



A University of Sussex PhD thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details



**Inverse Problems for Hyperbolic
Conservation Laws**

A Bayesian approach

Duc Lam Duong

Supervisor

Masoumeh Dashti

Submitted for the degree of Doctor of Philosophy

University of Sussex

May 2020

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree. This thesis is comprised entirely of my own research, conducted under the supervision of Masoumeh Dashti. Work due to other authors will be made clear and cited where appropriate.

Signature:

Duc Lam DUONG

UNIVERSITY OF SUSSEX

DUC LAM DUONG, DOCTOR OF PHILOSOPHY

INVERSE PROBLEMS FOR HYPERBOLIC CONSERVATION LAWS

A BAYESIAN APPROACH

SUMMARY

This thesis contributes to the development of the Bayesian approach to inverse problems for hyperbolic conservation laws. Inverse problems in the context of hyperbolic conservation laws are challenging due to the presence of shock waves and their irreversibility. The first main contribution of this work is the study of the forward problems where we develop a theory of particle trajectories for scalar conservation laws. Motivated from a model of traffic flow, we consider some ordinary differential equations with their right-hand side depending on solutions of scalar conservation laws. Despite the presence of discontinuities in the entropy solutions, it is showed that the trajectories are well-posed by using Filippov theory. Moreover, we prove that the approximate trajectory generated by either the front tracking approximation method or the vanishing viscosity method converges uniformly to the trajectory corresponding to the entropy solution of the scalar conservation law. For certain flux functions, illustrated by traffic flow, we are able to obtain the convergence rate for the approximate trajectory with respect to changes in the initial field or the flux function by combining the front tracking method with Filippov theory.

As the second main contribution of the thesis, we study some Bayesian inverse problems for scalar conservation laws and establish several well-posedness and approximation results. Specifically, we consider two types of inverse problems: the inverse problem of recovering the upstream field and the inverse problem of finding the flux function, both from observations of appropriate functions of the entropy solutions of scalar conservation laws. Based on the theory of trajectories developed in the first part of the thesis and the Bayesian inversion theory developed by Stuart *et.al.*, we prove that the statistical solutions to these inverse problems are well-posed and stable with respect to changes in the forward model. Rates of convergence of the approximate posteriors are also given for certain inverse problems.

“If I have been able to see further, it was only because I stood on the shoulders of giants.”

Sir Isaac Newton (1643 - 1727)

“ $e^{i\pi} + 1 = 0.$ ”

Leonhard Euler (1707 - 1783)

“It is impossible to be a mathematician without being a poet in soul.”

Sofia Kovalevskaya (1850 - 1891)

Acknowledgements

First and foremost, I am greatly grateful to my supervisor, Dr Masoumeh Dashti, for her constant support and encouragement during the last four years. Thank you for having the immense patience and the willingness to answer any of my naive questions, for setting me back on the road when I got lost sometimes, and also for giving me enough freedom to work on my own. I really appreciate not only the time you spent with me for the mathematical discussions but also the advice and sharing about many other things in life. Through your guidance, I have learnt and grown a lot, both professionally and personally. I would like to say that I am proud to be your (first) PhD student.

I am deeply indebted to Prof. Babis Makridakis for partly funding me so that I had the chance to come to Sussex, and for accepting to be my second supervisor. During my PhD, I was able to travel to many conferences due to the generous support of ModCompShock Network – of which Prof. Makridakis is the coordinator, the Department of Mathematics, the London Mathematical Society and the Society for Industrial and Applied Mathematics (SIAM).

I would like to thank Prof. Constantine Dafermos for a few short-but-precious discussions in two conferences in Oxford, for the encouragement and for showing me some results regarding the fine properties of shocks. In particular, I thank Prof. Gui-Qiang Chen for organising these conferences and making the opportunities possible.

I wish to thank Dr Kostas Koumatos and Dr Aretha Teckentrup for many useful comments, discussions and feedback that have helped to improve greatly the presentation of this thesis.

I am grateful to Dr Andrew Duncan for the discussions we had during the first two years of my PhD and for continuing to support me in one way or another. I also thank Dr Miroslav Chlebik for the advice and for taking the time to do the annual review with me.

I thank other staff members of the MPS School and the Department of Mathematics, some of which I have had the opportunity to chat with, either about mathematics or life, or share the lunch table with. In particular, I would like to thank Prof. Enrico Scalas and Prof. Michael Melgaard for their support over the years. I would also like to thank the

MPS School Office team, especially Mr Richard Chambers, Ms Rebecca Gibson-Foster, Ms Lily King, Ms Sinead Rance and Ms Amelia Redman for their support and for helping me with various paperwork-related issues.

Reaching this point in my life, I am grateful to many of my former mentors and teachers from the Hanoi National University of Education, the Institute of Mathematics-Hanoi and the Laboratory J.A. Dieudonne in Nice who have influenced me in my career development. In particular, I would like to thank Prof. Cung The Anh, Prof. Tran Dinh Ke, Prof. Le Van Hien, Prof. Nguyen Quang Dieu, Prof. Pham Hoang Hiep, Prof. Nam Q. Le, Prof. Phung Ho Hai, Prof. Nguyen Viet Dung, Prof. Dinh Nho Hao, Prof. Le Van Thanh, Prof. Vu Hoang Linh, Prof. David Chiron, Prof. Cedric Bernardin and Prof. Thierry Goudon. I would also like to send my thanks to Quoc Hung Nguyen and Nguyen Thac Dung for the discussions and encouragement, to Quốc Anh Ngô, Phan Thành Nam, Minh-Binh Tran and Bao Quoc Tang for their sincere career advice.

I am very much thankful to my colleagues and friends for sharing such a colourful journey with me during my time in the UK, especially the last two years with many unforgettable moments. In particular, I want to mention James, Stelios, Aleksandra, Rosie, Francesco, Andreas, Sandra, Rabab, Blake, Mariana, Alex, Davide, Farzad, Salam, Awss, Nóra and Omar. My special thanks go to James for all his support, sharing and caring, like a big brother, and to Sandra for her kindness and encouragement (I owe you both, for organising a wonderful surprise birthday party for me that I feel I never deserved!). Thanks to Tiến and Hương for their great company, to Sam (from the table tennis team), Thảo and cháu Dịu for the sharing, to anh Chí Thành and Tú Anh for kindly treating me as an unofficial member in their beautiful little family. Thanks to you all, the last few years of my PhD has been really memorable¹! Part of this thesis is written in a strange time when the world is under lockdown due to coronavirus. I thank my landlord David and housemate Mohamed for their sharing and understanding during this difficult time.

Finally, I am greatly thankful to my beloved family. Con cảm ơn bố mẹ vì những hi sinh, vất vả để con có ngày hôm nay. Cảm ơn anh Nga, em Trúc và út Hiền. Cảm ơn gia đình ta đã luôn ủng hộ và đồng hành cùng con trong suốt chặng đường đã qua.

I dedicate this thesis to my late grandmother, who passed away while I was away doing my PhD, with whom I spent most of my childhood. I am so thankful for each and every precious moment I had with you. You will be greatly missed.

¹It took a pandemic (COVID-19) to stop me from enjoying your company!

Contents

List of Tables	ix
List of Figures	x
1 Introduction	1
1.1 Thesis outline	6
1.2 Notations and symbols	7
2 Scalar conservation laws	8
2.1 Derivation and examples	9
2.1.1 Burgers equation	10
2.1.2 Traffic flow	10
2.2 Shock wave phenomena and weak solutions	11
2.2.1 The formation of shocks	11
2.2.2 Weak solutions	12
2.3 Entropy conditions	14
2.4 Existence and uniqueness of entropy solutions	17
2.4.1 Front tracking method	19
2.4.2 Vanishing viscosity method	21
2.5 Some regularity results	23
2.6 Summary and outlook	24
3 Inverse problems and Bayesian approach	25
3.1 Inverse problems	25
3.1.1 Mathematical formulation	26
3.1.2 Ill-posedness	27
3.1.3 Classical regularization approach	29
3.2 Bayesian approach to inverse problems	31

3.2.1	Well-posedness	32
3.2.2	Approximations	35
3.3	Summary and notes	37
4	Filippov theory for discontinuous differential equations	38
4.1	Equations with discontinuous right-hand sides	38
4.2	Carathéodory differential equations	40
4.3	Filippov solution: Existence and uniqueness	41
4.3.1	Existence	42
4.3.2	Uniqueness	44
4.4	Summary and notes	46
5	Particle trajectories for scalar conservation laws	47
5.1	Definition and Well-posedness	48
5.2	Stability properties of the particle trajectories	56
5.2.1	The case of traffic flow	56
5.2.2	Stability with respect to changes in flux functions	64
5.3	Summary and outlook	66
6	Bayesian inverse problems in scalar conservation laws	67
6.1	Recovering the upstream field with trajectory data	68
6.1.1	Well-posedness	68
6.1.2	Approximations	69
6.2	The case of traffic flow: rate of convergence	72
6.2.1	Recovering upstream density	72
6.2.2	Recovering the flux function	74
6.3	Inverse problems with direct measurement data	75
6.4	Further discussions	79
6.4.1	A modified framework	79
6.4.2	Applications	82
6.5	Summary and notes	84
7	Conclusion and Outlook	85
7.1	Conclusion	85
7.2	Outlook and future work	86
7.2.1	The consistency problem	86

7.2.2	The non-convex flux case	86
7.2.3	Rate of convergence for vanishing viscosity approximation	87
7.2.4	Numerical experiments	87
7.2.5	Other related questions	87
	Bibliography	88
	Index	99

List of Figures

- 5.1 The Riemann problem. The particles starts at x_0 , after a while they catch up with the shocks in front of them, then right after that, they continue to travel with lower speeds. 63

Chapter 1

Introduction

Many physical phenomena such as traffic flow, wave propagation, gas dynamics and elastodynamics can be modelled via conservation laws. The mathematical description of scalar conservation laws in one space dimension is given by the first order partial differential equation (PDE)

$$\partial_t v(x, t) + \partial_x f(v(x, t)) = 0, \quad x \in \mathbb{R}, t > 0, \quad (1.1)$$

where $v : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$ is the *conserved quantity* (which we often refer to as the *flow field* in this thesis), while $f : \mathbb{R} \rightarrow \mathbb{R}$ is called the *flux function* (or *flux*). The equation (1.1) is often coupled with initial data (often termed as *upstream field*)

$$v(x, 0) = v_0(x), \quad x \in \mathbb{R}, \quad (1.2)$$

where $v_0 : \mathbb{R} \rightarrow \mathbb{R}$ is some bounded function. It is well known (see, for instance, [Smo83]) that a solution v of the conservation law (1.1)-(1.2) can develop discontinuities, called *shock waves* (or *shocks*), no matter how smooth the flux f and the initial data v_0 are. Global solutions therefore can only be sought in the space of discontinuous functions and one needs to work with weak solutions. This difficulty is further exacerbated by the fact that the weak solutions of conservation laws are not unique unless some additional restrictions, called *entropy conditions*, are imposed. The resulting solution is then called the *entropy solution*. For scalar conservation laws, v is an entropy solution to (1.1) if it satisfies

$$\partial_t \eta(v(x, t)) + \partial_x q(v(x, t)) \leq 0,$$

in the sense of distribution, for all entropy-entropy flux pairs (η, q) with

$$\eta \text{ convex, and } q'(v) = \eta'(v)f'(v).$$

A large body of works have been carried out regarding the existence, uniqueness and regularity of entropy solutions (given appropriate flux and initial data) for (1.1)-(1.2).

We refer to the book by Dafermos [Daf16] for a comprehensive treatise on this profound subject (see also the monographs by Bressan [Bre00] and Holden-Risebro [HR15]).

In this work, we are interested in the inverse problems of recovering the upstream field v_0 (with f given) or the flux function f (with v_0 given) from observations of the solution v (or a function of the solution v). These kinds of inverse problems have many applications, depending on how we interpret v in the model (1.1). An outstanding example is the traffic flow on a highway (the Lighthill-Whitham-Richards (LWR) model, see [LW55]-[Ric56]). In that model, $v(x, t)$ represents the density of cars, that is, the number of cars per unit length at some location x and at time t . The quantity $f(v(x, t))$ is the flux of cars across the point x on the road at time t . The inverse problems we are considering in this case correspond to the problems of determining the upstream car density or the car flux given finite observations of (a function of) car density at later times.

In the language of mathematics, these inverse problems can be written as

$$y = G(u),$$

where u denotes the *unknown* (that is, either $v(\cdot, t_0)$ or f), y is the *observed data* and G is the *parameter-to-observation map* (or *observation map*) which is defined via the solution of the forward problem. One typical example of the observation map is given by the pointwise measurements of the solution,

$$G(u) = \{v(x_i, t_j)\}_{i \in I, j \in J},$$

for some (mostly) finite index sets I, J . In the context of conservation laws, due to the discontinuities in v , this way of making measurements may lead to great fluctuations in observed data, and hence causing large errors in recovering the unknown of the inverse problems.

To get round this, we use an alternative way of making observations by tracking the trajectories of particles moving along the flow. To this end, we need to study a stability theory of particle trajectories for (1.1)-(1.2). Denote by w the velocity of the flow. In most cases, w is some smooth function of the conserved quantity, that is, $w = w(v)$ (if v represents the velocity field in (1.1) then we take $w = v$). The trajectory $z(t)$ of a particle starting from an initial position x_0 at t_0 is then defined by the ordinary differential equation

$$\frac{dz(t)}{dt} = w(v(z(t), t)), \quad (1.3)$$

subject to the initial condition $z(t_0) = x_0$. However, since v is typically a discontinuous function of both z and t , the standard Cauchy-Lipschitz theory is not applicable, and a

solution z to (1.3) has to be understood in an appropriate way. In this thesis, we employ the Filippov theory of differential equations [Fil88] and define an absolutely continuous function z to be a solution to (1.3) if z solves (1.3), viewed as a differential inclusion, almost everywhere. We need to establish appropriate stability properties of this particle trajectory z . The existing theory (using Oleinik’s decay estimate, see, for instance, [Leg11]) proves the existence and uniqueness of z , but the method cannot be used to prove the stability of z with respect to changes in the velocity field. In this work, we are able to prove the stability and provide suitable rates in particular situations by combining the Filippov theory and the front tracking method. The idea of front tracking is to approximate the initial function v_0 by a step function v_0^N and the flux f by its piecewise linear interpolation f^N . Then the approximate solution v^N will be given by solving a (finite) set of so-called *Riemann problems* ([HR15]). By using Filippov theory and some structure of solutions of conservation laws (developed in [Daf77], [BL99]), we prove that the approximate trajectory z^N corresponding to the approximate solution v^N given by the front tracking method converges to z uniformly in time. Given that solutions of scalar conservation laws generally develop shocks and can only be stable in L^1 , this may serve as a different perspective to the well-posedness theory of hyperbolic conservation laws.

We shall see now how this continuity property of z may help to recover the initial field or flux functions in our inverse problems. We note that, in general, inverse problems are ill-posed, meaning solutions may not exist, may not be unique or may depend sensitively on data. In hyperbolic conservation laws, the situation is even more complicated by the fact that the physically relevant solutions are often irreversible. This irreversibility property, induced by the entropy condition (see [Daf16]), renders severe difficulties in inversion: if $v(x, t)$ is an entropy solution, $v(-x, -t)$ is no longer an entropy solution unless $v(x, t)$ is a classical one (and has in particular no shocks). Finding the right techniques to tackle inverse problems in hyperbolic conservation laws are therefore challenging.

Nevertheless, by appropriate regularisation one is able to find some estimation of the missing information. We employ the Bayesian approach as one way of regularising and provide some sort of “solution” to our inverse problems, given in the form of a probability distribution, called the posterior. The Bayesian inverse problems for unknown functions have been studied extensively in the last decade, in particular, for nonlinear models involving PDEs; see [Stu10, DS16], the early paper [Fra70], and for a more applied and computational overview, [KS05]. In the Bayesian approach, the data and the unknown are treated as random variables, and the regularisation enters the framework in the form

of a given prior probability distribution on the unknown. The posterior may then be derived through the Bayes' theorem, as the updated prior using the data. Interestingly, the well-posedness of the posterior can be monitored through the properties of the forward model [Stu10]. Therefore, exploring the regularity (such as continuity and stability) of solutions of the forward problem is key in this approach.

The convergence results of the particle trajectories mentioned above suggest that if we consider the observation map as¹

$$G(u) = \{z(t_j)\}_{j \in J},$$

then the collected data are stable in some sense to be made precise later on, giving some regularity property for the observation map. Thanks to the approximation theory of Bayesian inverse problems, developed by Stuart and others ([Stu10]), the approximate posterior is continuous in appropriate metrics, giving the well-posedness for the solutions of our inverse problems.

We also note that, in many situations, making observations by tracking particle trajectories is probably more practical and more economical than measuring the flow field itself. Consider for example the traffic flow passing through a tunnel, where measuring the density of the cars inside the tunnel might not be easy, one may instead track the position of a marked car over time. In practice, this can be done easily via a GPS device mounted on the car. If necessary, at the same time one can track more cars to have a more accurate picture.

We will treat the traffic flow model, an important example that motivates our work, in great detail. In that model, ρ denotes the car density and w denotes the car speed. The equation reads

$$\partial_t \rho + \partial_x(\rho w(\rho)) = 0. \tag{1.4}$$

Some observations regarding shock speed are made (Lemma 5.2.2). We then prove that the convergence of the particle trajectory given by the front tracking method leads us to an even stronger result. We provide (in Theorem 5.2.3) a rate for the convergence of the approximate trajectory in terms of the rate of changes in the initial density for any

¹We note that, in a similar context in fluid mechanics, the particle trajectory z is normally termed the *Lagrangian representation* – while its counterpart v is called the *Eulerian representation* – of the solution (or *Lagrangian solution* and *Eulerian solution* for short). For this reason, the terms *Lagrangian data* and *Eulerian data* are occasionally used to denote the different types of data collected. Note also that the term *Lagrangian representation* for scalar conservation laws is used in quite a different situation by Bianchini and others (see, for example, [BM17]).

decreasing function w . More precisely, we show that if two positive upstream densities (at some time $t_0 > 0$) are close in L^1 ,

$$\|\rho(\cdot, t_0) - \bar{\rho}(\cdot, t_0)\|_{L^1(\mathbb{R})} \leq \varepsilon,$$

then, for every $T > 0$, the corresponding particle trajectories are close in L^∞ ,

$$\|z(\cdot) - \bar{z}(\cdot)\|_{L^\infty([t_0, T])} \leq C\sqrt{\varepsilon},$$

for some constant C depending on t_0 and T . The main idea of the proof is to first consider the piecewise constant initial conditions $\rho^N(\cdot, t_0)$ and $\bar{\rho}^N(\cdot, t_0)$ which are approximating $\rho(\cdot, t_0)$ and $\bar{\rho}(\cdot, t_0)$ and construct the corresponding approximate trajectories z^N and \bar{z}^N , defined for some small time, by solving a series of Riemann problems. The nature of the front tracking method allows this procedure to extend further in time. We prove that if $\rho^N(\cdot, t_0)$ and $\bar{\rho}^N(\cdot, t_0)$ are close then the resulting trajectories from the corresponding Riemann problems remain close. This result together with the uniform convergence result for general scalar conservation laws (established in Theorem 5.1.6) of each z^N and \bar{z}^N give the stability estimate as stated.

These strong stability properties on the trajectory z will be of great value for our inverse problems since they translate to the rate of convergence for the approximations of the corresponding posteriors. We note that obtaining such stability estimates for the forward map in the situation where pointwise measurements are made directly on the entropy solution itself is not possible, due to shock wave phenomena. See Section 6.3 for a more detailed discussion.

We will also consider the approximate trajectories arising from the method of vanishing viscosity. Let v^ϵ be the solution of the viscous version of (1.1), that is

$$\partial_t v^\epsilon(x, t) + \partial_x f(v^\epsilon(x, t)) = \epsilon \partial_{xx} v^\epsilon(x, t).$$

Consider the particle trajectory z^ϵ starting at $z^\epsilon(0) = x_0$ that solves

$$\dot{z}^\epsilon(t) = w(v^\epsilon(z^\epsilon(t), t)).$$

Then, at the limit when ϵ goes to 0, the trajectory z^ϵ converges in L^∞ to the Filippov solution z of (1.3). To prove this result, we make the assumption that the trajectory $z(t)$, even though it may cross the shock curves of v , never lies on any shock curve for a positive period of time (Assumption 5.1.8). We then verify that this assumption is at least satisfied for the case of traffic flow when there is significant number of cars on the road.

The notion of Filippov solutions to discontinuous right-hand side differential equations was employed by Dafermos in [Daf77] to build the theory of generalized characteristics for hyperbolic conservation laws and has been an efficient method for studying the regularity of solutions. It is worth noting that the theory of particle trajectories that we consider here differs from the one of generalized characteristics considered by Dafermos. The speed of the generalized characteristics is either classical characteristic speed or shock speed, while the particle speed considered here, given on the right-hand side of (1.3), is the speed of the flow itself. The approach that we follow here is motivated by the traffic flow model and addressed, for instance, in [CM03]. However, the framework in this thesis is more general than the one in [CM03], and we are able to provide strong stability results that have not been obtained in the literature.

Most of results in the thesis are established for genuinely nonlinear flux functions. The approximation theory for particle trajectories when the flux function has inflection points will be considered in our future work.

1.1 Thesis outline

The thesis is organised as follows. In Chapter 2, we give a brief overview of the mathematical theory of scalar conservation laws. After some derivation examples, we introduce some of the most important concepts of conservation laws, including shock wave phenomena and entropy solutions. Two of the most popular methods in studying conservation laws, front tracking and vanishing viscosity, are discussed. Chapter 3 presents inverse problems and their ill-posedness, together with some regularization methods. We also give a short introduction to the Bayesian approach to inversion. The Filippov theory for ordinary differential equations with discontinuous right-hand side is introduced in Chapter 4. The most important contributions of the thesis lie in Chapter 5 and Chapter 6. Chapter 5 presents a theory of particle trajectories for scalar conservation laws. The stability properties via the convergence of various approximations of the particle trajectories are studied. The Bayesian approach to the inverse problems of recovering the initial field or flux function is considered in Chapter 6. In the last two sections of Chapter 6, we will discuss a number of different ways of collecting data, including direct measurements of the entropy solution itself and some situations where these measurements work. We make some concluding remarks and present the thesis outlook, including future work, in Chapter 7.

1.2 Notations and symbols

Throughout the thesis, the letter v will be used to denote the unknown of the forward problem, which is a scalar conservation law for the majority of the time. The unknown of the inverse problem will be denoted by u . The letter z is saved to denote the particle trajectory. The parameter-to-observation map for a general inverse problem is denoted by \mathcal{G} , and by G for the case of scalar conservation laws. Normally $\mathcal{G} = \mathcal{O} \circ S$ where \mathcal{O} denotes the evaluation operator while S is the forward map. Other standard notations are listed below.

- X (Banach) space of parameters (where the forward map is mapping from)
- Y (Banach) space of data (finite or infinite)
- $\|\cdot\|_X$ norm in X
- $|\cdot|$ Euclidean norm in \mathbb{R}^n
- ξ observable noise
- $\mathcal{B}(X)$ space of Borel measures on X
- d_{Hell} Hellinger distance
- d_{Wass} Wasserstein distance
- $L^1_\mu(X; \mathbb{R})$ Lebesgue space on $(X, \mathcal{B}(X), \mu)$
- TV total variation
- BV space of functions of bounded variation
- BV_{loc} space of functions of locally bounded variation
- ∇ differential (del) operator
- Δ Laplacian operator
- div divergence operator
- grad gradient vector

Chapter 2

Scalar conservation laws

In this chapter, we give a brief overview of the theory of scalar conservation laws. One of the fundamental features of scalar conservation laws (and hyperbolic conservation laws in general) that distinguish them from other types of partial differential equations is the formation of shocks. Starting with even smooth initial data, the wave-like solutions of conservation laws will generally get steeper and steeper, and eventually break, leading to the development of shock waves. These waves then continue to travel at their own speeds. This striking phenomenon makes the study of conservation laws challenging. Many modern techniques of functional analysis for partial differential equations are not applicable. The solutions are only sought in the space of discontinuous functions and therefore, working with weak solutions is inevitable. The difficulty is exacerbated by the fact that weak solutions of conservation laws are not unique. To single out the right solution, some additional restrictions, called entropy conditions, are imposed. These are essentials in the mathematical theory of hyperbolic conservation laws.

