s, x, y, p, \sigma, \hbar) ds dp d\sigma,$$

where $a \in \mathcal{A}^{\ell',m'}_{classical}$, $\ell' = \ell + 1/2$, $m' = m - 1/2$. The Schwartz kernel of the adjoint is given by

$$K^*(x, y) = \overline{K(y, x)} = \frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}((x-y)p - sp_1) + is\sigma} \overline{a(-s, y, x, p, \sigma, \hbar)} ds dp d\sigma,$$

where $a(s, x, p, \sigma, \hbar)$ was replaced by $\overline{a(-s, y, x, p, \sigma, \hbar)}$. Taking $\overline{x} = (x, (p_1 = 0, p'))$, $\overline{y} = \phi^P_s \overline{x} = ((x_1 + s, x'), (p_1 = 0, p'))$, the symbol on the flow-out is given by

$$\sigma_1(A^*)(\phi^P_s \overline{x}, \overline{x}) = \sqrt{2\pi} \int \overline{a_{-\ell'}(-s, y, x, p, \sigma)} e^{is\sigma} d\sigma \sqrt{dx dy_1 dp'}\Big|_{x'=y', p_1=0, s=x_1-y_1}$$

$$= \sqrt{2\pi} \overline{\int a_{-\ell'}(-s, y, x, p, \sigma) e^{-is\sigma} d\sigma \sqrt{dx dy_1 dp'}\Big|_{x'=y', p_1=0, y_1-x_1=-2}} = \overline{\sigma_1(A)(\overline{x}, \phi^P_s \overline{x})}$$

The proof is now clear, since the symbol $\sigma_1(A^*)$ intertwines the variables and takes

the conjugate of $\sigma_1(A)$. $\square$

## 7.6 Conclusions

The algebra is defined in the manifold case assuming that the cut region $X_c$ is fibrating and Bohr-Sommerfeld. A global symbolic calculus is defined using the $\hbar$-Transform introduced in [PU95]. The two principal symbols have singularities in the intersection of the diagonal and the flow-out. However, they blow up in a controlled way, and there exists a symbolic compatibility condition in the intersection. We expose this condition in a natural setting after identifying the symbol on the flow-out with families of classical pseudodifferential operators acting on half-densities on the fibers of the foliation. Other results like the principal symbol of the composition and the principal symbol of the adjoint seems to be expressed more naturally under the identification described in this chapter.

# CHAPTER VIII

# Asymptotics of the trace

## 8.1 The trace in case $m' \geq 0$.

We now assume that $X_c$ is compact and $\partial X_c$ is fibrating, as in section VII. We will prove:

**Theorem VIII.1.** *Let $A$ be an operator in the class $\mathcal{A}_{X_c}$. Then, if $m > 1/2$,*

$$\mathrm{Tr}(A) = (2\pi)^{-n}\hbar^{-\ell-m-n}\int_{X_c}\sigma_0(x,x,p,-p)dxdp + O(\hbar^{-\ell-m-n+1}), \tag{8.1}$$

*where $\sigma_0$ is the symbol of $A$ on the diagonal. If $m = 1/2$,*

$$\mathrm{Tr}(A) = (2\pi)^{-n}\hbar^{-\ell-m-n}\int_{X_c}\sigma_0(x,x,p,-p)dxdp + O(\hbar^{-\ell-m-n}\ \hbar\ \log(1/\hbar)). \tag{8.2}$$

The rest of this subsection is devoted to a proof of this result, which we break into a series of lemmas. Note that it is enough to estimate the integral along the diagonal of the kernel of $A$ in the model case with $\ell' = 0$. Consider therefore a semiclassical kernel of the form:

$$u_\hbar(x,y) = \frac{1}{(2\pi\hbar)^n}\int e^{i\hbar^{-1}[(x_1-y_1-s)p_1+(x'-y')p']+is\sigma}a(s,x,y,p,\sigma)dsdpd\sigma, \tag{8.3}$$

where $p = (p_1, p')$, $x = (x_1, x')$, $y = (y_1, y')$ and $a$ is a classical symbol in $\sigma$ of degree $m' = m - 1/2$, compactly supported in $(s, x, y, p)$.

**Lemma VIII.2.** *Let $\mu_0 > 0$ be large enough so that $a|_{\{|p_1|>\mu_0/2\}} = 0$, and let $\rho \in C_0^\infty$ be a smooth function with compact support which is equal to one on $[-\mu_0/2, \mu_0/2]$ and is supported on $[-\mu_0, \mu_0]$. Then*

$$u_\hbar(x,x) = \frac{1}{(2\pi\hbar)^n} \int e^{-isp_1/\hbar + is\sigma} a(s,x,x,p,\sigma)\rho(\hbar\sigma) ds dp d\sigma + O(\hbar^\infty), \tag{8.4}$$

*uniformly in $x$.*

*Proof.* The rigorous definition of $u_h$ when the integral in $\sigma$ diverges is given in equation (6.3): If $K >> 0$ ($K \geq m'/2+1$), then $u_h$ is equal to the absolutely convergent integral

$$\frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}[(x_1-y_1-s)p_1+(x'-y')p']+is\sigma} \frac{1}{(1+\sigma^2)^K} D^K \left[ e^{-is\hbar^{-1}p_1} a(s,x,y,p,\sigma) \right] ds dp d\sigma,$$

and therefore

$$u_\hbar(x,y) = \frac{1}{(2\pi\hbar)^n} \sum_{j=0}^{K} \int e^{i\hbar^{-1}[(x-y)p - sp_1] + is\sigma} \left(-i\hbar^{-1}p_1\right)^{2K-j} \frac{a_j(s,x,y,p,\sigma)}{(1+\sigma^2)^K} ds dp d\sigma,$$

where the last expression is obtained by expanding the action of $D^K$. Note that, $\forall j$, $a_j$ consists of linear combinations of derivatives of $a$ with respect to $s$ (and therefore $a_j \in \mathcal{A}^{0,m'}$). Using this we get that the remainder in equation (8.4) is equal to

$$\sum_{j=0}^{K} \frac{1}{(2\pi\hbar)^n} \int e^{-isp_1/\hbar + is\sigma} \frac{a_j(s,x,x,p,\sigma)}{(1+\sigma^2)^K} \left(-i\hbar^{-1}p_1\right)^{2K-j} (1-\rho(\hbar\sigma))\, ds dp d\sigma.$$

Let $b_j(s,x,p,\sigma \ ; \ \hbar) = \frac{a_j(s,x,x,p,\sigma)}{(1+\sigma^2)^K}(-i\hbar^{-1}p_1)^{2K-j}(1-\rho(\hbar\sigma))$. We will show that for each $j$,

$$B_j := \frac{1}{(2\pi\hbar)^n} \int e^{-isp_1\hbar^{-1} + is\sigma} b_j \, ds dp d\sigma$$

is $O(\hbar^\infty)$

Starting with the change of variables $\mu = \hbar\sigma$, $\omega = -p_1 + \mu$, we obtain that

$$B_j = \frac{\hbar^{-1}}{(2\pi\hbar)^n} \int e^{is\omega/\hbar}\, b_j(s, x, (-\omega + \mu, p'), \hbar^{-1}\mu\ ; \hbar)\, ds d\omega dp' d\mu$$

$$= \frac{1}{(2\pi\hbar)^n} \int e^{-i\hbar\xi_1\xi_2}\, c_j(x, \mu, p'\ ;\ \hbar\ ;\ \xi)\, d\xi dp' d\mu,$$

where $\xi = (\xi_1, \xi_2)$ are the dual variables to $(s, \omega)$, and

$$c_j(x, \mu, p'\ ;\ \hbar\ ;\ \xi) = \frac{1}{2\pi} \int e^{-i(s,\omega)\cdot\xi}\, b_j(s, x, (-\omega + \mu, p'), \hbar^{-1}\mu\ ;\ \hbar) ds d\omega$$

is the Fourier transform of $b_j$ in the $(s, \omega)$ variables. Using the inequality

$$\left| e^{it} - \sum_{k=0}^{N-1} \frac{(it)^k}{k!} \right| \leq \frac{|t|^N}{N!}, \tag{8.5}$$

we obtain that for each $N > 0$,

$$\int e^{-i\hbar\xi_1\xi_2} c_j(x, \mu, p'\ ;\ \hbar\ ;\ \xi) d\xi - \sum_{k=0}^{N-1} \int \frac{(-i\hbar\xi_1\xi_2)^k}{k!} c_j(x, \mu, p'\ ;\ \hbar\ ;\ \xi) d\xi = \int R_N(\xi\ ;\ \hbar) c_j d\xi,$$

where $|R_N(\xi\ ;\ \hbar)| \leq \frac{\hbar^N |\xi_1\xi_2|^N}{N!}$. Moreover, for each $k = 0, \ldots, N-1$,

$$\frac{(i\hbar)^k}{k!} \int (-\xi_1\xi_2)^k c_j d\xi = \frac{2\pi(-i\hbar)^k}{k!} \left( \frac{\partial^2}{\partial s \partial p_1} \right)^k b_j(s, x, p, \hbar^{-1}\mu\ ;\ \hbar)\Big|_{s=0, p_1=\mu}$$

$$= \frac{2\pi(-i\hbar)^k\ (1 - \rho(\mu))}{(1 + (\hbar^{-1}\mu)^2) k!} \left( \frac{\partial^2}{\partial s \partial p_1} \right)^k a_j(s, x, (\mu, p'), \hbar^{-1}\mu)\Big|_{s=0, p_1=\mu} = 0,$$

since $\rho(\mu)$ is equal to one in the support of $a$. It follows that

$$\left| \int e^{-i\hbar\xi_1\xi_2} c_j d\xi \right| \leq \frac{\hbar^N}{N!} \int \left| (\xi_1\xi_2)^N\, c_j \right| d\xi$$

$$= \frac{\hbar^N}{2\pi N!} \left|\left| \int e^{-i(s,\omega)\cdot\xi} \left(\frac{\partial^2}{\partial s \partial \omega}\right)^N \left(b_j(s,x,(-\omega+\mu,p'),\hbar^{-1}\mu\ ;\ \hbar)\right) ds d\omega \right|\right|_{L^1_\xi},$$

where $\|\cdot\|_{L^1_\xi}$ denotes the $L^1$-norm of a function of the $\xi$ variables. The well-known inequality

$$||\hat{v}|\,|_{L^1(\mathbb{R}^d)} \leq \sum_{|\alpha|\leq d+1} ||\partial^\alpha v|\,|_{L^1(\mathbb{R}^d)} \tag{8.6}$$

implies that the above bound is in turn bounded by

$$\frac{\hbar^N}{2\pi N!} \sum_{|\alpha|\leq 3} \left|\left| \partial^\alpha \left(\frac{\partial^2}{\partial s \partial \omega}\right)^N \left[b_j(s,x,(-\omega+\mu,p'),\hbar^{-1}\mu\ ;\ \hbar)\right] \right|\right|_{L^1_{s,\omega}}$$

$$= \frac{\hbar^N}{2\pi N!} \sum_{|\alpha|\leq 3} \left|\left| \frac{(1-\rho(\mu))(-i\hbar^{-1})^{2K-j}}{(1+(\hbar^{-1}\mu)^2)^K} \partial^\alpha \left(\frac{\partial^2}{\partial s \partial \omega}\right)^N \right.$$

$$\left. \left[a_j(s,x,x,(-\omega+\mu,p'),\hbar^{-1}\mu)(-\omega+\mu)^{2K-j}\right] \right|\Big|_{L^1_{s,\omega}},$$

where $\|\cdot\|_{L^1_{s,\omega}}$ is defined similarly, and $\partial^\alpha = \frac{\partial^{\alpha_1}}{\partial s^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial \omega^{\alpha_2}}$, $\alpha = (\alpha_1, \alpha_2)$. Therefore $B_j$ is bounded above by

$$\frac{\hbar^{-n+N-2K+j}}{(2\pi)^{n+1} N!} \sum_{|\alpha|\leq 3}$$

$$\int \left| \partial^\alpha \left(\frac{\partial^2}{\partial s \partial \omega}\right)^N \left[\frac{(-\omega+\mu)^{2K-j} a_j(s,x,(-\omega+\mu,p'),\hbar^{-1}\mu)}{(1+(\hbar^{-1}\mu)^2)^k}(1-\rho(\mu))\right] \right| ds d\omega dp' d\mu$$

$$= \frac{\hbar^{-n+N+1-2K+j}}{(2\pi)^{n+1} N!} \sum_{|\alpha|\leq 3} \int \left| \partial^\alpha \left(\frac{\partial^2}{\partial s \partial p_1}\right)^N \left[\frac{p_1^{2K-j} a_j(s,x,p,\sigma)}{(1+\sigma^2)^k}(1-\rho(\hbar\sigma))\right] \right| ds dp d\sigma.$$

Finally, notice that the integrand is of order $O\left(\sigma^{m'-2K}\right)$ and that $1-\rho(\hbar\sigma)$ has support in $|\sigma| \geq \hbar^{-1}\mu_0/2$. Therefore the above upper bound is less than a constant

times

$$\hbar^{-n+N+1-2K+j} \int_{\hbar^{-1}\mu_0/2}^{\infty} \sigma^{m'-2K} d\sigma = \hbar^{-n+N+1-2K+j} \frac{\sigma^{m'-2K+1}}{m'-2K+1} \Big|_{\hbar^{-1}\mu_0/2}^{\infty}$$
$$= O\left(\hbar^{j-n-m'+N}\right).$$

Since this is true for all positive integers $N$ we are done. □

**Lemma VIII.3.** *If $m' \geq 0$,*

$$\int u_\hbar(x,x)\,dx = \frac{1}{(2\pi\hbar)^n} \int 2\pi\ a(0,x,x,\mu,p',\hbar^{-1}\mu)\,dx\,dp'\,d\mu + O(\hbar^{-n-m'+1}). \quad (8.7)$$

*Proof.* By (8.4), it suffices to estimate

$$\tilde{u}_\hbar(x,x) = \frac{1}{(2\pi\hbar)^n} \int e^{-isp_1/\hbar + is\sigma} a(s,x,x,p,\sigma)\rho(\hbar\sigma)dsdpd\sigma.$$

By an argument identical to the one used in the proof of the Lemma,

$$\tilde{u}_\hbar(x,x) = \frac{1}{(2\pi\hbar)^n}\ \hbar^{-1} \int e^{is\omega/\hbar} a(s,x,x,(-\omega+\mu,p'),\hbar^{-1}\mu)\rho(\mu)dsdp'd\omega d\mu.$$

We will apply the method of stationary phase to the $(s,\omega)$ integral, before integrating with respect to $p'$ and $\mu$. To this end introduce the notation

$$u_\hbar(x,x,p',\mu) = \frac{1}{(2\pi\hbar)^n}\ \hbar^{-1} \int e^{is\omega/\hbar} a(s,x,x,(-\omega+\mu,p'),\hbar^{-1}\mu)\rho(\mu)dsd\omega,$$

so that the left-hand side of (8.7) is equal to $\int u_\hbar(x,x,p',\mu)\,dxdp'd\mu$ modulo $O(\hbar^\infty)$. We also have that

$$u_\hbar(x,x,p',\mu) = \frac{1}{(2\pi\hbar)^n} \int e^{-i\hbar\xi_1\xi_2} c(x,\mu,p';\ \hbar\ ;\ \xi)d\xi,$$

where $\xi = (\xi_1, \xi_2)$ are the dual variables to $(s, \omega)$, and

$$c(x, \mu, p';\ \hbar\ ;\ \xi) = \frac{1}{2\pi} \int e^{-(s,\omega)\cdot\xi} a(s, x, x, (-\omega + \mu, p'), \hbar^{-1}\mu)\rho(\mu) ds d\omega$$

is the Fourier transform of $\rho a$ in $(s, \omega)$. Note that

$$\int c(x, \mu, p';\ \hbar;\ \xi) d\xi = 2\pi\ a(0, x, x, (\mu, p'), \hbar^{-1}\mu),$$

and therefore (by (8.5))

$$u_\hbar(x, x, p', \mu) - \frac{1}{(2\pi\hbar)^n}\ 2\pi\ a(0, x, x, (\mu, p'), \hbar^{-1}\mu) = \frac{1}{(2\pi\hbar)^n} \int R(\xi, \hbar) c(\xi) d\xi, \quad (8.8)$$

where $|R(\xi, \hbar)| \leq \hbar\, |\xi_1 \xi_2|$. Integrating (8.8) we see that the error term in (8.7) is bounded by

$$\frac{1}{(2\pi\hbar)^n} \int\limits_K \|Rc\|_{L^1_\xi}(x, p', \mu)\, dx\, dp'\, d\mu, \quad (8.9)$$

where $K$ is a compact set containing the support of the left-hand side of (8.8). Using (8.6) again,

$$\begin{aligned} \hbar\, ||\xi_1 \xi_2 c||_{L^1_\xi} &= \frac{\hbar}{2\pi} \left|\left| \int e^{-i(s,\omega)\cdot\xi} \frac{\partial^2}{\partial s \partial \omega} a(s, x, x, (-\omega + \mu, p'), \hbar^{-1}\mu)\rho(\mu) ds d\omega \right|\right|_{L^1} \\ &\leq \frac{\hbar}{2\pi} \sum_{|\alpha| \leq 3} \left|\left| \partial^\alpha \frac{\partial^2}{\partial s \partial \omega} a(s, x, x, (-\omega + \mu, p'), \hbar^{-1}\mu)\rho(\mu) \right|\right|_{L^1_{(s,\omega)}}. \end{aligned}$$

Since $a(s, x, y, p, \sigma)$ is a classical symbol in $\sigma$ of order $m'$ and compactly supported in the rest of the variables, for all $\alpha$ there exists a constant $C = C(\alpha)$ such that

$$\begin{aligned} &\left| \tfrac{\hbar}{2\pi} \partial^\alpha \tfrac{\partial^2}{\partial s \partial \omega} a(s, x, x, (-\omega + \mu, p'), \hbar^{-1}\mu)\rho(\mu) \right| \leq C(\alpha)\, (1 + \hbar^{-1}|\mu|)^{m'}\, \hbar \\ &= C(\alpha)\, (\hbar + |\mu|)^{m'}\, \hbar^{-m'+1} \end{aligned} \quad (8.10)$$

for all $(s, x, y, p)$. Integrating (8.10) with respect to $(s, \omega)$ over a sufficiently large

compact set we obtain that

$$\|Rc\|_{L^1_\xi}(x,p',\mu) \leq C\,(\hbar+|\mu|)^{m'}\,\hbar^{-m'+1}.$$

We now integrate this inequality over the compact set, $K$, in (8.9), to obtain that the error term in (8.7) is bounded above by a constant times

$$\hbar^{-m'-n+1}\int\limits_{|\mu|\leq\mu_1}(\hbar+|\mu|)^{m'}\,d\mu \tag{8.11}$$

for some $\mu_1>0$ independent of $\hbar$, and this is $O(\hbar^{-m'-n+1})$ when $m'\geq 0$. □

*Remark* VIII.4. Lemma VIII.3 implies the Theorem (with a better error estimate in case $m'=0$) if the amplitude $a$ is homogeneous in the variable $\sigma$.