In the first section, we give a derivation for scalar conservation laws and highlight a few examples of equations of this type. These examples will be considered in later chapters. Section 2.2 reflects the most significant property of scalar conservation laws, the formation of shocks, as well as defines the weak formulation of solutions. The Rankine-Hugoniot condition on jumps is discussed here. Entropy conditions and entropy solutions, with motivations and related discussions are contents of Section 2.3. In the next section, Section 2.4, we present the existence and uniqueness of entropy solutions. Two of the most popular and fundamental methods for constructing solutions of conservation laws, the front tracking method (Section 2.4.1) and the vanishing viscosity method (Section 2.4.2), are also discussed here. In the last section, Section 2.5, we highlight a few results regarding fine structures of the entropy solution, which will be useful later on.

Many results in this chapter do not include proofs. However, we will refer to appropriate references where relevant. Throughout this chapter, we assume that $v_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ is a measurable bounded function. That is, there exists some constant M such that

$$\|v_0\|_{L^\infty(\mathbb{R}^n)} \leq M.$$

The function f is assumed to be smooth enough so that all the calculations make sense (for example, f is differentiable almost everywhere). Some additional conditions on f might be imposed in specific situations.

For simplicity, some results are stated in one (space) dimension, even though most of them can be extended to higher dimensions.

2.1 Derivation and examples

Scalar conservation laws take the form as follows

$$\frac{\partial v}{\partial t}(x, t) + \nabla \cdot f(v(x, t)) = 0, \quad (2.1)$$

where $v : \mathbb{R}^n \times [0, \infty) \rightarrow \mathbb{R}$ is the *conserved quantity*. The function $f = (f_1, \dots, f_n) : \mathbb{R} \rightarrow \mathbb{R}^n$ is called *flux*. The notation $\nabla \cdot f$ stands for the divergence of f with respect to the spacial variable x , that is $\nabla \cdot f = \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}$. If we assign an initial condition for v at $t = 0$,

$$v(x, 0) = v_0(x), \quad (2.2)$$

then we obtain an *initial value problem* (or *Cauchy problem*).

To have an idea where does the name “conservation law” come from, we consider in one dimension a (sufficiently smooth) quantity v that satisfies (2.1). By integrating (2.1) in x over $[a, b]$ ones get

$$\begin{aligned} \frac{d}{dt} \int_a^b v(x, t) dx &= \int_a^b f'(v(x, t)) dx \\ &= f(v(a, t)) - f(v(b, t)). \end{aligned} \quad (2.3)$$

The term $\int_a^b v(x, t) dx$ represents the total amount of conserved quantity v inside the interval $[a, b]$. The right-hand side of (2.3) describes the quantity flow going in and going out of the interval $[a, b]$ (which explains the name “flux” for f). The equation (2.3) therefore presents a law of conservation: the rate of change in the total amount of conserved quantity over time in a domain equals the amount of quantity flowing in and out through the boundary of that domain.

2.1.1 Burgers equation

Consider the Burgers equation

$$\partial_t u(x, t) + \partial_x \left(\frac{1}{2} u^2(x, t) \right) = 0. \quad (2.4)$$

This equation was introduced by Bateman [Bat15] and Burgers [Bur48] (in the viscous form) and studied thoroughly by Hopf [Hop50]. It is normally known as a simplification of the Navier-Stokes equations and is fundamental in the theory of fluid mechanics. Moreover, it is also a prototype example of scalar conservation laws due to the fact that it reserves every typical feature of scalar conservation laws. One way to see the formation of Burgers' equation is to look at some particle $x(t)$ moving with velocity $u(x, t)$ for which we know that the following relationship holds

$$\dot{x}(t) = u(x(t), t).$$

Assume that u is smooth enough, we have

$$\ddot{x}(t) = \frac{d}{dt} u(x(t), t) = u_x \dot{x}(t) + u_t = uu_x + u_t.$$

If there are no external forces then the Newton law of motion ensures that the acceleration $a = \ddot{x}(t)$ equals zero, leading to

$$u_t + uu_x = 0, \quad (2.5)$$

which is equivalent to (2.4) since u is smooth. The equation (2.5) is normally termed the *quasilinear Burgers equation*.

2.1.2 Traffic flow

Another typical example of scalar conservation laws is the so-called LWR model of traffic flow on the highway. This model is first described in the seminal paper [LW55] by Lighthill and Whitham where the authors develop it via the method of kinetic waves. One year later, Richards [Ric56] independently comes up with a similar model.

The key idea of the LWR model initiates from an observation that the traffic streams at a macroscopic level behave like fluid streams, where the distribution of cars can be replaced by a “continuous” function. Let ρ be the density and v be the average velocity of cars. The LWR assumes that the traffic velocity v depends only on its density, $v = v(\rho)$. Then $\rho v(\rho)$ represents the car flux. Since the change in the total amount of cars in a road

segment, say $[a, b]$, is given by the flux of cars flowing in at a subtracting the flux of cars flowing out at b (assuming that there are no entries and exits on this road segment), then

$$\begin{aligned} \int_a^b \partial_t \rho(x, t) dx &= \frac{d}{dt} \int_a^b \rho(x, t) dx = \rho(a, t)v(\rho(a, t)) - \rho(b, t)v(\rho(b, t)) \\ &= \int_a^b \partial_x(\rho(x, t)v(\rho(x, t))) dx. \end{aligned}$$

Therefore, since a and b are arbitrary,

$$\rho_t + (\rho v(\rho))_x = 0. \quad (2.6)$$

The flux function in this case is $f(\rho) = \rho v(\rho)$. In the simplest case, the velocity may be assumed to depend linearly on the density, that is

$$v(\rho) = v_{max} \left(1 - \frac{\rho}{\rho_{max}} \right),$$

where v_{max} and ρ_{max} denote the maximum velocity and density of cars, respectively. The flux is then given by

$$f(\rho) = v_{max} \rho \left(1 - \frac{\rho}{\rho_{max}} \right),$$

which is a strictly concave and decreasing function. Moreover, if we make a scaling $\rho_{max} = 1, v_{max} = 1$ then $f(\rho) = \rho(1 - \rho)$ and we obtain the normalised traffic flow equation

$$\rho_t + (\rho(1 - \rho))_x = 0. \quad (2.7)$$

This normalised model is often used as a toy model in studying scalar conservation laws.

We note that, in this case, $v = 1 - \rho$ satisfies the following extra conservation law

$$v_t + (v(v - 1))_x = 0, \quad (2.8)$$

with a strictly convex flux function $\tilde{f}(v) = v(v - 1)$. We note, moreover, that by setting $\rho = \frac{1-u}{2}$ we recover the Burgers equation (2.4).

2.2 Shock wave phenomena and weak solutions

2.2.1 The formation of shocks

To describe shock wave phenomena and how shocks are formed, we consider the quasilinear Burgers equation

$$u_t + uu_x = 0, \quad (2.9)$$

with some smooth initial data

$$u(x, 0) = u_0(x).$$

The left hand side of (2.9) can be interpreted as a directional derivative of u in the direction $\vec{v} = (1, u)$,

$$u_t + uu_x = 1 \cdot u_t + u \cdot u_x = \partial_{\vec{v}=(1,u)} u(x, t).$$

Therefore, (2.9) ensures that u is constant in the direction \vec{v} , or along the characteristic lines

$$t \mapsto (t, x + tu_0(x)).$$

Some straightforward calculations lead to

$$u(x, t) = u_0(x - ut).$$

Differentiation with respect to x and t gives

$$u_x = u'_0(1 - u_x t), \quad u_t = u'_0(-u - u_t t).$$

Thus,

$$u_x = \frac{u'_0}{1 + tu'_0}, \quad u_t = \frac{-u'_0 u}{1 + tu'_0}.$$

If $u'_0(x) < 0$ for some x then u_x and u_t will blow up when $t = t_0 = -\frac{1}{u'_0}$. Thus, u must break down in finite time and wave breaking will essentially happen, forming *shock waves*, regardless of the smoothness of u_0 . Therefore, global solutions for (2.9), and hence for scalar conservation laws, can only be found in the space of discontinuous functions.

2.2.2 Weak solutions

Definition 2.2.1. A function $v \in L^\infty(\mathbb{R}^n \times [0, \infty))$ is called a weak solution for (2.1)-(2.2) if the identity

$$\int_{\mathbb{R}^n} \int_0^\infty [v(x, t)\phi_t + f(v(x, t)) \cdot \nabla \phi] dt dx + \int_{\mathbb{R}^n} v_0(x)\phi(x, 0) dx = 0 \quad (2.10)$$

holds for any test function $\phi \in C_0^\infty(\mathbb{R}^n \times [0, \infty))$.

Note that v is only required to be in $L^\infty(\mathbb{R}^n \times [0, \infty))$, which includes discontinuous functions, for (2.10) to make sense. However, not every discontinuous function is permitted since (2.10) does impose severe restrictions on the curves of discontinuity as we shall see from the following result, known as the *Rankine–Hugoniot jump condition*.

Theorem 2.2.2. Let $\varphi = \varphi(t)$ be a discontinuity curve (shock curve) of the weak solution v . Then φ satisfies the following constraint

$$\varphi'(t) = \frac{f(v_+) - f(v_-)}{v_+ - v_-}, \quad (2.11)$$

where v_+ and v_- are the right and the left states of v on the shock curve φ .

We shall call $s := \varphi'(t)$ the *shock speed*. For a proof of this theorem, see [Eva10]. This result has a great impact on gas dynamics and the theory of hyperbolic conservation laws [Daf16]. The formula (2.11) can be used to find explicitly weak solutions of some initial value problems that are not classically solvable. The following example is taken from [Smo83].

Example 2.2.3. Consider the Burgers equation (2.4) with initial data

$$u_0(x) = \begin{cases} 1 & \text{if } x < 0 \\ 1 - x & \text{if } 0 \leq x < 1 \\ 0 & \text{if } x \geq 1. \end{cases} \quad (2.12)$$

A weak solution u for (2.4) can be calculated up to $t < 1$ by using characteristics (see Section 2.1.1),

$$u(x) = \begin{cases} 1 & \text{if } x < t \\ \frac{1-x}{1-t} & \text{if } t \leq x < 1 \\ 0 & \text{if } x \geq 1. \end{cases} \quad (2.13)$$

For $t \geq 1$, by applying the Rankine-Hugoniot jump condition for $u_+ = 1, u_- = 0$, we have that $s = \frac{1}{2}$. Thus we can define

$$u(x) = \begin{cases} 1 & \text{if } x < 1 + \frac{1}{2}(t - 1) \\ 0 & \text{if } x \geq 1 + \frac{1}{2}(t - 1), \end{cases} \quad (2.14)$$

to be a weak solution of (2.4) for $t \geq 1$.

A serious problem with weak solutions is that they are not unique as the following example shows.

Example 2.2.4 (Non-uniqueness of weak solutions, [Daf16]). We consider again the Burgers equation (2.4) with initial data

$$u_0(x) = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{if } x > 0. \end{cases} \quad (2.15)$$

Then for any $\alpha \in [0, 1]$, each u_α provides a solution for (2.4) where

$$u_\alpha(x) = \begin{cases} -1 & \text{if } x < -t \\ x/t & \text{if } -t < x \leq -\alpha t \\ -\alpha & \text{if } -\alpha t < x \leq 0 \\ \alpha & \text{if } 0 < x \leq \alpha t \\ x/t & \text{if } \alpha t < x \leq t \\ 1 & \text{if } t < x. \end{cases} \quad (2.16)$$

Hence, there exists a continuum set of weak solutions to (2.4)-(2.15)! It is therefore required to impose more restrictions on weak solutions to single out the correct one.

2.3 Entropy conditions

One of the criteria for selecting solutions is coming from the fact that the desired solution should be physically meaningful. In other words, it should be the limit state (in a sense to be made precisely) of a physical model where one adds an extra diffusion term to the right-hand side of (2.1),

$$\frac{\partial v}{\partial t}(x, t) + \nabla \cdot f(v(x, t)) = \epsilon \Delta v(x, t), \quad (2.17)$$

where Δ is the Laplacian operator with respect to spacial variables, that is, $\Delta v(x, t) = \sum_{i=1}^n \frac{\partial^2 v}{\partial x_i^2}$. The equation (2.17) may reflect a more physically meaningful model. Let us consider the traffic flow equation

$$\rho_t + (\rho v(\rho))_x = 0.$$

One may argue that in reality, the car drivers would naturally reduce their speed as they are approaching a traffic jam (or as they are seeing a relatively larger density of cars ahead). In that case, $v(\rho)$ may be replaced by

$$\tilde{v}(\rho) = v(\rho) - \epsilon \frac{\rho_x}{\rho},$$

and hence the equation becomes

$$\rho_t + (\rho v(\rho))_x = \epsilon \rho_{xx}.$$

This example shows the natural appearance of a diffusion term on the right-hand side of (2.17). It does not imply however that this should be the case for other situations. In

general, one may consider (2.17) even without any physical explanations to the diffusive term on the right-hand side. This may be because its presence could be beneficial for analytical or computational purposes. Therefore, as mentioned by Dafermos in his famous book [Daf16], the term $\epsilon\Delta v$ should be viewed as “artificial viscosity”.

We denote by η some auxiliary function, $\eta : \mathbb{R} \rightarrow \mathbb{R}$, and define the associated function $q : \mathbb{R} \rightarrow \mathbb{R}^n$ by

$$q(u) = \int^u \eta'(s)f'(s)ds, \quad (2.18)$$

that is, $q'(u) = \eta'(u)f'(u)$.

Remark 2.3.1. *The pair (η, q) satisfying (2.18) will be called an entropy-entropy flux pair.*

Now since (2.17) is parabolic, there exists a unique smooth solution (with f regular enough, see [Eva10]), denoted by v^ϵ , of the initial value problem (2.17)-(2.2). Multiplying both sides of (2.17) by $\eta'(v^\epsilon(x, t))$ and using (2.18), we get

$$\partial_t \eta(v^\epsilon) + \partial_x q(v^\epsilon) = \epsilon \Delta \eta(v^\epsilon) - \epsilon \eta''(v^\epsilon) |\nabla v^\epsilon|^2. \quad (2.19)$$

If η is convex, $\eta'' \geq 0$, then we can deduce from (2.19) that

$$\partial_t \eta(v^\epsilon) + \partial_x q(v^\epsilon) \leq \epsilon \Delta \eta(v^\epsilon).$$

Thus if we let $\epsilon \rightarrow 0$ and assume that,

$$v^\epsilon \rightarrow v \quad \text{almost everywhere,}$$

then v satisfies

$$\partial_t \eta(v) + \partial_x q(v) \leq 0. \quad (2.20)$$

We just arrive at the following definition (where we relax the regularity of η and q a bit and allow them to be merely Lipschitz continuous).

Definition 2.3.2. *A function $v \in L^\infty(\mathbb{R}^n \times [0, \infty))$ is called an entropy solution for (2.17)-(2.2) if (2.20) holds in the sense of distributions, that is*

$$\int_{\mathbb{R}^n} \int_0^\infty \eta(v(x, t)) \phi_t + q(v(x, t)) \phi_x dt dx + \int_{\mathbb{R}^n} \phi(x, 0) \eta(v_0(x)) dx \geq 0 \quad (2.21)$$

for any convex entropy-entropy flux pair (η, q) and any non-negative test function $\phi \in C_0^\infty(\mathbb{R}^n \times [0, \infty))$.

This entropy condition is sometimes called *Kruzkov's entropy condition*, to recognise the celebrated work by Kruzkov in [Kru70], where the entropy condition is stated in an abstract form as above. In that work, Kruzkov shows that we can actually choose the entropy-entropy flux pair to be very specific, namely

$$\eta(v) = |v - k|, \quad q(v) = \text{sign}(v - k)(f(v) - f(k)). \quad (2.22)$$

for any $k \in \mathbb{R}$. This will be referred to as the *Kruzkov entropy-entropy flux pair*.

When dealing with discontinuous solutions, it may be convenient (for example in numerical approximations) to test the entropy condition at every discontinuity curve. A weak solution will be an entropy solution if each of its shock curves satisfies certain admissibility conditions. When f is genuinely nonlinear, meaning $f'' \neq 0$, if a shock curve $\varphi = \varphi(t)$ (with speed $s = \dot{\varphi}(t)$) satisfies the following condition

$$f'(v_-) > s > f'(v_+), \quad (2.23)$$

then it will be called an *admissible shock* and (2.23) will be called *Lax's shock admissibility condition*, a criterion introduced in the seminal paper [Lax57]. Equivalently, if f is convex, then (2.23) reduces to

$$v_- > v_+.$$

In other words, one only accepts a shock curve if the wave from the left moves faster than the wave from the right of the curve. This admissibility criterion may be illustrated in the traffic flow model, where shocks are present as traffic jams, the car flow approaching a traffic jam moves faster than the one after joining in the traffic jam.

With more general f (not necessarily convex), for v to be an entropy solution, one requires the following condition, known as *Oleinik's shock admissibility condition* [Ole59], along every shock curve,

$$\frac{f(v_+) - f(v)}{v_+ - v} \leq \frac{f(v_+) - f(v_-)}{v_+ - v_-} \leq \frac{f(v_-) - f(v)}{v_- - v}, \quad (2.24)$$

for v in between v_- and v_+ .

It is easy to see that Oleinik's shock admissibility condition implies Lax's shock admissibility condition. However, if the flux function is genuinely nonlinear, they are equivalent. Moreover, we have

Theorem 2.3.3. *Assume that f is convex. Then the Kruzkov entropy condition, the Lax entropy condition and the Oleinik entropy condition are equivalent.*

Discussions leading to a proof of this result can be found, for instance, in Chapter 2 in [HR15] and in Chapter VIII in [Daf16].

Remark 2.3.4. • *Even though Lax or Oleinik entropy condition may be easier to work with, they do not indicate whether or not the considered solution is a weak solution. The Kruzkov notion of entropy condition is convenient in the sense that it includes the notion of weak solutions. Indeed, by applying (2.21) with $\eta(v) = \pm v$ and $q(v) = \pm f(v)$ we derive Definition 2.2.1 of weak solutions.*

- *There are other entropy conditions, such as Liu entropy condition, that are particularly useful in the case of systems of hyperbolic conservation laws, hence not discussed here. For an extensive presentation of this topic, see [Daf16].*

Remark 2.3.5 (Entropy, irreversibility and entropy production). *The entropy condition (2.20) has some connection to the Second Law of Thermodynamics (which says that the entropy¹ of a system is always increasing in time and the process is irreversible). In fact, (2.20) induces a time irreversibility condition on solutions. Indeed, assume that $v(x, t)$ is an entropy solution to (2.1) and denote*

$$\tilde{v}(x, t) := v(-x, -t).$$

Then it can be easily verified that \tilde{v} still satisfies (2.1) in the weak sense. However, \tilde{v} is not an entropy solution except when it is a classical solution, since the inequality (2.20) reverses its sign. When v develops shocks, the left hand side of (2.20) (viewed in the sense of distributions)

$$\mu := \partial_t \eta(v) + \partial_x q(v)$$

is a strictly negative measure. The measure defined as above is called the entropy production measure. An interesting property of this measure is that it concentrates on the set of points of jump discontinuity² (see, for instance, [DLR03, DLOW03, COW08, BM17]). We refer to [Daf16, Section 4.5, 3.3] for further discussions on this topic.

2.4 Existence and uniqueness of entropy solutions

We first need some definitions regarding functions of bounded variations which play a fundamental role in the theory of hyperbolic conservation laws.

¹We note that the “entropy” notion η used in the context of hyperbolic conservation laws takes an opposite sign with the counterpart “entropy” used in physics.

²This is still a conjecture for scalar conservation laws in \mathbb{R}^n with $n > 1$. See recent work [Sil19] for more information.

Definition 2.4.1. Let $\Omega \subset \mathbb{R}^n$ be open. The total variation of $u \in L^1(\Omega)$ is given by

$$TV(u, \Omega) := \sup \left\{ \int_{\Omega} u(x) \operatorname{div} \psi(x) dx : \psi \in C_c^\infty(\Omega, \mathbb{R}^n), \|\psi\|_{L^\infty(\Omega)} \leq 1 \right\}.$$

(We use the notation $TV(u)$ instead of $TV(u, \Omega)$ when Ω is clear from the context.) We say that u is a function of bounded variation if $TV(u) < \infty$, and we denote

$$BV(\Omega) := \{u : \Omega \rightarrow \mathbb{R}, TV(u) < \infty\}, \quad (2.25)$$

the space of functions of bounded variation on Ω . If $u \in BV(\Omega')$ for all $\Omega' \subset\subset \Omega$ (i.e., Ω' is compactly contained in Ω), we say that u is of locally bounded variation and use $BV_{\text{loc}}(\Omega)$ to denote the space of all these functions.

Remark 2.4.2. • If $u \in L^1(\Omega)$ is a function of bounded variation then its distributional derivative, denoted by Du or $\operatorname{grad} u$, is a finite (vector-valued) signed measure, and the total variation of this measure, denoted by $|Du|$, equals the total variation of u . We refer to [AFP00] for detailed discussion and treatment of BV functions.

- In one dimension, the above definition of functions of bounded variation agrees with the classical one, that is, u has bounded variation in an interval $I \subset \mathbb{R}$ if

$$TV(u) := \sup \left\{ \sum_{i=1}^{N-1} |u(x_{i+1}) - u(x_i)| : N \geq 2, x_1 < x_2 < \dots < x_N, x_i \in I \right\} < \infty. \quad (2.26)$$

We note however that in one dimension one does not normally require u to be integrable. Nevertheless, if I is bounded then u is automatically integrable since it must be bounded thanks to (2.26). Again, we refer to [AFP00, Section 3.2] (see also [Leo17, Chapter 2 and 7]) for details.

Given the entropy conditions, one recovers the uniqueness.

Theorem 2.4.3 (Kruzkov [Kru70]). For every $v_0 \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$, there exists a unique entropy solution v to (2.1)-(2.2) in $C([0, \infty); L^1_{\text{loc}}(\mathbb{R}^n))$ that satisfies, for every $t > 0$,

$$\|v(\cdot, t)\|_{L^\infty(\mathbb{R}^n)} \leq \|v_0(\cdot)\|_{L^\infty(\mathbb{R}^n)}.$$

Moreover, if \bar{v} is the entropy solution to (2.1) corresponding to the initial data \bar{v}_0 with $\bar{v}_0 \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$, then

$$\|v(\cdot, t) - \bar{v}(\cdot, t)\|_{L^1(\mathbb{R}^n)} \leq \|v_0(\cdot) - \bar{v}_0(\cdot)\|_{L^1(\mathbb{R}^n)}, \quad (2.27)$$

for all $t > 0$.

The inequality (2.27), often called L^1 -contraction property, implies immediately that the entropy solution is unique³. It also implies another important property of the entropy solution: if v_0 is a function of bounded variation, then $v(\cdot, t)$ is also a function of bounded variation, for all $t > 0$. To see this, we consider the total variation of v , regarded as the total variation of the (\mathbb{R}^n -valued) measure $\text{grad}v$ and notice that

$$\begin{aligned} |\text{grad}_\alpha(v(\cdot, t))|(\mathbb{R}^n) &= \lim_{h \rightarrow 0} \frac{1}{h} \int_{\mathbb{R}^n} |v(x + he_\alpha, t) - v(x, t)| dx \\ &\leq \lim_{h \rightarrow 0} \frac{1}{h} \int_{\mathbb{R}^n} |v_0(x + he_\alpha) - v_0(x)| dx = |\text{grad}_\alpha v_0|(\mathbb{R}^n), \end{aligned}$$

for $\alpha = 1, \dots, n$, thanks to (2.27).

The existence theory for scalar conservation laws is well developed. In the next sections we outline two of the most used methods to construct an entropy solution for (2.1).

2.4.1 Front tracking method

The front tracking approximation method (also known as *Dafermos's method*) was first studied by Dafermos in [Daf72] to show the existence of solutions for 1D scalar conservation laws. It was then developed further by Holden, Holden and Hoegh-Krohn in [HHHK88] as a numerical tool.

The idea is to approximate the initial function by a step function and the flux by a piecewise linear function. An approximate solution of (2.1)-(2.2) is then given by solving a finite set of so-called *Riemann problems*, with each initial data consists of a single discontinuity wave. More precisely, let v_0 be a function of bounded variation and M be a positive number such that $\|v_0\|_{L^\infty(\mathbb{R})} \leq M$. For some natural number N , define the set $A_N \subset [-M, M]$ by

$$A_N := \{v_{j,N} = j \frac{M}{2^N}, j = -2^N, \dots, 2^N\}.$$

Then define the piecewise linear function f^N by

$$\begin{aligned} f^N(s) &= f(v_{j,N}) + \frac{2^N}{M}(s - v_{j,N})(f(v_{j+1,N}) - f(v_{j,N})), \\ &\text{for } s \in (v_{j,N}, v_{j+1,N}], \quad j = -2^N, \dots, 2^N. \end{aligned} \tag{2.28}$$

Let $v_{0,N}$ be a step function taking values in the set A_N such that

$$\|v_{0,N} - v_0\|_{L^1(\mathbb{R})} \rightarrow 0, \quad \text{as } N \rightarrow \infty,$$

and consider the Cauchy problem

$$v_t + f^N(v)_x = 0, \quad v(x, 0) = v_{0,N}(x). \tag{2.29}$$

³The original result of Kruzkov proves the uniqueness for initial data in $L^\infty(\mathbb{R}^n)$ where (2.27) has a slightly stronger version.

Since $v_{0,N}$ is a step function with break points at, say $\{x_j\}_{j=1}^J$, a local solution to (2.29), defined for sufficiently small $t > 0$, is given by solving a finite number of Riemann problems, each of them with initial data of the form

$$v(x, 0) = \begin{cases} v_l & \text{if } x < x_j \\ v_r & \text{if } x > x_j, \end{cases} \quad (2.30)$$

where v_l and v_r denote the values of v_0 at the left and the right limits $v_{0,N}(x_j-), v_{0,N}(x_j+) \in A_N$. One claims that Riemann problems obey a maximum principle, meaning the solution to (2.29) with initial data (2.30) remains in between v_l and v_r , and that the solution of (2.29) will take values in the set $\{v_{j,N}\} \cup \{\text{break points of } f^{\frown N} \text{ or } f^{\smile N}\}$, where f^{\frown} (or f^{\smile}) denotes the convex envelope (or concave envelope) of f . This solution is defined up to some time $t = t_1$ where two or more jump discontinuities (coming from nearby Riemann problems) collide, forming new Riemann problems. The above procedure continues for new Riemann problems and the solution is prolonged up to some new collision time $t = t_2$, and so on. Luckily enough, this process does not go on forever thanks to a fact that the number of interactions is finite, see for example [Bre00] for a proof. This is understandable because each time a new collision forms, two or more discontinuities collapse to produce a single discontinuity. The wave pattern is hence simplified as the number of jump discontinuities is decreasing over time. In fact, if the initial condition v_0 is a non-negative differentiable function with compact support and the flux function is smooth and uniformly convex, then after a certain time, all shocks will finally be merged and continue as a single shock (see [KT05], also [Whi75]).

The method of front tracking approximations provides an effective tool in establishing the existence of entropy solutions of (1.1)-(1.2). Moreover, one can use this method to derive the stability property of the entropy solution with respect to the flux function, as the following result shows. Hereafter we denote

$$\|f\|_{Lip} := \sup_{u \neq v} \left| \frac{f(u) - f(v)}{u - v} \right|, \quad (2.31)$$

the Lipschitz constant for f .

Theorem 2.4.4. *Let $v_0 \in L^1(\mathbb{R}) \cap BV(\mathbb{R})$ and v, \bar{v} be the entropy solutions to (1.1)-(1.2) with respect to Lipschitz continuous flux functions f, g (respectively), then there exists a constant $C > 0$ such that*

$$\|v(\cdot, t) - \bar{v}(\cdot, t)\|_{L^1(\mathbb{R})} \leq Ct \|f - g\|_{Lip}, \quad (2.32)$$

for every $t > 0$.

The constant C in (2.32) depends only on v_0 and can be chosen to be $TV(v_0)$. This result was first obtained in [Luc86] (using Kuznetsov’s approximation theory for solutions). A proof using the front tracking method can be found in [HR15, Chapter 2]. We note that the assumption $v_0 \in BV(\mathbb{R})$ is needed in establishing *a priori* a bound on $TV(v(\cdot, t))$ which is an essential part in the proof of existence (typically via Helly’s theorem, see [Bre00]).

The front tracking method is considerably more intricate for higher dimensions, we refer to the monographs [HR15] and [Bre00] for the detailed treatment.

2.4.2 Vanishing viscosity method

Let ϵ be a positive number. Consider the “viscous” equation

$$\frac{\partial v}{\partial t}(x, t) + \nabla \cdot f(v(x, t)) = \epsilon \Delta v(x, t), \quad (2.33)$$

with some initial data

$$v(x, 0) = v_0(x), \quad (2.34)$$

where $v_0 \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$. The discussions carried in Section 2.3 lead to the following result.

Theorem 2.4.5. *Let v^ϵ be the solution of (2.33)-(2.34). Assume that for some sequence $\{\epsilon_k\}$, with $\epsilon_k \rightarrow 0$ as $k \rightarrow \infty$,*

$$v^{\epsilon_k} \rightarrow v \quad \text{boundedly almost everywhere on } \mathbb{R}^n \times [0, \infty).$$

for some function v . Then v is an entropy solution of (2.1)-(2.2).

Note that since (2.33) is a parabolic equation, the Cauchy problem (2.33)-(2.34) always has a unique smooth solution (see, for example, [Eva10] or [Smo83]). Moreover, this solution also enjoys the L^1 -contraction property.

Theorem 2.4.6. *Let v^ϵ and \bar{v}^ϵ be solutions of (2.33) with respect to initial data v_0 and \bar{v}_0 that are in $L^1(\mathbb{R}^n)$ and bounded. Then*

$$\|v^\epsilon(\cdot, t) - \bar{v}^\epsilon(\cdot, t)\|_{L^1(\mathbb{R}^n)} \leq \|v_0(\cdot) - \bar{v}_0(\cdot)\|_{L^1(\mathbb{R}^n)},$$

for all $t > 0$.

The result of Theorem 2.4.5 implies that the definition of entropy solution agrees with the notion of “viscosity” solution obtained via vanishing viscosity method. Finally, the following theorem shows that this viscosity solution does exist.

Theorem 2.4.7. *The solution v^ϵ of (2.33)-(2.34) converges as $\epsilon \rightarrow 0$ boundedly almost everywhere on $\mathbb{R}^n \times [0, \infty)$ to the unique entropy solution v of (2.1)-(2.2).*

The following lemma establishes an *equicontinuity in average* property (some sort of “integral” equicontinuity) for $\{v^\epsilon\}$ that comes in useful in proving Theorem 2.4.7 where a compactness argument in L^p is needed.

Lemma 2.4.8. *Let v^ϵ be the solution of (2.33)-(2.34). Then there exist a constant c , independent of ϵ , and a nondecreasing function ω on $[0, \infty)$, with $\omega(r) \rightarrow 0$ as $r \rightarrow 0$, such that the followings hold for any $t > 0$,*

$$\int_{\mathbb{R}^n} |v^\epsilon(x+y, t) - v^\epsilon(x, t)| dx \leq \omega(|y|), \quad y \in \mathbb{R}^n,$$

$$\int_{\mathbb{R}^n} |v^\epsilon(x, t+h) - v^\epsilon(x, t)| dx \leq c\omega(h^{1/3}), \quad h > 0.$$

Proof of Theorem 2.4.7. The sequence $\{v^\epsilon(\cdot, t)\}$ is equicontinuous in average thanks to Lemma 2.4.8. Hence by the Kolmogorov-Riesz theorem (a version of Arzelà-Ascoli’s theorem in L^p , see [Bre10, Theorem 4.26] or [HOH10]), $\{v^\epsilon(\cdot, t)\}$ lies in a compact set of $L^1_{\text{loc}}(\mathbb{R})$. Therefore, for every sequence $\{\epsilon_k\}$, $\epsilon^k \rightarrow 0$ as $k \rightarrow \infty$, there exists a subsequence, still denoted by $\{\epsilon_k\}$, and a function $v \in C^0([0, \infty); L^1_{\text{loc}}(\mathbb{R}^n))$, such that

$$v^{\epsilon_k}(\cdot, t) \rightarrow v(\cdot, t), \quad \text{in } L^1_{\text{loc}}(\mathbb{R}^n),$$

uniformly in $t \in [0, T]$ for any $T > 0$. This implies that, by passing to a further subsequence if necessary, v^{ϵ_k} converges to v boundedly almost everywhere on $\mathbb{R}^n \times [0, \infty)$. By Theorem 2.4.5, v is an entropy solution of (2.1)-(2.2). Due to the uniqueness of entropy solutions, the whole family $\{v^\epsilon\}$ must converge to v as $\epsilon \rightarrow 0$. This completes the proof. \square

The vanishing viscosity method in the context of scalar conservation laws is due to Kruzkov [Kru70]. Proofs of Theorem 2.4.6 and Lemma 2.4.8 can be found in there, or in [Daf16] (where the proof of Theorem 2.4.7 is adapted from) with somewhat simplified versions. Also in [Kru70], the ingenious *doubling of variables* technique was introduced in order to establish the contraction property (2.27). This technique can also be used, for instance, in establishing the uniqueness of entropy solutions of Hamilton-Jacobi equations in [CEL84], see also [Eva10].

2.5 Some regularity results

We highlight in this section a few fine properties of scalar conservation laws that will be useful later on. Consider the one-space dimensional conservation law

$$\partial_t v(x, t) + \partial_x f(v(x, t)) = 0, \quad (x, t) \in \mathbb{R} \times [0, \infty). \quad (2.35)$$

Assume that f is genuinely nonlinear, $f''(v) > 0$, $v \in \mathbb{R}$. First of all, v , being the entropy solution to (2.35), instantaneously becomes a function of locally bounded variation and enjoys the one-sided Lipschitz property, even if v_0 is merely in $L^\infty(\mathbb{R})$.

Theorem 2.5.1. *Let $v_0 \in L^\infty(\mathbb{R})$. Then the entropy solution v of (2.35), with initial data v_0 , belongs to $BV_{\text{loc}}(\mathbb{R} \times (0, \infty))$. Moreover, for every $t > 0$ and $x > y$, we have*

$$\frac{v(x, t) - v(y, t)}{x - y} \leq \frac{1}{\kappa t}, \quad (2.36)$$

where $\kappa = \min\{f''(v) : |v| \leq \|v\|_{L^\infty(\mathbb{R})}\}$.

The inequality (2.36) is first established by Oleinik [Ole57] and is known in the literature as the *Oleinik decay estimate*. A reverse result of this is also true, if a weak solution of (2.35) satisfies the one-sided Lipschitz property (2.36) then it is an entropy solution. Because of this reverse property, the inequality (2.36) is sometimes used as an admissibility condition to single out or characterise the entropy solution. Some extensions of this nice result to the case of non-convex flux, multi-dimensional space and convex balance laws can be found in [Gla08, AGN12, Hof83].

Theorem 2.5.1 records some sort of *regularising effect* ([COW08]) that is due to the genuine nonlinearity: entropy solutions generally become more regular than their initial data⁴. This regularising phenomenon shows up again in the following result, whose proof can be found in [Daf16], saying that the continuity is upgraded automatically to Lipschitz continuity once v becomes an entropy solution.

Theorem 2.5.2. *Assume that the set \mathcal{C} of continuity points of an entropy solution v has nonempty interior \mathcal{C}^0 . Then v is locally Lipschitz on \mathcal{C}^0 .*

Finally, we mention here a *generic property* of scalar conservation laws in one space dimension that features the finiteness of shocks.

⁴This is only applicable if the initial data is relatively rough (i.e., in L^∞), but not anymore when the initial data is smoother as we have seen previously, smooth initial data may result in discontinuous, hence less smooth, solutions!

Theorem 2.5.3. *Generically, entropy solutions of (2.35) with smooth initial data are piecewise smooth. More precisely, there exists a set $\mathcal{S} \in C^k(\mathbb{R})$, $k \geq 3$, of first category such that for each initial data in $v_0 \in C^k(\mathbb{R}) \setminus \mathcal{S}$, the solution v of (2.35) has only a finite number of shock curves and v is C^k in the complementary of the shocks.*

It is worth noting that, the interior set of $C^k(\mathbb{R}) \setminus \mathcal{S}$ is an open dense set in $C^k(\mathbb{R})$. In other words, for most initial data, the total number of shocks is finite. This result was first established, in a weaker version, by Schaeffer in [Sch73] and then improved by Dafermos in [Daf85] and in [Daf16], to which we refer for the proof. See also Tadmor and Tassa [TT93, Section 4] for a characterization of the set \mathcal{S} . We note however that this generic property fails to hold for the case of system, see Caravenna and Spinolo [CS17].

2.6 Summary and outlook

The theory of scalar conservation laws, especially in one-space dimension, is well-developed. This chapter only highlights a few aspects of this rich theory. In particular, we discuss the entropy conditions and the well-posedness theory of entropy solutions. Two strategies for the existence theory, the front tracking approximation and the vanishing viscosity method, are presented. Other methods for constructing entropy solutions include the finite difference scheme [Smo83], the nonlinear contraction semigroup [CL71, Cra72], the kinetic formulation theory [LPT94, TT07], to name a few. We refer to the monographs [Daf16, Bre00, LeF02, Ser99] for more information.

For the case of systems, much less is known. Some results in this chapter can be extended to the case of systems, but mostly not trivial. The first global existence result was established in the landmark paper by Glimm [Gli65] for initial data with small total variation, in which the so-called *random choice method* was introduced. Later on, DiPerna [DiP76] extended the front tracking method in the scalar case of Dafermos to 2×2 systems. Based on these works, Bressan [Bre92] and Risebro [Ris93] extended the front tracking method for systems of any size, providing an alternative existence proof. The construction of entropy solutions using the vanishing viscosity method (the extension of results in Section 2.4.2, in particular, Theorem 2.4.7) has been obtained only recently by Bressan and Bianchini in [BB05].

The general theory of multidimensional systems of conservation laws is *terra incognita*⁵, with several obstacles and open problems. We refer to [BGS07] for some aspects of the theory, with emphasis on the boundary-valued problems.

⁵a term used by Dafermos [Daf16].

Chapter 3

Inverse problems and Bayesian approach

Inverse problems arise in different fields and have numerous real-world applications. These can be seen in medical imaging, signal processing, seismology, oil recovery, weather prediction, machine learning, and many other examples in science, engineering and everyday life. Due to their wide range of applications, the study of inverse problems is an important task in both pure and applied mathematics. There are several monographs addressing the study of these problems, see, for example, [Gro93, EHN96, Vog02, Isa06, Tar05, Kir11], to mention only a few.

This chapter presents the mathematical theory of inverse problems. Section 3.1 gives an overview to inverse problems and their properties and in particular, introduces the notion of ill-posedness. We present a mathematical setting with an emphasis on inverse problems for functions, where the unknown lies in some function spaces. This is in particular applicable to inverse problems whose underlying physical processes are given by (partial) differential equations. Classical approaches to inverse problems, including Tikhonov regularization, are discussed here. Section 3.2 describes the Bayesian approach to inversion, which we employ to deal with inverse problems considered later in the thesis. Some general well-posedness and approximation theory will be given.

3.1 Inverse problems

Many problems in science and engineering can be described as follows: given a model and some causes, find the consequences of these causes over time, through the model. Such problems are called *direct* or *forward problems*.

On the other hand, in many applications, one is required to solve problems in a “reverse way”: starting with known consequences (given in some form of measurement data), one wants to find the causes, the model, or (parts of) both of them. The problems of determining causes or missing information in these situations are called *inverse problems*.

An outstanding feature of inverse problems is that they are often not well-posed, meaning that their solutions may not be unique, may be sensitive with respect to changes in data, or may not even exist. Dealing with inverse problems hence requires special techniques, called *regularizations*.

3.1.1 Mathematical formulation

An inverse problem of finding the unknown $u \in X$ from some observation $y \in Y$ can be written in the form

$$y = \mathcal{G}(u), \tag{3.1}$$

where

- $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces
- $u \in X$, the unknown parameters (input)
- $y \in Y$, the observation or measurement data (output)
- $\mathcal{G} : X \rightarrow Y$, the parameter-to-observation map.

The choice of Banach spaces for the input and output is due to the fact that they are natural spaces for most of problems in mathematical physics. In practice, or for analytical convenience, one may work only with certain subspaces of Banach spaces where some desired structures (such as separability) are required. The mapping \mathcal{G} is typically a composition of the *evaluation operator* \mathcal{O} and the *forward map* S , that is

$$\mathcal{G}(u) = \mathcal{O} \circ S(u).$$

Considering a model involving some partial differential equation (PDE) for example, the parameter-to-observation map then evaluates the value of the solution of the PDE, which is in turn a result of the forward map, defined as the mapping of parameters to the solution of the PDE.

If one takes into account the effect of noise then the problem can be written as

$$y = \mathcal{G}(u) + \xi, \tag{3.2}$$

where ξ denotes the noise. The correct value of noise is typically not known, but some statistical information may be given. The noise is present in the problem due to various reasons, but it is mostly contributed from the following three factors.

- Device errors. In practice, the evaluation process is performed using some device. Therefore, the errors caused by imperfect devices are almost always unavoidable. These errors contribute to the measurement noise.
- Numerical errors. Most PDEs cannot be solved exactly. If the equations governing the model can be solved numerically, then the numerical errors can be considered as the measurement noise.
- Modelling errors. Many physical phenomena are complex. Mathematical modelling of these phenomena is often simplified or approximated to more simple ones to make them easier to deal with. Therefore, there are errors caused by the imperfect modelling of the underlying physical process of mapping u to y .

The presence of noise also motivates the statistical consideration of inverse problems with a Bayesian perspective. We will discuss this approach in detail in Section 3.2.

3.1.2 Ill-posedness

Mathematical modelling of physical problems sometimes can be a tough task. It is often desirable for the resulting models to satisfy certain requirements. For instance, one requires that the model is solvable, if it represents the correct physical phenomenon and if it is stable with respect to changes in the model parameters. In other words, for any model of a physical phenomenon, one asks the following questions

- Existence: does there exist at least one solution?
- Uniqueness: is there only one solution?
- Stability: does the solutions depend continuously on the initial data?

If the answers are positive for all three questions, one says that the model is *well-posed*, a term coined by Jacques Hadamard [Had02] at the beginning of the last century. According to Hadamard, a mathematical model of any physically relevant problem should be well-posed to be considered meaningful. Many forward PDE problems are well-posed¹. This well-posedness concept is also adapted to inverse problems.

¹It is also assumed that the counterpart forward problem which the considered inverse problem is based on is well-posed.

Definition 3.1.1. *The inverse problem (3.2) is said to be well-posed if it satisfies the following three conditions*

- *for each $y \in Y$, there exists $u \in X$, called a solution, for which the relation (3.2) holds;*
- *there is only one solution; and*
- *the solution is stable with respect to changes in the data: if $y_n = \mathcal{G}(u_n) + \xi$ and $y_n \rightarrow y$ then $u_n \rightarrow u$ (in appropriate topologies).*

If at least one of these conditions is not satisfied, the problem is called ill-posed.

In contrast with forward problems, inverse problems are typically ill-posed, as the following example shows.

Example 3.1.2. *Consider a linear inverse problem of finding u from y where*

$$y = Ku \tag{3.3}$$

for some linear operator K . The solution u may be formally given as

$$u = K^{-1}y.$$

This inversion may be infeasible since

- *K^{-1} may be unbounded. This happens for instance, when K is a compact operator and X is infinite-dimensional (see, e.g. [Bre10]).*
- *y may not belong to $\text{Im}(K)$ when noise is taken into account, that is, when the data is given as $y = Ku + \xi$.*

Therefore, this inversion in general fails to produce any sort of reasonable solution of the parameter u . The situation is even worse in most mathematical physics models, when K is nonlinear (or highly nonlinear, where K is given implicitly via a forward nonlinear PDE), this inversion task is simply impossible.

This example, in fact, brings out two types of ill-posedness of inverse problems that one encounters in practice. The first one is when there is not enough data to reflect all aspects of the parameters. This is especially the case when the unknown is parameters of some PDE, hence typically lies in infinite-dimensional spaces, while data obtained by measurement devices is mostly finite. This leads to the inverse problem being highly *underdetermined*.

The other type of ill-posedness is when the parameter-to-observation map \mathcal{G} is not surjective, and the data is perturbed by noise. Note that specific instances of noise are largely unknown and one cannot just subtract it.

Several techniques have been developed to overcome difficulties due to ill-posedness and provide reasonable solutions to inverse problems. In the next section, we review some of the most used techniques.

3.1.3 Classical regularization approach

Basics of regularizations

By solving an inverse problem, one aims to find $u \in X$ such that the forward map given via the forward problem best fits the observation data, that is, find u such that it solves the minimization problem

$$u = \arg \min_{v \in X} J(v), \quad J(v) = \frac{1}{2} \|y - \mathcal{G}(v)\|_Y^2. \quad (3.4)$$

For $Y = \mathbb{R}^n$ with Euclidean norm, this is known as the *least squares method*. However, if infimizing sequences do not converge in X , then there does not exist a minimizer to the problem (3.4). Also, there might be more than one minimizers to (3.4), leading to multiple solutions to the inverse problem. To get round this, one needs some regularization techniques.

The general idea of regularization is, instead of trying to solve (3.1) exactly, one solves a “nearby problem” that has more desirable properties. To illustrate this, consider the linear inverse problem (3.3) where the inverse operator K^{-1} is unbounded. A regularization strategy is to introduce an approximation R_α of the inverse operator K^{-1} where

$$R_\alpha : Y \rightarrow X, \quad \alpha > 0,$$

such that

$$\lim_{\alpha \rightarrow 0} R_\alpha K u = u, \quad \text{for all } u \in X. \quad (3.5)$$

In other words, the operator $R_\alpha K$ converges pointwise to the identity operator. Note however that, we do not expect this convergence to be uniform, since otherwise K^{-1} will be a bounded operator.

So far in this strategy, we have not taken into account the influence of noise. When incorporating noise, the data y is only known up to an error of, say, $\delta > 0$. That is, the measured data is given as y^δ where

$$\|y - y^\delta\|_Y \leq \delta.$$

The approximate solution u^δ of u solves the perturbed equation

$$y^\delta = Ku^\delta.$$

Consider the total error

$$\begin{aligned} \|R_\alpha y^\delta - u\|_X &\leq \|R_\alpha y^\delta - R_\alpha y\|_X + \|R_\alpha y - u\|_X \\ &\leq \delta \|R_\alpha\|_{\mathcal{L}(Y,X)} + \|R_\alpha Ku - u\|_X. \end{aligned} \tag{3.6}$$

As α tends to zero, the second term in (3.6) tends to zero thanks to (3.5). The first term is, however, not bounded. Indeed, if there exists $c > 0$ such that $\|R_\alpha\|_{\mathcal{L}(Y,X)} \leq c$ for all $\alpha > 0$, then since $R_\alpha y \rightarrow K^{-1}y$ for all $y \in \mathcal{R}(K)$ (the range of K) and $\|R_\alpha y\|_X \leq c\|y\|_Y$ we conclude that $\|K^{-1}y\|_X \leq c\|y\|_Y$ for every $y \in \mathcal{R}(K)$, a contradiction to the assumption that K^{-1} is unbounded. Therefore, to find a best fit solution to the inverse problem, one needs to keep both these terms minimal. In other words, one has to minimize the following sum

$$\delta \|R_\alpha\|_{\mathcal{L}(Y,X)} + \|R_\alpha Ku - u\|_X,$$

in terms of α .

For more detailed discussions, in particular, how to carry the minimization procedures and how to construct regularization operators, we refer to [EHN96, Kir11], see also the notes [Bal12]. For other considerations and techniques, especially for inverse problems in PDEs, see [Isa06, FSU19].

Tikhonov regularization

One of the most used regularization techniques is the so-called *Tikhonov regularization* where one considers the following minimization problem

$$u_\alpha = \arg \min_v J_\alpha(u), \quad J_\alpha(u) = \|y - \mathcal{G}(v)\|_Y + \alpha \|v\|_X. \tag{3.7}$$

Compared to the original least squares minimization problem (3.4), there is a newly added term $\alpha \|v\|_X$, which is called the *penalty term*. The constant α is referred to as the *regularization parameter*.

The penalty term induces some form of prior information to the inverse problem. The following theorem ensures the well-posedness of the Tikhonov regularization.