*Proof of Theorem VIII.1.* Let us first assume that $m'>0$. Since $a$ is a symbol in $\sigma$ of degree $m'>0$, there exists $\tilde{a}(x,p)$ and a constant $C$ such that

$$\Big|a(0,x,x,p,\sigma)-\sigma^{m'}\tilde{a}(x,p)\chi(\sigma)\Big|\leq C\;(1+|\sigma|)^{m'-1},$$

where $\chi(\sigma)$ is smooth in $\sigma\neq 0$, and homogeneous of degree zero in $\sigma$. Then, in particular,

$$\left|\frac{2\pi}{(2\pi\hbar)^n}\;a(0,x,x,\mu,p',\hbar^{-1}\mu)-\frac{2\pi}{(2\pi\hbar)^n}\;\hbar^{-m'}\mu^{m'}\chi(\mu)\tilde{a}(0,x,(\mu,p'))\right|$$

$$\leq(2\pi)^{1-n}C\;(1+\hbar^{-1}|\mu|)^{m'-1}\hbar^{-n}.$$

The left-hand side of this inequality is supported in $\mu\in[-\mu_0,\mu_0]$. After integrating with respect to $\mu$, the remainder is bounded by constant times

$$\hbar^{-n}\int\limits_{-\mu_0}^{\mu_0}(1+\hbar^{-1}|\mu|)^{m'-1}d\mu=2\hbar^{-n-m'+1}\int\limits_0^{\mu_0}(\hbar+\mu)^{m'-1}d\mu$$

$$= 2\hbar^{-n-m'+1}\left((\mu_0+\hbar)^{m'}-\hbar^{m'}\right)/m' = O(\hbar^{-n-m'+1}) \text{ since } m'>0.$$

Therefore, for any $\ell'$ and $m'>0$

$$\int u_\hbar(x,x)\,dx = (2\pi)^{-n}\hbar^{-\ell-m-n}\int 2\pi\ p_1^{m'}\tilde{a}(0,x,p)dxdp + O(\hbar^{-\ell-m-n+1}),$$

where $2\pi\ p_1^{m'}\tilde{a}(0,x,p)$ is the principal symbol on the diagonal.

Now let us assume $m'=0$. Since $a$ is a symbol in $\sigma$ of degree zero, there exists $\tilde{a}(x,p)$ such that,

$$|a(0,x,x,p,\sigma)-\tilde{a}(x,p)\chi(\sigma)| < \frac{C}{1+|\sigma|}$$

for some constant $C>0$.Then

$$\left|\frac{2\pi}{(2\pi\hbar)^n}a(0,x,x,\mu,p',\hbar^{-1}\mu)-\frac{2\pi}{(2\pi\hbar)^n}\tilde{a}(x,(\mu,p'))\chi(\mu)\right| < \frac{(2\pi)^{1-n}C}{1+\hbar^{-1}\,|\mu|}\ \hbar^{-n}$$

Again, since the left hand side is supported in the set $\{|\mu|\le\mu_0\}$, after integrating with respect to $\mu$, the remainder is bounded by a constant times

$$\hbar^{-n}\int_0^{\mu_0}\frac{1}{1+\hbar^{-1}\mu}d\mu = \hbar^{-n+1}\int_0^{\mu_0}\frac{1}{\mu+\hbar}d\mu$$

$$= \hbar^{-n+1}\,(log(\mu_0+\hbar)-log(\hbar)) = O(\hbar^{-n+1}\ log(1/\hbar)).$$

Therefore, for $m=1/2$

$$\int u_\hbar(x,x)\,dx = (2\pi)^{-n}\hbar^{-\ell-m-n}\int 2\pi\ \tilde{a}(0,x,p)dpdx + O(\hbar^{-\ell-m-n+1}\ \log(1/\hbar)),$$

where $2\pi\ \tilde{a}(0,x,p)$ is the principal symbol of the family on the diagonal.

## 8.2 The trace in case $m' \leq -4$.

**Theorem VIII.5.** *With the previous notation, if $m \leq -7/2$, and $A$ an operator in $\mathcal{A}_{X_c}$, then each classical $\Psi DO$ $\sigma_1(A)_s$ obtained from the symbol of $A$ in the flow out on each orbit is of trace class, and*

$$\mathrm{Tr}(A) = (2\pi)^{-n+1/2}\hbar^{-n-\ell+1/2} \int_S \mathrm{Tr}\left(\sigma_1(A)_s\right) ds + O(\hbar^{-n-\ell+3/2}). \tag{8.12}$$

**Lemma VIII.6.** *If $m' < 0$ and $\ell' = 0$,*

$$\int u_\hbar(x,x)\, dx = \frac{1}{(2\pi\hbar)^n} 2\pi \int a(0,x,x,\mu,p',\hbar^{-1}\mu)\, dx\, dp'\, d\mu + O(\hbar^{-n+2}). \tag{8.13}$$

*Proof.* This follows immediately from (8.11) (which was derived under no assumptions on $m'$). □

*Proof.* Starting with equation (8.3), since $m' \leq -4$

$$a_2(s,x,y,p) := \int e^{is\sigma} a(s,x,y,p,\sigma) d\sigma$$

is absolutely convergent and can be extended to a $C^2$ function of $s$. In addition, $a_2$ is compactly supported in $s, x, y, p$. Using the stationary phase theorem for $C^{2k}$ amplitudes (Theorem 7.7.5 in [Hör83]), we get

$$\int e^{i\hbar^{-1}sp_1} a_2(s,x,x,-p_1,p') ds dp_1 \sim 2\pi\, \hbar\, a_2(0,x,x,0,p'),$$

and then

$$\int u_\hbar(x,x)\, dx = (2\pi)^{-n+1/2}\hbar^{-n-\ell+1/2} \int \sqrt{2\pi}\; a_2(0,x,x,0,p') dx dp' + O(\hbar^{-n-\ell+3/2}),$$

where $\sqrt{2\pi}\; a_2(0,x,x,0,p') = \sqrt{2\pi} \int e^{is\sigma} a(s,x,x,p) d\sigma \big|_{s=0,p_1=0}$ is the extension of the

symbol $\sigma_1$ to the intersection of the Lagrangians. This is the desired result in the model case. □

## 8.3 Conclusions

In this chapter we stated and proved the asymptotic of the trace of elements of the algebra with compact microsupport on $X_c$. We showed how the contribution to the principal term in the expansion of the trace can either come from the principal symbol on the diagonal or from the principal symbol on the flow-out depending on the bi-degree of the operator.

# CHAPTER IX

# Projectors and "cut" quantum observables

In this section we will prove that, under a very mild additional condition on $\partial X_c$, the algebra $\mathcal{A}_{X_c}$ contains orthogonal projectors. We will also prove that, in case there exists an semiclassical pseudodifferential operator on $M$, $\hat{P}$, such that:

1. The spectrum of $\hat{P}$ is discrete and is contained in $\hbar\mathbb{Z}$, and
2. $X_c = P^{-1}(I)$ for $I \subset \mathbb{R}$ a closed interval,

then the spectral projector of $\hat{P}$ associated to the interval $I$ is in $\mathcal{A}_{X_c}$

## 9.1 On the Existence of Projectors

In addition to the assumptions on $\partial X_c$ that we have been making throughout, let us now assume that $\partial X_c$ is of contact type. Recall that this means that there exists a one-form $\beta$ on $\partial X_c$ such that (a) $d\beta$ is the pull-back of the symplectic form to $\partial X_c$, and (b) $\beta \rfloor \Xi_P$ is constant, where $P$ is a defining function of $X_c$ with periodic flow on $\partial X_c$.

Following the proof of Lemma 5 in ([Doz97]), one obtains:

**Lemma IX.1.** *There exists an smooth function, which will be called again* $P : T^*M \to \mathbb{R}$*, such that*

*(a) $P$ is bounded from below and tends to $\infty$ at infinity in the cotangent directions,*

*(b) $\partial X_c = P^{-1}(0)$, and*

*(c) There exists a neighborhood $\mathcal{W}$ of $\partial X_c$ such that the Hamilton flow of $P$ is $2\pi$-periodic in $\mathcal{W}$.*

Next we recall (see [HR84] Proposition 3.8) how to obtain a quantum version of the previous result:

**Lemma IX.2.** *Let $\widehat{P}(\hbar)$ be a semiclassical pseudodifferential operator with principal symbol $P$ and vanishing sub-principal symbol. Let $\mu$ be the Maslov index of the trajectories of $\Phi^P$ (the Hamilton flow of $P$) in $\mathcal{W}$. Assume the Bohr-Summerfeld conditions* (7.1)*. Then there exists a semiclassical pseudodifferential operator $\widehat{R}_2(\hbar)$ of order $-2$ such that for $\epsilon << 1$*

$$Spec\left(\widehat{P} - \frac{\mu}{4}\,\hbar - \widehat{R}_2(\hbar)\right) \cap [-\epsilon/3, \epsilon/3] \subset \hbar\mathbb{Z}$$

*when we restrict $\hbar$ to the sequence $\hbar = 1/N$ with $N$ large.*

*Proof.* Pick $\epsilon > 0$ such that $P^{-1}[-\epsilon, \epsilon] \subset \mathcal{W}$. Let $\rho$ be a smooth function with support in $[-\epsilon, \epsilon]$, such that $\rho \equiv 1$ on $[-\epsilon/2, \epsilon/2]$. Let

$$\gamma = \frac{1}{2\pi} \int_0^{2\pi} p(t)\dot{x}(t) - P(x(t), p(t))dt$$

be the (common) action of the trajectories of the Hamilton flow of $P$ in $\mathcal{W}$. Then $e^{-2\pi i\hbar^{-1}\left(\widehat{P} - \frac{\mu}{4}\,\hbar - \gamma\right)}$ is microlocally in $\mathcal{W}$ a pseudodifferential operator with symbol identically equal to one, and thus one can write

$$\rho(\widehat{P})e^{-2\pi i\hbar^{-1}\left(\widehat{P} - \frac{\mu}{4}\,\hbar - \gamma\right)} = \rho(\widehat{P})(I + \hbar\widehat{R}(\hbar)),$$

where $\widehat{R}(\hbar)$ is a zeroth order $\hbar$-$\Psi$DO. Recall that $\gamma$ is an integer in $\partial X_c$, and therefore, for $\hbar = 1/N$ we obtain

$$\rho(\widehat{P})e^{-2\pi i\hbar^{-1}\left(\widehat{P}-\frac{\mu}{4}\ \hbar\right)} = \rho(\widehat{P})(I + \hbar\widehat{R}(\hbar)),$$

Since $I + \hbar\widehat{R}(\hbar)\rho(\widehat{P})$ has spectrum close to 1 for $\hbar << 1$, one can then define for $\hbar$ small

$$\widehat{R}_2 = -\frac{\hbar}{2\pi i}\log(I + \hbar\widehat{R}(\hbar)\rho(\widehat{P})),$$

and since $\widehat{R}_2$ commutes with $\widehat{P}$, we obtain

$$\rho(\widehat{P})e^{-2\pi i\hbar^{-1}\left(\widehat{P}-\frac{\mu}{4}\ \hbar-\widehat{R}_2\right)} = \rho(\widehat{P})\left(I + \hbar\widehat{R}(\hbar)\right)\left(I + \hbar\widehat{R}(\hbar)\rho(\widehat{P})\right)^{-1}.$$

Since $\rho \equiv 1$ on $[-\epsilon/2, \epsilon/2]$, the spectral theorem guarantees that the above operator is the identity on any eigenfunction of $\hat{P}$ with eigenvalue in $[-\epsilon/2, \epsilon/2]$. Since, for $\hbar$ small enough, the spectrum of $\widehat{P} - \frac{\mu}{4}\ \hbar - \widehat{R}_2$ in $[-\epsilon/3, \epsilon/3]$ corresponds to eigenfunctions of $\hat{P}$ with eigenvalues in $[-\epsilon/2, \epsilon/2]$, the result follows. □

Let us define $\widehat{P}_2 := \widehat{P} - \frac{\mu}{4}\ \hbar - \widehat{R}_2$, whose principal symbol continues to be $P$. Let $\chi$ be the characteristic function on $(-\infty, 0]$, and define the projector

$$\Pi = \chi\left(\widehat{P}_2\right)$$

**Theorem IX.3.** *The projector* $\Pi$ *defined above belongs to the class* $J^{-1/2,1/2}(M \times M, \Delta, \mathcal{F}\partial X_c)$.

*Proof.* Let $\rho$ be the cut-off function of the previous proof. We decompose

$$\Pi = (1-\rho)\chi\left(\widehat{P}_2\right) + \rho\chi\left(\widehat{P}_2\right)$$

Clearly $(1-\rho)\chi\left(\widehat{P}_2\right)$ is a semiclassical pseudodifferential operator. Therefore we

need to prove that $\rho\ \chi\left(\widehat{P}_2\right)$ belongs to $J^{-1/2,1/2}$. This operator has microsupport in $\mathcal{W}$.

We take $\mathbb{R}^{n-1} \times S^1$ as the model case, with coordinates $(x, \theta)$ and $T^*M$ with coordinates $(x, \theta; p, \tau)$. Let $\Pi^E = \Pi^E_N$ be the projector on eigenfunctions of $\widehat{P}^n = \hbar D_\theta = \frac{\hbar}{i}\frac{\partial}{\partial\theta}$ with eigenvalues greater than or equal to $E/N$, where $E \in \mathbb{Z}$ is a constant. (Here $\theta$ is the $2\pi$-periodic variable in $S^1$.) Let $T_s = e^{-is\hbar^{-1}\widehat{P}_2}$, and let

$$T^n_s = e^{-i\hbar^{-1}s\ \widehat{P}^n}$$

be the translation representation on $L^2(\mathbb{R}^{n-1} \times S^1)$. Let $(x_0, p_0) \in \mathcal{W}$. As in [Gui94, GL02], there exist an $S^1$-invariant neighborhood $\mathcal{U} \subset \mathcal{W}$ of $(x_0, p_0)$ (the circle action given by the Hamilton flow of $P$), an $S^1$-invariant open set $\mathcal{U}^n \subset T^*(\mathbb{R}^{n-1} \times S^1)$, and an $S^1$-equivariant canonical transformation

$$\phi : \mathcal{U} \to \mathcal{U}^n,$$

which sends $\partial X_c \cap \mathcal{U}$ into $\left\{(x, \theta; p, \tau) \in T^*\left(\mathbb{R}^{n-1} \times S^1\right) \ \middle|\ \tau = E\right\}$. As in [Gui94, GL02], one can show that there exists a semiclassical zeroth order Fourier integral operator

$$F : L^2(M) \to L^2(\mathbb{R}^{n-1} \times S^1),$$

with microsupport on $\mathcal{U} \times \mathcal{U}^n$ such that

$$F^*F = I_{\mathcal{U}^n},\ FF^* = I_{\mathcal{U}},$$

and

$$F\rho\ \chi(\widehat{P}_2) = \rho\left(\Pi^E\right) F,$$

This reduces the proof to the model case. It suffices to show that $\Pi\widehat{Q}$ is on the algebra,

for any zeroth order compactly supported semiclassical pseudodifferential operator $\widehat{Q}$ in $\mathbb{R}^{n-1} \times S^1$. Note that

$$\Pi^E \ \widehat{Q} = \frac{1}{2\pi} \int_0^{2\pi} e^{-iNs\widehat{P}^n} e^{iNsE} \ \widehat{Q} \ \frac{1}{1-e^{is}} ds$$

The operator $e^{-iNs\widehat{P}^n} e^{iNsE} \widehat{Q}$ is a semiclassical FIO with Lagrangian

$$\left\{ ((x, \theta \ ; \ p, \tau = E), (y = x, \alpha + s = \theta \ ; \ -p, -\tau = -E)) \, \middle| x \in \mathbb{R}^{n-1}, \theta, \alpha \in [0, 2\pi] \right\}.$$

Therefore, in local coordinates, the Schwartz kernel of $\Pi^E \widehat{Q}$ can be written as

$$\frac{1}{(2\pi\hbar)^n} \frac{1}{2\pi} \int e^{iN( \ (x-y)p + (\theta - \alpha - s)(\tau - E) \ )} q(x, \theta, y, \alpha, p, \tau, \hbar) \frac{1}{1-e^{is}} dp d\tau ds, \tag{9.1}$$

where $q$ is symbol with expansion in $\hbar$. Notice that $\frac{1}{1-e^{is}}$ is a conormal distribution in $s = 0$ and equation (9.1) shows the hybrid nature of the amplitude of the projector. Equation (9.1) also proves that $\Pi^E \widehat{Q} \in J^{-1/2,1/2}$, with principal symbol $\chi_{X_c} Q(x, \theta \ ; \ p, \tau)$ in the diagonal, and

$$\frac{1}{\sqrt{2\pi}} \frac{Q(\Phi_s^P(x, \theta \ ; \ p, \tau))}{1 - e^{is}}$$

in the flow-out. □

*Remark* IX.4. If one has an semiclassical pseudodifferential operator $\widehat{P}$ with discrete spectrum such that

$$\mathrm{Spec}(\widehat{P}(\hbar)) \subset \hbar\mathbb{Z},$$

then the previous proof shows that, for any given $E_1, E_2 \in \mathbb{Z}$ such that $E_1 < E_2$, and

for $j = 1, 2$:

the trajectories on $P^{-1}(E_j)$ satisfy the Bohr-Sommerfeld condition (7.1).

Then the orthogonal projector onto

$$\mathcal{H}_N = \text{span of eigenvectors of } \widehat{P}(\hbar) \text{ with eigenvalues in } [E_1, E_2],$$

is in the algebra $J^{-1/2,1/2}\left(M \times M; \Delta, \mathcal{F}\partial X_c\right)$, associated to

$$X_c = \left\{\overline{x} \in T^*M \;\middle|\; E_1 \leq P(\overline{x}) \leq E_2\right\}.$$

## 9.2 Cut quantum observables and Generalized Töplitz Matrices

In this section we fix a projector $\Pi$ as in §5.1, and consider "cut" quantum observables, by which we mean operators of the form

$$\Pi\widehat{Q}\Pi$$

where $\widehat{Q}$ is a semiclassical pseudodifferential operator on $M$. By Theorem IX.3 these operators are in $J^{-1/2,1/2}\left(M \times M; \Delta, \mathcal{F}\partial X_c\right)$. The symbolic properties of these operators are summarized by the following Proposition:

**Proposition IX.5.** *Let $\widehat{Q}(\hbar)$ be a zeroth-order semiclassical pseudodifferential operator with compact microsupport. Then $\Pi\widehat{Q}\Pi$ is in the class $J^{-1/2,1/2}(M\times M\ ;\ \Delta, \mathcal{F}\partial X_c)$. Its symbols, ignoring Maslov factors, are as follows:*

$$\sigma_0(\Pi\widehat{Q}\Pi)(\overline{x}, \overline{x}) = \chi_{X_c}(\overline{x})Q(\overline{x})\ \sqrt{dx \wedge dp}, \tag{9.2}$$

$$\sigma_1\left(\Pi\widehat{Q}\Pi\right)_s = \Pi_{F_s} M_{|_{Q_{F_s}}} \Pi_{F_s} \tag{9.3}$$

*where* $\chi_{X_c}$ *is the characteristic function of* $X_c$, $\overline{x} = (x,p) \in T^*M$, $F_s$ *is the fiber above* $s \in S$, $Q_{|_{F_s}}$ *is the restriction of* $Q$ *to* $F_s$, $M_{Q_{|_{F_s}}}$ *is the operator "multiplication by* $Q_{|_{F_s}}$*", and* $\Pi_{F_s}$ *is the Szegö projector in the orbit* $F_s$*, i.e., for* $u : F_s \to \mathbb{C}$ *smooth,* $u(\Phi_s^P(\bar{y})) = \sum_j u_j(\bar{y}) \frac{e^{ijs}}{\sqrt{2\pi}}$,

$$[\Pi_{F_s} u](\bar{x}) = \sum_{j\geq 0} \frac{u_r(\overline{x})}{\sqrt{2\pi}}.$$

*Remark* IX.6. Notice that the Szegö projector $\Pi_{F_s}$ is a classical pseudodifferential operator with principal symbol $\chi_{(T^*F_s)^+}$, where $(T^*F_s)^+$ is the part with positive momentum variable, in the direction of the Hamilton flow. This function is smooth in $T^*F_s \setminus 0$.

*Remark* IX.7. The symbol of $\Pi_{|_{F_s}} M_{Q_{|_{F_s}}} \Pi_{|_{F_s}}$ is $\chi_{(T^*F_s)^+} Q_{|_{F_s}}$, which agrees with the symbol on the diagonal, restricted to the intersection. This is the symbolic compatibility condition referred to in Theorem VII.18.

*Proof.* The first part was proven in Theorem IX.3. The principal symbol on the diagonal is clear. Using the relation (0.2) in [AU85], we obtain that for $\overline{y} \in \partial X_c$, $s \neq 0$,

$$\begin{aligned}\sigma_1(\Phi_s^P(\overline{y}) \ ; \ \overline{y}) &= \frac{1}{\sqrt{2\pi}} \frac{1}{2\pi} \int_0^{2\pi} \frac{Q(\Phi_{s-\tilde{s}}^P(\overline{y}))}{1-e^{i(s-\tilde{s})}} \frac{1}{1-e^{i\tilde{s}}} d\tilde{s} \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{2\pi} \int_0^{2\pi} \frac{Q(\Phi_{\tilde{s}}^P(\overline{y}))}{1-e^{i\tilde{s}}} \ \frac{1}{1-e^{i(s-\tilde{s})}} d\tilde{s}.\end{aligned}$$

Since $Q(\Phi_{\tilde{s}}^P(\overline{y}))$ is a smooth $2\pi$-periodic function in $\tilde{s}$, there exists a sequence of functions $\{Q_j(\overline{y})\}_{j=-\infty}^{\infty}$ such that

$$Q(\Phi_{\tilde{s}}^P(\overline{y})) = \sum_{j=-\infty}^{\infty} Q_j(\overline{y}) \frac{e^{ij\tilde{s}}}{\sqrt{2\pi}},$$

and the symbol becomes:

$$\begin{aligned}\sigma_1(\Phi_s^P(\overline{y})\ ;\ \overline{y}) &= \frac{1}{\sqrt{2\pi}}\frac{1}{2\pi}\int\left(\sum_j \frac{Q_j(\overline{y})}{\sqrt{2\pi}}e^{ij\tilde{s}}\right)\left(\sum_{k\geq 0} e^{ik\tilde{s}}\right)\left(\sum_{k'\geq 0} e^{ik'(s-\tilde{s})}\right)d\tilde{s}\\ &= \frac{1}{2\pi}\sum_{k,\ k'\geq 0,\ j\in\mathbb{Z}} \frac{Q_j(\overline{y})}{2\pi}\int e^{i\tilde{s}(j+k-k')}e^{ik's}d\tilde{s} = \sum_{k'\geq 0,\ k'\geq j} Q_j(\overline{y})\frac{e^{ik's}}{2\pi}\\ &= \frac{1}{\sqrt{2\pi}}\sum_{j\in\mathbb{Z}} Q_j(\overline{y})\left(\sum_{k'\geq max(0,j)}\frac{e^{ik's}}{\sqrt{2\pi}}\right) = \frac{1}{\sqrt{2\pi}}\sum_{j\in\mathbb{Z}}\frac{Q_j(\overline{y})}{\sqrt{2\pi}}\frac{e^{max(0,j)\ is}}{1-e^{is}}\\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{1-e^{is}}\left[\sum_{j\geq 0} e^{ijs}\frac{Q_j(\overline{y})}{\sqrt{2\pi}} + \sum_{j<0}\frac{Q_j(\overline{y})}{\sqrt{2\pi}}\right].\end{aligned}$$

We now interpret this as the kernel of an operator acting on the fibers of $\partial X_c \to S$. Let us consider a fiber $F_s \subset \partial X_c$, and a function $u : F_s \to \mathbb{C}$. For every fixed $\overline{y} \in \partial X_c$, let $u_r(\overline{y})$ and $Q_j(\overline{y})$ be the corresponding Fourier coefficients in each decomposition

$$u(\Phi_s^P(\overline{y})) = \sum_{r=-\infty}^{\infty} u_r(\overline{y})\frac{e^{irs}}{\sqrt{2\pi}},\ Q(\Phi_s^P(\overline{y})) = \sum_{j=-\infty}^{\infty} Q_j(\overline{y})\frac{e^{ijs}}{\sqrt{2\pi}}.$$

As a pseudodifferential operator, the symbol on the flow-out of $\Pi\widehat{Q}\Pi$ acting the function $u$, is given by

$$\left[\sigma_1\left(\Pi\widehat{Q}\Pi\right)_s\ u\right](\overline{x}) = \int_{F_s}\sigma_1(\Pi\widehat{Q}\Pi)(\overline{x},\overline{y})u(\overline{y})d\overline{y} = \int_{S^1}\sigma_1(\Pi\widehat{Q}\Pi)(\overline{x},\Phi_s^P(\overline{x}))u(\Phi_s^P(\overline{x}))ds$$

$$\begin{aligned}&= \int\frac{1}{\sqrt{2\pi}}\frac{1}{1-e^{-is}}\left[\sum_{j\geq 0}\frac{Q_j(\overline{x})}{\sqrt{2\pi}} + \sum_{j<0}\frac{Q_j(\overline{x})}{\sqrt{2\pi}}e^{ijs}\right]\sum_r u_r(\overline{x})\frac{e^{irs}}{\sqrt{2\pi}}ds\\ &= \sum_{k\geq 0,j\geq 0} u_k(\overline{x})\frac{Q_j(\overline{x})}{\sqrt{2\pi}} + \sum_{k\geq 0,j<0} u_{k-j}(\overline{x})\frac{Q_j(\overline{x})}{\sqrt{2\pi}}\\ &= \sum_{k\geq 0,j\geq 0} u_k(\overline{x})\frac{Q_j(\overline{x})}{\sqrt{2\pi}} + \sum_{r\geq -j,j<0} u_r(\overline{x})\frac{Q_j(\overline{x})}{\sqrt{2\pi}}\end{aligned}$$

$$= \sum_{r\geq 0,j\geq 0,j\geq -r} u_r(\overline{x})\frac{Q_j(\overline{x})}{\sqrt{2\pi}} + \sum_{r\geq 0,j<0,j\geq -r} u_r(\overline{x})\frac{Q_j(\overline{x})}{\sqrt{2\pi}} = \sum_{r\geq 0,j\geq -r} u_r(\overline{x})\frac{Q_j(\overline{x})}{\sqrt{2\pi}} \qquad (9.4)$$

On the other hand,

$$\left[\Pi_{F_s} M_{|_{Q_{F_s}}} \Pi_{F_s}\ (u)\right](\overline{x}) = \Pi_{F_s} M_{|_{Q_{F_s}}} \left(\sum_{r\geq 0} \frac{u_r(\overline{y})}{\sqrt{2\pi}}\right)(\overline{x}) = \Pi_{F_s}\left(Q(\overline{y})\sum_{r\geq 0}\frac{u_r(\overline{y})}{\sqrt{2\pi}}\right)(\overline{x})$$

$$= \Pi_{S^1}\left(Q(\Phi_s^P(\overline{x}))\sum_{r\geq 0}\frac{u_r(\Phi_s^P(\overline{x}))}{\sqrt{2\pi}}\right)\Big|_{s=0} = \Pi_{S^1}\left(\sum_{j\in\mathbb{Z},r\geq 0}\frac{Q_j(\overline{x})}{\sqrt{2\pi}}u_r(\overline{x})\frac{e^{i(j+r)s}}{\sqrt{2\pi}}\right)\Big|_{s=0}$$

$$= \sum_{r\geq 0, j\geq -r} Q_j(\overline{x})\frac{u_r(\overline{x})}{\sqrt{2\pi}},$$

which agrees with equation (9.4). This proves that

$$\sigma_1\left(\Pi\widehat{Q}\Pi\right)_s = \Pi_{F_s} M_{|_{Q_{F_s}}} \Pi_{F_s},$$

which yields (9.3) after applying Proposition VII.19.

□

Operators of the form $\Pi_N \widehat{Q}_N \Pi_N : \mathcal{H}_N \to \mathcal{H}_N$ generalize Toeplitz matrices, and this is reflected in its principal symbol on $\mathcal{F}\partial X_c$. (In case $M = S^1$ and $\Pi_N$ the projector onto the span of $\{e^{ij\theta},\ j = 0,\ldots,N\}$, the $\Pi_N\widehat{Q}_N\Pi_N$ are to leading order the generalized Toeplitz matrices of [GS58], page 84.)

### 9.2.1 Applications: A symbolic proof of the Szegö Limit Theorem

We begin with the functional calculus, the heart of which is the following

**Lemma IX.8.** *Let $\widehat{Q}$ be a self-adjoint pseudodifferential operator of order zero on $M$. Then*

$$\Pi\, e^{-it\Pi\widehat{Q}\Pi} \in J^{-1/2,1/2}\left(M\times M;\Delta,\mathcal{F}\partial X_c\right).$$

*Proof.* Let us define

$$W(t) = \Pi e^{-it\Pi\widehat{Q}\Pi}.$$

It is the solution of the problem

$$\begin{cases} \frac{1}{i}\frac{\partial}{\partial t}W(t) + \Pi\widehat{Q}\Pi W(t) = 0 \\ W(t)\Big|_{t=0} = \Pi \end{cases} \tag{9.5}$$

The idea in the following proof is to construct a solution which will be in the algebra and will make the right-hand side of first equation (9.5) of order $O(\hbar^\infty)$.

As a first approximation we take

$$\widetilde{W_0} = \Pi e^{-it\widehat{Q}},$$

which satisfies

$$\frac{1}{i}\frac{\partial}{\partial t}\widetilde{W_0} + \Pi\widehat{Q}\Pi\ \widetilde{W_0} = -\Pi\left[\Pi, \widehat{Q}\right] e^{-it\widehat{Q}}.$$

We will prove below that $\left[\Pi, \widehat{Q}\right] \in \text{sc-}I^{-1/2}(M\times M; \mathcal{F}\partial X_c)$ (see Section 9.2.2). Therefore, we obtain

$$\begin{cases} \frac{1}{i}\frac{\partial}{\partial t}\widetilde{W_0}(t) + \Pi\widehat{Q}\Pi\widetilde{W_0}(t) =: \widetilde{R}_0(t) \in \text{sc-}I^{-1/2}\left(M \times M; \mathcal{F}\partial X_c\right) \\ \widetilde{W_0}(t)\Big|_{t=0} = \Pi \end{cases}$$

We will now modify $W_0$ so as to make the right hand side $O(\hbar^\infty)$ instead of an operator in $\text{sc-}I^{-1/2}(M \times M; \mathcal{F}\partial X_c)$. For the rest of the proof, it will be convenient to identify symbols in the flow-out with corresponding families of smoothing operators acting on functions on the fibers $F_s$ for each $s \in S$. The symbol $R_0(t)$ has a corresponding family operator $\{\widetilde{\mathcal{R}}_{0,s}(t)\}_{s\in S}$. Let us consider the following problem,

$$\begin{cases} \frac{1}{i}\frac{\partial}{\partial t}\widetilde{\mathcal{V}}_{0,s} + \Pi_{F_s} M_{Q_{|_{F_s}}} \Pi_{F_s} \circ \widetilde{\mathcal{V}}_{0,s} = -\widetilde{\mathcal{R}}_{0,s}, \\ \widetilde{\mathcal{V}}_{0,s}\Big|_{t=0} = 0, \end{cases}$$

whose solution is

$$\widetilde{\mathcal{V}}_{0,s}(t) = -i \int_0^t e^{i(\tilde{t}-t)\Pi_{F_s} M_{Q_{|_{F_s}}} \Pi_{F_s}} \widetilde{\mathcal{R}}_{0,s}(\tilde{t}) d\tilde{t}.$$

Notice that $\widetilde{\mathcal{V}}_{0,s}$ is a smoothing operator. Let us call $V_0$ the smooth symbol on $\mathcal{F}\partial X_c$ given by the operator $\widetilde{\mathcal{V}}_{0,s}$. Take $\widetilde{V}_0 \in \text{sc-}I^{-1/2}(M \times M; \mathcal{F}\partial X_c)$ with symbol $V_0$. By construction

$$\frac{1}{i}\frac{\partial}{\partial t}\left(\widetilde{W}_0 + \widetilde{V}_0\right) + \Pi\widehat{Q}\Pi\left(\widetilde{W}_0 + \widetilde{V}_0\right) = \hbar\widetilde{R}_1(t) \in \text{sc-}I^{-3/2}\left(M \times M; \mathcal{F}\partial X_c\right).$$

Proceeding inductively one can find a sequence of operators $\tilde{V}_j$ such that for all $J$

$$\begin{aligned} &\frac{1}{i}\frac{\partial}{\partial t}\left(\widetilde{W}_0 + \textstyle\sum_{j=0}^{J} \hbar^j\ \widetilde{V}_j\right) + \Pi\widehat{Q}\Pi\left(\widetilde{W}_0 + \textstyle\sum_{j=0}^{J} \hbar^j\ \widetilde{V}_j\right) \\ &= \hbar^{J+1}\widetilde{R}_{J+1} \in \text{sc-}I^{-3/2-J}\left(M \times M; \mathcal{F}\partial X_c\right) \end{aligned} \tag{9.6}$$

Finally, take an operator $\widetilde{V} \in \text{sc-}I^{-1/2}(M \times M; \mathcal{F}\partial X_c)$ such that $\widetilde{V} \sim \sum_{j=0}^{\infty} \widetilde{V}_j$, and define $\widetilde{W} = \widetilde{W}_0 + \widetilde{V}$. Then

$$\frac{1}{i}\frac{\partial}{\partial t}\widetilde{W} + \Pi\widehat{Q}\Pi\widetilde{W} = O(\hbar^\infty).$$

A standard application of Duhamel's principle finishes the proof. □

**Proposition IX.9.** *Let $\widehat{Q}$ be a self-adjoint semiclassical pseudodifferential operator. Then for any smooth function $f$, $\Pi\, f(\Pi\widehat{Q}\Pi)$, is in the class $J^{-1/2,1/2}(M \times M\ ;\ \Delta, \mathcal{F}\partial X_c)$. The symbols, ignoring Maslov factors, are as follows:*

$$\sigma_0\left(\Pi\, f(\Pi\widehat{Q}\Pi)\right)(\overline{x},\overline{x}) = \chi_{X_c}(\overline{x}) f\left(Q(\overline{x})\right)\ \sqrt{dx \wedge dp},\ \textit{and} \tag{9.7}$$

$$\sigma_1\left(\Pi\, f(\Pi\widehat{Q}\Pi)\right)\Big|_{F_s} = \Pi_{F_s}\, f\left(\Pi_{F_s} M_{Q_{|_{F_s}}} \Pi_{F_s}\right), \tag{9.8}$$

*where $F_s$,$Q_{|_{F_s}}$, $M_{Q_{|_{F_s}}}$, and $\Pi_{F_s}$ are as in Proposition IX.5*

*Proof.* We have:

$$\Pi\, f\left(\Pi\widehat{Q}\Pi\right) = \frac{1}{\sqrt{2\pi}} \int \Pi\, e^{-it\Pi\widehat{Q}\Pi} \check{f}(t)dt, \tag{9.9}$$

where $\check{f}(t) = \frac{1}{\sqrt{2\pi}} \int e^{is} f(s)ds$. By the previous lemma we can conclude that $\Pi f(\Pi\widehat{Q}\Pi) \in J^{-1/2,1/2}$. Moreover

$$\sigma_0(\Pi f(\Pi\widehat{Q}\Pi))(\overline{x},\overline{x}) = \frac{1}{\sqrt{2\pi}} \int e^{-itQ(\overline{x})} \chi_{X_c}(\overline{x}) \check{f}(t)dt = f(Q(\overline{x}))\ \chi_{X_c}(\overline{x}),$$

and

$$\sigma_1\left(\Pi f(\Pi\widehat{Q}\Pi)\right)_s = \frac{1}{\sqrt{2\pi}} \int \Pi_{F_s} e^{-it\Pi_{F_s} M_{Q_{F_s}} \Pi_{F_s}} \check{f}(t)dt = \Pi_{F_s} f(\Pi_{F_s} M_{Q_{F_s}} \Pi_{F_s}).$$

□

As an immediate corollary of Theorems VIII.1 and IX.5, we obtain the following Szegö limit theorem:

**Corollary IX.10.** *Assume that $X_c$ is compact. Then for any smooth function $f$*

$$\mathrm{Tr}\left(\Pi_N f(\Pi_N \widehat{Q}_N \Pi_N)\right) = (2\pi)^{-n}\ N^n \int\limits_{X_c} f\circ Q\ \frac{\omega^n}{n!} + O(N^{n-1}\log(N)).$$

### 9.2.2 Commutators

We now describe another property of the projector. Let $\widehat{Q}$ be a semiclassical pseudodifferential operator as above. The projector $\Pi$ behaves microlocally as the identity on the interior $X_c$, suggesting that $\left[\Pi, \widehat{Q}\right]$ is microlocally $O(\hbar^\infty)$ on the diagonal. Using Proposition VII.15, we anticipate that $\left[\Pi, \widehat{Q}\right] \in \text{sc-}I^{-1/2}\left(M\times M; \mathcal{F}\partial X_c\right)$. We now prove that this is indeed the case, and compute the principal symbol of the commutator.