Theorem 3.1.3. *Let $K : X \rightarrow Y$ be a linear and bounded operator and $\alpha > 0$. Then the Tikhonov functional J_α has a unique minimizer $u_\alpha \in X$.*

We refer to [EHN96, Vog02] for a proof of this result and for other discussions.

3.2 Bayesian approach to inverse problems

Consider the inverse problem of finding u from perturbed noise y , given by the following relation

$$y = \mathcal{G}(u) + \xi. \quad (3.8)$$

We assume that the noise ξ is *observable*, meaning that its probability distribution is known. The Bayesian approach to inverse problems initiates with the observation that, since the data y is perturbed by the noise ξ , which is a random variable, it is then natural to treat the unknown u in (3.8) as a random variable too. In this view, the relation (3.8) defines a conditioned random variable $y|u$ (y given u). Hence, by this logic, the conditioned random variable $u|y$ (u given y) should provide some form of a “solution” to the inverse problem (3.8).

The starting point of the Bayesian approach, in finite dimension, is the following usual Bayes’ theorem for conditional probability

$$\begin{aligned} \mathbb{P}(u|y) &= \frac{\mathbb{P}(y|u)\mathbb{P}(u)}{\mathbb{P}(y)} \\ &\propto \mathbb{P}(y|u)\mathbb{P}(u), \end{aligned} \quad (3.9)$$

where \propto denotes the proportionality, if $\mathbb{P}(y) > 0$. In this formula, $\mathbb{P}(y|u)$ is the *likelihood* of data which indicates the relationship between the data y and the forward model $\mathcal{G}(u)$. The probability $\mathbb{P}(u)$ is the *prior*, it represents the belief of which values that u is more likely to take (before the data y is taken into account). The formula (3.9) expresses that the conditional distribution of u given y , usually called the *posterior*, can be computed if some knowledge of the prior and the likelihood is known. An equivalent form of (3.9) for densities reads

$$\begin{aligned} \rho(u|y) &= \frac{\rho(y|u)\rho(u)}{\rho(y)} \\ &\propto \rho(y|u)\rho(u). \end{aligned} \quad (3.10)$$

Note that this formula only makes sense when the *normalization constant* $\rho(y)$, which may be expressed as

$$\rho(y) = \int_X \rho(y|u)\rho(u)du,$$

is positive.

In many inverse problems, in particular the ones involving PDEs, the unknown u belongs to some function space of infinite dimension. The above formulae then need to be adapted to an infinite-dimensional setting, where densities in normal sense (i.e., with respect to the Lebesgue measure) do not exist. We therefore characterise the prior $\rho(u)$ and

posterior $\rho(u|y)$ as probability measures, μ_0 and μ^y , respectively. And the link between them may be given via Radon-Nikodym theorem

$$\frac{d\mu^y}{d\mu_0}(u) = \frac{1}{Z(y)} \exp(-\Phi(u; y)), \quad (3.11)$$

with some function, so-called *potential*, Φ and the normalization constant given by

$$Z(y) = \int_X \exp(-\Phi(u; y)) \mu_0(du). \quad (3.12)$$

To know how the Bayesian approach works in the context of inverse problems, we will have to answer the following questions

- (1) Is the posterior μ^y well-defined?
- (2) Is the posterior μ^y stable under small changes in data y ?

Positive answers to these two questions would give some form of well-posedness for solutions of Bayesian inverse problems. Moreover, since we are interested in inverse problems involving PDEs, it is natural to ask a third question

- (3) Is the posterior μ^y stable under disturbances in the forward model?

In the next two sections, we will present the ideas of Bayesian inversion rigorously in a general setting and discuss these questions in detail. Before going forward, we need some notion of distance on measure spaces.

Definition 3.2.1. *Let μ and μ' be two probability measures on $(X, \mathcal{B}(X))$ that are both absolutely continuous with respect to μ_0 . Then*

$$d_{\text{Hell}}(\mu, \mu') = \sqrt{\frac{1}{2} \int_X \left(\sqrt{\frac{d\mu}{d\mu_0}} - \sqrt{\frac{d\mu'}{d\mu_0}} \right)^2 d\mu_0},$$

defines a distance between two measures μ and μ' and is called the Hellinger distance.

For the properties of Hellinger distance and its relation to other measures, we refer to [DS16, PC19, GS02].

3.2.1 Well-posedness

Assume that the unknown u of the inverse problem (3.8) lies in a set X and the data y is given in Y , with X, Y Banach spaces. We suppose that a prior probability measure μ_0 on $(X, \mathcal{B}(X))$ is given and $\xi \sim \mathbb{Q}_0$ with \mathbb{Q}_0 known. We assume also that ξ and u are

independent with $y|u \sim \mathbb{Q}_u$ which for some potential $\Phi : X \times Y \rightarrow \mathbb{R}$ (obtained through equation (3.8)) satisfies

$$\frac{d\mathbb{Q}_u}{d\mathbb{Q}_0} = \exp(-\Phi(u; y)).$$

Let the measure ν be defined by $\nu(du, dy) := \mu_0(du)\mathbb{Q}_u(dy)$. Then we have the following version of Bayes' theorem.

Theorem 3.2.2 ([DS11]). *Assume that $\Phi : X \times Y \rightarrow \mathbb{R}$ is $\mu_0 \otimes \mathbb{Q}_0$ measurable and that, for y \mathbb{Q}_0 -a.s.,*

$$Z(y) := \int_X \exp(-\Phi(u; y))\mu_0(du) \in (0, \infty). \quad (3.13)$$

Then the conditional distribution of $u|y$ exists under ν , and is denoted by μ^y . Furthermore $\mu^y \ll \mu_0$ and, for $y \in Y$, it holds ν -a.s.,

$$\frac{d\mu^y}{d\mu_0}(u) = \frac{1}{Z(y)} \exp(-\Phi(u; y)). \quad (3.14)$$

We refer to [DS16] for a proof of this theorem and related discussions. In the following, it is showed that the finiteness and positivity condition (3.13) of the normalization constant Z are satisfied for some general conditions of Φ . Under these conditions, the posterior will also be stable.

Assumption 3.2.3. The function $\Phi : X \times Y \rightarrow \mathbb{R}$, for X, Y are separable Banach spaces, satisfies the following properties.

(i) For every $y \in Y$, $\Phi(\cdot, y)$ is a measurable function with respect to u .

(ii) For every $r > 0$, there exists a constant $M(r)$ such that

$$\Phi(u; y) \leq M(r), \quad \forall u \in B_X(0, r), \forall y \in B_Y(0, r).$$

(iii) For every $r > 0$, there exists a function $M_1(r, \cdot) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that

$$\Phi(u; y) \geq -M_1(r, \|u\|_X), \quad \forall u \in X, \forall y \in B_Y(0, r).$$

(iv) For every $r > 0$, there exists a function $M_2(r, \cdot) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that

$$|\Phi(u; y_1) - \Phi(u; y_2)| \leq M_2(r, \|u\|_X)\|y_1 - y_2\|_Y, \quad \forall u \in X, \forall y_1, y_2 \in B_Y(0, r).$$

Theorem 3.2.4 (Well-definedness). *Let Assumption 3.2.3 holds. Assume that μ_0 is a probability prior measure on $(X, \mathcal{B}(X))$ and for each $r > 0$,*

$$\exp(M_1(r, \|u\|_X)) \in L^1_{\mu_0}(X; \mathbb{R}). \quad (3.15)$$

Then, for every $y \in Y$, $Z(y)$ given by (3.12) is positive and finite, and μ^y given by (3.11) defines a posterior probability measure on X .

Proof. For fixed $y \in Y$, we need to show that $0 < Z(y) < \infty$. The boundedness from above is followed from Assumption 3.2.3 (iii) and (3.15),

$$Z(y) \leq \int_X \exp(M_1(r, \|u\|_X)) \mu_0(du) < \infty.$$

To prove the positivity of Z , let $r > 0$ and using Assumption 3.2.3 (ii), we see that

$$\begin{aligned} Z(y) &\geq \int_{B_X(0,r)} \exp(-\Phi(u; y)) \mu_0(du) \\ &\geq \int_{B_X(0,r)} \exp(-M(r)) \mu_0(du) \\ &= \exp(-M(r)) \mu_0(B_X(0, r)), \end{aligned}$$

which is positive provided that $\mu_0(B_X(0, r)) > 0$. Since $\mu_0(X) = 1$, there must be some $r > 0$ large enough such that $\mu_0(B_X(0, r)) > 0$. This completes the proof. \square

Theorem 3.2.5 (Stability). *Let Assumption 3.2.3 holds. Assume that μ_0 is a probability prior measure on $(X, \mathcal{B}(X))$ and for each $r > 0$,*

$$\exp(M_1(r, \|u\|_X)) (1 + M_2(r, \|u\|_X)^2) \in L^1_{\mu_0}(X; \mathbb{R}). \quad (3.16)$$

Then, for all $y, y' \in B_Y(0, r)$, there exists $C = C(r)$ such that

$$d_{\text{Hell}}(\mu^y, \mu^{y'}) \leq C \|y - y'\|_Y.$$

Sketch of Proof. The condition (3.16) implies the condition (3.15), thus implies the well-definedness of μ^y and $\mu^{y'}$. It follows from the local Lipschitz of the exponential function, Assumption 3.2.3 and (3.16) that

$$\begin{aligned} |Z(y) - Z(y')| &\leq \int_X \exp(M_1(r, \|u\|_X)) |\Phi(u; y) - \Phi(u; y')| \mu_0(du) \\ &\leq C \|y - y'\|_Y. \end{aligned}$$

Now from the definition of Hellinger distance

$$\left(d_{\text{Hell}}(\mu^y, \mu^{y'})\right)^2 \leq I_1 + I_2,$$

where, by using similar arguments,

$$\begin{aligned} I_1 &= \frac{1}{Z(y)} \int_X \left(\exp\left(-\frac{\Phi(u; y)}{2}\right) - \exp\left(-\frac{\Phi(u; y')}{2}\right) \right)^2 \mu_0(du) \\ &\leq \frac{1}{4Z(y)} \int_X \exp(M_1(r, \|u\|_X)) |\Phi(u; y) - \Phi(u; y')|^2 \mu_0(du) \\ &\leq C \|y - y'\|_Y^2, \end{aligned}$$

and

$$\begin{aligned} I_2 &= \left| Z(y)^{-\frac{1}{2}} - Z(y')^{-\frac{1}{2}} \right|^2 \int_X \exp(-\Phi(u; y')) \mu_0(du) \\ &\leq C |Z(y) - Z(y')|^2 \leq C \|y - y'\|_Y^2, \end{aligned}$$

which completes the proof. \square

The assumption leading to the well-posedness results, Assumption 3.2.3, is proposed in [CDRS09, Stu10, CDS10] in a weaker version than the one above, where Φ is assumed to be Lipschitz continuous with respect to u . It is upgraded to continuity in [DS16]. Looking closely at the proof in [DS16], we can see that only the measurability (of Φ in u) is enough to derive the same results, as stated and proved above (see also [Sul17]).

When the data space Y is finite-dimensional and the observational noise ξ is non-degenerate Gaussian, then μ^y is continuous in y under very mild conditions.

Theorem 3.2.6 ([Lat20]). *Let $Y := \mathbb{R}^J$, ξ have Gaussian distribution $\mathcal{N}(0, \Gamma)$ with Γ symmetric positive definite, and $\mathcal{G} : X \rightarrow \mathbb{R}$ be μ_0 -measurable. Then for any $y \in Y$ and $\{y^k\} \subset Y$ with $y^k \rightarrow y$ as $k \rightarrow \infty$, we have*

$$d_{\text{Hell}}(\mu^y, \mu^{y^k}) \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

This result establishes a so-called *continuity well-posedness* (in contrast to the above *Lipschitz well-posedness*) property of the posterior with respect to data.

3.2.2 Approximations

We now investigate the stability property of the posterior with respect to the forward map. Typically, this forward map is defined as a solution map of a PDE. In practice, solving a PDE normally carries some sort of approximations. The question of whether the perturbed posterior arising from the approximation of the forward model converges to the posterior given by the exact one is then importantly necessary to study. Let μ (we drop the variable y since it does not have any explicit role in this task) be the solution of the Bayesian inverse problem (3.8), which is given by

$$\frac{d\mu}{d\mu_0}(u) = \frac{1}{Z} \exp(-\Phi(u)), \quad (3.17)$$

$$Z = \int_X \exp(-\Phi(u)) \mu_0(du). \quad (3.18)$$

Let μ^N be the measure defined by

$$\frac{d\mu^N}{d\mu_0}(u) = \frac{1}{Z^N} \exp(-\Phi^N(u)), \quad (3.19)$$

$$Z^N = \int_X \exp(-\Phi^N(u)) \mu_0(du), \quad (3.20)$$

where $\Phi^N(u)$ is some approximation of $\Phi(u)$. We need some general assumptions on this approximation.

Assumption 3.2.7. *Let Φ and Φ^N satisfy Assumption 3.2.3 and assume that there exists a measurable function $M_3 : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that*

$$|\Phi(u) - \Phi^N(u)| \leq M_3(\|u\|_X) \psi(N),$$

for some $\psi(N) \rightarrow 0$ as $N \rightarrow \infty$.

Theorem 3.2.8 (Approximation). *Let Assumption 3.2.7 holds. Assume that μ_0 is a probability prior measure on $(X, \mathcal{B}(X))$ and*

$$\exp(M_1(\|u\|_X)) (1 + M_3(\|u\|_X)^2) \in L_{\mu_0}^1(X; \mathbb{R}). \quad (3.21)$$

Then there exists $C > 0$ such that

$$d_{\text{Hell}}(\mu, \mu^N) \leq C\psi(N).$$

The proof of Theorem 3.2.8 is very similar to the one of Theorem 3.2.5 and can be found in [DS16].

If the data is finite and the noise is Gaussian, $\xi \sim \mathcal{N}(0, \Gamma)$, then the potential Φ takes the form

$$\Phi(u) = \frac{1}{2} |y - \mathcal{G}(u)|_{\Gamma}^2. \quad (3.22)$$

We define Φ^N as

$$\Phi^N(u) := \frac{1}{2} |y - \mathcal{G}^N(u)|_{\Gamma}^2. \quad (3.23)$$

where $\mathcal{G}^N(u)$ is some approximation of $\mathcal{G}(u)$. Then the approximation $\mathcal{G}^N(u)$ of $\mathcal{G}(u)$ may translate to the approximation μ^N of μ , as the following consequence of Theorem 3.2.8 shows.

Corollary 3.2.9. *Assume that there exists a measurable function $M(\cdot) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that $\mathcal{G}(u), \mathcal{G}^N(u) : X \rightarrow \mathbb{R}^J$ satisfy the following condition for all $u \in X$,*

$$|\mathcal{G}(u) - \mathcal{G}^N(u)| \leq M(\|u\|_X) \psi(N), \quad (3.24)$$

where $\psi(N) \rightarrow 0$ as $N \rightarrow \infty$. Suppose in addition that μ_0 is a probability measure on X such that

$$\exp(M(\|u\|_X)) \in L_{\mu_0}^1(X; \mathbb{R}).$$

Then, there exists $C > 0$ such that

$$d_{\text{Hell}}(\mu, \mu^N) \leq C\psi(N),$$

for all N sufficiently large.

3.3 Summary and notes

This chapter introduces some fundamental notions of inverse problems and presents the basics of regularisation methods in solving inverse problems with the focus on the Bayesian approach. We outline the main points of the well-posedness and approximation theories for solutions of the Bayesian approach. See [DLSV13] for the connection between the Bayesian approach and Tikhonov regularization (in particular, the Maximum A Posteriori (MAP) of the solution of the Bayesian approach forms a minimization problem). The Bayesian framework allows us to obtain much more information. However, the computational cost for the Bayesian approach is more expensive. See [Bal12] for further discussions.

The origin of regularization methods dates back to the forties, starting with the pioneering work of Tikhonov [Tik43]. It was subsequently developed further during the sixties, mostly by the Soviet community (see, e.g., [Iva62, Tik63, Tik66, Bak67]). These methods were then described systematically in the book [TA77] and later on in [EHN96]. For a recent review, we refer to [BB18]. The presentation of regularisation methods here is adapted from [Kir11].

The Bayesian inverse problems for functions was first described for the linear case in [Fra70], followed by [Man84] and [LPS89]. The approach was developed further by the Finnish School, see for instance, [LS04], [KS07], [Las07], [LSS09]. In the last ten years, this approach was developed systematically and studied intensively in the general setting of Banach spaces, started with [CDRS09], where the authors proposed an infinite-dimensional framework to Bayesian inverse problems, with applications in fluid mechanics. The well-posedness theory of Bayesian inverse problems under numerical approximations of the forward model was studied in [CDS10]. See also [Lat20], [Spr20], [ST18] and [Sul17] for recent developments in the well-posedness and approximation theory. The applications of the Bayesian approach to inverse problems are vast, we mention a sample of recent works [BTGMS13, ILS14, ILS16, DIS17, GTSA17, JYPG18, AN19, SW20].

For more comprehensive details and background information on the Bayesian approach for inverse problems, we refer to the book by Kaipio and Somersalo [KS05], the survey by Stuart [Stu10] and the notes by Dashti and Stuart [DS16].

Chapter 4

Filippov theory for discontinuous differential equations

In this chapter, we present the fundamentals of differential equations whose right-hand sides are discontinuous functions. This type of equations arises from a large number of applications in mechanics, the automatic control theory and dynamical system [AVK66]. We consider two common scenarios of the discontinuities. The first is for the right-hand side function that is discontinuous in t but is continuous in x , where the notion of Carathéodory solutions is introduced. The second allows the function on the right-hand side to be discontinuous in both x and t , where we present Filippov theory [Fil88]. Starting in Section 4.1 with a few typical examples, these theories will be described in Section 4.2 and 4.3. We first provide appropriate definitions of solutions. We then discuss the existence theory and some sufficient conditions for the uniqueness for both Carathéodory solutions and Filippov solutions. We also mention some classical results in the theory of ordinary differential equations which still hold true for equations with discontinuous right-hand sides.

4.1 Equations with discontinuous right-hand sides

Consider the initial value problem

$$\frac{dx}{dt} = f(x, t), \tag{4.1}$$

$$x(t_0) = x_0. \tag{4.2}$$

When f is a Lipschitz continuous function of x uniformly in t , the Cauchy-Lipschitz theory (see, for instance, [Tes12]) implies the existence and uniqueness of solutions. However, in applications, there are situations where one encounters differential equations with a

discontinuous function on the right-hand side. New definitions of solutions are required and justified for these applications. Below we consider a few examples, the first two are classical and taken from [Fil88], the third example appears in traffic flow model and will be studied in great detail later on.

Example 4.1.1. Consider the following equation

$$\dot{x}(t) = \operatorname{sgn}(t).$$

The function on the right-hand side $f(x, t) = \operatorname{sgn}(t)$ is continuous in x and discontinuous in t . A reasonable solution may be given as $x(t) = |t| + c$, however x is not differentiable at the point $t = 0$. In the (x, t) plane, the corresponding vector field is joining the line $t = 0$ from the left and leaving it from the other side.

Example 4.1.2. Consider now the equation

$$\dot{x}(t) = 1 - 2\operatorname{sgn}(x(t)).$$

The function on the right-hand side $f(x, t) = 1 - 2\operatorname{sgn}(x)$ is discontinuous in x . For $x < 0$, $\dot{x} = 3$ and a solution is given by $x(t) = 3t + c$. For $x > 0$, a solution is given by $x(t) = -t + c$. When t increases, each solution reaches the line $x = 0$. It does not indicate how the solution will continue however. Indeed, because of the direction field, if $x > 0$ then $\dot{x} < 0$ and the solution tends to go down, but once $x < 0$, $\dot{x} > 0$ and the solution has to go up. Thus, the solution cannot leave the line $x = 0$. However, the solution cannot stay along this line either, since otherwise $x(t) = 0$ and this solution does not satisfy the equation. This renders some difficulty in expressing the solution further in time if it is understood in the classical sense.

Example 4.1.3. Consider a practical example when one wants to observe the trajectory of a car travelling with speed v on the highway. Its trajectory satisfies the equation

$$\dot{x}(t) = v(x(t), t).$$

The function on the right-hand side is expected to be discontinuous due to shocks (which may be seen in the presence of traffic jams or traffic lights) as discussed in Chapter 2, in both x and t . The set of discontinuities consists of (at most countable) shock curves and (also at most countable) irregular points which are formed by the collision of shock curves.

In the next two sections, we will consider different concepts of “solutions” which cover the above situations, first when the function f is discontinuous in t (Carathéodory solution), and then when f is discontinuous in both x and t (Filippov solution), with emphasis on the latter.

4.2 Carathéodory differential equations

Definition 4.2.1. *The function f is said to satisfy the Carathéodory conditions in the domain D if in this domain*

- $f(x, t)$ is continuous in x for almost all t ;
- $f(x, t)$ is measurable in t for every x ; and
- $|f(x, t)| \leq m(t)$ with some locally integrable function m .

The equation $\dot{x} = f(x, t)$ (where x is a scalar or a vector) is called a Carathéodory equation if f satisfies the Carathéodory conditions.

Definition 4.2.2 (Carathéodory solution). *A function $x(t)$ defined on an (open or closed) interval I is called a solution of the Carathéodory equation if $x(t)$ is absolutely continuous on each closed interval in I and satisfies this equation almost everywhere. In other words, it solves*

$$x(t) = x(t_0) + \int_{t_0}^t f(x(s), s) ds,$$

for some $t_0 \in I$.

Theorem 4.2.3. *Let the function f satisfy the Carathéodory conditions in the domain D with $(x_0, t_0) \in D$. Then there exists a Carathéodory solution of the problem (4.1)-(4.2).*

Moreover, if there exists an integrable function $l(t)$ such that for any (x, t) and (y, t) in D ,

$$|f(x, t) - f(y, t)| \leq l(t)|x - y|, \tag{4.3}$$

then (4.1)-(4.2) has exactly one solution in the domain D .

Proof. The existence follows from a standard iteration with the help of Arzelà-Ascoli's theorem. For details we refer to [Fil88, Theorem 1, §1]. For the uniqueness, since the technique is quite useful and will be used later we present it here. Assume that $x(t)$ and $y(t)$ are solutions of the problem (4.1)-(4.2). Let $z(t) = x(t) - y(t)$, then for $t \geq t_0$,

$$\begin{aligned} \frac{1}{2} \frac{d|z(t)|^2}{dt} &= z(t) \cdot \frac{dz(t)}{dt} = (x - y) \cdot (f(x, t) - f(y, t)) \\ &\leq |x - y| \cdot |f(x, t) - f(y, t)| \\ &\leq l(t)|x - y|^2, \end{aligned}$$

almost everywhere, thanks to (4.3). This implies that

$$\frac{d}{dt} \left(|z(t)|^2 \exp \left(-2 \int_{t_0}^t l(s) ds \right) \right) \leq 0.$$

Together with $z(t_0) = x(t_0) - y(t_0) = 0$, it follows that $z(t) = 0$ for $t \geq t_0$. The case $t \leq t_0$ is proved by the substitution of $-t$ for t . \square

Many properties of solutions of equations with continuous right-hand sides continue to hold for the case of Carathéodory equations. In particular, the existence theorem can be extended to a maximal interval (up to the boundary of the domain), and the uniform limit of a sequence of solutions is a solution to the same Carathéodory equation (see [Fil88, Chapter 1]).

4.3 Filippov solution: Existence and uniqueness

In this section, we consider the differential equation (4.1) where the right-hand side is allowed to be discontinuous in both t and x . The definition of Carathéodory solution is no longer applicable for most equations of this type for which the right-hand sides are discontinuous on a smooth curve or surface (see Example 4.1.2, where the right-hand side function is discontinuous on the line $x(t) = 0$). It is then necessary to have a general notion of solutions that are flexible enough to cover these situations. At the same time this general notion must also be restrictive enough so that the general solutions coincide with the usual ones when the equation becomes classical, i.e., when the right-hand side is Lipschitz continuous.

Definition 4.3.1 (Filippov solution). *Let a function f in a domain D be discontinuous on a set M of measure zero. A solution in the sense of Filippov (or a Filippov solution) of (4.1) is an absolutely continuous function $x(t)$ defined on an interval I such that the following differential inclusion*

$$\dot{x}(t) \in F(x, t), \tag{4.4}$$

holds almost everywhere on I , where F is a set-valued function defined as follows

- *if the point (x, t) is a continuity point of the function f , then $F(x, t) \equiv \{f(x, t)\}$;*
- *if the point (x, t) is a discontinuity point of the function f , then $F(x, t)$ is the smallest convex closed set containing all the limit values of the function $f(x^*, t)$ for $(x^*, t) \notin M, x^* \rightarrow x, t = \text{const}$.*

It is easy to see that the set $F(x, t)$ defined above is upper semicontinuous in x . The set of the limit values of $f(x^*, t)$ may be replaced, so that F is also upper semicontinuous in t , by the set of the limit values of $f(x^*, t^*)$ for $(x^*, t^*) \notin M, x^* \rightarrow x, t^* \rightarrow t$. It turns

out however that this definition is equivalent to the one in Definition 4.3.1 in a wide class of equations with piecewise continuous right-hand sides [Fil88, §6].

If the function f is discontinuous on a smooth curve S given by the equation $x = \chi(t)$ such that S separates the domain into two domain D^- and D^+ , then at any point (x, t) on this curve $F(x, t)$ is simply given as

$$F(x, t) = \begin{cases} [f(x-, t), f(x+, t)] & \text{if } f(x-, t) < f(x+, t) \\ [f(x+, t), f(x-, t)] & \text{if } f(x-, t) > f(x+, t), \end{cases}$$

where $f(x-, t) = \lim_{y \in D^-, y \rightarrow x} f(y, t)$ and $f(x+, t) = \lim_{y \in D^+, y \rightarrow x} f(y, t)$.

4.3.1 Existence

Definition 4.3.2. We say that $F(x, t)$ satisfies the basic conditions on D if for every (x, t) in the domain D ,

- (i) $F(x, t)$ is a nonempty, closed and bounded, convex set; and
- (ii) the function F is upper semicontinuous in x, t .

The following existence theory applies to any differential inclusion

$$\dot{x} \in F(x, t) \tag{4.5}$$

for some set-valued function F that satisfies the basic conditions and needs not arise from f (as in Definition 4.3.1).

When dealing with the existence of solutions, one often needs to consider approximate solutions. For some set K , we denote by K^δ the closed δ -neighbourhood of K , that is,

$$K^\delta := \left\{ x : \inf_{y \in K} |y - x| \leq \delta \right\}.$$

When K reduces to a point k , the set k^δ reduces to the set $\{x : |x - k| \leq \delta\}$. Following this notation we define $F(x^\delta, t^\delta)$ to be the set of $F(x_1, t_1)$ for all $x_1 \in x^\delta, t_1 \in t^\delta$. We will use the following notion of approximate solutions.

Definition 4.3.3. We call $y(t)$ a δ -solution of the inclusion (4.4) if $y(t)$ is absolutely continuous and almost everywhere,

$$\dot{y}(t) \in F_\delta(y(t), t), \quad F_\delta(y, t) := [\text{co}F(y^\delta, t^\delta)]^\delta, \tag{4.6}$$

where $\text{co}A$ denotes the smallest convex set containing A .

The following lemma, whose proof can be found in [Fil88], will be useful in proving the existence of solutions.

Lemma 4.3.4. *Assume that $F(x, t)$ satisfies the basic conditions in D . Let $x_k(t)$ be a uniformly convergent sequence of δ_k -solutions of the inclusion (4.4) with $\delta_k \rightarrow 0$ as $k \rightarrow \infty$. Then the limit $x(t) = \lim_{k \rightarrow \infty} x_k(t)$ is a solution of this inclusion.*

Theorem 4.3.5. *Assume that F satisfies the basic conditions in D . Then for every $(x_0, t_0) \in D$ there exists a solution of the problem*

$$\dot{x} \in F(x, t), \quad x(t_0) = x_0.$$

Proof. We reproduce the constructive proof given in [Fil88] here. Assume that G is a closed and bounded set in D that contains (x_0, t_0) and let $M = \sup_G |F(x, t)| < \infty$. For $k = 1, 2, \dots$, define

$$t_{k,i} = t_0 + \frac{ai}{k}, \quad i = 0, 1, \dots, k,$$

for some small number¹ a . We construct a sequence x_k as follows

$$x_k(t_{k,0}) = x_0,$$

and on each interval $[t_{k,i}, t_{k,i+1}]$, $i = 0, 1, \dots, k-1$,

$$x_k(t) = x_k(t_{k,i}) + (t - t_{k,i})a_{k,i},$$

for some $a_{k,i} \in F(t_{k,i}, x_k(t_{k,i}))$. It follows from this construction that

$$|x_k(t) - x_0| \leq M|t - t_0|, \quad t_{k,i} < t \leq t_{k,i+1}. \quad (4.7)$$

Obviously, $x_k(t)$ is absolutely continuous. Moreover,

$$\dot{x}_k(t) = a_{k,i} \in F(t_{k,i}, x_k(t_{k,i})), \quad |t - t_{k,i}| \leq \frac{a}{k}, \quad |x_k(t) - x_{k,i}| \leq M\frac{a}{k},$$

hence, $x_k(t)$ is a δ_k -solution of the inclusion (4.5), where

$$\delta_k = \max \left\{ \frac{a}{k}, M\frac{a}{k} \right\} \rightarrow 0, \quad \text{as } k \rightarrow \infty.$$

Now thanks to (4.7) and $|\dot{x}_k(t)| \leq M$, the sequence $\{x_k(t)\}$ is uniformly bounded and equicontinuous. The Arzelà-Ascoli theorem assures that there exists a subsequence that converges to some function $x(t)$, which is also a solution of the inclusion (4.5) thanks to Lemma 4.3.4. The initial condition $x(t_0) = x_0$ is implied from (4.7). \square

Remark 4.3.6. *As in the case of solutions for Carathéodory equations, Filippov solutions can be extended to the boundary of D , see [Fil88, §7, Theorem 2].*

¹This is to make sure the we do not go out of the (open) domain D !

4.3.2 Uniqueness

Consider the equation

$$\dot{x} = f(x, t) \tag{4.8}$$

with some discontinuous function f on the domain D .

Definition 4.3.7. *We say that the right uniqueness for (4.8) holds on D if at each point $(x_0, t_0) \in D$, there exists $t_1 > t_0$ (t_1 can be ∞) such that each two solutions of (4.8) satisfying the condition $x(t_0) = x_0$ coincide on $[t_0, t_1]$.*

Theorem 4.3.8. *Let a function f in a domain D be discontinuous on a set M of measure zero. Assume that there exists an integrable function $l(t)$ such that for almost all points (x, t) and (y, t) in D we have $|f(x, t)| \leq l(t)$ and*

$$(x - y) \cdot (f(x, t) - f(y, t)) \leq l(t)|x - y|^2. \tag{4.9}$$

Then (4.8) has right uniqueness in the domain D , where solutions are understood in the sense of Filippov.

Proof. Since (4.9) holds for almost all t , then for those t one has

$$(x - y) \cdot (v - w) \leq l(t)|x - y|^2,$$

where v and w are arbitrary values in the sets V and W of the limit values of the function $f(x^*, t)$ for $x^* \rightarrow x$ and of the function $f(y^*, t)$ for $y^* \rightarrow y$, respectively. By taking the limit it is easy to see that this inequality is also true if one replaces V by $\text{co}V = F(x, t)$ and then W by $\text{co}W = F(y, t)$. Now for any two Filippov solutions $x(t)$ and $y(t)$ in the domain D , for almost all t , one has

$$\frac{1}{2} \frac{d}{dt} |x(t) - y(t)|^2 = (x(t) - y(t))(\dot{x}(t) - \dot{y}(t)) \leq l(t)|x(t) - y(t)|^2.$$

This implies that

$$\frac{d}{dt} \left(|x(t) - y(t)|^2 \exp \left(-2 \int_{t_0}^t l(s) ds \right) \right) \leq 0.$$

Therefore, the right uniqueness follows since $x(t_0) = y(t_0)$. This completes the proof. \square

Remark 4.3.9. *The sufficient condition for uniqueness (4.9) comes naturally in some applications. For instance, as we see in Chapter 2, this condition is satisfied for f being given by the entropy solution of scalar conservation laws, thanks to the famous Oleinik decay estimate (Theorem 2.5.1).*

We will now see that Filippov solutions of (4.8), whether unique or not, depend continuously not only on initial data but also on the right-hand side of the equation or inclusion in the sense of the following theorem.

Theorem 4.3.10. *Let $F(x, t)$ satisfy the basic conditions on the domain $D \ni (x_0, t_0), t_0 \in [a, b]$. Suppose that for $t \geq t_0$ all Filippov solutions of the problem*

$$\dot{x}(t) \in F(x, t), \quad x(t_0) = x_0, \quad (4.10)$$

are defined on $[a, b]$. Then for any $\varepsilon > 0$, there exists $\delta > 0$ such that for any $t_0^ \in [a, b]$, x_0^* and $F^*(x, t)$ satisfying*

$$|t_0^* - t_0| \leq \delta, \quad |x_0^* - x_0| \leq \delta, \quad F^*(x, t) \subset [\text{co}F(x^\delta, t^\delta)]^\delta, \quad (4.11)$$

and the basic conditions for every (x, t) in D , each solution of the problem

$$\dot{x}^* \in F^*(x^*, t), \quad x^*(t_0^*) = x_0^*, \quad (4.12)$$

exists on $[a, b]$ and differs from some certain solution of (4.10) by not more than ε .

It follows from this theorem that if the uniqueness holds for (4.8) then the solution depends continuously on the initial data and the right-hand side of the inclusion.

Corollary 4.3.11. *Let $F(x, t)$ satisfy the basic conditions on the domain $D \ni (x_0, t_0), t_0 \in [a, b]$. Suppose that for $t \geq t_0$ the problem (4.10) has a unique Filippov solution on $[t_0, b]$. Then for any $\varepsilon > 0$, there exists $\delta > 0$ such that for any $t_0^* \in [a, b]$, x_0^* and $F^*(x, t)$ satisfying (4.11) and the basic conditions in D each solution of the problem (4.12) on $[t_0, b]$ exists and differs from $x(t)$ by less than ε .*

Sketch of Proof of Theorem 4.3.10. By contradiction arguments. Assume that for some relatively small $\varepsilon > 0$ there exists a sequence of solutions $\{x_i(t)\}_i$ of (4.12) for which the condition (4.11) is satisfied for any sequence $\delta_i \rightarrow 0$, and the solution $x_i(t)$, defined on $[a_i, b_i]$, either does not extend to the whole $[a, b]$ or

$$\max_{a \leq t \leq b} |x_i(t) - x(t)| > \varepsilon, \quad (4.13)$$

for each solution $x(t)$ of (4.10). Now choose an appropriate subsequence $\{x_{i_k}\}$ defined on $[a_{i_k}, b_{i_k}]$ such that $(a_{i_k}, x_{i_k}(a_{i_k})) \rightarrow (\bar{a}_1, \bar{x}_1)$, $(b_{i_k}, x_{i_k}(b_{i_k})) \rightarrow (\bar{b}_2, \bar{x}_2)$. Then it can be shown that, up to a subsequence, $x_{i_k}(t)$ converges uniformly to $x(t)$ defined on $[\bar{a}_1, \bar{b}_2] \supset [a, b]$, where again $x(t)$ can be shown to be the solution of the problem (4.10). But this is not possible because of (4.13). \square

4.4 Summary and notes

This chapter briefly introduces two notions of solutions for ordinary differential equations with discontinuous right-hand sides and their well-posedness in different settings. Our presentation here is adapted from [Fil88]. We somewhat simplified the theory and reproduced some key proofs, which are either important on their own or will be useful later on in the thesis.

The theory of Carathéodory solutions is also treated thoroughly in the classical book [CL55]. Note that Filippov's theory only allows discontinuities in the derivatives but not in the state itself. When discontinuities are allowed in the state, the notion of differential inclusion is extended to the so-called *measure differential inclusion*, see [Ste11] and recent review [BT20].

Chapter 5

Particle trajectories for scalar conservation laws

Let v be the unique entropy solution of the scalar conservation law

$$\partial_t v + \partial_x f(v) = 0, \quad (5.1)$$

with initial data $v_0(\cdot) = v(\cdot, 0) \in [-M, M]$. We assume that the flux function f is Lipschitz continuous and is *genuinely nonlinear*, that is, f is either strictly convex or strictly concave¹, on $[-M, M]$.

In the equation (5.1), the variable of interest v , given as a function of position x and time t , represents the conserved quantity such as density field, velocity field, pressure, temperature and energy, or a combination of these quantities.

In this chapter, we study the conservation law (5.1) from a different point of view. We assume that the physical system that equation (5.1) describes is comprised of particles moving with the flow field. Now instead of looking at v , we study the trajectories of these particles². Let us consider a single particle starting at t_0 from a point x_0 . Denote by $z(t)$ its position at time t (and, by abuse of language, we will also denote the particle itself by z). In the general model, the velocity field of the flow is not necessarily v but a function $w(v)$ of v . Then z satisfies the following equation

$$\dot{z}(t) = w(v(z(t), t)), \quad (5.2)$$

since both sides refer to the velocity of the particle at time t . We will assume that the function w satisfies the following assumption.

¹ $f'' \neq 0$ if f is differentiable.

² In a similar context in fluid mechanics, this is often called the *Lagrangian representation* of the flow field.

Assumption 5.0.1. *The velocity field w is a bounded and continuous function of v which is non-increasing if f is strictly concave, and non-decreasing if f is strictly convex.*

This assumption is satisfied, in particular, for the traffic flow model (see Section 2.1.2) which will be studied in great detail later on. In that model v represents the density of cars, where a typical example of the velocity of the car flow w is given by $w(v) = w_{max}(1 - v/v_{max})$ where w_{max} and v_{max} denote the maximum value of car speed and car density, respectively (see the discussions in Section 5.2.1).

The equation (5.2) can be viewed as an ODE of $z(t)$. If we take the initial position x_0 of the particle into account then we have the following Cauchy initial data to (5.2),

$$z(t_0) = x_0. \quad (5.3)$$

We note however that, since v is a solution of the conservation law (5.1), the function $\alpha(x, t) := w(v(x, t))$ on the right-hand side of the ODE (5.2) might be discontinuous in both x and t , hence the Cauchy–Lipschitz theory for ODEs does not apply. Therefore more attention should be paid to define the trajectory z as a solution of (5.2) in an appropriate way. Here we use the definition of solution in the sense of Filippov [Fil88], detailed in Chapter 4.

It is worth noting that, the particle trajectories theory that we consider in this work differs from the theory of generalized characteristics developed by Dafermos [Daf77], even though both approaches use Filippov theory to deal with discontinuous right-hand side differential equations. In particular, there is a fundamental difference in the use of the term *speed* (that is, the direction field in classical ODEs). The *speed* in the generalized characteristics is either the classical characteristic speed or shock speed, while the *speed* in our approach is the speed of the flow itself. See Section 5.3 for more information and bibliographical notes.

Throughout this chapter, we assume that the flux function f is genuinely nonlinear. The approximation theory for the particle trajectories when the flux has inflection points will be considered in our future work.

5.1 Definition and Well-posedness

Definition 5.1.1. *A function $z : [t_0, T] \rightarrow \mathbb{R}$ is called a solution to (5.2) in the sense of Filippov if it is absolutely continuous on $[t_0, T]$ and it satisfies the differential inclusion*

$$\dot{z}(t) \in [w(v(z(t)\pm, t)), w(v(z(t)\mp, t))], \quad \text{for almost every } t \in [t_0, T], \quad (5.4)$$

for strictly convex and strictly concave flux function f respectively, where $v(x\pm, t)$ denote the one-sided limits of v at x .

The traces $v(x\pm, t)$ exist thanks to a classical result that $v(\cdot, t) \in BV_{\text{loc}}(\mathbb{R})$ for all $t > 0$ (even if v_0 is merely in L^∞ , see Theorem 2.5.1). Moreover, it follows from the Lax entropy condition that

$$v(x+, t) \leq v(x-, t),$$

for almost all $t > 0$ and all $x \in \mathbb{R}$, when f is strictly convex and with the reverse inequality when f is strictly concave. Hence, bear Assumption 5.0.1 in mind, the right-hand side of (5.4) makes sense. The following theorem is an easy application of Filippov theory and ensures the existence and uniqueness of $z(t)$. Similar results were already obtained in [CM03] and, in a slightly different setting, in [Leg11].

Theorem 5.1.2. *Let v be a unique entropy solution of (1.1)-(1.2) with f being strictly convex or strictly concave. Then for $(x_0, t_0) \in \mathbb{R} \times (0, \infty)$, there exists a unique absolutely continuous function $z : [t_0, \infty) \rightarrow \mathbb{R}$ satisfying (5.2)-(5.3) in the sense of Filippov.*

Proof. First consider the convex flux f . Denote

$$V(z, t) := [w(v(z+, t)), w(v(z-, t))]. \quad (5.5)$$

The existence of z follows from the Filippov existence theory (see Chapter 4, Theorem 4.3.5).

For the uniqueness, in view of Theorem 4.3.8 and thanks to the Lipschitz property of w we aim at the following estimate

$$(x - y)(v(x, t) - v(y, t)) \leq l(t)(x - y)^2, \quad (5.6)$$

for some $l \in L^1([t_0, T])$. But this follows from Oleinik's decay estimate (see Theorem 2.5.1), that is, there exists $C > 0$ such that

$$v(x, t) - v(y, t) \leq \frac{x - y}{Ct}, \quad \text{for } x > y, t > 0. \quad (5.7)$$

Indeed, if $x > y$, the condition (5.6) is satisfied by multiplying both sides of (5.7) by $x - y > 0$, with $l(t) = 1/(Ct) \in L^1([t_0, T])$ since $t_0 > 0$. The case $x < y$ is easily implied from (5.7) by swapping the role of x and y . Now if f is concave, by sending $x \mapsto -x$ we obtain a conservation law with convex flux $-f$, the result then follows. \square

Remark 5.1.3. *By using the technique in the proof of Theorem 5.1.2, one can prove the following stability-like estimate*

$$|x(t) - y(t)|^2 \leq |x_0 - y_0|^2 \left(\frac{t}{t_0}\right)^C, \quad (5.8)$$

for x, y being the Filippov solution of (5.2) with respect to initial positions x_0, y_0 , respectively, where C depends only on f and the Lipschitz constant of w . However, this proof cannot be used to derive the continuity or stability of the particle trajectories with respect to changes in the velocity field v itself as a result of the changes in the initial field which we study here. This is because now the two trajectories do not solve the same ODE anymore.

In the rest of this section, for convenience, we will assume that f is strictly convex and notice that the analysis applies to the case of concave flux functions as well. The velocity field $w(v)$ is then, by Assumption 5.0.1, a non-decreasing function of v .

We now turn to investigate the continuity properties of the particle trajectories with respect to changes in the solution field v . These changes may be a result of perturbations in the initial field or from any kind of approximations of the forward model incurred in a computational process. In particular, we consider two of the most popular approximations used in scalar conservation laws, the front tracking approximation and the vanishing viscosity approximation. We prove that an approximate particle trajectory arising from a small perturbation of the initial field (including the front tracking method) converges to the unique particle trajectory of the original equation. We prove a similar stability property when the approximation is a result of the vanishing viscosity method, however, with some restriction on the shock speed.

In this context, it is natural to consider some appropriate notion of approximate solutions. Following Filippov, for some set K , we denote by K^δ the following set

$$K^\delta := \left\{ x : \inf_{y \in K} |y - x| \leq \delta \right\},$$

that is, a closed δ -neighbourhood of K . For the convenience, we recall the following definition of δ -solution which has already been introduced in Chapter 4 (Definition 4.3.3) with now a specific term on the right-hand side.

Definition 5.1.4. *We call $y(t)$ a δ -solution of the inclusion (5.4) if $y(t)$ is absolutely continuous and we have, almost everywhere,*

$$\dot{y}(t) \in V(y(t)^\delta, t)^\delta, \tag{5.9}$$

where $V(y, t) := [w(v(y+, t)), w(v(y-, t))]$.

The following lemma is an immediate consequence of Lemma 4.3.4, and will be of later use.

Lemma 5.1.5. *Let $x_k(t)$ be a uniformly convergent sequence of δ_k -solutions of the inclusion (5.4) with $\delta_k \rightarrow 0$ as $k \rightarrow \infty$. Then the limit $x(t) = \lim_{k \rightarrow \infty} x_k(t)$ is a solution of this inclusion.*

We now consider a sequence of exact solutions³ (or front tracking approximations) v^N that converges to the solution v in L^1 . The Filippov solution to (5.2) with v replaced by v^N is denoted by z^N , that is, z^N satisfies

$$\dot{z}^N(t) \in [w(v^N(z^N(t)+, t)), w(v^N(z^N(t)-, t))], \quad (5.10)$$

for almost every t . We assign $z^N(t_0) = z(t_0) = x_0$ and study the behavior of z^N when N is large. The following result establishes the uniform convergence for z^N .

Theorem 5.1.6. *Let v^N be a sequence of exact solutions (or front tracking approximations) of bounded variation converging in L^1 to the entropy solution v of (5.1). Let z^N be defined as (5.10). Then z^N converges to z uniformly on $[t_0, T]$ for every $T > t_0 > 0$, as $N \rightarrow \infty$, where z is the Filippov solution of (5.2).*

The proof of this theorem, which can be found below, uses structural stability properties of the entropy solution established by Bressan-LeFloch [BL99]. Some of the proof techniques are based on [DMG14]. However, in [DMG14] the authors studied Caratheodory solutions. We adapt the argument to obtain convergence results for Filippov solutions. The following corollary is an immediate consequence of the above theorem.

Corollary 5.1.7. *The particle trajectory z defines a map from the space of initial fields into C^0 ,*

$$L^1 \ni v_0 \mapsto z \in C^0,$$

which is a continuous map.

Proof of Theorem 5.1.6. We first consider the case where $\{v^N\}$ is a front tracking approximation of v . Fix $T > 0$. According to (5.2) and the boundedness of v^N , and since w is bounded and continuous by Assumption 5.0.1, there exists a constant C such that, for any $t \in [t_0, T]$,

$$|\dot{z}^N(t)| \leq C.$$

This implies that $\{z^N\}$ is uniformly bounded and also

$$|z^N(t) - z^N(s)| \leq C|t - s|,$$

which means $\{z^N\}$ is an equicontinuous sequence. By Arzelà-Ascoli theorem, there exists a subsequence, still denoted by z^N , such that

$$z^N(\cdot) \rightarrow z(\cdot) \quad \text{in } C^0([t_0, T]), \quad (5.11)$$

³i.e., solutions obtained from a sequence of initial data.

for some Lipschitz continuous function z . We claim that z solves (5.2)-(5.3) in the sense of Filippov, that is, for a.e.- t ,

$$\dot{z}(t) \in V(z(t), t) = [w(v(z(t)+, t)), w(v(z(t)-, t))].$$

Indeed, from the definition of z^N we have, for a.e.- t ,

$$\dot{z}^N(t) \in V^N(z^N(t), t) := [w(v^N(z^N(t)+, t)), w(v^N(z^N(t)-, t))].$$

We prove that, for a.e. $t \in [t_0, T]$,

$$v^N(z^N(t)+, t) \rightarrow v^+(t) := v(z(t)+, t), \quad \text{as } N \rightarrow \infty. \quad (5.12)$$

Indeed, by extracting a further subsequence if needed, v^N converges a.e. to v , there exists a sequence $\bar{z}^N \geq z^N(t)+$ such that $\bar{z}^N \rightarrow z(t)+$ and $v^N(\bar{z}^N, t) \rightarrow v^+(t)$.

For a.e. t , the point $(z(t), t)$ is either a point of continuity of v or it belongs to a discontinuity curve (represented by $z(t)$). For any fixed $\varepsilon_0 > 0$, assume that $TV(v(\cdot, t) : (z(t) - \delta, z(t) + \delta)) \leq \varepsilon_0$, for some $\delta > 0$. We will show that

$$TV(v^N(\cdot, t) : (z(t) - \delta, z(t) + \delta)) \leq 2\varepsilon_0, \quad (5.13)$$

for large enough N . Indeed, it is obvious if $(z(t), t)$ is a continuity point of v . If $(z(t), t)$ belongs to a discontinuity curve of v with strength $|v(z(t)+, t) - v(z(t)-, t)|$ small but positive, [BL99, Lemma 15. (2)] implies that the shock strength of v^N converges to the shock strength of v , leading to (5.13). Therefore

$$|v^N(z^N(t)+, t) - v^+(t)| \leq |v^N(z^N(t)+, t) - v^N(\bar{z}^N, t)| + |v^N(\bar{z}^N, t) - v^+(t)| \leq 3\varepsilon_0,$$

for large enough N . Since ε_0 can be chosen to be arbitrary small, it implies (5.12).