**Proposition IX.11.** *For any zeroth order compactly supported semiclassical pseudodifferential operator $\widehat{Q}$, $\left[\Pi, \widehat{Q}\right] \in$ sc-$I^{-1/2}\left(M \times M; \mathcal{F}\partial X_c\right)$ is a semiclassical Fourier integral operator, with (smooth) principal symbol*

$$\sigma\left(\left[\Pi, \widehat{Q}\right]\right)\left(\overline{x} = \Phi_s^P(\overline{y})\ ;\ \overline{y}\right) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{Q(\overline{y}) - Q(\overline{x})}{1 - e^{is}} & \text{if } \overline{x} \neq \overline{y} \\ \frac{1}{i\sqrt{2\pi}} \{P, Q\}(\overline{x}) & \text{if } \overline{x} = \overline{y}. \end{cases} \tag{9.10}$$

*Furthermore, if the principal symbol of $\widehat{Q}$ is constant along the orbits in the flow-out, and the Poisson bracket $\{P, Q\}$ vanishes on $\partial X_c$ to second order, and the subprincipal symbol of $Q$ vanishes on $\partial X_c$ (Levi condition), then $\left[\Pi, \widehat{Q}\right] \in I^{-5/2}\left(M \times M; \mathcal{F}\partial X_c\right)$.*

*Remark* IX.12. The Hamilton flow of the principal symbol of an operator that commutes with $\Pi$ preserves the region $X_c$. However, the converse is not true; the invariance of the region $X_c$ is a much weaker condition than the commuting property. As can be seen in the next chapter, the Poisson bracket vanishes at the boundary to second order if and only if the Hamilton flows of $P$ and $Q$ commute on the boundary, which is still a mild condition.

*Proof.* It is enough to prove it in the model case $\mathbb{R}^{n-1} \times S^1$ with coordinates $(x, \theta)$, and $T^*(\mathbb{R}^{n-1} \times S^1)$ with coordinates $(x, \theta\,;\, p, \tau)$. We consider only one energy level, say $E$ so that $X_c = \{\tau \geq E\}$. For simplicity, let us consider zeroth order semiclassical pseudodifferential operators of the form:

$$\widehat{Q}(\hbar) f(x, \theta) = \sum_m e^{im\theta} \int q(x, p, \theta, \hbar m) e^{ixp/\hbar} \hat{f}(p, m, \hbar) dp, \text{ where}$$

$$\hat{f}(p, m, \hbar) = \frac{1}{(2\pi\hbar)^{n+1}} \int e^{-iyp/\hbar} e^{-im\alpha} f(y, \alpha) dy d\alpha,$$

and $q(x, p, \theta, s)$ is the full symbol. Decompose $q$ in its Fourier modes,

$$q_k(x, p, \theta, s) = \sum e^{ik\theta} q_k(x, p, s), \text{ where } q_k(x, p, s) = \frac{1}{2\pi} \int q(x, p, \theta, s) e^{-ik\theta} d\theta$$

Then $\widehat{Q} = \sum \widehat{Q}_k$, where

$$\widehat{Q}_k(\hbar) f = \sum_m e^{im\theta} \int e^{ik\theta} q_k(x, p, \hbar m) e^{ixp/\hbar} \hat{f}(p, m, \hbar) dp \tag{9.11}$$

$$= \int f(y, \alpha) \left[ \frac{1}{(2\pi\hbar)^{n+1}} \sum_m e^{i(m+k)\theta - im\alpha} \int e^{i(x-y)p/\hbar} q_k(x, p, \hbar m) dp \right] dy d\alpha.$$

is a sc-$\Psi DO$ with symbol $q_k(x, p, s)$.

The kernel of $Q_k$ is

$$K_{Q_k}(x, y, \theta, \alpha) = \sum_m e^{i(k+m)\theta - im\alpha} p_k(x, y, m, \hbar),$$

where

$$p_k(x, y, m, \hbar) = \frac{1}{(2\pi\hbar)^{n+1}} \int e^{i(x-y)p/\hbar} q_k(x, p, \hbar m) dp.$$

A calculation shows that for $k > 0$,

$$K_{[\Pi_N, \widehat{Q}_k(1/N)]}(x, \theta, y, \alpha)$$

$$= \frac{N^{n+1}}{(2\pi)^{n+1}} \int e^{iN(x-y)p} e^{iNE(\theta-\alpha)} \left[ \sum_{0<j\leq k} e^{-ij(\theta-\alpha)} e^{ik\theta} \; q_k(x, p, E - j/N) \right] dp.$$

Notice that the amplitude in the integral above (in brackets) has an expansion in powers of $\hbar$, and in fact is a semiclassical symbol. The phase parametrizes the flow-out of $\{\tau = E\}$ by the canonical $S^1$ action, which is $\mathcal{F}\partial X_c$. This proves that the commutator is in the corresponding class, and the principal symbol is

$$\frac{1}{\sqrt{2\pi}} \sum_{0<j\leq k} e^{-ij(\theta-\alpha)} e^{ik\theta} q_k(x, p, E) = \frac{q_k(x, p, E) e^{ik\alpha} - q_k(x, p, E) e^{ik\theta}}{1 - e^{-i(\alpha-\theta)}}.$$

The case $k < 0$ is similar, and taking the sum over $k$, we obtain (9.10).

For the last part, assume that the principal symbol is constant in the fibers of

$\partial X_c \to S$, which implies that $\left[\Pi, \widehat{Q}\right] \in \text{sc-}I^{-3/2}(M \times M; \mathcal{F}\partial X_c)$. Assuming the Levi condition, we will show next that the principal symbol (corresponding to the degree $-3/2$) vanishes again. Take $\overline{x}_0, \overline{y}_0 \in \mathcal{F}\partial X_c$, $\overline{x}_0 \neq \overline{y}_0$. Consider two zeroth order semiclassical pseudodifferential operators $\widehat{T}_1$, $\widehat{T}_2$ of disjoint compact microsupport, such that their principal symbol is 1 in a neighborhood $\overline{x}$, $\overline{y}$ respectively. Notice that

$$\widehat{T}_1 \left[\Pi, \widehat{Q}\right] \widehat{T}_2 = \left[\widehat{T}_1 \Pi \widehat{T}_2, \widehat{Q}\right] + \widehat{T}_1 \Pi \left[\widehat{Q}, \widehat{T}_2\right] + \left[\widehat{Q}, \widehat{T}_1\right] \Pi \widehat{T}_2. \tag{9.12}$$

Near $(\overline{x}, \overline{y})$, the symbol of $\widehat{T}_1 \left[\Pi, \widehat{Q}\right] \widehat{T}_2$ and $\left[\Pi, \widehat{Q}\right]$ coincide. Consider, on the other hand, the first term on the right-hand side of equation (9.12). First, by the assumption on the microsupports of $T_1$ and $T_2$, the operator $\widehat{T}_1 \Pi \widehat{T}_2$ does not have wave-front set along the diagonal, and therefore it is in $\text{sc-}I^{-1/2}(M \times M; \mathcal{F}\partial X_c)$. We can then apply Proposition X.3 below to compute the symbol of the commutator $\left[\widehat{T}_1 \Pi \widehat{T}_2, \widehat{Q}\right]$. Near $(\overline{x}_0, \overline{y}_0)$ the symbol $\widehat{T}_1 \Pi \widehat{T}_2$ is equal to the symbol of $\Pi$, and clearly the (diagonal) Lie derivative of this symbol with respect to the Hamilton flow of $Q$ is zero. Therefore the symbol of this commutator (as an operator of order $-3/2$) is zero.

The principal symbols of the last two terms in (9.12) also vanish because the principal symbols of $T_1, T_2$ are constant near $\overline{x}_0, \overline{y}_0$, respectively. Therefore, the $(-3/2)$ principal symbol of $\left[\Pi, \widehat{Q}\right]$ vanishes off the diagonal, and therefore everywhere on $\mathcal{F}\partial X_c$ by continuity. This concludes the proof. □

## 9.3 Conclusions

Here we showed that under very mild conditions the algebra admits projectors. We defined the cut quantum observables, which can be combined with the theorem on the trace to provide a symbolic proof for a generalized Szegö limit theorem. We analyzed the commutator of the projector with semiclassical pseudodifferential operators to be used in the next section on the *propagator*.

# CHAPTER X

# On some propagators $e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi}$

## 10.1 The classical counterpart

We begin by considering classical Hamiltonians $Q : T^*M \to \mathbb{R}$ with the property that their Hamilton field is tangent to $\partial X_c$, that is, $\Xi_Q(\overline{x}) \in T_{\overline{x}}\partial X_c$ for all $\overline{x} \in \partial X_c$. It is easy to see that this occurs if and only if the Poisson bracket satisfies $\{P, Q\}_{|X_c} = 0$, and a as consequence, $Q$ is constant on the fibers of $\partial X_c \to S$. The Hamilton flow of such a $Q$ preserves the region $X_c$, that is, it defines a classical flow in the symplectic manifold with boundary $X_c$. The restriction of $Q$ to $\partial X_c$ descends to a smooth function $Q_S : S \to \mathbb{R}$, which in turn defines a Hamilton flow on $S$. However, the following diagram (where $\Phi_t^Q$ denotes the Hamilton flow of $Q$, etc.) does *not* commute in general:

$$\begin{array}{ccc} \partial X_c & \stackrel{\Phi_t^{Q}|_{\partial X_c}}{\longrightarrow} & \partial X_c \\ \downarrow & & \downarrow \\ S & \stackrel{\Phi_t^{Q_S}}{\longrightarrow} & S \end{array} \tag{10.1}$$

**Lemma X.1.** *Assume that $\Xi_Q$ is tangent to $\partial X_c$, and let $P$ be a defining function of $X_c$, as above. Then the Poisson bracket of $P$ and $Q$ vanishes to second order in the boundary $\partial X_c$ of $X_c$ if and only if the Hamilton flows of $\Phi^P$ and $\Phi^Q$ of $P$ and $Q$*

*commute on $\partial X_c$, i.e., for every $\overline{x} \in \partial X_c$, and any $s, t \in \mathbb{R}$,*

$$\Phi_t^Q \circ \Phi_s^P(\overline{x}) = \Phi_s^P \circ \Phi_t^Q(\overline{x}).$$

*Remark* X.2. In particular, if the conditions of the lemma above are satisfied, then the diagram (10.1) is commutative.

*Proof.* Since $\Xi_Q(\overline{x}) \in T_{\overline{x}}\partial X_c$ for all $\overline{x} \in \partial X_c = P^{-1}(0)$, then $\{Q, P\}(\overline{x}) = dP_{\overline{x}}(\Xi_Q) = 0$ for all $\overline{x} \in \partial X_c$. Therefore we can write $\{Q, P\} = FP$ for some smooth function $F$, and

$$[\Xi_Q, \Xi_P] = \Xi_{\{Q,P\}} = F\Xi_P + P\Xi_F$$

will vanish on $\partial X_c$ if and only if $F$ vanishes on $\partial X_c$. As a result,

$$\Phi_t^Q \circ \Phi_s^P(\overline{x}) = \Phi_s^P \circ \Phi_t^Q(\overline{x})$$

$\forall \overline{x} \in \partial X_c$ if and only if $\{Q, P\}$ vanishes to second order on $\partial X_c$. □

## 10.2 A Symbolic Description of The Propagator $e^{-it\ h^{-1}\Pi\widehat{Q}\Pi}$

Throughout this section $Q$ will denote a smooth function such that the Poisson bracket $\{P, Q\}$ vanishes to second order at $\partial X_c$ (c.f. Lemma X.1). As we saw in the previous section we then obtain a classical flow $\Phi_t^Q|_{X_c}$ that descends to a continuous flow on the cut space $Y$. In this section we analyze the quantum mechanical propagator $e^{-it\ h^{-1}\Pi\widehat{Q}\Pi}$, where $\widehat{Q}$ is a pseudodifferential operator on $M$ with symbol $Q$.

Before we state the main result, let us start with a proposition:

**Proposition X.3.** *For each semiclassical Fourier integral operator $\widetilde{V}(\hbar) \in sc\text{-}I^{-1/2}(M \times M; \mathcal{F}\partial X_c)$, the commutator $\left[\widehat{Q}(\hbar), \widetilde{V}(\hbar)\right]$ is in*

*sc-$I^{-3/2}(M \times M; \mathcal{F}\partial X_c)$. Its principal symbol is*

$$\sigma_{[\widehat{Q},\widetilde{V}]}(\bar{x},\bar{y}) = \frac{\hbar}{i}\mathfrak{L}_{\Xi_Q} V(\bar{x},\bar{y}),$$

*where $V$ is the principal symbol of $\widetilde{V}$, and $\mathfrak{L}_{\Xi_Q}$ is the Lie derivative obtained by letting the Hamilton flow of $Q$ act diagonally on $\mathcal{F}\partial X_c$.*

*Proof.* Write the Schwartz kernel of $\widehat{Q}$ in the model case as

$$K(\widehat{Q})(y,x) = \frac{1}{(2\pi)^n}\int e^{i(y-x)p/h} q(y,p,\hbar)dp,$$

where

$$q(x,p,\hbar) \sim q_0(x,p) + \hbar q_1(x,p) + \ldots,$$

and the Schwartz kernel of $\widetilde{V} \in \text{sc-}I^{-1/2}(M \times M\ ;\ \mathcal{F}\partial X_c)$ as

$$K(\widetilde{V})(z,y) = \frac{1}{(2\pi\hbar)^n}\int e^{i(z'-y')\omega/\hbar} v(z,y,\omega,\hbar)d\omega,$$

where

$$v(z,y,\omega,\hbar) \sim v_0(z,y,\omega) + \hbar v_1(z,y,\omega) + \ldots$$

The Schwartz kernel of $\widetilde{V} \circ \widehat{Q}$ is

$$K(\widetilde{V}\circ\widehat{Q})(z,x) = \frac{1}{(2\pi\hbar)^n}\int e^{i\hbar^{-1}(z'-x')}\Big[\frac{1}{(2\pi\hbar)^n} \int e^{i\hbar^{-1}((y_1-x_1)p_1+(y'-x')(p'-\omega'))} v(z,y,\omega,\hbar)q(y,p,\hbar)dydp\Big]d\omega$$

Applying the stationary phase method to the amplitude above, one gets

$$\frac{1}{(2\pi\hbar)^n}\int e^{i\hbar^{-1}((y_1-x_1)p_1+(y'-x')(p'-\omega'))} v(z,y,\omega,\hbar)q(y,p,\hbar)dydp$$

$$\sim v_0(z,x,\omega)q_0(x,p) - \tfrac{\hbar}{i}\tfrac{\partial v_0(z,x,\omega)}{\partial x}\tfrac{\partial q_0(x,p)}{\partial p}\Big|_{p=(0,\omega)} - \tfrac{\hbar}{i}v_0(z,x,\omega)\tfrac{\partial q_0(x,p)}{\partial x \partial p}\Big|_{p=(0,\omega)} \qquad (10.2)$$
$$+\hbar v_1(z,x,\omega)q_0(x,(0,\omega)) + \hbar v_0(z,x,\omega)q_1(x,(0,\omega)$$

The Schwartz kernel of $\widehat{Q}\circ\widetilde{V}$ can be written as

$$K(\widehat{Q}\circ\widetilde{V})(z,x) = \frac{1}{(2\pi\hbar)^n}\int e^{i(z'-x')\omega/\hbar}\Big[\frac{1}{(2\pi\hbar)^n}$$
$$\int e^{i\hbar^{-1}((y_1-z_1)(-p_1)+(y'-z')(-p'+\omega))}q(z,p,\hbar)v(y,x,\omega,\hbar)dpdy\Big]d\omega$$

Applying stationary phase to the amplitude above, one gets

$$\frac{1}{(2\pi\hbar)^n}\int e^{i\hbar^{-1}((y_1-z_1)(-p_1)+(y'-z')(-p'+\omega))}q(z,p,\hbar)v(y,x,\omega,\hbar)dpdy$$

$$\sim q_0(z,(0,\omega))v_0(z,x,\omega) + \tfrac{\hbar}{i}\tfrac{\partial q_0(z,p)}{\partial p}\Big|_{p=(0,\omega)}\tfrac{\partial v_0(z,x,\omega)}{\partial z} \qquad (10.3)$$
$$+\hbar q_1(z,(0,\omega))v_0(z,x,\omega) + \hbar q_0(z,(0,\omega))v_1(z,x,\omega)$$