If $(z(t), t)$ belongs to a discontinuity curve of v with strength $|v(z(t)+, t) - v(z(t)-, t)| \geq \varepsilon_0$, then again, since the shock strength of v^N converges to the shock strength of v ,

$$|v^N(z^N(t)+, t) - v^N(z^N(t)-, t)| \geq \frac{\varepsilon_0}{2},$$

for N large enough. We now proceed as in [BL99, Proof of Lemma 15]) to prove that, for each $\varepsilon > 0$, there exists $\delta > 0$ such that for all n large enough we get

$$|v^N(x, s) - v^N(z^N(t)+, t)| < \varepsilon, \quad \text{for } |s - t| \leq \delta, |x - z(t)| \leq \delta, x > z^N(s). \quad (5.14)$$

Indeed, if (5.14) does not hold, there will be $\varepsilon > 0$ and sequences $t_N \rightarrow t$, $\delta_N \rightarrow 0$ such that

$$TV(v^N(\cdot, t_N) : (z^N(t_N), z^N(t_N) + \delta_N)) \geq \varepsilon.$$

By the genuine nonlinearity of f , this would generate a uniformly positive amount of interactions in an arbitrarily small neighbourhood of $(z(t), t)$, against the assumption. Hence (5.14) holds and therefore, for N large enough,

$$|v^N(z^N(t)+, t) - v^+(t)| \leq |v^N(z^N(t)+, t) - v^N(\bar{z}^N, t)| + |v^N(\bar{z}^N, t) - v^+(t)| \leq 2\varepsilon,$$

which proves (5.12).

Similarly, we will have, for a.e. $t \in [t_0, T]$,

$$v^N(z^N(t)-, t) \rightarrow v(z(t)-, t) \quad \text{as } N \rightarrow \infty. \quad (5.15)$$

From (5.12) and (5.15), for N big enough, there exists a sequence $\delta_N \rightarrow 0$ as $N \rightarrow \infty$ such that a.e. $t \in [t_0, T]$,

$$\dot{z}^N \in [V((z(t)), t)]^{\delta_N},$$

In other words, z^N is a δ_N -solution of

$$\dot{z}(t) \in V(z(t), t). \quad (5.16)$$

Hence, thanks to (5.11), we have a uniformly convergent sequence of δ_N -solution $z^N(t)$ of the inclusion (5.16). As a result, $z(t)$ is also a solution of this inclusion, thanks to Lemma 5.1.5.

Finally, since the particle trajectory is unique (Theorem 5.1.2), the whole sequence z^N must converge to z . Indeed, if there exists a subsequence z^{N_i} of z^N such that

$$z^{N_i} \rightarrow y \neq z, \quad (\text{strongly}) \text{ uniformly in } [t_0, T],$$

then, by extracting further subsequences if needed, the same argument as above shows that y also satisfies (5.2)-(5.3) in the sense of Filippov. This contradicts the conclusion of Theorem 5.1.2.

The proof is complete for the case of front tracking approximations. For the case where $\{v^N\}$ is a sequence of exact solutions, one can first approximate each v^N with a sequence of front tracking approximations $v^{N,n}$, and then use diagonalisation arguments by working with a suitable subsequence $v^{N,n(N)}$. \square

We now move on to investigate the approximate particle trajectory arising from the vanishing viscosity approximation. Consider the parabolic equation

$$v_t + f(v)_x = \varepsilon v_{xx}, \quad (5.17)$$

where ϵ is some small positive number. In many cases, the “artificial” diffusive term added on the right-hand side may be devoid of any physical reasoning but just for the sake of analytical or computational convenience.

Since (5.17) is a parabolic equation, the Cauchy problem (5.17) coupled with some bounded initial value v_0 always provides a unique smooth solution v^ϵ . Now consider the trajectory z^ϵ of a particle starting from x_0 and moving along the flow. Assume that the speed of the flow w is a smooth function (of v^ϵ). Then it follows that

$$\dot{z}^\epsilon(t) = w(v^\epsilon(z^\epsilon(t), t)), \quad z^\epsilon(t_0) = x_0. \quad (5.18)$$

We investigate the behaviour of z^ϵ when ϵ is small and compare it with the trajectory z , given by (in the sense of Filippov)

$$\dot{z}(t) = w(v(z(t), t)), \quad z(t_0) = x_0, \quad (5.19)$$

which has been studied in the previous chapter. Our aim is to establish a convergence result of z^ϵ to z as ϵ goes to 0. We make the following assumption on the shock speed of the original equation (5.1).

Assumption 5.1.8. *The shock speed is always smaller or greater than the speeds of the left and right flows, that is*

$$s < \min\{w(v_l), w(v_r)\} \quad \text{or} \quad s > \max\{w(v_l), w(v_r)\},$$

where s is the shock speed,

$$s = \frac{f(v_l) - f(v_r)}{v_l - v_r},$$

with v_l and v_r the left and right limits of v at the shock, and the flow velocity w satisfy Assumption 5.0.1.

The restriction on the shock speed is fulfilled in some applications such as traffic flow, see Lemma 5.2.2 and later sections for more information.

The following theorem provides a convergence result for z^ϵ .

Theorem 5.1.9. *Assume that Assumption 5.1.8 holds at every shock curve of the entropy solution v to (5.1) with initial data $v_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. For each $\epsilon > 0$, let z^ϵ be the solution of (5.18) where v^ϵ is the solution of the viscous scalar conservation law (5.17) with $v^\epsilon(\cdot, 0) = v_0$. Then $z^\epsilon(\cdot)$ converges strongly to the Filippov solution $z(\cdot)$ of (5.19).*

Proof. We first observe that, since $|\dot{z}^\epsilon| \leq C$ for some $C > 0$ due to the smoothness of w and the boundedness of v^ϵ , the Arzelà-Ascoli theorem will ensure that there exists some

absolutely continuous function $z : [t_0, \infty) \rightarrow \mathbb{R}$ such that, up to a subsequence,

$$z^\epsilon(\cdot) \rightarrow z(\cdot), \quad \text{uniformly in } L^\infty[t_0, T], \quad (5.20)$$

for every $T > t_0$.

We prove that z is a Filippov solution to (5.19). By Theorem 2.4.7, as $\epsilon \rightarrow 0$, we have that

$$v^\epsilon(x, t) \rightarrow v(x, t) \quad \text{boundedly almost everywhere on } \mathbb{R} \times [t_0, \infty). \quad (5.21)$$

Now thanks to Assumption 5.1.8, for almost every $t \in [t_0, \infty)$, the point $(z(t), t)$ is a continuity point of v . This, together with (5.21), ensures that

$$v^\epsilon(z(t), t) \rightarrow v(z(t), t) \quad \text{boundedly almost everywhere on } [t_0, \infty). \quad (5.22)$$

Thanks to Lemma 2.4.8,

$$\int_{\mathbb{R}} |v^\epsilon(x + y^\epsilon, t) - v^\epsilon(x, t)| dx \leq w(|y^\epsilon|) \rightarrow 0,$$

as $\epsilon \rightarrow 0$, uniformly for $t \in [t_0, T]$. Then up to a subsequence

$$|v^\epsilon(x + y^\epsilon, t) - v^\epsilon(x, t)| \rightarrow 0,$$

as $\epsilon \rightarrow 0$, almost everywhere on $\mathbb{R} \times [t_0, T]$. Letting $x = z(t)$ where $(z(t), t)$ is a point of continuity of v and

$$y^\epsilon = z^\epsilon(t) - z(t),$$

we have, as $\epsilon \rightarrow 0$,

$$|v^\epsilon(z^\epsilon(t), t) - v^\epsilon(z(t), t)| \rightarrow 0. \quad (5.23)$$

Finally by writing

$$|v^\epsilon(z^\epsilon(t), t) - v(z(t), t)| \leq |v^\epsilon(z^\epsilon(t), t) - v^\epsilon(z(t), t)| + |v^\epsilon(z(t), t) - v(z(t), t)|, \quad (5.24)$$

and using the estimates (5.22) and (5.23), we conclude that

$$v^\epsilon(z^\epsilon(t), t) \rightarrow v(z(t), t), \quad \text{almost everywhere in } [t_0, T], \quad (5.25)$$

for every $T > t_0$. This ensures that, for every $\epsilon > 0$ there exists $\delta_\epsilon > 0$, such that $\delta_\epsilon \rightarrow 0$ as $\epsilon \rightarrow 0$, and

$$\dot{z}^\epsilon(t) \in [v(z(t), t) - \delta_\epsilon, v(z(t), t) + \delta_\epsilon], \quad (5.26)$$

almost everywhere. Thanks to (5.21) and Lemma 4.3.4, z is a Filippov solution to (5.19). Due to the uniqueness of z , the convergence (5.20) applies to the whole sequence. This completes the proof. \square

Remark 5.1.10. *We will see in the next section that for the traffic flow described by equation (5.27), Assumption 5.1.8 is fulfilled for instance if the road is significantly crowded (see Lemma 5.2.2). When the speed of particles $\alpha(z(t), t) := w(v(z(t), t))$ has such a property, then $\alpha(x, t)$ for $x \in \mathbb{R}$ and $t > t_0$ satisfies the Carathéodory conditions.*

5.2 Stability properties of the particle trajectories

5.2.1 The case of traffic flow

We shall focus our attention in this section on the traffic flow. This is one of the most popular applications of the scalar conservation law in one dimension. We switch to the traditional notation, ρ , to denote the quantity of interest in this case, which is the density of cars at given space and time. The velocity of the flow is still denoted by w . The LWR model for traffic flow is derived under the general assumptions that the vehicle length is negligible, the road is flat and has only one lane and overtaking is not allowed (see Section 2.1.2). The Cauchy problem now reads

$$\partial_t \rho + \partial_x(\rho w(\rho)) = 0, \quad (5.27)$$

$$\rho(x, 0) = \rho_0(x). \quad (5.28)$$

Denote by ρ_{max} and w_{max} the maximum density and maximum speed of the traffic. We make the following assumption on the car speed and the flux.

Assumption 5.2.1. *The car speed $w : [0, \rho_{max}] \rightarrow [0, w_{max}]$ is a Lipschitz continuous and strictly decreasing function of ρ with $w(\rho_{max}) = 0$. The flux $f(\rho) = \rho w(\rho)$ is a strictly concave function.*

This assumption is reasonable as we expect the car to go at its maximum speed when there are only a few cars on the road, and to slow down when the car density increases. A typical example for w is that w depends linearly on ρ ,

$$w(\rho) = w_{max} \left(1 - \frac{\rho}{\rho_{max}} \right). \quad (5.29)$$

By scaling we can assume that $w_{max} = 1, \rho_{max} = 1$. Note that if we replace ρ by $1 - w(\rho)$ in (5.27) and denote $v(x, t) := w(\rho(x, t))$ then v follows the following conservation law

$$v_t + [v(v - 1)]_x = 0, \quad (5.30)$$

with now a strictly convex flux $\tilde{f}(v) = v(v - 1)$. We can see that working with (5.30) is not less general than working with (5.27) when the car speed is given as (5.29). Note also

that, by setting

$$\rho = 1 - v = \frac{1 - \tilde{v}}{2}, \quad (5.31)$$

we obtain the familiar Burgers equation

$$\tilde{v}_t + \left(\frac{\tilde{v}^2}{2} \right)_x = 0, \quad (5.32)$$

for \tilde{v} , which is not the velocity here but a function of it as described in (5.31).

For traffic flow where trapping in the queue is not allowed, then shock speed satisfies Assumption 5.1.8, as showed by the following lemma. It follows easily from the Rankine–Hugoniot jump condition, yet will be useful in establishing the stability estimates.

Lemma 5.2.2. *Assume that the car density before each time a shock happens is always positive, then shocks travel more slowly than the flows right before and after the shocks. Consequently, trajectories do not lie on shock curves except at countably many points.*

Proof. From the Rankine–Hugoniot condition, the speed of a shock curve at the point with left limit ρ_l and right limit ρ_r is given as

$$\begin{aligned} s &= \frac{\rho_l w(\rho_l) - \rho_r w(\rho_r)}{\rho_l - \rho_r} \\ &= \rho_l \cdot \frac{w(\rho_l) - w(\rho_r)}{\rho_l - \rho_r} + w(\rho_r) \\ &= \rho_r \cdot \frac{w(\rho_l) - w(\rho_r)}{\rho_l - \rho_r} + w(\rho_l) \\ &< \min\{w(\rho_l), w(\rho_r)\}, \end{aligned}$$

since $\rho_l, \rho_r > 0$, and w is a strictly decreasing function of ρ . \square

This property applies, in particularly, for traffic flow with significant cars present at any point on the road. Whenever this is the case, a given car generally interacts with (countably many) shocks that are coming from the opposite direction toward the car. After hitting a shock, the car continues to travel with its own speed that is different with the shock speed, hence, never lies on any shock curve for a positive period of time.

We now consider the trajectory z of a car passing through some point x_0 at $t = t_0 > 0$ and travelling at speed w . From the previous section, z is the unique Filippov solution to

$$\dot{z}(t) = w(\rho(z(t), t)), \quad z(t_0) = x_0. \quad (5.33)$$

The aim is to obtain a suitable convergence rate for some approximation of z with respect to changes in upstream density ρ_0 , for fixed flux f . In the following the Lipschitz constant of w is denoted by

$$L_w := \sup_{u \neq v} \left| \frac{w(u) - w(v)}{u - v} \right|.$$

Theorem 5.2.3. *Let $0 < m_\rho < 1$ and $T > 0$ be given. Suppose that ρ and $\bar{\rho}$ are solutions of (5.27)-(5.28) with initial data $\rho_0, \bar{\rho}_0 \in L^1 \cap BV(\mathbb{R}; [m_\rho, 1])$ respectively, satisfying*

$$\|\rho_0 - \bar{\rho}_0\|_{L^1 \cap L^\infty([-2L_w T, 3L_w T])} \leq \varepsilon. \quad (5.34)$$

Let z and \bar{z} be the corresponding particle trajectories with the same initial position x_0 (so they solve (5.33) in the sense Filippov, for ρ and $\bar{\rho}$, respectively). Then

$$\|z - \bar{z}\|_{L^\infty([t_0, T])} \leq C_\rho \varepsilon^{1/2},$$

with the constant C_ρ is defined as

$$C_\rho = 1 + (T - t_0) \left(1 + \frac{2}{m_\rho}\right) L_w + \frac{1}{m_\rho} (\|\rho_0\|_{BV} + \|\bar{\rho}_0\|_{BV}).$$

This result is, although not surprising, the first stability result for the particle trajectory with a specific rate of convergence given in the context of scalar conservation laws, even in traffic flow. Note that no L^∞ stabilities (in fact, no L^p stabilities with $p \neq 1$) are expected for the entropy solution ρ itself, except when it is classical.

Our strategy is to use the front tracking method to discretise the initial fields and construct the approximate trajectories, then use the convergence theorem, Theorem 5.1.6, as a glueing tool to connect separate paths together. Thanks to the nature of the front tracking method, the proof is constructive and may be used as a numerical tool. In the following, we define

$$\Lambda := \{w : [0, 1] \rightarrow \mathbb{R}, w \text{ is strictly decreasing, } w(1) = 0, \text{ and } \|w\|_{\text{Lip}} < \infty\}, \quad (5.35)$$

where $\|\cdot\|_{\text{Lip}}$ is defined as (2.31). We will also refer to “cars” as “particles”.

Proof of Theorem 5.2.3. For a given $N \in \mathbb{N}$, we start with constructing two simple functions ρ_0^N and $\bar{\rho}_0^N$ approximating ρ_0 and $\bar{\rho}_0$ as follows. We let

$$\begin{aligned} A_{N,j} &:= \left\{x \in \mathbb{R} : \rho_0(x) \in \left[\frac{j-1}{2^N}, \frac{j}{2^N}\right)\right\} \quad \text{and} \\ \bar{A}_{N,j} &:= \left\{x \in \mathbb{R} : \bar{\rho}_0(x) \in \left[\frac{j-1}{2^N}, \frac{j}{2^N}\right)\right\} \end{aligned}$$

for $j = 1, \dots, 2^N$ and define

$$\rho_0^N = \sum_{j=1}^{2^N} \frac{j-1}{2^N} \mathbb{1}_{A_{N,j}} \quad \text{and} \quad \bar{\rho}_0^N = \sum_{j=1}^{2^N} \frac{j-1}{2^N} \mathbb{1}_{\bar{A}_{N,j}}.$$

We choose $N > \hat{N}$ with \hat{N} large enough so that $1/2^N < \varepsilon < 1$ and

$$\|\rho_0^N - \bar{\rho}_0^N\|_{L^1 \cap L^\infty([-2L_w T, 3L_w T])} \leq 2\varepsilon.$$

We use the front tracking method to construct approximations of ρ , denoted by ρ^N and $\bar{\rho}^N$ corresponding to ρ_0^N and $\bar{\rho}_0^N$. That is, we consider $f(\rho) = \rho w(\rho)$ to be approximated by a piecewise linear function f^N whose graph is inscribed by graph of f and $f(j/2^N) = f^N(j/2^N)$ for all $j \in \{1, \dots, 2^N\}$. Then, ρ^N and $\bar{\rho}^N$ are solutions of

$$\zeta_t + (f^N(\zeta))_x = 0 \quad (5.36)$$

with $\zeta(x, 0) = \rho_0^N(x)$ and $\zeta(x, 0) = \bar{\rho}_0^N(x)$ respectively. We then define z^N and \bar{z}^N as

$$\dot{z}^N = w(\rho^N(z^N, t)) \quad \text{and} \quad \dot{\bar{z}}^N = w(\bar{\rho}^N(\bar{z}^N, t)),$$

with $z^N(t_0) = \bar{z}^N(t_0) = x_0$. We first find an upper bound for $\|z^N - \bar{z}^N\|_{L^\infty(t_0, T)}$.

STEP 1 (Convergence rate for front tracking approximations). In this part for notational convenience we drop the superscript N in ρ^N , $\bar{\rho}^N$, z^N and \bar{z}^N . Consider $0 < \beta < 1$ and $t_1 > t_0$ to be the smallest time at which

$$\lim_{t \rightarrow t_1^+} |\rho(z(t), t) - \bar{\rho}(\bar{z}(t), t)| > \varepsilon^\beta.$$

This implies that at t_1 , one of the particles, which without loss of generality we assume to be z , coincides with a shock point. We call this a_1 and denote the closest next shock point at this instance to \bar{z} by \bar{a}_1 . We then let the time instance where \bar{z} hits \bar{a}_1 to be $\tau_1 > t_1$. Define

$$\begin{aligned} \iota_1 &= \lim_{t \rightarrow t_1^+} \rho(z(t), t) - \lim_{t \rightarrow t_1^-} \rho(z(t), t), \quad \text{and} \\ \bar{\iota}_1 &= \lim_{t \rightarrow t_1^+} \bar{\rho}(\bar{z}(t), t) - \lim_{t \rightarrow t_1^-} \bar{\rho}(\bar{z}(t), t). \end{aligned}$$

i) Let $\iota_1 > 0$ (that is when we have an up-jump at the shock). We consider the Riemann problem around a_1 and \bar{a}_1 over the time interval (t_1, t) with $t > \tau_1$ and less than the instant when the next shock is hit by one of the particles. By Lemma 5.2.4, noting that $a_1 = z(t_1)$ and $\bar{w} = w$, we have

$$\begin{aligned} |\bar{z}(t) - z(t)| &\leq |w(\bar{\rho}(\bar{z}(t), t)) - w(\rho(z(t), t))|(t - t_1) \\ &\quad + \frac{\iota_1}{m_\rho} |\bar{a}_1 - a_1| + \frac{\lim_{t \rightarrow t_1^-} \rho(z(t), t)}{\lim_{t \rightarrow t_1^+} \rho(z(t), t)} |\bar{z}(t_1) - z(t_1)|. \end{aligned} \quad (5.37)$$

We then note that $\lim_{t \rightarrow t_1^-} \rho(z(t), t) < \lim_{t \rightarrow t_1^+} \rho(z(t), t)$ (in the up-jump case), and since

$$\sup_{t_0 < t < t_1} |\bar{\rho}(\bar{z}(t), t) - \rho(z(t), t)| \leq \varepsilon^\beta$$

we have

$$|\bar{z}(t_1) - z(t_1)| \leq (t_1 - t_0)L_w \varepsilon^\beta.$$

Furthermore by [Daf16, Theorem 6.2.3],

$$\|\rho(\cdot, t_1) - \bar{\rho}(\cdot, t_1)\|_{L^1(0, L_w t_1)} \leq \|\rho_0^N - \bar{\rho}_0^N\|_{L^1(-2L_w T, 3L_w T)} \leq 2\varepsilon.$$

Indeed in the case considered here where ρ and $\bar{\rho}$ are bounded by 1, constant s in equation (6.2.22) of [Daf16] is bounded by $2L_w$. Hence

$$|a_1 - \bar{a}_1| \varepsilon^\beta \leq \|\rho(\cdot, t_1) - \bar{\rho}(\cdot, t_1)\|_{L^1(0, L_w t_1)} \leq 2\varepsilon$$

implying that

$$|a_1 - \bar{a}_1| \leq \varepsilon^{1-\beta}.$$

We therefore conclude that

$$\begin{aligned} |\bar{z}(t) - z(t)| &\leq (t - t_1)L_w \varepsilon^\beta + \frac{\iota_1}{m_\rho} \varepsilon^{1-\beta} + (t_1 - t_0)L_w \varepsilon^\beta \\ &= (t - t_0)L_w \varepsilon^\beta + \frac{\iota_1}{m_\rho} \varepsilon^{1-\beta}. \end{aligned} \quad (5.38)$$

The above estimate is valid for $t < t_2$ where t_2 is the next time instant at which

$$\lim_{t \rightarrow t_2^+} |\rho(z(t), t) - \bar{\rho}(\bar{z}(t), t)| > \varepsilon^\beta.$$

We also note for the first term in the right-hand side of (5.38) we have assumed that over (τ_1, t) , ρ and $\bar{\rho}$ are within distance ε^β . If this is not the case \bar{a} can be replaced at the start of the argument by the next shock after which ρ and $\bar{\rho}$ are within distance ε^β . Then a similar argument considering $\lim_{t \rightarrow t_1^-} \bar{\rho}(\bar{z}(t), t)$ as the value for $\bar{\rho}$ over (t_1, τ_1) gives the same estimate.

To derive (5.38) we supposed that z hits the shock first, that is $a_1 = z(t_1)$. If we have instead $\bar{z}(t_1) = \bar{a}_1$, then the only change in (5.38) would be the replacement of ι_1 by $\bar{\iota}_1$. Hence, in the situation of an up-jump, regardless of which particle hits the shock at t_1 , we obtain for $\tau_1 < t < t_2$,

$$|\bar{z}(t) - z(t)| \leq (t - t_0)L_w \varepsilon^\beta + \frac{\max\{\iota_1, \bar{\iota}_1\}}{m_\rho} \varepsilon^{1-\beta}. \quad (5.39)$$

ii) If $\iota < 0$, that is when we have a down-jump, it has to be of size $2^{-N} < \varepsilon$. Indeed, for a front tracking solution, all down-jumps after the initial time turn into fans of small shocks, and we note that since $\|\rho_0 - \bar{\rho}_0\|_{L^\infty} < 2\varepsilon$, for small enough ε , $t_1 > 0$. Hence by Lemma 5.2.4 we have for $\tau_1 < t < t_2$

$$\begin{aligned} |\bar{z}(t) - z(t)| &\leq |w(\bar{\rho}(\bar{z}(t), t)) - w(\rho(z(t), t))|(t - t_1) \\ &\quad + \frac{\varepsilon}{m_\rho} \max\{a_1 - z(t_1), \bar{a}_1 - \bar{z}(t_1)\} + |\bar{z}(t_1) - z(t_1)| \\ &\leq (t - t_0)L_w \varepsilon^\beta + \frac{\varepsilon}{m_\rho} \max\{a_1 - z(t_1), \bar{a}_1 - \bar{z}(t_1)\}, \end{aligned} \quad (5.40)$$

where we have used the same argument as part (i) to bound $|\bar{z}(t_1) - z(t_1)|$.

As discussed above, estimates (5.39) and (5.40) remain valid upto t_2 . Then at t_2 at least one of the particles is at a shock point and the other one is about to hit one, we denote the position of these shocks at t_2 for z and \bar{z} , by a_2 and \bar{a}_2 respectively. Defining τ_2 , ι_2 and $\bar{\iota}_2$ similar to above we obtain

$$\begin{aligned} |\bar{z}(t) - z(t)| &\leq (t - t_0)L_w\varepsilon^\beta + \frac{\varepsilon^{1-\beta}}{m_\rho} \sum_{j=1}^2 \max\{\iota_j, \bar{\iota}_j, 0\} \\ &\quad + \frac{\varepsilon}{m_\rho} \sum_{j=1}^2 \max\{a_j - z(t_j), \bar{a}_j - \bar{z}(t_j)\} \end{aligned}$$

for $t \leq t_3$, with t_3 the next instant after which ρ and $\bar{\rho}$ are again more than ε^β apart. Continuing in this fashion until final time T and noting that for a given ε , there are finite k number of such shocks (k can depend on ε) [BL99], in the interval $(0, L)$ (with $L < T$), we obtain

$$\begin{aligned} |\bar{z}(T) - z(T)| &\leq (T - t_0)L_w\varepsilon^\beta + \frac{\varepsilon^{1-\beta}}{m_\rho} \sum_{j=1}^k \max\{\iota_j, \bar{\iota}_j, 0\} \\ &\quad + \frac{\varepsilon}{m_\rho} \sum_{j=1}^k \max\{a_j - z(t_j), \bar{a}_j - \bar{z}(t_j)\}. \end{aligned}$$

By [Daf16, Theorem 6.2.6] and noting that by Lemma 5.2.2 z and \bar{z} do not encounter any of the initial shocks more than once, we have

$$\sum_{j=1}^k \max\{\iota_j, \bar{\iota}_j, 0\} \leq \sum_{j=1}^k \max\{\iota_j + \bar{\iota}_j, 0\} \leq \|\rho_0\|_{BV} + \|\bar{\rho}_0\|_{BV}.$$

Moreover, by definition of t_j , a_j and \bar{a}_j

$$\sum_{j=1}^k \max\{a_j - z(t_j), \bar{a}_j - \bar{z}(t_j)\} \leq 2L_w(T - t_0).$$

We hence conclude that

$$\begin{aligned} \|z - \bar{z}\|_{L^\infty(t_0, T)} &= |\bar{z}(T) - z(T)| \\ &\leq (T - t_0)L_w\varepsilon^\beta + \frac{\varepsilon^{1-\beta}}{m_\rho} (\|\rho_0\|_{BV} + \|\bar{\rho}_0\|_{BV}) + \frac{2L_w(T - t_0)\varepsilon}{m_\rho} \end{aligned}$$

as z and \bar{z} are increasing functions of t . The best rate is then achieved when $\beta = 1 - \beta$, that is $\beta = 1/2$. We hence have, after reinstating superscript N ,

$$\|z^N - \bar{z}^N\|_{L^\infty(t_0, T)} \leq C_0\sqrt{\varepsilon},$$

with $C_0 = (T - t_0)(1 + \frac{2}{m_\rho})L_w + \frac{1}{m_\rho}(\|\rho_0\|_{BV} + \|\bar{\rho}_0\|_{BV})$.

STEP 2 (The general case). Now for any $\rho_0, \bar{\rho}_0 \in L^1 \cap BV(\mathbb{R}; [m_\rho, 1])$ such that

$$\|\rho_0 - \bar{\rho}_0\|_{L^1([-2L_w T, 3L_w T])} \leq \varepsilon,$$

let $\bar{\rho}^N$ and ρ^N be the the front tracking approximations of ρ and $\bar{\rho}$, with \bar{z}^N and z^N the corresponding trajectories all constructed as in STEP 1. We have just proved that,

$$\|\bar{z}^N(t) - z^N\|_{L^\infty(t_0, T)} \leq C_0 \sqrt{\varepsilon}. \quad (5.41)$$

Thanks to Theorem 5.1.6, as $N \rightarrow \infty$ we have

$$\|\bar{z}^N - \bar{z}\|_{L^\infty([t_0, T])} \rightarrow 0,$$

$$\|z^N - z\|_{L^\infty([t_0, T])} \rightarrow 0,$$

for any $t_0 > 0$. Hence, there exist constants N_1, N_2 big enough such that

$$\|\bar{z}^N - \bar{z}\|_{L^\infty([t_0, T])} \leq \frac{\sqrt{\varepsilon}}{2}, \quad \text{for all } N > N_1,$$

$$\|z^N - z\|_{L^\infty([t_0, T])} \leq \frac{\sqrt{\varepsilon}}{2}, \quad \text{for all } N > N_2.$$

Choose $N > \max\{\hat{N}, N_1, N_2\}$, together with (5.41) we obtain

$$\begin{aligned} \|\bar{z} - z\|_{L^\infty([t_0, T])} &\leq \|\bar{z}^N - \bar{z}\|_{L^\infty([t_0, T])} + \|z^N - z\|_{L^\infty([t_0, T])} + \|\bar{z}^N - z^N\|_{L^\infty([t_0, T])} \\ &\leq (1 + C_0)\sqrt{\varepsilon}. \end{aligned}$$

This completes the proof. □

Lemma 5.2.4 (Estimates for Riemann problems). *Let $w, \bar{w} \in \Lambda$, and $\rho_l < \rho_r, \bar{\rho}_l < \bar{\rho}_r$. Let ζ and $\bar{\zeta}$ satisfy*

$$\begin{aligned} \zeta_t + (f(\zeta))_x &= 0, & \bar{\zeta}_t + (\bar{f}(\bar{\zeta}))_x &= 0, \\ \zeta(x, 0) &= \begin{cases} \rho_l & \text{if } x < a \\ \rho_r & \text{if } x > a, \end{cases} & \bar{\zeta}(x, 0) &= \begin{cases} \bar{\rho}_l & \text{if } x < \bar{a} \\ \bar{\rho}_r & \text{if } x > \bar{a}, \end{cases} \end{aligned}$$

respectively and suppose that $f(\rho_i) = \rho_i w(\rho_i)$ and $\bar{f}(\bar{\rho}_i) = \bar{\rho}_i \bar{w}(\bar{\rho}_i)$ for $i = l, r$.

Let $z : [t_0, T] \rightarrow \mathbb{R}$ and $\bar{z} : [t_0, T] \rightarrow \mathbb{R}$ be solutions of

$$\dot{z} = w(\zeta(z, t)) \quad \text{and} \quad \dot{\bar{z}} = \bar{w}(\bar{\zeta}(\bar{z}, t)),$$

respectively with $z(t_0) = z_0 < a$ and $\bar{z}(t_0) = \bar{z}_0 < \bar{a}$.

Then, for any $t > t_0$ with $t - t_0$ relatively small,

$$\bar{z}(t) - z(t) = (\bar{w}(\bar{\rho}_r) - w(\rho))(t - t_0) + \frac{\bar{\rho}_r - \bar{\rho}_l}{\bar{\rho}_r}(\bar{a} - \bar{z}_0) - \frac{\rho_r - \rho_l}{\rho_r}(a - z_0) + \bar{z}_0 - z_0.$$

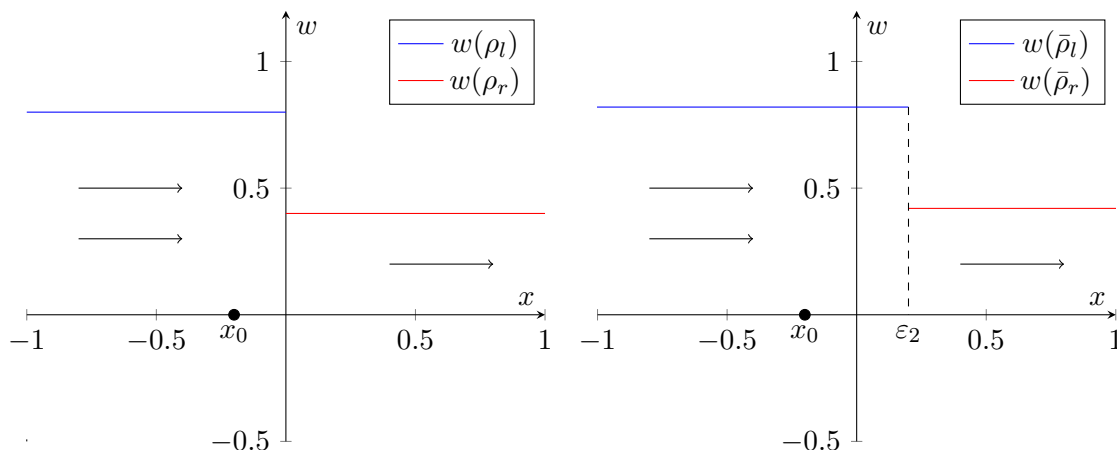


Figure 5.1: The Riemann problem. The particles starts at x_0 , after a while they catch up with the shocks in front of them, then right after that, they continue to travel with lower speeds.

Proof. We have

$$\zeta(x, t) = \begin{cases} \rho_l & \text{if } x < a + \lambda t \\ \rho_r & \text{if } x > a + \lambda t, \end{cases}$$

with

$$\lambda = \frac{f(\rho_l) - f(\rho_r)}{\rho_l - \rho_r} = \frac{\rho_l w(\rho_l) - \rho_r w(\rho_r)}{\rho_l - \rho_r},$$

and

$$\bar{\zeta}(x, t) = \begin{cases} \bar{\rho}_l & \text{if } x < \bar{a} + \bar{\lambda} t \\ \bar{\rho}_r & \text{if } x > \bar{a} + \bar{\lambda} t, \end{cases}$$

with

$$\bar{\lambda} = \frac{\bar{\rho}_l \bar{w}(\bar{\rho}_l) - \bar{\rho}_r \bar{w}(\bar{\rho}_r)}{\bar{\rho}_l - \bar{\rho}_r}.$$

The particle z first travels with speed $w(\rho_l)$ until it hits the shock, then travels with speed $w(\rho_r)$. The hitting time τ can be calculated as $w(\rho_l)\tau + z_0 = \lambda\tau + a$, thus

$$\tau = \frac{a - z_0}{w(\rho_l) - \lambda}.$$

The trajectory of z is given by

$$z(t) = \begin{cases} w(\rho_l)(t - t_0) + z_0 & \text{if } t - t_0 < \tau \\ w(\rho_r)(t - t_0 - \tau) + w(\rho_l)\tau + z_0 & \text{if } t - t_0 \geq \tau. \end{cases}$$

On the other hand, the particle \bar{z} first travels with speed $\bar{w}(\bar{\rho}_l)$ until it hits the shock and

travels with speed $\bar{w}(\bar{\rho}_r)$ after that. Similarly,

$$\bar{z}(t) = \begin{cases} \bar{w}(\bar{\rho}_l)(t - t_0) + \bar{z}_0 & \text{if } t - t_0 < \bar{\tau} \\ \bar{w}(\bar{\rho}_r)(t - t_0 - \bar{\tau}) + \bar{w}(\bar{\rho}_l)\bar{\tau} + \bar{z}_0 & \text{if } t - t_0 \geq \bar{\tau}, \end{cases} \quad (5.42)$$

where $\bar{\tau}$ is the hitting time

$$\bar{\tau} = \frac{\bar{a} - \bar{z}_0}{\bar{w}(\bar{\rho}_l) - \bar{\lambda}}. \quad (5.43)$$

Hence after the particle has passed both shock points, that is $t > \max\{\tau, \bar{\tau}\} + t_0$, we have, setting $w_i := w(\rho_i)$ and $\bar{w}_i := \bar{w}(\bar{\rho}_i)$ for $i = r, l$,

$$\bar{z}(t) - z(t) = (\bar{w}_r - w_r)(t - t_0) + (\bar{w}_l - \bar{w}_r)\bar{\tau} - (w_l - w_r)\tau + \bar{z}_0 - z_0.$$

Noting that

$$(w_l - w_r)\tau = (w_l - w_r) \frac{(\rho_r - \rho_l)(a - z_0)}{\rho_r(w_l - w_r)} = \frac{\rho_r - \rho_l}{\rho_r}(a - z_0),$$

and similarly

$$(\bar{w}_l - \bar{w}_r)\bar{\tau} = \frac{\bar{\rho}_r - \bar{\rho}_l}{\bar{\rho}_r}(\bar{a} - \bar{z}_0),$$

we have

$$\bar{z}(t) - z(t) = (\bar{w}_r - w_r)(t - t_0) + \frac{\bar{\rho}_r - \bar{\rho}_l}{\bar{\rho}_r}(\bar{a} - \bar{z}_0) - \frac{\rho_r - \rho_l}{\rho_r}(a - z_0) + \bar{z}_0 - z_0,$$

and the result follows. \square

5.2.2 Stability with respect to changes in flux functions

In this section, we study the stability of the particle trajectories with respect to small changes in the flux function (with the initial field being fixed). We define the set Λ as (5.35). Let $\bar{w} \in \Lambda$ and $\bar{\rho}$ satisfy

$$\partial_t \bar{\rho} + \partial_x(\bar{\rho}\bar{w}(\bar{\rho})) = 0, \quad (5.44)$$

$$\bar{\rho}(x, 0) = \rho_0(x). \quad (5.45)$$

We have the following stability estimate.

Theorem 5.2.5. *Assume that ρ and $\bar{\rho}$ are the entropy solutions to (5.27)-(5.28) and (5.44)-(5.45) respectively with the same initial data $\rho_0 \in L^1 \cap BV(\mathbb{R}; [m_\rho, 1])$, $0 < m_\rho < 1$, and with $w, \bar{w} \in \Lambda$ satisfying*

$$\|w - \bar{w}\|_{\text{Lip}} \leq \varepsilon.$$

Then the corresponding particle trajectories z and \bar{z} satisfy

$$\|z - \bar{z}\|_{L^\infty((t_0, T])} \leq C_w \varepsilon^{1/2},$$

for any given $T > 0$, and with

$$C_w = 1 + 2(T - t_0)\left(1 + \frac{2}{m_\rho}\right)\|w\|_{\text{Lip}} + \frac{2}{m_\rho}\|\rho_0\|_{BV}.$$

Proof. We argue as in the proof of Theorem 5.2.3 and first consider ρ^N and $\bar{\rho}^N$, the front tracking approximations of ρ and $\bar{\rho}$, and their corresponding trajectories z^N and \bar{z}^N respectively. We follow a similar argument to STEP 1 of proof of Theorem 5.2.3 and employ Lemma 5.2.4, the only difference here is that in the right-hand side of inequalities (5.37) and (5.40) $w(\bar{\rho})$ is replaced with $\bar{w}(\bar{\rho})$ and we write instead

$$\begin{aligned} |\bar{w}(\bar{\rho}(\bar{z}(t), t)) - w(\rho(z(t), t))| &\leq |w(\bar{\rho}(\bar{z}(t), t)) - w(\rho(z(t), t))| + |\bar{w}(\bar{\rho}(\bar{z}(t), t)) - w(\bar{\rho}(z(t), t))| \\ &\leq L_w \varepsilon^\beta + \|w - \bar{w}\|_{L^\infty} \\ &\leq L_w \varepsilon^\beta + \varepsilon, \end{aligned}$$

as $\rho \leq 1$ and hence $\|w - \bar{w}\|_{L^\infty} \leq \|w - \bar{w}\|_{\text{Lip}}$. We therefore obtain

$$\begin{aligned} |\bar{z}^N(T) - z^N(T)| &\leq (T - t_0)(L_w \varepsilon^\beta + \varepsilon) + \frac{\varepsilon^{1-\beta}}{m_\rho} \sum_{j=1}^k \max\{\iota_j, \bar{\iota}_j, 0\} \\ &\quad + \frac{\varepsilon}{m_\rho} \sum_{j=1}^k \max\{a_j - z(t_j), \bar{a}_j - \bar{z}(t_j)\}. \end{aligned}$$

Since, again by [Daf16, Theorem 6.2.6] and as by Lemma 5.2.2 z and \bar{z} do not encounter any of the initial shocks more than once, we have

$$\sum_{j=1}^k \max\{\iota_j, \bar{\iota}_j, 0\} \leq \sum_{j=1}^k \max\{\iota_j + \bar{\iota}_j, 0\} \leq 2\|\rho_0\|_{BV},$$

and noting that

$$\begin{aligned} \sum_{j=1}^k \max\{a_j - z(t_j), \bar{a}_j - \bar{z}(t_j)\} &\leq \sum_{j=1}^k (a_j - z(t_j)) + (\bar{a}_j - \bar{z}(t_j)) \\ &\leq \|w\|_{\text{Lip}}(T - t_0) + \|\bar{w}\|_{\text{Lip}}(T - t_0) \\ &\leq (2L_w + \varepsilon)(T - t_0). \end{aligned}$$

we obtain

$$\begin{aligned} |\bar{z}^N(T) - z^N(T)| &\leq (T - t_0)(L_w \varepsilon^\beta + \varepsilon) + \frac{2\|\rho_0\|_{BV}}{m_\rho} \varepsilon^{1-\beta} + \frac{\varepsilon}{m_\rho} (2L_w + \varepsilon)(T - t_0) \\ &\leq 2(T - t_0)\left(1 + \frac{2}{m_\rho}\right) \varepsilon^\beta + \frac{2\|\rho_0\|_{BV}}{m_\rho} \varepsilon^{1-\beta} \end{aligned}$$

as $\beta < 1$, and we get the best rate in ε when we choose $\beta = 1/2$. We hence have

$$\|\bar{z}^N - z^N\|_{L^\infty} \leq C_0 \sqrt{\varepsilon}$$

with $C_0 = 2(T - t_0)(1 + \frac{2}{m_\rho}) + \frac{2\|\rho_0\|_{BV}}{m_\rho}$. Then, the same argument as STEP 2 of Theorem 5.2.3 gives the result. \square

5.3 Summary and outlook

This chapter studies the particle trajectories for scalar conservation laws and their properties. Several new results on the stability of the particle trajectories with respect to changes in the initial field are established.

The study of the differential equations when their vector field depends on solutions to hyperbolic conservation laws has been a subject of a number of works in literature, see for instance [BS98, Bre03, CM03, Mar04, LMP11, DMG14] and references therein. The study of particle trajectories in the context of traffic flow starts with the work by Colombo and Marson in [CM03]. In that work, the existence of the particle trajectories and their stability with respect to the initial position of the particles are obtained. The continuity property of the mapping from initial fields to trajectories is also studied, but their convergence with specific rates had not been known until our work [DD20].

In a more general context, differential equations of the form $\dot{z}(t) = b(z(t), t)$ when the vector field b is not necessarily continuous is the subject of an active research direction, starting with [DL89] (thus, it is usually called DiPerna-Lions flow). In that seminal paper, the authors study the well-posedness of z for Sobolev vector fields with bounded divergence in connection with some continuity equation (which is closely related to the connection between (5.1) and (5.2)). This work has been extended to BV coefficients with bounded divergence in [Amb04].

Chapter 6

Bayesian inverse problems in scalar conservation laws

In this chapter, we investigate some inverse problems arising from scalar conservation laws in one-space dimension by employing a Bayesian approach. Consider again the scalar conservation law

$$\partial_t v(x, t) + \partial_x f(v(x, t)) = 0, \quad (x, t) \in \mathbb{R} \times (0, \infty), \quad (6.1)$$

with some initial value, which we will mostly refer to as the *upstream field*¹,

$$v(x, 0) = v_0(x), \quad x \in \mathbb{R}, \quad (6.2)$$

Specially, we will be looking at the following two problems

- Problem 1. The model (6.1) is known. Finding the upstream field v_0 based on (finite) observations of the (function of) solution v .
- Problem 2. The upstream field v_0 is known. Finding the flux function f of the model (6.1), again based on (finite) observations of the (function of) solution v .

Throughout this chapter, when we say the solution v of the forward problem (6.1)-(6.2), we mean the unique entropy solution. These two inverse problems are important in applications. Consider the case in traffic flow for example. Problem 1 may be helpful in predicting the future behaviour of traffic flow by finding its state $v(x, t_0)$ at some time $t_0 \geq 0$, and then input this data into some device or toy model to observe the traffic state at a later time. Problem 2, especially when the flux is partly unknown, may be used to

¹We avoid using the term *initial data* here to not confuse with the *data* given by the inverse problems.

detect road conditions, obstacles or other information that cannot be observed directly by the observer. A specific example is to detect obstructions due to car accidents or other events that happened inside a tunnel.

We will be employing the Bayesian approach (see Chapter 3) to our inverse problems and investigate their well-posedness and approximation. Both inverse problems are written in the same form: finding unknown u from observed data y with

$$y = G(u) + \xi, \quad (6.3)$$

where ξ denotes the *observable noise* and $G : X \rightarrow Y$ is the parameter-to-observation² map (or observation map if no confusion arises) with Banach spaces X, Y . Given a prior measure μ_0 on X that stores already known information of u , the solution of (6.3) is understood as a posterior measure μ^y . This measure updates the information of u when the data is taken into account.

Throughout this chapter, two kinds of observation maps will be considered, corresponding to two kinds of data. The first one is the trajectory data³, given when the observations are made by observing the particle trajectory $z(t)$ of the forward problem (6.1). In this case, the forward map is given by

$$S_t : u \mapsto z(t). \quad (6.4)$$

The second kind of data is direct measurement data⁴, obtained when the observations are made by observing directly the entropy solution v of the forward problem (6.1). The forward map is given in this case by

$$S_e : u \mapsto v(x, t). \quad (6.5)$$

We will consider these two forms of observation maps for both inverse problems, that is, u may be the initial field or the flux function. Based on the discussions in Chapter 3, we study the well-posedness and approximation of the probability posterior μ^y .

6.1 Recovering the upstream field with trajectory data

6.1.1 Well-posedness

Suppose that we make finite number of noisy observations $y = (y_1 \ y_2 \ \dots \ y_J)^T$ from the particle trajectory

$$y_j = z(t_j) + \xi_j, \quad j = 1, \dots, J, \quad (6.6)$$

²Note there is a change in notation, we use G here instead of the general \mathcal{G} in Chapter 3.

³occasionally called Lagrangian data.

⁴also called Eulerian data.

with z is a particle trajectory of (6.1)-(6.2), defined as the Filippov solution of (5.2)-(5.3). We suppose that ξ_j are independent and Gaussian distributed

$$\xi_j \sim \mathcal{N}(0, \gamma^2).$$

Here we are interested in recovering the upstream field. We let

$$\begin{aligned} u &:= v(\cdot, 0) \\ G(u) &:= (G_1(u), \dots, G_J(u)), \text{ with } G_j(u) := z(t_j), \end{aligned} \tag{6.7}$$

where $t_J > t_{J-1} > \dots > t_1 > 0$. We have the following well-posedness result.

Theorem 6.1.1. *Suppose $y \in \mathbb{R}^J$ is given by (6.6). Let $\mu_0(X) = 1$ where $X = L^1(\mathbb{R}) \cap BV(\mathbb{R})$. Then the posterior measure μ^y given by*

$$\frac{d\mu^y}{d\mu_0}(u) = \frac{1}{Z} \exp\left(-\frac{1}{2\gamma^2}|y - G(u)|^2\right),$$

with

$$Z = \int_X \exp\left(-\frac{1}{2\gamma^2}|y - G(u)|^2\right) d\mu_0,$$

is well-defined and continuous in Hellinger distance with respect to y .