Therefore, the commutator $[\widehat{Q},\widetilde{V}]$ has a leading amplitude

$$v_0(z,x,\omega)\left(q_0(z,(0,\omega)-q_0(x,(0,\omega))\right).$$

This vanishes at $x'=z'$ which corresponds the flow-out. Therefore $[\widehat{Q},\widetilde{V}]\in$ sc-$I^{-3/2}(M\times M\ ;\ \mathcal{F}\partial X_c)$. In order to compute the principal symbol there, we notice that

$$q_0(z,(0,\omega)-q_0(x,(0,\omega)) = (z'-x')d(z,x,\omega),$$

where

$$d(z,x,\omega)\big|_{x'=z'} = \frac{\partial q_0}{\partial x'}(x,(0,\omega))\big|_{x'=z'}.$$

Therefore

$$\frac{1}{(2\pi\hbar)^n}\int e^{i(z'-x')\omega/\hbar}\left(q_0(z,(0,\omega)-q_0(x,(0,\omega))\right)v_0(z,x,\omega)d\omega$$

$$= \frac{-\hbar}{i(2\pi\hbar)^n} \int e^{i(z'-x')\omega/\hbar} \sum_{j=1}^{n-1} \frac{\partial}{\partial \omega_j} \left( d(z,x,\omega) v_0(z,x,\omega) \right) d\omega \tag{10.4}$$

The principal symbol of $[\widehat{Q}, \widetilde{V}]$ in sc-$I^{-3/2}(M \times M \; ; \; \mathcal{F}\partial X_c)$ can be then computed taking all the contributions from equations (10.2), (10.3) and (10.2). Since $q_0(x,p)$ is constant along the orbits, then

$$\frac{\partial^2 q_0(x,p)}{\partial x_j \partial p_j} \Big|_{p=(0,\omega)} = \frac{\partial^2 q_0(z,p)}{\partial x_j \partial p_j} \Big|_{\substack{x'=z' \\ p=(0,\omega)}}, \text{ for any } j \geq 2$$

Since $\{Q, P\}$ vanishes at second order on $\partial X_c$, then $\frac{\partial^2 q_0(x,p)}{\partial x_1 \partial x_1}\Big|_{p_1=0} = 0$. Assuming all these conditions, and the fact that the subprincipal symbol

$$\text{Sub } \widehat{Q}(x,p) = q_1(x,p) - \frac{1}{2i} \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j \partial p_1} \Big|_{p=(0,\omega)} = 0$$

vanishes on $\partial X_c$, the principal symbol of the commutator reduces to

$$\frac{\partial q_0(z,p)}{\partial p} \Big|_{\substack{x'=z' \\ p=(0,\omega)}} \frac{\partial v_0(z,x,\omega)}{\partial z} \Big|_{x'=z'} + \frac{\partial q_0(x,p)}{\partial p} \Big|_{\substack{x'=z' \\ p=(0,\omega)}} \frac{\partial v_0(z,x,\omega)}{\partial x} \Big|_{x'=z'}$$

$$- \sum_{j=2}^{n} \frac{\partial q_0(x,p)}{\partial x_j} \Big|_{\substack{x'=z' \\ p=(0,\omega)}} \frac{\partial v_0(z,x,\omega)}{\partial \omega} \Big|_{x'=z'},$$

which can be checked to be the Lie derivative of $V$ with respect to the Hamiltonian field $\Xi_Q$. □

Notice that there is a way to define classes $J^{\ell,m}$ for a general pair of admissible Lagrangian submanifolds that intersect cleanly (see [GU81, MU79]), and not only for the diagonal and the flow-out of $\partial X_c$. The main result of this section is the following:

**Theorem X.4.** *Suppose $\widehat{Q}$ is a zeroth-order semiclassical pseudodifferential operator whose principal symbol satisfies the conditions of Lemma X.1. Assume sub$\widehat{Q}(\hbar) = 0$.*

*Then*

$$\Pi e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi}\Pi \in J^{-1/2,1/2}\left(M\times M;\Delta(t),\mathcal{F}\partial X_c(t)\right),$$

*where*

$$\Delta(t)=\left\{(\overline{x},\overline{y})\,\middle|\,\overline{x},\overline{y}\in T^*(M\times M),\overline{x}=\Phi_t^Q(\overline{y})\right\} \tag{10.5}$$

$$\mathcal{F}\partial X_c(t)=\left\{(\overline{x},\overline{y})\,\middle|\,\overline{x},\overline{y}\in\partial X_c,\ \exists s\in\mathbb{R}\ \textit{such that}\ \overline{x}=\Phi_s^P\Phi_t^Q(\overline{y})\right\} \tag{10.6}$$

*Remark* X.5. In this statement $t$ is a parameter, but we could also consider $t$ as a variable (in which case the kernel of the operator would be a family of functions on $\mathbb{R}\times M\times M$). Also, the symbols of $\Pi e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi}$ can easily be computed.

*Proof.* Let us define the following operator:

$$W(t):=\Pi e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi}\,e^{it\hbar^{-1}\widehat{Q}}. \tag{10.7}$$

We first prove the following:

**Lemma X.6.** *$W(t)\in J^{-1/2,1/2}\left(M\times M;\Delta,\mathcal{F}\partial X_c\right)$, and the principal symbol in the diagonal is $\sigma_0=\chi_{X_c}$.*

*Proof.* Let us define $D_t=\frac{1}{i}\frac{\partial}{\partial t}$. $W(t)$ satisfies the following equation

$$\begin{cases}\hbar D_tW(t)+\left[\widehat{Q},W(t)\right]+\left[\Pi,\widehat{Q}\right]W(t)=0\\[2ex] W\Big|_{t=0}=\Pi\end{cases} \tag{10.8}$$

Similarly to the proof of Lemma IX.8, using the symbol calculus we will construct a sequence of approximate solutions of equation (10.8). This construction makes the right-hand side of order $O(\hbar^\infty)$, and an application of Duhamel's principle concludes the proof.

As a first approximation we take $\widetilde{W}_0 = \Pi$. This is a sensible choice since

$$\hbar^2 \widetilde{S}_2 := \hbar D_t \widetilde{W}_0 + \left[\widehat{Q}, \widetilde{W}_0\right] + \left[\Pi, \widehat{Q}\right] \widetilde{W}_0 = -\left[\Pi, \widehat{Q}\right] (I - \Pi) \in \text{sc-}I^{-5/2}\left(M \times M; \mathcal{F}\partial X_c\right),$$

by Proposition IX.11.

We now modify $\widetilde{W}_0$ by elements in $I(M \times M; \mathcal{F}\partial X_c)$ to lower the order of the remainder. It is easy to see that the symbol of the correction term is the solution to the problem

$$\begin{cases} \frac{\partial V_1(\bar{x},\bar{y},t)}{\partial t} + \mathfrak{L}_{\Xi_Q} V_1(\bar{x}, \bar{y}, t) = -iS_2, \\ \\ V_1\big|_{t=0} = 0 \end{cases}$$

where $S_1$ is the principal symbol of $\widetilde{S}_1 \in \text{sc-}I^{-1/2}(M \times M; \mathcal{F}\partial X_c)$. The solution is

$$V_1(\bar{x}, \bar{y}, t) = -i \int_0^t e^{-(t-s)\mathfrak{L}_{\Xi_Q}} S_2(\bar{x}, \bar{y}, s) ds.$$

Let $\widetilde{V}_1 \in \text{sc-}I^{-1/2}(M \times M; \mathcal{F}\partial X_c)$ be an operator with this as symbol, and let $\widetilde{W}_1 = \widetilde{W}_0 + \hbar \widetilde{V}_1$.

Since $\left[\Pi, \widehat{Q}\right] \in \text{sc-}I^{-5/2}(M \times M; \mathcal{F}\partial X_c)$ by Proposition IX.11, we get

$$\hbar^3 \widetilde{S}_3 := \hbar D_t \widetilde{W}_1 + \left[\widehat{Q}, \widetilde{W}_1\right] + \left[\Pi, \widehat{Q}\right] \widetilde{W}_1 \in I^{-7/2}(M \times M; \mathcal{F}\partial X_c)$$

Proceeding inductively in this fashion, we obtain an infinite sequence $\{\widetilde{V}_j\}$ such that for all $J$

$$\hbar D_t \left(\widetilde{W}_0 + \sum_{j=1}^{J} \hbar^j \widetilde{V}_j\right) + \left[\widehat{Q}, \widetilde{W}_0 + \sum_{j=1}^{J} \hbar^j \widetilde{V}_j\right] + \left[\Pi, \widehat{Q}\right] \left(\widetilde{W}_0 + \sum_{j=1}^{J} \hbar^j \widetilde{V}_j\right)$$

$$= \hbar^{J+2} \widetilde{S}_{J+2} \in I^{-5/2-J}(M \times M; \mathcal{F}\partial X_c).$$

Next we take an operator $\widetilde{V} \in \text{sc-}I^{-1/2}(M \times M; \mathcal{F}\partial X_c)$ such that $\widetilde{V} \sim \sum_{j=1}^{\infty} \hbar^j \widetilde{V}_j$, and define $\widetilde{W} = \widetilde{W}_0 + \widetilde{V}$. □

Going back to the proof of the theorem, notice that

$$\Pi e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi} = W(t)e^{-itN\widehat{Q}}.$$

The Lagrangian $\Delta(t)$ intersects $\Delta$ and $\mathcal{F}\partial X_c$ transversally. Using a variation of the Proposition 4.1 in [GU81], we conclude that composing elements in $J^{-1/2,1/2}(M \times M; \Delta, \mathcal{F}\partial X_c)$ with $e^{-it\hbar^{-1}\widehat{Q}}$ gives elements in $J^{-1/2,1/2}(M \times M; \Delta(t), \mathcal{F}\partial X_c(t))$.

□

## 10.3 An Egorov-type Theorem

We can easily prove the following corollary.

**Corollary X.7.** *Let $\widehat{Q}$ be a zeroth order semiclassical pseudodifferential operator satisfying the conditions of Lemma X.1, and the Levi condition. Then for any zeroth order semiclassical pseudodifferential operator $\widehat{A}(\hbar)$, we have*

$$\widetilde{B}(t) := e^{it\hbar^{-1}\Pi\widehat{Q}\Pi}\Pi\widehat{A}\Pi e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi} \in J^{-1/2,1/2}(M \times M, \Delta, \mathcal{F}\partial X_c),$$

*with the following principal symbols:*

$$\sigma_0(\widetilde{B}(t))(\overline{x}, \overline{x}) = \chi_{X_c}(\overline{x}) \ \left(a \circ \Phi_t^Q(\overline{x})\right) \quad \textit{for } (\overline{x}, \overline{x}) \in \Delta \setminus \Sigma, \textit{and}$$

$$\sigma_1(\widetilde{B}(t))\Big|_{F_s} = \Pi_{F_s} M_{(a\circ\Phi_t^Q)_{|F_s}} \Pi_{F_s},$$

*where $a$ is the principal symbol of $\widehat{A}$, $\Phi_t^Q$ the Hamilton flow of $Q$, $F_s$ is an orbit in $\partial X_c$, and $\Pi_{F_s}$ is the Szegö projector of $F_s$.*

*Proof.* Let us consider $W(t)$ as in equation (10.7). Notice that

$$\widetilde{B}(t) = e^{it\hbar^{-1}\Pi\widehat{Q}\Pi}\Pi\widehat{A}\Pi e^{-itN\Pi\widehat{Q}\Pi} = W(-t)e^{it\widehat{Q}}\widehat{A}e^{-it\widehat{Q}}W(-t)^*,$$

which proves it belongs to the algebra. Since $e^{it\hbar^{-1}\widehat{Q}}\widehat{A}e^{-it\hbar^{-1}\widehat{Q}}$ is a semiclassical pseudodifferential operator with symbol $a\circ\Phi_t^Q$, the symbol on the diagonal can be trivially obtained. To compute the symbol on the flow-out, we note that the principal symbols of $W$ are exactly those of $\Pi$ and we use Proposition VI.7:

$$\sigma_1(\widetilde{B}(t))(\overline{y}=\Phi_s^P(\overline{x}),\overline{x}) = \frac{1}{\sqrt{2\pi}}\int \sigma_1\left(W(-t)\right)(\overline{y},\Phi_{\tilde{s}}^P(\overline{x}))\sigma_1\left(e^{it\widehat{Q}}\widehat{A}e^{-it\widehat{Q}}W(-t)^*\right)(\Phi_{\tilde{s}}^P(\overline{x}),\overline{x})d\tilde{s}$$

$$= \frac{1}{\sqrt{2\pi}}\int \frac{1}{\sqrt{2\pi}}\frac{1}{1-e^{i(s-\tilde{s})}}a(\Phi_t^Q\Phi_{\tilde{s}}^P(\overline{x}))\frac{1}{\sqrt{2\pi}}\frac{1}{1-e^{i\tilde{s}}}d\tilde{s}.$$

As a pseudodifferential operator on the fibers, this is $\Pi_{F_s}M_{(a\circ\Phi_t^Q)_{|F_s}}\Pi_{F_s}$. □

## 10.4 Conclusions

The goal in this chapter is the study of some singular propagators. We provided a symbolic description of the singular propagator $\Pi e^{-i\hbar^{-1}\Pi\widehat{Q}\Pi}$ in cases when the principal symbol $Q$ of $\widehat{Q}$ preserves the region $X_c$. An Egorov-type theorem was also proved.

# CHAPTER XI

# A numerical study of propagation

Let $\widehat{Q}$ be a zeroth order semiclassical pseudodifferential operator with symbol $Q$. It is well-known that, if $\psi_{(x_0,p_0)}$ is a coherent state with center at $(x_0, p_0)$, then $e^{-it\hbar^{-1}\widehat{Q}}(\psi_{(x_0,p_0)})$ is a coherent state (appropriately "squeezed") with center at $\Phi_t^Q(x_0, p_0)$, where $\Phi^Q$ is the Hamilton flow of $Q$. If the flow $\Phi^Q$ preserves $X_c$ and the center $(x_0, p_0)$ is in the interior of $X_c$, then the same conclusion holds for the propagation $e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi}(\psi_{(x_0,p_0)})$ of the coherent state by $\Pi\widehat{Q}\Pi$, as the trajectory of the center will remain away from the boundary $\partial X_c$ and everything is as if we were in the boundaryless case.

In this section we present results of a numerical calculation of $e^{-it\hbar^{-1}\Pi\widehat{Q}\Pi}(\psi_{(x_0,p_0)})$ in an example where the Hamilton flow of $Q$ does not preserve $X_c$, that is, trajectories of $\Phi^Q$ cross the boundary $\partial X_c$.

## 11.1 Numerical results for the Harmonic oscillator in Bargmann space

We consider the Harmonic oscillator $\widehat{P} = \frac{1}{2}(x^2 - \hbar^2\partial_x^2)$ in $\mathbb{R}^1$, and the corresponding projector $\Pi$ onto the span of its eigenfunctions of with eigenvalues less than or equal to one. We take $Q = x^2 - p^2$, and $\widehat{Q}$ the obvious quantization of $Q$. Figure 11.1 shows some energy levels of $Q$. Notice that the energy levels cross the boundary of

$X_c = \{x^2 + p^2 \leq 2\}$. The operators consider in the previous chapter have principal symbol that preserve $X_c$ and the energy levels of those principal symbols do not cross $\partial X_c$.

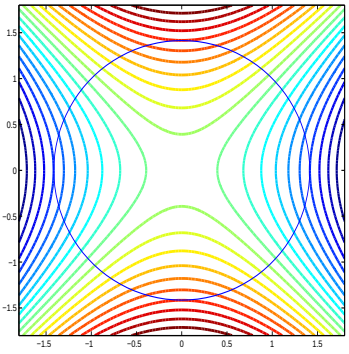



Figure 11.1: Energy levels of $Q$ and the boundary $\partial X_c = \{P = 1\}$.

We now take a coherent state centered inside the interior of $X_c$, and numerically compute its propagation under $\Pi\widehat{Q}\Pi$. We do the calculation in Bargmann space, for symplicity.

We recall that the Bargmann space is defined as the Hilbert space

$$\mathcal{B} = \left\{ f : \mathbb{C} \to \mathbb{C} : f \text{ entire and } \int |f(z)|^2 e^{-\frac{|z|^2}{\hbar}} dm_z < \infty \right\},$$

where $dm_z = dxdp$ and $z = \frac{x-ip}{\sqrt{2}}$, with the Hermitian inner product

$$\langle f, g \rangle = \int f(z)\overline{g(z)} e^{-\frac{z\overline{z}}{\hbar}} dm_z.$$

The Harmonic oscillator in Bargmann space is given by

$$\widehat{P} = \hbar z \frac{\partial}{\partial z} + \frac{\hbar}{2},$$

with principal symbol $P(z,\overline{z}) = z\overline{z}$ and eigenbasis

$$\left\{ b_n = \frac{z^n}{\sqrt{\hbar^n n!}} \right\}_{n\in\mathbb{N}\cup\{0\}}, \ \widehat{P}b_n = \hbar\left(n+\frac{1}{2}\right)b_n = \lambda_n b_n.$$

The quantization of $Q = x^2 - p^2$ in Bargmann space is the operator $\widehat{Q} = \hbar^2\frac{\partial^2}{\partial z^2} + z^2$. Applying $\widehat{Q}$ to the eigenbasis, we get

$$\widehat{Q}(b_n) = \hbar\sqrt{n(n-1)}b_{n-2} + \hbar\sqrt{(n+1)(n+2)}b_{n+2},$$

which gives a "generalized" Toeplitz matrix for $\Pi_N\widehat{Q}\Pi_N$ for each positive integer $N$, where $\hbar = \frac{1}{N}$. The (normalized) coherent state in Bargmann space, with center at $w$, is given by the simple formula

$$\Psi_w(z) = e^{\frac{z\overline{w}}{\hbar}}e^{-\frac{w\overline{w}}{2\hbar}}.$$

We apply the propagator $e^{-it\hbar^{-1}\Pi_N\widehat{Q}\Pi_N}$ to the projected coherent state

$$\Psi_w(z,N) = \Pi_N\Psi_w(z) = \sum_{n=0}^{N}\frac{\overline{w}^n}{\sqrt{n!\hbar^n}}e^{-\frac{\overline{w}w}{2\hbar}}b_n.$$

In Bargmann space, we measure the concentration in phase space of any semiclassical family $\psi$ by taking the absolute value of the family times the square root of the Bargmann weight, namely, by forming the Husimi density:

$$|\psi|_H(z) := |\psi(z)|\, e^{-z\overline{z}/2\hbar}.$$

We took as initial data a projected coherent state with center at $w = -0.25 - 0.6i$, which corresponds to $(x,p) = \sqrt{2}\,(-0.25, 0.6)$. Figure 11.2 consists of plots of the Husimi density in the $z$ variable of the initial projected coherent state, its propagation at $t = 0.25$ (approximately when the center of the coherent state hits the boundary),

and at time $t = 0.5$. We observe that after the collision time the coherent state splits into two localized states with centers on inward trajectories and with the same initial energy. Here we took $N = 100$. The splitting happens immediately after the

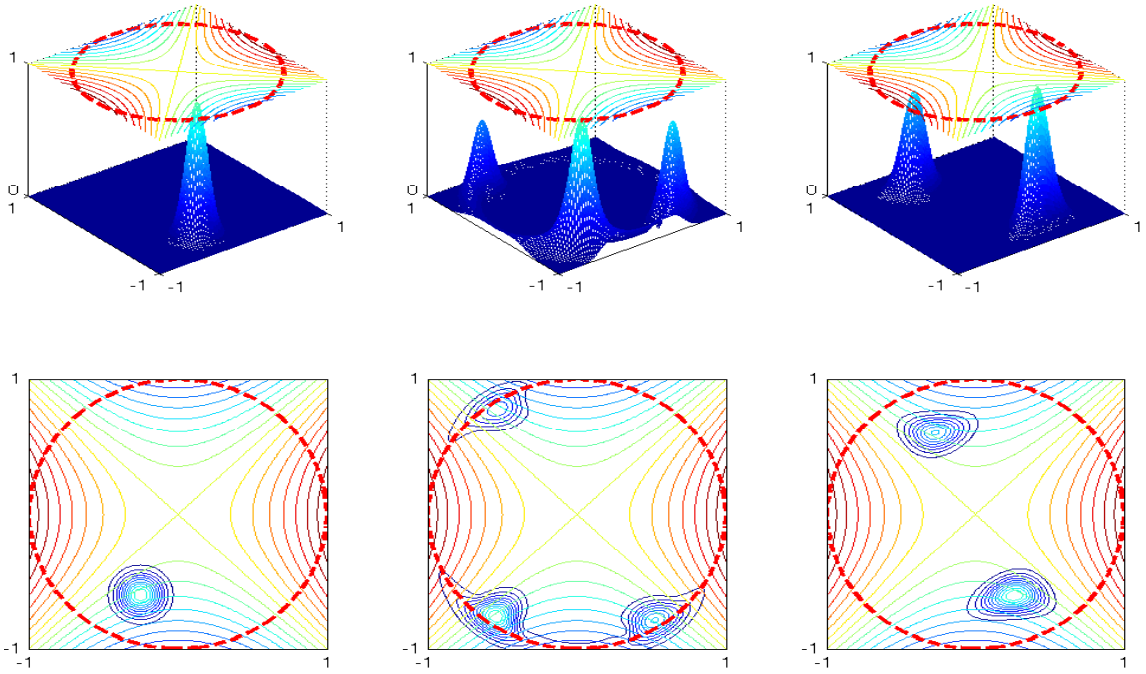


Figure 11.2:
A coherent state is propagated by $e^{-itN\Pi_N \hat{Q} \Pi_N}$ ($Q = x^2 - p^2$) at time $t = 0$ (left), $t = .25$ (middle) and $t = 0.5$ (right). The contour plots of the Husimi densities at each time and of $Q$ are also given (bottom).

center collides with the boundary; thus one can speak of infinite-propagation speed along the boundary. Note that the evolution is time-reversible, so that in some cases the opposite phenomenon will occur, namely, two localized states with same classical energy will hit the boundary at the same time and combine into one.

## 11.2 The cylinder case: $T^*S^1$

Consider $S^1$, and the cylinder $T^*S^1$ as its cotangent bundle. Let us define the operator $\widehat{P} = \hbar D_\theta$, $D_\theta = \frac{1}{i}\frac{\partial}{\partial \theta}$, with principal symbol $P(x, p) = p$. Suppose we make the symplectic cut from $p = 0$ to $p = 1$ in the momentum variable $p$. On the classical

side, it corresponds to the region

$$X_c = \left\{(x,p) \in T^*S^1 \ ; \ p \in [0,1]\right\}.$$

On the quantum side, it corresponds in the Hilbert space $L^2\left(S^1\right) = span\{\frac{e^{ik\theta}}{\sqrt{2\pi}}\}_{k\in\mathbb{Z}}$ to the image of the projector

$$\Pi_N : L^2\left(S^1\right) \longrightarrow \mathcal{H}_N, \text{ where } \mathcal{H}_N := \text{span } \left\{\frac{e^{ik\theta}}{\sqrt{2\pi}}\right\}_{0\leq k\leq N},$$

and $\hbar = \frac{1}{N}$ is the Planck's constant. As it was shown in Chapter IX, this projector belongs to the algebra $J^{-1/2,1/2}(S^1 \times S^1; \Delta, \mathcal{F}\partial X_c)$.

Let us consider a "coherent state" $u_N$ centered at a point on the cotangent bundle $X = T^*S^1$ and project it into $\mathcal{H}_N$. Let $\widehat{Q}$ be a zeroth order semiclassical pseudodifferential operator acting on smooth half-densities defined on $S^1$. As discussed in previous chapters, if $\widehat{Q}$ commutes with $\Pi$, the symbol of $Q$ is constant on the boundary of $X_c$. Classically, the propagated particle does not cross the boundary of $X_c$. The right-hand side of Figure 11.3 shows the energy levels in the reduced space $X_{cut}/\sim$, and we observe that they do not cross the poles.

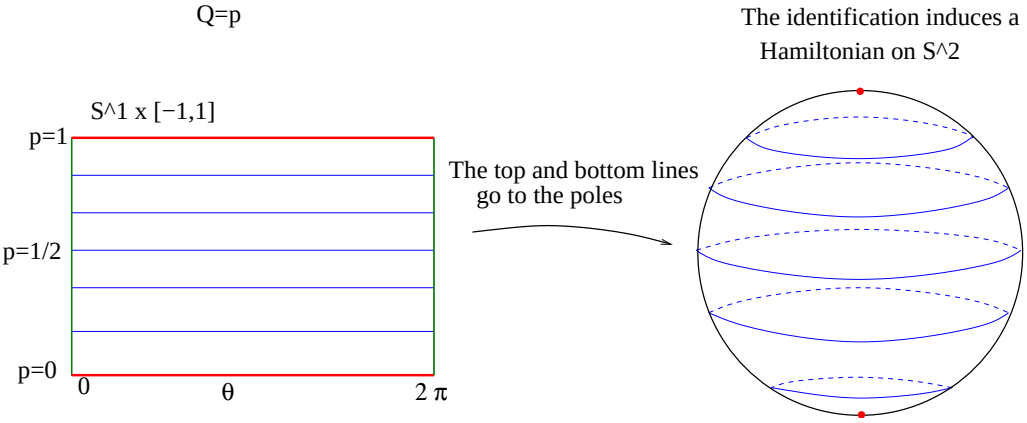



Figure 11.3: Examples of Hamiltonians that commute with the projector $\Pi_N$

When the operator $\widehat{Q}$ does not commute with the projector $\Pi$, the principal symbol

of $Q$ may not be constant in the boundary $\partial X_c$. In such a case, the energy levels of $Q$ may cross the boundary, and the induced Hamiltonian on $X_c/\sim$ becomes singular at the poles, as shown in Figure 11.4. The question of how to solve the Hamilton equations of a singular Hamiltonian arises. Our intuition indicates that a propagated coherent state passing through one pole may choose more than one trajectory to continue after the collision time. Which trajectory will it choose? Figure 11.4 shows the induced Hamiltonian when $Q = -(2p-1)\cos(\theta)$.

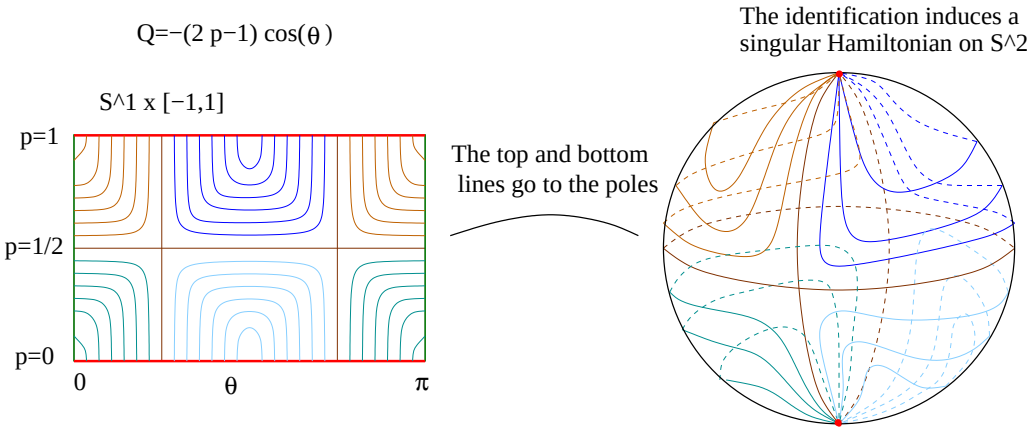



Figure 11.4: Examples of Hamiltonians no commuting with the projector Π. The identification induces a singular Hamiltonian where more than one classical trajectory may intersect at the poles.

### 11.2.1 Numerical discretization

In this section we describe the numerical scheme used to get the results we referred to above. Let $Q(\theta, p)$ be a Hamiltonian defined on $T^*S^1 \cong S^1 \times \mathbb{R}$, which can be identified with a $2\pi$-periodic function in the $\theta$ variable. So, $Q(\theta, p)$ can be written in the form

$$Q(\theta,p) = \sum_n a_n(p)e^{in\theta} = \sum Q_n(\theta,p), \qquad Q_n(\theta,p) = a_n(p)e^{in\theta}.$$

The Weyl quantization of $Q_n$ is given by

$$Op_\hbar^W Q_n(u)\,(\theta,\hbar) = \frac{1}{2\pi\hbar}\int e^{i\hbar^{-1}(\theta-\alpha)p} a_n(p) e^{i\frac{n}{2}(\theta+\alpha)} u(\alpha) d\alpha dp$$

$$= e^{i\frac{n}{2}\theta}\frac{1}{2\pi\hbar}\int e^{i\hbar^{-1}(\theta-\alpha)p} a_n(p)\left(e^{i\frac{n}{2}\alpha}u(\alpha)\right) d\alpha dp = e^{i\frac{n}{2}\theta} a_n^W\left(\hbar D_\theta\right)\left(e^{i\frac{n}{2}\theta}u(\theta)\right).$$

Moreover, we can compute the Fourier coefficient of each $Op_\hbar^W Q_n(u)$ as

$$\begin{aligned}\left\langle Op_\hbar^W\left(Q_n\right)(u)(\theta,\hbar), \frac{e^{im\theta}}{\sqrt{2\pi}}\right\rangle_{L^2(S^1)} &= \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} a_n^W\left(\hbar D_\theta\right)\left(e^{i\frac{n}{2}\theta}u(\theta)\right) e^{-i\left(m-\frac{n}{2}\right)\theta} d\theta \\ &= \frac{1}{\sqrt{2\pi}}\langle e^{i\frac{n}{2}\theta}u(\theta), \bar{a}_n^W\left(\hbar D\right)\left(e^{i\left(m-\frac{n}{2}\right)\theta}\right)\rangle_{L^2(S^1)}.\end{aligned}$$

Assume for simplicity that $a_n^W\left(\hbar D\right)$ is a differential operator. Then $a_n$ is a polynomial in $p$, and

$$\bar{a}_n^W\left(\hbar D\right)\left(e^{i\left(m-\frac{n}{2}\right)\theta}\right) = \bar{a}_n\left(\hbar\left(m-\frac{n}{2}\right)\right) e^{i\left(m-\frac{n}{2}\right)\theta}.$$

Therefore

$$\left\langle Op_\hbar^W\left(Q_n\right)(u)(\theta,\hbar), \frac{e^{im\theta}}{\sqrt{2\pi}}\right\rangle_{L^2(S^1)} = \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} e^{i\frac{n}{2}\theta}u(\theta) a_n\left(\hbar(m-\frac{n}{2})\right) e^{-i(m-\frac{n}{2})\theta} d\theta$$

$$= a_n\left(\hbar(m-\frac{n}{2})\right)\int_0^{2\pi} u(\theta)\frac{e^{-i(m-n)\theta}}{\sqrt{2\pi}} d\theta = a_n\left(\hbar\left(m-\frac{n}{2}\right)\right) u_{m-n},$$

where $u(\theta) = \sum_k u_k \frac{e^{ik\theta}}{\sqrt{2\pi}}$. Then, the $N+1$ by $N+1$ matrix associated to $\Pi_N \widehat{Q}_n \Pi_N$ is given by

$$T_N = \left(a_{m-k}\left(\frac{1}{N}\left(\frac{m+k}{2}\right)\right)\right)_{m,k=0}^{N}. \tag{11.1}$$

Now consider a delta function $\delta_{e^{i\alpha}}$, whose microsupport is given by the vertical line $\left\{(\alpha,p)\middle| p\in\mathbb{R}\right\}$. As a result, the projected delta functions $\Pi_N\delta_{e^{i\alpha}}$ have microsupport

$\{(\alpha, p)\big|0 \leq p \leq 1\}$, and

$$\left\langle \delta_{e^{i\alpha}}, \frac{e^{ik\theta}}{\sqrt{2\pi}} \right\rangle_{L^2(S^1)} = \int \delta_{e^{i\alpha}} \frac{e^{-ik\theta}}{\sqrt{2\pi}} d\theta = \frac{1}{\sqrt{2\pi}} e^{-ik\alpha}.$$

Therefore

$$\Pi_N \left(\delta_{e^{i\alpha}}\right) = \sum_{k=0}^{N} \frac{1}{\sqrt{2\pi}} e^{-ik\alpha} \frac{1}{\sqrt{2\pi}} e^{ik\theta}.$$

One discretizes the interval $[0, 2\pi]$ with a grid of size $N+1$, and ignore the last entry due to periodicity. So consider $\alpha_m = \frac{2\pi m}{N+1}$, $0 \leq m \leq N$. The change of variables matrix from $\left\{\frac{e^{ik\theta}}{\sqrt{2\pi}}\right\}_{k=0}^{N}$ to $\{\Pi_N \delta_{e^{i\alpha_m}}\}_{m=0}^{N}$ is given by the matrix

$$F_N = (F_N(k,m))_{m,k=0}^{N} = \left(\frac{1}{\sqrt{2\pi}} e^{-i\frac{2\pi mk}{N+1}}\right)_{k,m=0}^{N},$$

i.e, $\Pi_N \delta_{e^{i\alpha_k}} = F_N \cdot \Pi_N \frac{e^{ik\theta}}{\sqrt{2\pi}}$. As a result, we can express the projected eigenfunctions $\Pi_N \frac{e^{ik\theta}}{\sqrt{2\pi}}$ in terms of the projected Dirac deltas $\Pi_N \frac{e^{ik\theta}}{\sqrt{2\pi}} = F_N^{-1} \cdot \Pi_N \delta_{e^{i\alpha_k}}$.

*Remark* XI.1. The matrix $F_N$ is known as the *Discrete Fourier Transform*.

### 11.2.2 Numerical frequency set and numerical coherent states

A family of projected vector functions $u_N$ on the subspace $\mathcal{H}_N = span\left\{\frac{e^{ik\theta}}{\sqrt{2\pi}}\right\}_{k=0}^{N}$ is represented by an $N+1$-dimensional vector $f_N$ depending on $\hbar = \frac{1}{N}$. The entries of $f_N$ are the Fourier coefficients. Then, $F_N^{-1} f_N$ are the coefficients when the family of projected functions are expressed in terms of the projected delta functions. The propagation of $u_N$ by $T_N = \Pi_N \widehat{Q}_N \Pi_N$ is represented by the vector $e^{-itNT_N} f_N$, where $T_N$ is the $N+1$ by $N+1$ matrix defined as in 11.1.

The frequency set of a family of projected functions is found numerically as follows. We recall that the eigenfunctions $\frac{1}{\sqrt{2\pi}} e^{ik\theta}$ have microsupport at the level $p = \hbar k = \frac{k}{N}$, and the projected dirac deltas $\Pi_N \delta_{e^{i\alpha_m}}$ at $\{(\theta = \alpha_m, p)\big|0 \leq p \leq 1\}$. With this principle in mind, we conclude that for any family of projected $\hbar$-dependent functions

$f_N$, the $k$-th entry gives an indication of the concentration of the family at $p = \frac{k}{N}$, and the $m$-th entry of the vector $g_N = F_N^{-1} f_N$ gives the concentration at the vertical line passing trough $\alpha_m$. Thus, the product $|f_{N,k}|^2 |g_{N,m}|^2$ provides the concentration of the semiclassical projected family at the point $\left(\theta = \alpha_m, p = \frac{k}{N}\right)$ in phase space, where $0 \leq k, m \leq N$.

We use the following family of function as a coherent state:

$$u_N = \sum_{m=0}^{N} e^{-\left|\alpha_m - \alpha_{m_0}\right|^{3/2} N^{1/2}} \Pi_N \left(\delta_{e^{i\alpha_m}} e^{ik_0\theta}\right),$$

This coherent state has center at $\left(\theta_0 = \alpha_{m_0}, p_0 = \frac{k_0}{N}\right)$, $0 \leq k_0, m_0 \leq N$, and corresponds to the vector whose $k$-th entry is

$$\sum_{m=0}^{N} e^{-\left|\alpha_m - \alpha_{m_0}\right|^{3/2} N^{1/2}} \frac{1}{\sqrt{2\pi}} e^{-(k-k_0)\alpha_{m_0}}.$$

### 11.2.3 Numerical results in the cylinder case

We consider the propagation under the singular operator $e^{-itN\Pi_N \widehat{Q}_N \Pi_N}$, where $\widehat{Q}$ is the Weyl quantization of the Hamiltonian

$$Q = -(2p-1)\cos(\theta).$$

This operator does not commute with the projection $\Pi_N$ as it does not leave the subspace $\mathcal{H}_N$ invariant. Figure 11.5 shows the concentration of the propagated coherent state at different times (top) together with a top view (bottom). The center of the coherent state is $(\alpha = 1.8\pi, p = 0.8)$. We observe that after the collision time, the center of the coherent state jumps to another point in the boundary, and takes an inward curve with the same energy.

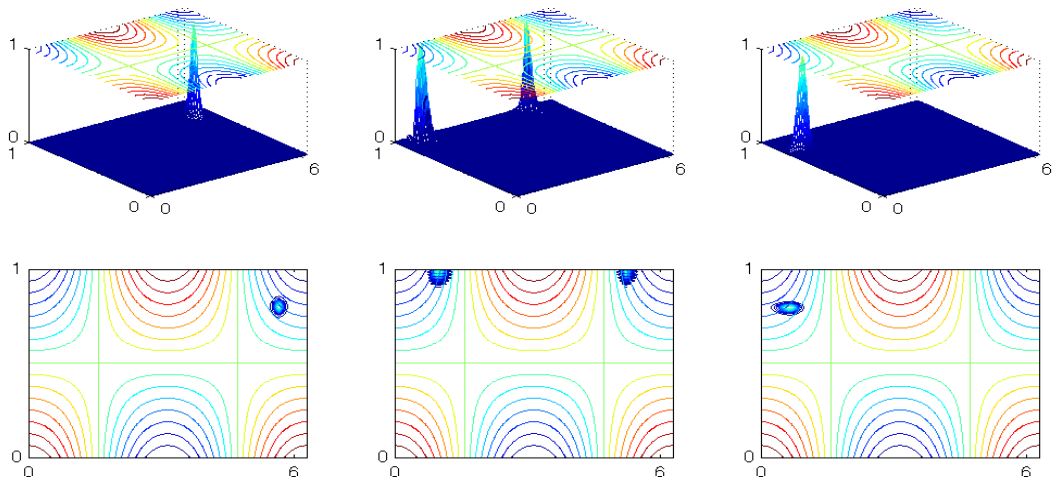


Figure 11.5: A coherent state concentrated at $(\alpha = 1.8\pi, p = 0.8)$ is propagated by $e^{-it\hbar^{-1}\Pi_N \widehat{Q} \Pi_N}$ $(Q = -(2p-1)\cos(\theta))$ at time $t = 0$ (left), $t = .353$ (middle) and $t = 0.706$ (right).

## 11.3 Conclusions

The goal in the last chapter is the numerical study of singular propagators that are not covered theoretically in the previous chapter. We did numerical analysis on the propagation of coherent states under the present propagators. We first considered the case $M = R^1$ and used the Bargmann space for simplicity. We observed that the center of the coherent state may split into more than one as it approaches the boundary $\partial X_c$. A similar analysis was done for the case $M = S^1$. Here we described how we measured the frequency set of projected functions in a numerical way using the discrete Fourier transform. We observed the same numerical phenomena in both cases.

# APPENDICES

# APPENDIX A

# Eigenvectors for the Reduced System (3.15)

## Eigenvectors in Different Formulations

The eigenvectors of different formulations of the hyperbolic system (3.6) are related through the Jacobian matrix of the nonlinear transformation between the respective sets of the dependent variables. Here, we obtain the eigenvectors of (3.15) by an equivalent but more direct calculation.

We first note that across the $k^{th}$ simple wave, $\delta W \sim r_k$, where $\delta$ denotes the total differential. This implies, for example that across a 1-wave, a small change in the density, $\delta\tilde{\rho}$, is accompanied by a $(u-c)\delta\tilde{\rho}$ change in the momentum, $\tilde{\rho}u$, and a $(h-uc)\delta\tilde{\rho}$ change in the energy $\tilde{E}$. Alternatively, across a 1-wave, using the eigenvectors for the primitive formulation (3.7), a small change in the density $\delta\rho$ is accompanied by a small change in velocity $\delta u = -(c/\rho)\delta\rho$ and a small change in pressure $\delta p = c^2\delta\rho$, but no change in porosity $\delta\phi = 0$. A similar interpretation applies to the other waves.

To obtain $r_1^{NC}$ for the nonconservative set of variables $W^{NC} = (\tilde{\rho}, \eta, \tilde{E})$, we observe that

$$\delta\tilde{\rho} = \delta\phi\rho = \phi\delta\rho + \rho\delta\phi = \phi\delta\rho$$

and using logarithmic differentiation of $\eta = p\rho^{-\gamma}$ we get

$$\frac{\delta\eta}{\eta} = \frac{\delta p}{p} - \gamma\frac{\delta\rho}{\rho} = \frac{c^2\delta\rho}{p} - \gamma\frac{\delta\rho}{\rho} \equiv 0$$

implying that a 1-wave does not carry any changes in entropy. The first eigenvector for the nonconservative formulation (3.15) is therefore $r_1^{NC} = (1, 0, h - uc)$. The calculation for the 3-wave is practically identical. For the 2-wave, a similar calculation gives

$$\frac{\delta\eta}{\eta} = \frac{\delta p}{p} - \gamma\frac{\delta\rho}{\rho} = -\gamma\frac{\delta\rho}{\rho} = -\gamma\frac{\delta\tilde{\rho}}{\tilde{\rho}}$$

implying that across a 2-wave

$$\delta\eta = -\frac{\gamma\eta}{\tilde{\rho}}\delta\tilde{\rho}.$$

Combining results together, we obtain

$$R^C = \begin{pmatrix} 1 & 1 & 1 \\ u-c & u & u+c \\ h-uc & \frac{1}{2}u^2 & h+uc \end{pmatrix} \qquad R^{NC} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -\frac{\gamma\eta}{\tilde{\rho}} & 0 \\ h-uc & \frac{1}{2}u^2 & h+uc \end{pmatrix}$$

The wave strengths $\alpha_k$ are obtained as follows. Away from the porosity jump, the wave strengths for the conservative formulation are given by the familiar expressions

$$\alpha_1^C = \frac{\delta\tilde{p} - \tilde{\rho}c\ \delta u}{2c^2},\ \alpha_2^C = \frac{c^2\delta\tilde{\rho} - \delta\tilde{p}}{c^2},\ \alpha_3^C = \frac{\delta\tilde{p} + \tilde{\rho}c\delta u}{2c^2}$$

Near the interface, we express the wave strengths in terms of the Riemann invariants $\tilde{\rho}u$, $\eta$ and $h$. This guarantees the preservation of interface data.

$$A^{NC}\delta W^{NC} = \begin{pmatrix} \delta\left(\tilde{\rho}u\right) \\ u\delta\eta \\ \delta\left(\tilde{\rho}uh\right) \end{pmatrix} = \begin{pmatrix} \delta\left(\tilde{\rho}u\right) \\ u\delta\eta \\ \tilde{\rho}u\delta h \ + \ h\delta\tilde{\rho}u \end{pmatrix} = \sum \alpha_k \lambda_k r_k^{NC}$$

A simple calculation gives

$$\begin{aligned} z_1 = \alpha_1\lambda_1 &= \frac{1}{2}\delta\left(\tilde{\rho}u\right) + \frac{\tilde{p}\left(c + (\gamma-1)u\right)}{2(\gamma-1)\ \eta c^2}\delta\eta - \frac{\tilde{\rho}}{2c}\delta h \\ z_2 = \alpha_2\lambda_2 &= -\frac{\tilde{p}}{\eta c^2}\ u\delta\eta \\ z_3 = \alpha_3\lambda_3 &= \frac{1}{2}\delta\left(\tilde{\rho}u\right) - \frac{\tilde{p}\left(c - (\gamma-1)u\right)}{2(\gamma-1)\ \eta c^2}\delta\eta + \frac{\tilde{\rho}}{2c}\delta h, \end{aligned}$$

as desired.

The eigenvectors for both the conservative and non-conservative formulation for the full system (3.19) and (3.20) are obtain in an analogous way, and are given in section 3.4.

# APPENDIX B

# Lagrangian Distributions and Fourier Integral Operators

## Lagrangian Distributions and Phase Functions

Here we review briefly the notion of Lagrangian distributions and their principal symbols. See [Dui96, Hör85] for all the details.

Let $\mathcal{D}'(X)$ be the set of distributions on $X$, $X$ open in $\mathbb{R}^n$, and let $u \in \mathcal{D}'$. According to the Paley-Wiener theorem, $u$ is $C^\infty$ in a neighborhood of $x \in X$ ($x \notin$ sing supp$u$) if and only if there exists $\varphi \in C_0^\infty$ such that $\varphi(x) \neq 0$ and

$$\mathcal{F}(\varphi u)(\xi) = O(|\xi|^{-N}) \text{ for } |\xi| \to \infty, \text{ and all } N,$$

where $\mathcal{F}$ is the (non semiclassical) Fourier transform

$$\mathcal{F}(u)(\xi) = \widehat{u}(\xi) := \frac{1}{(2\pi)^{n/2}} \int e^{-ix\xi} u(x) dx.$$

The singular support (sing supp$u$) describes where the distribution has singularities. The wave front set describes the "position and direction" of the singularities in

phase space.

**Definition B.1.** If $u \in \mathcal{D}'(X)$, then the wave front set $WF(u)$ of $u$ is defined as the complement in $X \times (\mathbb{R}^n \setminus \{0\})$ of the collection of all $(x_0, \xi_0) \in X \times (\mathbb{R}^n \setminus \{0\})$ such that for some neighborhood $U$ of $x_0$, $V$ of $\xi_0$ we have that for $\varphi \in C_0^\infty(U)$ and each $N$,

$$\mathcal{F}(\varphi u)(\tau\xi) = O(\tau^{-N}) \text{ for } \tau \to \infty, \text{ uniformly in } \xi \in V.$$

**Definition B.2.** Let $V$ be a conic manifold, $\mu, \rho \in \mathbb{R}$, $0 \leq \rho \leq 1$. A symbol on $V$ of order $\mu$ and type $\rho$ is a function $a \in C^\infty(V)$ such that

$$\tau^*(L_k \cdot \ldots \cdot L_1 a) = O(\tau^{\mu - k\rho}) \text{ for } \tau \to \infty,$$

locally uniformly in $V$ and for all $C^\infty$ vector fields $L_1, \ldots, L_k$ in $V$ that are homogeneous of order $-1$. Here $\tau^* f(v) = f(\tau v)$ for $v \in V$. The space of these symbols is denoted by $S_\rho^\mu(V)$.

A real function $\phi \in C^\infty\left(X \times \mathbb{R}^N \setminus \{0\}\right)$ is called a phase function if it is homogeneous of degree 1 ($\phi(x, \tau\theta) = \tau\phi(x, \theta)$), and $d_{x,\theta}\phi(x, \theta) \neq 0$ for all $(x, \theta) \in X \times \mathbb{R}^N$ (i.e. $\phi$ has no critical points). The integral

$$I_\phi(au) = \int\int e^{i\phi(x,\theta)} a(x, \theta) u(x) dx d\theta, \ u \in C_0^\infty(X),$$

is absolutely convergent if $a \in S_\rho^\mu(X \times \mathbb{R}^N)$, and $\mu + N < 0$. The mapping $u \to I_\phi(au)$ is continuous on $C_0^0(X)$ and therefore defines a distribution of order zero. The following is a series of theorems from [Dui96] that summarizes the theory behind Lagrangian distributions.

**Theorem** *Suppose $\rho > 0$. Then the mapping $a \to I_\phi(au)$, defined for symbols $a$ that vanish for large $|\theta|$, can for every $u \in C_0^\infty(X)$ be extended to $S_\rho^\infty(X \times \mathbb{R}^N)$ such that is continuous on $S_\rho^\mu(X \times \mathbb{R}^N)$ for every $\mu$. Moreover, for every $a \in S_\rho^\mu(X \times \mathbb{R}^N)$ the*

*linear form $A : u \to I_\phi(au)$ is a distribution of order $k$ is $\mu - k\rho + N < 0$.*

**Theorem** *Let $a \in S^m_\rho$, $\rho > 0$ and $\phi$ as above. Then*

$$WF(A) \subset \left\{(x, d_x\phi(x,\theta)) \in T^*(X) \;\middle|\; (x,\theta) \in \textit{ess-supp}(a),\ d_\theta\phi(x,\theta) = 0\right\}.$$

*Here ess-supp$(a)$ is defined as the smallest conic subset of $X \times \mathbb{R}^N \setminus \{0\}$ outside of which $a$ is of class $S^{-\infty}$.*

**Definition B.3.** A symbol $a \in S^m$ is called classical if it has an asymptotic expansion

$$a \sim a_m + a_{m-1} + \ldots,$$

where each $a_j$ is an homogeneous symbol of degree $j$, smooth at $X \times \mathbb{R}^N \setminus 0$.

**Definition B.4.** Let $\Gamma$ be a cone in $X \times \mathbb{R}^N \setminus \{0\}$. The phase function $\phi$ is called nondegenerate on $\Gamma$ if

$$d_\theta\phi(x,\theta) = 0, (x,\theta) \in \Gamma$$

implies that $d_{(x,\theta)}\frac{\partial\phi(x,\theta)}{\partial\theta_j}$ are linearly independent for $j = 1, \ldots, N$.

**Lemma** *If $\phi$ is a non-degenerate phase function, then*

$$T^{(\phi)} : (x,\theta) \mapsto (x, d_x\phi(x,\theta))$$

*is an immersion: $C_\phi \to T^*(X) \setminus 0$, commuting with the multiplication with positive numbers in the fibers. So its image $\Lambda_\phi$ is an immersed $n$-dimensional conic submanifold of $T^*(X) \setminus 0$.*

**Theorem** *Suppose $\phi(x,\theta)$ and $\tilde\phi(x,\tilde\theta)$ are nondegenerate phase functions at $(x,\theta_0) \in X \times \mathbb{R}^N \setminus \{0\}$ and at $(x_0,\tilde\theta_0) \in X \times \mathbb{R}^{\tilde N} \setminus \{0\}$, respectively. Let $\Gamma$ and $\tilde\Gamma$ be open conic neighborhoods of $(x_0,\theta_0)$ and $(x_0,\tilde\theta_0)$ such that $T_\phi : C_\phi \to \Lambda_\phi$ and $T_{\tilde\phi} : C_{\tilde\phi} \to \Gamma_{\tilde\phi}$ are injective, respectively. If $\Lambda_\phi = \Lambda_{\tilde\phi}$, then for any Fourier integral $A$, defined for*

*the phase function $\phi$ and amplitude $s \in S^{\mu}_{\rho}(X \times \mathbb{R}^N)$, $\rho > 1/2$, with ess-supp $(a)$ contained in a sifficiently small conic neighborhood of $(x_0, \theta_0)$, is equal to a Fourier integral defined by the phase function $\tilde{\phi}$ and an amplitude $\tilde{a} \in S^{\mu+\frac{1}{2}(N-\tilde{N})}_{\rho}(X \times \mathbb{R}^{\tilde{N}})$.*

Lagrangian distributions can be invariantly defined on $C^\infty$ manifolds $M$, and the class is denoted by $I^m(M;\Lambda)$, where $\Lambda \subset T^*M \setminus \{0\}$ is a Lagrangian conic submanifold, and $m$ is the order. Elements in the class $I^m(M;\Lambda)$ are locally sums of Fourier integrals with phase functions defining the Lagrangian submanifold $\Lambda$. The theorem above says that changing to a different phase function defining the same Lagrangian manifold gives the same class, changing the amplitude. Even though the amplitude changes if we impose a change of coordinates, the *principal symbol* of the operator stays invariant. Due to technicalities, the principal symbol has to be interpreted as a half-density on the Lagrangian manifold $\Lambda$ tensored with a Maslov bundle.

## The Principal Symbol

The following is Proposition 25.1.5 from [Hör85].

**Proposition** *Let $\phi(x,\theta)$ be a non-degenerate phase function on an open conic neighborhood of $(x_0,\theta_0) \in \mathbb{R}^n \times (\mathbb{R}^N \setminus 0)$ wich parametrizes the Lagrangian manifold $\Lambda$ in a neighborhood of $(x_0,\xi_0)$, $\xi_0 = \phi'_x(x_0,\theta_0)$, $\phi'_\theta(x_0,\theta_0) = 0$. If $a \in S^{m+(n-2N)/4}(\mathbb{R}^n \times \mathbb{R}^N)$ has support in the interior of a sufficiently small conic neighborhood $\Gamma$ of $(x_0,\theta_0)$, then the oscillatory integral*

$$u(x) = (2\pi)^{-(n+2N)/4} \int e^{\phi(x,\theta)} a(x,\theta) d\theta$$

*defines a distribution $u \in I^m(\mathbb{R}^n,\Lambda)$. If $\Lambda = \{H'(\xi),\xi\}$ where $H$ is a real valued function in $C^\infty(\mathbb{R}^n \setminus 0)$, homogeneous of degree 1, then $e^{iH(\xi)}\hat{u}(\xi) =: (2\pi)^{n/4} v(\xi) \in$*

$S^{m-n/4}$, *and (for* $|\xi| > 1$*)*

$$e^{iH(\xi)}\widehat{u}(\xi) - (2\pi)^{n/4}a(x,\theta)e^{\pi i/4 sign(\Phi)}|det\Phi|^{-1/2} \in S^{m-n/4-1}, \tag{B.1}$$

*where* $(x,\theta)$ *is determined by* $\phi'_\theta(x,\theta) = 0, \phi'_x(x,\theta) = \xi$, *and*

$$\Phi = \begin{pmatrix} \phi''_{xx} & \phi''_{x\theta} \\ \phi''_{\theta x} & \phi''_{\theta\theta} \end{pmatrix}$$

We will show that the second term in equation (B.1), interpreted as as a 1/2-density on $\Lambda$, is independent of coordinates when we tensor it with a Maslov bundle of $\Lambda$. This will be true even if we change the amplitude and phase function for the same distribution, as in the proposition above.

For an open subset $\mathcal{U} \subset \mathbb{R}^k$, and $W : \mathcal{U} \to \mathbb{R}^N, u \to \omega, \mathcal{U} \subset \mathbb{R}^k$ smooth, assume that zero is a regular value. Then $W^{-1}(0)$ has a natural 1-density on it, namely the quotient of $|dz|$ in $\mathbb{R}^N$ by $d\omega$ in $\mathbb{R}^k$. That is, let $\{e_1, \ldots, e_{k-N}\}$ be a basis of $T_z W^{-1}(0)$. Let Let $\{f_1, \ldots, f_N\}$ be vectors in $\mathbb{R}^k$ such that $\{e_1, \ldots, e_{k-N}, f_1, \ldots, f_N\}$ is a basis of $\mathbb{R}^k$, then the density on $W^{-1}(0)$ on $e_1, \ldots, e_{k-N}$ is defined as

$$\frac{|e_1, \ldots, e_{k-N}, f_1, \ldots, f_N|_{\mathbb{R}^k}}{|dw(f_1), \ldots, dw(f_N)|_{\mathbb{R}^N}}.$$

**Definition B.5.** If $\phi$ is a non-degenerate phase function, denote by $\mathcal{D}_\phi$ the density on $C_\phi$ defined by the function

$$C_\phi \xrightarrow{d_\theta \phi} \Lambda$$

**Lemma** *If* $\Lambda = \{(\nabla\xi, \xi)\}$, *then* $\mathcal{D}_\phi = \frac{|d\xi|}{|detQ|}$. *Hence*

$$v(\xi)|d\xi|^{1/2} \equiv a(x,\theta)e^{\pi i/4 sign(\Phi)}\sqrt{\mathcal{D}_\phi} mod S^{m+/4-1}(\Lambda, \Omega^{1/2}),$$

*where $\Omega^{1/2}$ is the half-density bundle on $\Lambda$.*

If we now introduce new coordinates $\tilde{x}$ and transform $\tilde{u}$ as a half-density, that is, $\tilde{u}(\tilde{x}) = |Dx/D\tilde{x}|^{1/2}u(x)$, then

$$\tilde{u}(\tilde{x}) = (2\pi)^{-(n+2N)} \int e^{i\tilde{\phi}(\tilde{x})}\tilde{a}(\tilde{x},\theta)d\theta,$$

$$\tilde{\phi}(\tilde{x},\theta) = \phi(x,\theta),\ \tilde{a}(\tilde{x},\theta) = |Dx/D\tilde{x}|^{1/2}a(x,\theta)$$

and

$$\tilde{a}\sqrt{\mathcal{D}_{\tilde{\phi}}} = a\sqrt{\mathcal{D}_{\phi}},$$

which proves that the half-density is invariant under change of coordinates. If, in addition, we change the phase function

$$u(x) = (2\pi)^{-(n+2\tilde{N})/4} \int e^{i\tilde{\phi}(x,\tilde{\theta})}\tilde{a}(x,\tilde{\theta})d\theta,$$

then

$$\tilde{a}\sqrt{\mathcal{D}_{\tilde{\phi}}} - e^{\pi is/4}a\sqrt{\mathcal{D}_{\phi}} \in S^{m+n/4-1}(\Lambda,\Omega_{\Lambda}^{1/2}),$$

in the common domain of definition on $\Lambda$, if

$$s = sgn\phi''_{\theta\theta}(x,\theta) - sgn\tilde{\phi}''_{\tilde{\theta}\tilde{\theta}}(x,\tilde{\theta}) \tag{B.2}$$

This provides the definition of the principal symbol.

**Definition B.6.** For every $u \in I^m(X,\Lambda)$, we get a principal symbol in

$$S^{m+n/4}(\Lambda, M_{\Lambda}\otimes\Omega_{\Lambda}^{1/2})/S^{m+n/4-1}(\Lambda, M_{\Lambda}\otimes\Omega_{\Lambda}^{1/2}),$$

where $M_{\Lambda}$ is the locally constant (Maslov) line bundle. It is defined by a covering $\Lambda = \cup\Gamma_j$ of $\Lambda$ with open cones $\Gamma_j$ and transition functions which are just $e^{\pi is/4}$, where

$s$ is given by equation (B.2).

**Definition B.7.** Consider two $C^\infty$ manifolds $M_1$ and $M_2$. Let $C$ be a homogeneous canonical relation from $T^*M_1\backslash 0$ to $T^*M_2\backslash 0$, which is closed in $T^*(M_2\times M_2)\backslash 0$, and let $E_1, E_2$ be vector bundles on $M_1, M_2$ respectively. Then operators with Schwartz kernel belonging to $I^m(M_2 \times M_2, C'; \Omega^{1/2}_{M_2\times M_2} \otimes Hom(E_2, E_1))$ are called Fourier integral operators of order $m$ from sections of $F$ to sections of $E$, associated with the canonical relation $C$. Here the fiber of the vector bundle $Hom(E_2, E_1)$ at $(x, y) \in M_2 \times M_1$ consists of linear maps $E_{2,x} \to E_{1,y}$.

There exists a symbolic calculus that explains when two Fourier integral operators can be composed, and computes the principal symbol of the composition. See [Hör85] for more details.

# BIBLIOGRAPHY

# BIBLIOGRAPHY


[ABB+04] Emmanuel Audusse, François Bouchut, Marie-Odile Bristeau, Rupert Klein, and Benoît Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM J. Sci. Comput.*, 25(6):2050–2065 (electronic), 2004.

[AK01] Rémi Abgrall and Smadar Karni. Computations of compressible multifluids. *Journal of Computational Physics*, 169(2):594 – 623, 2001.

[AU85] José L. Antoniano and Gunther A. Uhlmann. A functional calculus for a class of pseudodifferential operators with singular symbols. In *Pseudodifferential operators and applications (Notre Dame, Ind., 1984)*, volume 43 of *Proc. Sympos. Pure Math.*, pages 5–16. Amer. Math. Soc., Providence, RI, 1985.

[AW04a] Nikolai Andrianov and Gerald Warnecke. On the solution to the Riemann problem for the compressible duct flow. *SIAM J. Appl. Math.*, 64(3):878–901 (electronic), 2004.

[AW04b] Nikolai Andrianov and Gerald Warnecke. The Riemann problem for the Baer-Nunziato two-phase flow model. *J. Comput. Phys.*, 195(2):434–464, 2004.

[BK09] Jorge Balbás and Smadar Karni. A central scheme for shallow water flows along channels with irregular geometry. *M2AN Math. Model. Numer. Anal.*, 43(2):333–351, 2009.

[BLMR02] Derek S. Bale, Randall J. Leveque, Sorin Mitran, and James A. Rossmanith. A wave propagation method for conservation laws and balance laws with spatially varying flux functions. *SIAM J. Sci. Comput.*, 24(3):955–978 (electronic), 2002.

[BN86] M.R. Baer and J.W. Nunziato. A two-phase mixture theory for the deflagration-to-detonation transition (ddt) in reactive granular materials. *International Journal of Multiphase Flow*, 12(6):861 – 889, 1986.

[Bou04] François Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2004.

[BS00] Martin Bojowald and Thomas Strobl. Group theoretical quantization and the example of a phase space $S^1 \times \mathbf{R}^+$. *J. Math. Phys.*, 41(5):2537–2567, 2000.

[BS03] Martin Bojowald and Thomas Strobl. Symplectic and projection quantization for non-holonomic constraints. *Internat. J. Modern Phys. D*, 12(4):713–725, 2003.

[CGRGV+04] Manuel J. Castro, José A. García-Rodríguez, José M. González-Vida, Jorge Macías, Carlos Parés, and M. Elena Vázquez-Cendón. Numerical simulation of two-layer shallow water flows through channels with irregular geometry. *J. Comput. Phys.*, 195(1):202–235, 2004.

[CMP01] Manuel Castro, Jorge Macías, and Carlos Parés. A $Q$-scheme for a class of systems of coupled conservation laws with source term. Application to a two-layer 1-D shallow water system. *M2AN Math. Model. Numer. Anal.*, 35(1):107–127, 2001.

[CPMP07] Manuel J. Castro, Alberto Pardo Milanés, and Carlos Parés. Well-balanced numerical schemes based on a generalized hydrostatic reconstruction technique. *Math. Models Methods Appl. Sci.*, 17(12):2055–2113, 2007.

[ČŽVS04] Nelida Črnjarić-Žic, Senka Vuković, and Luka Sopta. Balanced finite volume WENO and central WENO schemes for the shallow water and the open-channel flow equations. *J. Comput. Phys.*, 200(2):512–548, 2004.

[Doz97] Sandrine Dozias. Clustering for the spectrum of $h$-pseudodifferential operators with periodic flow on an energy surface. *J. Funct. Anal.*, 145(2):296–311, 1997.

[Dub99] Bruno Dubroca. Solveur de Roe positivement conservatif. *C. R. Acad. Sci. Paris Sér. I Math.*, 329(9):827–832, 1999.

[Dui96] J. J. Duistermaat. *Fourier integral operators*, volume 130 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1996.

[EB92] P. Embid and M. Baer. Mathematical analysis of a two-phase continuum mixture theory. *Continuum Mechanics and Thermodynamics*, 4:279–312, 1992. 10.1007/BF01129333.

[EZ03] L. C. Evans and M. Zworski. *Lectures on Semiclassical Analysis*. Unpublished Lecture Notes, Department of Mathematics, UC Berkeley, 2003.

[Geo08] David L. George. Augmented Riemann solvers for the shallow water equations over variable topography with steady states and inundation. *J. Comput. Phys.*, 227(6):3089–3113, 2008.

[Gio03] Dimitri Gioev. Lower order terms in Szegö type limit theorems on Zoll manifolds. *Comm. Partial Differential Equations*, 28(9-10):1739–1785, 2003.

[GL02] V. Guillemin and E. Lerman. Melrose-Uhlmann projectors, the metaplectic representation and symplectic cuts. *J. Differential Geom.*, 61(3):365–396, 2002.

[GNVC00] P. Garcia-Navarro and M. E. Vazquez-Cendon. On numerical treatment of the source terms in the shallow water equations. *Computers & Fluids*, 29(8):951 – 979, 2000.

[GO97a] V. Guillemin and K. Okikiolu. Spectral asymptotics of Toeplitz operators on Zoll manifolds. *J. Funct. Anal.*, 146(2):496–516, 1997.

[GO97b] V. Guillemin and K. Okikiolu. Subprincipal terms in Szegő estimates. *Math. Res. Lett.*, 4(1):173–179, 1997.

[God59] S. K. Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Mat. Sb. (N.S.)*, 47 (89):271–306, 1959.

[GS58] Ulf Grenander and Gabor Szegö. *Toeplitz forms and their applications*. California Monographs in Mathematical Sciences. University of California Press, Berkeley, 1958.

[GS77] Victor Guillemin and Shlomo Sternberg. *Geometric asymptotics*. American Mathematical Society, Providence, R.I., 1977. Mathematical Surveys, No. 14.

[GS79] Victor Guillemin and Shlomo Sternberg. Some problems in integral geometry and some related problems in microlocal analysis. *Amer. J. Math.*, 101(4):915–955, 1979.

[GS10] V. Guillemin and S. Sternberg. *Semi-classical analysis*. Preprint, MIT, 2010.

[GU81] V. Guillemin and G. Uhlmann. Oscillatory integrals with singular symbols. *Duke Math. J.*, 48(1):251–267, 1981.

[Gui94] Victor Guillemin. Gauged Lagrangian distributions II. *preprint (MIT)*, 1994.

[HDK10] Gerardo Hernández-Dueñas and Smadar Karni. Shallow water flows in channels. *Journal of Scientific Computing*, pages 1–19, 2010. 10.1007/s10915-010-9430-x.

[HH83] Ami Harten and James M. Hyman. Self-adjusting grid methods for one-dimensional hyperbolic conservation laws. *J. Comput. Phys.*, 50(2):235–269, 1983.

[Hör83] Lars Hörmander. *The analysis of linear partial differential operators. I*, volume 256 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1983. Distribution theory and Fourier analysis.

[Hör85] Lars Hörmander. *The analysis of linear partial differential operators. IV*, volume 275 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1985. Fourier integral operators.

[HR84] B. Helffer and D. Robert. Puits de potentiel généralisés et asymptotique semi-classique. *Ann. Inst. H. Poincaré Phys. Théor.*, 41(3):291–331, 1984.

[HV01] Andrew Hassell and András Vasy. Intersecting Legendrians and blow-ups. *Math. Res. Lett.*, 8(4):413–428, 2001.

[Jin01] Shi Jin. A steady-state capturing method for hyperbolic systems with geometrical source terms. *M2AN Math. Model. Numer. Anal.*, 35(4):631–645, 2001.

[Jos98] M. S. Joshi. A symbolic construction of the forward fundamental solution of the wave operator. *Comm. Partial Differential Equations*, 23(7-8):1349–1417, 1998.

[Kar94] Smadar Karni. Multicomponent flow calculations by a consistent primitive algorithm. *J. Comput. Phys.*, 112(1):31–43, 1994.

[Kar96] Smadar Karni. Hybrid multifluid algorithms. *SIAM J. Sci. Comput.*, 17(5):1019–1039, 1996.

[KHD09a] Smadar Karni and Gerardo Hernández-Dueñas. A hybrid scheme for flows in porous media. In *Hyperbolic problems: theory, numerics and applications*, volume 67 of *Proc. Sympos. Appl. Math.*, pages 715–724. Amer. Math. Soc., Providence, RI, 2009.

[KHD09b] Smadar Karni and Gerardo Hernández-Dueñas. A scheme for shallow water flow with area variation. *AIP Conference Proceedings*, 1168(1):1433–1436, 2009.

[KHD10] Smadar Karni and Gerardo Hernández-Dueñas. A hybrid algorithm for the Baer-Nunziato model using the Riemann invariants. *J. Sci. Comput.*, 45(1-3):382–403, 2010.

[KL02] Alexander Kurganov and Doron Levy. Central-upwind schemes for the Saint-Venant system. *M2AN Math. Model. Numer. Anal.*, 36(3):397–425, 2002.

[KP07] Alexander Kurganov and Guergana Petrova. A second-order well-balanced positivity preserving central-upwind scheme for the Saint-Venant system. *Commun. Math. Sci.*, 5(1):133–160, 2007.

[LeV92] Randall J. LeVeque. *Numerical methods for conservation laws*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 1992.

[LeV98] Randall J. LeVeque. Balancing source terms and flux gradients in high-resolution godunov methods: The quasi-steady wave-propagation algorithm. *Journal of Computational Physics*, 146(1):346 – 365, 1998.

[LeV02] Randall J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002.

[Low05] C. A. Lowe. Two-phase shock-tube problems and numerical methods of solution. *J. Comput. Phys.*, 204(2):598–632, 2005.

[LRS98] A. Laptev, D. Robert, and Yu. Safarov. Remarks on the paper of V. Guillemin and K. Okikiolu: "Subprincipal terms in Szegő estimates" [Math. Res. Lett. **4** (1997), no. 1, 173–179; MR1432819 (97k:58169)]. *Math. Res. Lett.*, 5(1-2):57–61, 1998.

[LS96] A. Laptev and Yu. Safarov. Szegő type limit theorems. *J. Funct. Anal.*, 138(2):544–559, 1996.

[LW60] Peter Lax and Burton Wendroff. Systems of conservation laws. *Comm. Pure Appl. Math.*, 13:217–237, 1960.

[Mar02] André Martinez. *An introduction to semiclassical and microlocal analysis*. Universitext. Springer-Verlag, New York, 2002.

[MU79] R. B. Melrose and G. A. Uhlmann. Lagrangian intersection and the Cauchy problem. *Comm. Pure Appl. Math.*, 32(4):483–519, 1979.

[NPPN06] Sebastian Noelle, Normann Pankratz, Gabriella Puppo, and Jostein R. Natvig. Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows. *J. Comput. Phys.*, 213(2):474–499, 2006.

[NXS07a] Sebastian Noelle, Yulong Xing, and Chi-Wang Shu. High-order well-balanced finite volume WENO schemes for shallow water equation with moving water. *J. Comput. Phys.*, 226(1):29–58, 2007.

[NXS07b] Sebastian Noelle, Yulong Xing, and Chi-Wang Shu. High-order well-balanced finite volume WENO schemes for shallow water equation with moving water. *J. Comput. Phys.*, 226(1):29–58, 2007.

[Pic09] Benedetto Piccoli. Flows on networks and complicated domains. In *Hyperbolic problems: theory, numerics and applications*, volume 67 of *Proc. Sympos. Appl. Math.*, pages 135–157. Amer. Math. Soc., Providence, RI, 2009.

[PS01] B. Perthame and C. Simeoni. A kinetic scheme for the Saint-Venant system with a source term. *Calcolo*, 38(4):201–231, 2001.

[PU95] T. Paul and A. Uribe. The semi-classical trace formula and propagation of wave packets. *J. Funct. Anal.*, 132(1):192–249, 1995.

[Roe81] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.*, 43(2):357–372, 1981.

[Roe87] P. L. Roe. Upwind differencing schemes for hyperbolic conservation laws with source terms. In *Nonlinear hyperbolic problems (St. Etienne, 1986)*, volume 1270 of *Lecture Notes in Math.*, pages 41–51. Springer, Berlin, 1987.

[Rus01] Giovanni Russo. Central schemes for balance laws. In *Hyperbolic problems: theory, numerics, applications, Vol. I, II (Magdeburg, 2000)*, volume 141 of *Internat. Ser. Numer. Math., 140*, pages 821–829. Birkhäuser, Basel, 2001.

[Rus05] Giovanni Russo. Central schemes for conservation laws with application to shallow water equations. In Salvatore Rionero and Giovanni Romano, editors, *Trends and Applications of Mathematics to Mechanics*, pages 225–246. Springer Milan, 2005.

[SWK06] D.W. Schwendeman, C.W. Wahle, and A.K. Kapila. The riemann problem and a high-resolution godunov method for a model of compressible two-phase flow. *Journal of Computational Physics*, 212(2):490 – 526, 2006.

[VC99] María Elena Vázquez-Cendón. Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry. *J. Comput. Phys.*, 148(2):497–526, 1999.

[VS03] Senka Vuković and Luka Sopta. High-order ENO and WENO schemes with flux gradient and source term balancing. In *Applied mathematics and scientific computing (Dubrovnik, 2001)*, pages 333–346. Kluwer/Plenum, New York, 2003.

[Wid73] Harold Widom. Toeplitz determinants with singular generating functions. *Amer. J. Math.*, 95:333–383, 1973.