Proof. We write $G = \mathcal{O} \circ S_l$ with $\mathcal{O} : X \rightarrow \mathbb{R}^J$ the point observation operator $z \mapsto y$ and S_l the forward operator mapping $u \in X$ to $z \in C_b([t_0, T])$ with C_b denoting the space of bounded continuous functions. Consider $\{u^N\} \subset L^1 \cap L^\infty$ and suppose that $u^N \rightarrow u$ in L^1 and let v^N be the solution of the conservation law for $t > t_0$ with $v^N(t_0) = u^N$. Then we have $v^N \rightarrow v$ in $L^1(0, T; L^1(\mathbb{R}))$. Let $z^N : [t_0, T] \rightarrow \mathbb{R}$ be the unique Filippov solution to

$$\dot{z}^N(t) \in [w(v^N(z^N(t)+, t)), w(v^N(z^N(t)-, t))], \quad \text{with } z^N(t_0) = z(t_0) = x_0.$$

By Theorem 5.1.6 we have $z^N \rightarrow z$ uniformly. This concludes the continuity of $S_l : u \mapsto z$. Since the point observation operator \mathcal{O} is continuous we have $G = \mathcal{O} \circ S_l : X \rightarrow \mathbb{R}^J$ is continuous. It is evident that $Z < 1$ and since G is bounded (due to the boundedness of z) we have $Z > 0$. The result follows by Theorem 3.2.2 and 3.2.6. \square

6.1.2 Approximations

In this section, we consider the approximation of the Bayesian inverse problem of recovering initial field from tracking the trajectory $z(t)$ of a given particle. The Bayesian formulation reads

$$y = G(u) + \xi, \tag{6.8}$$

where the observation map is given as $G(u) = \{z(t_j)\}_{j=1}^J$ and $t_j, j = 1, \dots, J$, denote the time slots of observations. Be aware that each $z(t_j)$ is a function of the unknown initial field u (to emphasise this fact, we will occasionally write $z(t)$ as $z(t, u)$ or even just $z(u)$). Consider the approximate version of (6.8)

$$y = G^N(u) + \xi, \quad (6.9)$$

with the observation map being

$$G^N(u) = \{z^N(t_j)\}_{j=1}^J, \quad (6.10)$$

where $z^N(t)$ denotes some approximation of the particle trajectory $z(t)$.

In the rest of this section we consider $y \in \mathbb{R}^J$, Γ a diagonal matrix with nonzero members all equal γ^2 , and μ and μ^N satisfying

$$\frac{d\mu}{d\mu_0}(u) = \frac{1}{Z} \exp\left(-\frac{1}{2\gamma^2}|y - G(u)|^2\right) \quad (6.11)$$

$$\frac{d\mu^N}{d\mu_0}(u) = \frac{1}{Z^N} \exp\left(-\frac{1}{2\gamma^2}|y - G^N(u)|^2\right) \quad (6.12)$$

with Z and Z^N given as

$$Z = \int_X \exp(-\Phi(u)) \mu_0(du), \quad (6.13)$$

$$Z^N = \int_X \exp(-\Phi^N(u)) \mu_0(du). \quad (6.14)$$

The following convergence results concern the approximate measure μ^N arising in approximations of initial field of the forward problem (giving a sequence of exact solutions) or front tracking approximations.

Theorem 6.1.2. *Let v^N be a sequence of exact solutions (or front tracking approximations) of bounded variation converging in L^1 to the entropy solution v of (6.1)-(6.2) with initial field v_0 taking value in a bounded interval $[-M, M]$. Let μ^N, μ be given as (6.11)-(6.12) where G, G^N are defined by (6.7)-(6.10). Then μ^N converges to μ in the sense that*

$$d_{\text{Hell}}(\mu^N, \mu) \rightarrow 0, \quad (6.15)$$

as $N \rightarrow \infty$.

Proof. From the definition of Hellinger distance, the formulations (6.11)-(6.12) and the basic inequality $(a + b)^2 \leq 2(a^2 + b^2)$ we have

$$\begin{aligned} d_{\text{Hell}}(\mu, \mu^N)^2 &= \frac{1}{2} \int_X \left(\sqrt{\frac{d\mu}{d\mu_0}} - \sqrt{\frac{d\mu^N}{d\mu_0}} \right)^2 \mu_0(du) \\ &\leq I_1 + I_2, \end{aligned}$$

where

$$I_1 = \frac{1}{Z} \int_X \left(e^{-\frac{1}{2}\Phi(u)} - e^{-\frac{1}{2}\Phi^N(u)} \right)^2 \mu_0(du), \quad (6.16)$$

$$I_2 = \left| \frac{1}{\sqrt{Z}} - \frac{1}{\sqrt{Z^N}} \right|^2 \int_X e^{-\Phi^N(u)} \mu_0(du). \quad (6.17)$$

Since the data is finite and the noise is Gaussian, $\xi \sim \mathcal{N}(0, \Gamma)$, then Φ and Φ^N can be defined as

$$\Phi(u) = \frac{1}{2\gamma^2} |y - G(u)|^2, \quad \Phi^N(u) = \frac{1}{2\gamma^2} |y - G^N(u)|^2, \quad (6.18)$$

where $G^N(u)$ is some approximation of $G(u)$. Basic calculations lead to

$$|\Phi(u) - \Phi^N(u)| \leq C |G(u) - G^N(u)|,$$

due to the boundedness of G, G^N . Together with the locally Lipschitz property of e^{-x} for $x > 0$, we have

$$\begin{aligned} I_1 &\leq \int_X |\Phi(u) - \Phi^N(u)|^2 \mu_0(du) \\ &\leq C \int_X |G(u) - G^N(u)|^2 \mu_0(du). \end{aligned}$$

Thanks to Theorem 5.1.6, for every $u \in X$, $G^N(u) \rightarrow G(u)$. By dominated convergence theorem, it follows that $I_1 \rightarrow 0$ as $N \rightarrow \infty$.

For I_2 , it is sufficient to show that $|Z - Z^N| \rightarrow 0$. We have

$$\begin{aligned} |Z - Z^N| &\leq \int_X |\exp(-\Phi(u)) - \exp(-\Phi^N(u))| \mu_0(du) \\ &\leq \int_X |\Phi(u) - \Phi^N(u)| \mu_0(du). \end{aligned}$$

The same arguments as above lead to $|Z - Z^N| \rightarrow 0$, therefore $I_2 \rightarrow 0$ as $N \rightarrow \infty$. This completes the proof of Theorem 6.1.2. \square

Remark 6.1.3. *The use of Hellinger distance here is mainly for the sake of convenience and also due to its popularity in the literature. Other choices of distances between measures may be possible. We mention here an important one, the Wasserstein distance, whose usage has been rising recently in the theory of optimal transport, statistics and machine learning (see for instance, [PC19]). A version of the convergence (6.15) in Wasserstein distance (of the first order) d_{Wass} can be proved. Indeed, thanks to the famous Kantorovich–Rubinstein duality theorem (see [Vil08]), we may write*

$$\begin{aligned} d_{\text{Wass}}(\mu, \mu^N) &= \sup_{\|h\|_{Lip} \leq 1} \left| \int_X h d\mu - \int_X h d\mu^N \right| \\ &= \sup_{\|h\|_{Lip} \leq 1} \left| \int_X h(u) \left(\frac{\exp(-\Phi(u))}{Z} - \frac{\exp(-\Phi^N(u))}{Z^N} \right) \mu_0(du) \right|, \end{aligned}$$

and then use the estimate

$$|h(u)| \leq \|h\|_{Lip} \|u\|_X + |h(0)| \leq C(1 + \|u\|_X) \quad (6.19)$$

to get rid of h . The rest of the proof can be carried out similarly, with the help of Fernique's theorem where appropriate. See also [Spr20]. In the rest of this chapter, all convergence results in Hellinger distance also apply to Wasserstein distance, with suitable modifications as we discussed above.

The convergence result of the approximate posterior when one uses the vanishing viscosity approximation for the forward problem is also given.

Theorem 6.1.4. *Assume that Assumption 5.1.8 holds at every shock curve of the entropy solution v to the equation (5.1). Let v^ϵ be the vanishing viscosity approximation of v (see (5.18)) and z^ϵ be the corresponding particle trajectory, defined as (5.19). Let μ be given as in (6.11) with G given in (6.7). Assume that μ^ϵ is defined as follows*

$$\begin{aligned} \frac{d\mu^\epsilon}{d\mu_0}(u) &= \frac{1}{Z^\epsilon} \exp\left(-\frac{1}{2\gamma^2}|y - G^\epsilon(u)|^2\right), \\ Z^\epsilon &= \int_X \exp\left(-\frac{1}{2\gamma^2}|y - G^\epsilon(u)|^2\right) d\mu_0, \end{aligned}$$

where $G^\epsilon(u) = \{z^\epsilon(x_j, t_j)\}_{j=1}^J$. Then μ^ϵ converges to μ in the sense that

$$d_{\text{Hell}}(\mu^\epsilon, \mu) \rightarrow 0, \quad (6.20)$$

as $\epsilon \rightarrow 0$.

The proof of Theorem 6.1.4 is entirely similar to the one of Theorem 6.1.2 with a recipe based on Theorem 5.1.9 and will be omitted here.

6.2 The case of traffic flow: rate of convergence

6.2.1 Recovering upstream density

This section is devoted to explore the convergence results for the Bayesian inverse problem (6.8) where the unknown is the upstream density $u = \rho_0$ of the traffic flow model (5.27)-(5.28). We still have G defined as in (6.7). We consider the approximate inverse problem (6.9) where the approximate observation map is given as

$$G^N(u) := G(u^N), \quad (6.21)$$

with $u^N \in X$ an approximation of $u \in X$ with $X = L^1 \cap BV(\mathbb{R}; (0, 1))$ (note that values of traffic densities lie in $[0, 1]$). The following theorem provides a convergence rate for the

approximation of the posterior in terms of the one of the upstream density. As we have discussed several times previously, such a specific rate of the convergence is extremely hard to get in the context of hyperbolic conservation laws if one's observations are based on the direct measurements of the entropy solution.

Theorem 6.2.1. *Let $X = L^1 \cap BV(\mathbb{R}; (0, 1))$. Suppose that for any $u \in X$, the approximating sequence $\{u^N\} \subset X$ satisfies*

$$\|u^N - u\|_{L^1 \cap L^\infty} \leq \psi(N), \quad \psi(N) \rightarrow 0 \text{ as } N \rightarrow \infty.$$

Assume that $\mu_0(X) = 1$, and

$$\int_X \frac{1 + \|u\|_{BV}^2}{m_u^2} \mu_0(du) < \infty \quad (6.22)$$

where $m_u = \inf_{x \in \mathbb{R}} u(x)$. Let μ and μ^N be given as in (6.11) and (6.12) with G^N defined by (6.21). Then

$$d_{\text{Hell}}(\mu^N, \mu) \leq C\sqrt{\psi(N)},$$

as $N \rightarrow \infty$.

Remark 6.2.2. *A prior satisfying the conditions of the above theorem can be constructed as follows. Let ν_0 be a Gaussian measure with $\nu_0(W^{1,1}) = 1$ where $W^{1,1}$ is the space of integrable functions with integrable derivatives on \mathbb{R} . We note that $W^{1,1}(\mathbb{R}) \subset L^1 \cap BV(\mathbb{R})$.*

Let $v \sim \nu_0$ and

$$u(x) = F(v(x)) := \begin{cases} \frac{1}{2}e^{v(x)}, & \text{if } v(x) \leq 0 \\ 1 - \frac{1}{2}e^{-v(x)}, & \text{if } v(x) > 0. \end{cases}$$

Consider $\mu_0 := \nu_0 \circ F^{-1}$. We have

$$m_u = \min_{x \in \mathbb{R}} u(x) \geq \frac{1}{2} e^{-\|v\|_{L^\infty}}$$

and

$$\|u\|_{BV} = \int_{\mathbb{R}} |u'| dx \leq \frac{1}{2} \int_{\mathbb{R}} |v'| dx \leq \frac{1}{2} \|v\|_{W^{1,1}}.$$

Hence

$$\int_X \frac{1 + \|u\|_{BV}^2}{m_u^2} \mu_0(du) \leq \int_{W^{1,1}} \|v\|_{W^{1,1}}^2 e^{2\|v\|_{W^{1,1}}} d\nu_0(dv) < \infty$$

since $\|v\|_{L^\infty} \leq \|v\|_{W^{1,1}}$ and by Fernique's theorem.

Proof of Theorem 6.2.1. By proceeding similarly as in the proof of Theorem 6.1.2, we arrive at

$$d_{\text{Hell}}(\mu, \mu^N)^2 \leq I_1 + I_2,$$

where I_1 and I_2 are given by (6.16) and (6.17). Therefore

$$I_1 \leq C \int_X |G(u) - G(u^N)|^2 \mu_0(du) \leq C\psi(N) \int_X \frac{1 + \|u\|_{BV}^2}{m_u^2} \mu_0(du) \leq C\psi(N),$$

by (6.22) and where in the second inequality we have used Theorem 5.2.3. A similar argument leads to

$$I_2 \leq C|Z - Z^N|^2 \leq C\psi(N).$$

The result then follows. \square

6.2.2 Recovering the flux function

We now consider the inverse problem of finding the velocity function w given finite Lagrangian trajectory data, with the upstream field is known. This is equivalent to finding the flux function since in traffic flow the flux function takes the form $f(\rho) = \rho w(\rho)$. As before, the Lagrangian data is collected as follows,

$$y_j = z(t_j) + \xi_j, \quad j = 1, \dots, J, \quad (6.23)$$

where $\xi_j \sim \mathcal{N}(0, \gamma^2)$ and are independent. Note that each $z(t_j)$ now depends implicitly on the unknown w . To be consistent with the notation of Theorem 3.2.2 and 3.2.6 we still set here $u = w$ and use the notation G for the mapping $u \mapsto y$. We have the following well-posedness result whose proof, thanks to the continuity of the forward map $w \mapsto z$ (Theorem 5.2.5), is very similar to that of Theorem 6.1.1 and will be omitted here.

Theorem 6.2.3. *Suppose $y \in \mathbb{R}^J$ is given by (6.23). Let $\mu_0(\Lambda) = 1$ where Λ with $Lip(\mathbb{R})$ is as given in (5.35). Then the posterior measure μ^y given by*

$$\frac{d\mu^y}{d\mu_0}(u) = \frac{1}{Z} \exp\left(-\frac{1}{2\gamma^2}|y - G(u)|^2\right),$$

with

$$Z = \int_{\Lambda} \exp\left(-\frac{1}{2\gamma^2}|y - G(u)|^2\right) d\mu_0,$$

is well-defined and continuous in Hellinger distance with respect to y .

We now show that μ^y is stable with respect to appropriate perturbations of the forward operator, provide that the shocks behave the same way as in the case of traffic flow.

Theorem 6.2.4. *Let the assumptions of Theorem 6.2.3 hold. Suppose also that for any $u \in \Lambda$, the approximating sequence $\{u^N\} \subset \Lambda$ satisfies*

$$\|u^N - u\|_{Lip} \leq \psi(N), \quad \psi(N) \rightarrow 0 \text{ as } N \rightarrow \infty.$$

Assume that $\int_{\Lambda} \|u\|_{\text{Lip}} \mu_0(du) < \infty$ and μ and μ^N are given as in (6.11) and (6.12) with $G^N = G(u^N)$. Then

$$d_{\text{Hell}}(\mu^N, \mu) \leq C\sqrt{\psi(N)},$$

as $N \rightarrow \infty$.

Proof. Proceed similarly as in the proof of Theorem 6.2.1 and of Theorem 6.1.2, where the corresponding stability result of the forward problem, Theorem 5.2.5, is used. \square

A related situation to the estimation of the flux function considered here is the problem considered in [HPR14] where they study the following inhomogeneous scalar conservation law

$$\partial_t \rho(x, t) + \partial_x(k(x)g(\rho(x, t))) = 0, \quad (6.24)$$

with some appropriate function $k : \mathbb{R} \rightarrow \mathbb{R}$. In the context of traffic flow, the function k represents external factors that influence the traffic flow. Such factors may be interpreted as road conditions or the presence of obstacles on the road. In particular, when k is constant, they show that one can recover, using Tikhonov regularisation, a piecewise linear interpolation g_ν that approximates g in the sense that the solutions corresponding to g and g_ν are close in L^1 . However, they assume that the whole solution is known at (almost) every point (or almost every point except some certain interval) and the initial data is chosen to be piecewise constant only.

6.3 Inverse problems with direct measurement data

In this section, we come back to our original inverse problem and suppose that we make finite number of noisy observations $y = (y_1 \ y_2 \ \dots \ y_J)^T$ from the solution v . We discuss the well-posedness and also the approximation of the Bayesian inverse problem (6.3). Denote by $u \in X$ the unknown of the inverse problems for the scalar conservation law

$$\partial_t v + \partial_x f(v) = 0, \quad v(x, 0) = v_0(x),$$

which is either the initial field v_0 or the flux function f . For the former, $X = L^1 \cap BV(\mathbb{R})$ with norm $\|\cdot\|_{L^1}$ and for the later

$$X = \{f : \mathbb{R} \rightarrow \mathbb{R}, f(0) = 0, f \text{ is convex}, \|f\|_{\text{Lip}} < \infty\},$$

where $\|\cdot\|_{\text{Lip}}$ is defined as (2.31). The observation maps of the solution and of the approximate solution are given as

$$G(u) = \mathcal{O} \circ S_e(u), \quad \text{and} \quad G^N(u) = \mathcal{O} \circ S_e^N(u), \quad (6.25)$$

where $\mathcal{O} : X \rightarrow \mathbb{R}^J$ is the evaluation operator and $S_e, S_e^N : X \rightarrow L^1 \cap BV(\mathbb{R})$ are the forward operator and its approximation. This approximation may be coming from any computational method involved in solving the forward problem (including front tracking and vanishing viscosity) or a result of some disturbance in the initial condition.

Using the ball evaluation operator

First of all, suppose that in some applications one can collect the data using the ball evaluation operator, that is, $\mathcal{O} = \{\mathcal{O}_j\}_{j=1}^J : X \rightarrow \mathbb{R}^J$, where

$$\mathcal{O}_j(v) := \int_{B_r(x_j)} v \, dx,$$

here we denote $B_r(x_j) = \{x : |x - x_j| < r\}$. Then, any small perturbation of u in X translates to a small perturbation of \mathcal{O} in L^∞ , thanks to Theorem 2.4.3 and Theorem 2.4.4. Indeed, if $\|u^N - u\|_X \leq \psi(N)$, then

$$\begin{aligned} |\mathcal{O}_j \circ S_e^N(u^N) - \mathcal{O}_j \circ S_e(u)| &\leq \int_{B_r(x_j)} |S_e^N(u^N) - S_e(u)| \, dx \\ &\leq \int_{\mathbb{R}} |S_e^N(u^N) - S_e(u)| \, dx \\ &\leq C\psi(N). \end{aligned}$$

Therefore, the condition (3.24) follows and so does Corollary 3.2.9. In fact, by using this method, we can estimate any unknown initial field $v_0 \in L^1 \cap L^\infty$ given a Lipschitz f , and any unknown Lipschitz continuous flux function for a given $v_0 \in L^1 \cap BV(\mathbb{R})$. We note that the flux function need not be convex here (see Theorem 2.4.4).

Using the point evaluation operator

Let us consider now the point evaluation operator

$$\mathcal{O}(v) = \{v(x_j, t_j)\}_{j=1}^J.$$

We show here that the Bayesian inverse problem is well-defined and continuous in data y because the measurability of the forward map still holds. Lack of continuity of the forward map however leads to weaker approximation properties of the posterior compared to the Lagrangian data case. We write

$$y_j = v(x_j, t_j) + \xi_j, \quad j = 1, \dots, J, \quad (6.26)$$

with $\xi_j \sim \mathcal{N}(0, \gamma^2)$ and independent, and let

$$\begin{aligned} u &:= v(\cdot, t_0) \\ G(u) &:= (G_1(u), \dots, G_J(u)), \quad \text{with } G_j(u) = v(x_j, t_j). \end{aligned} \quad (6.27)$$

We have the following well-posedness result.

Theorem 6.3.1. *Suppose $y \in \mathbb{R}^J$ is given by (6.26). Let $\mu_0(X) = 1$, with $X = L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Then the posterior measure μ^y given by*

$$\frac{d\mu^y}{d\mu_0}(u) = \frac{1}{Z} \exp\left(-\frac{1}{2\gamma^2}|y - G(u)|^2\right),$$

with $Z = \int_X \exp\left(-\frac{1}{2\gamma^2}|y - G(u)|^2\right) d\mu_0$ is well-posed.

Proof. The positivity and boundedness of G imply $0 < Z < 1$. As the likelihood is continuous in y it remains to show that $G : L^1 \cap L^\infty \rightarrow \mathbb{R}^J$ is measurable. We write $G = \mathcal{O} \circ S_e$ with $\mathcal{O} : X \rightarrow \mathbb{R}^J$ the point evaluation operator and $S_e : X \rightarrow X$ the forward operator of the PDE mapping u to v . Since \mathcal{O} is continuous, it is sufficient to show that S_e is measurable. By Theorem 2.4.3, $S_e : L^1_{\text{loc}} \rightarrow L^1_{\text{loc}}$ is continuous. Hence for an open set $A \subset L^1_{\text{loc}}$ we have that $S_e^{-1}(A)$ is open on L^1_{loc} . Then, since $S_e : L^\infty \rightarrow L^\infty$ is bounded, we have

$$L^\infty \cap S_e^{-1}(A \cap L^\infty) = S_e^{-1}(A) \cap L^\infty \text{ is open in } L^\infty.$$

Noting that $\mathcal{B}(L^\infty) = \{A \cap L^\infty : A \in \mathcal{B}(L^1_{\text{loc}})\}$, we conclude the measurability of S_e over L^∞ . Now all conditions of Theorem 3.2.2 and 3.2.6 are satisfied and the result follows. \square

The approximation of the posterior is much more complicated in this case. The general approximation theory of Bayesian inverse problems (see Chapter 3) fails since the stability (in L^∞) of the observation operator is no longer satisfied. Nevertheless, below we discuss possible situations where one still has some sort of convergence for the approximate posterior.

In practical applications, one may assume that the measurement device can detect the shocks of strength bigger than ε for some given $\varepsilon > 0$ and avoid them, where the shock strength is defined as

$$|v(x(t)+, t) - v(x(t)-, t)|,$$

for some shock curve $x = x(t)$. Then the measurements may be made in the areas where no shocks or only shocks of strength smaller than ε are present. We note that by [BL99, Theorem 3], for any $\varepsilon > 0$, shocks of strength greater than ε are finitely many and hence have measure zero. In the case that the initial field and flux function are smooth (C^k with $k \geq 3$), generically, solutions produce only a finite set of shocks in a given bounded domain, thanks to a regularity result discovered by Schaeffer in [Sch73] (see Theorem 2.5.3, see also [Daf85] for considerable improvements). Therefore, in reality, the chance of

having data from certain positions with no shocks or shocks of strength smaller than any given ε is likely.

Define

$$\mathcal{O}_\varepsilon := \{(x, t) \in \mathbb{R} \times [0, T] : |v(x+, t) - v(x-, t)| > \varepsilon\},$$

the set of shocks of strength larger than ε . For a small $\delta > 0$ let the $\mathcal{O}_\varepsilon^\delta$ denote the δ -neighbourhood of \mathcal{O}_ε . We have the following approximation result.

Theorem 6.3.2. *Let assumptions of Theorem 6.3.1 hold. Assume, for some small $\varepsilon, \delta > 0$, that the data $y = (y_1, \dots, y_J)$ is given as in (6.26) with (x_j, t_j) , $j = 1, \dots, J$, lying outside $\mathcal{O}_\varepsilon^\delta$. Let μ and μ^N be given by (6.11) and (6.12) respectively with G and G^N defined as in (6.25) and (6.27). Suppose that S_e and S_e^N in (6.25) satisfy, for any fixed $u \in X$, $\|S_e(u) - S_e^N(u)\|_{L^1} \rightarrow 0$ as $N \rightarrow \infty$. Then*

$$d_{\text{Hell}}(\mu^N, \mu) \leq C\varepsilon,$$

for N sufficiently large and with C depending on J and the covariance of the measurement noise.

Proof. For a given $u \in X$, by [BL99, Theorem 5], for large enough N , the shocks of strength larger than ε of $S_e^N(u)$ are in a δ -neighbourhood of shocks of strength larger than ε of $S_e(u)$. Therefore outside $\mathcal{O}_\varepsilon^\delta$ both $S_e(u)$ and $S_e^N(u)$ only have shocks of strength at most ε .

For each N , take an integrable, bounded and continuous function, denoted by $\tilde{S}^N(u)$, such that $\tilde{S}^N(u)$ lies in the ε -neighbourhood of $S_e^N(u)$. Since $S_e^N(u)$ converges to $S_e(u)$ almost everywhere along a subsequence, $\tilde{S}^N(u)$ converges everywhere, along a subsequence, to a bounded continuous function, denoted by $\tilde{S}(u)$, such that $\tilde{S}(u)$ lies in ε -neighbourhood of $S_e(u)$. We can always choose $\tilde{S}^N(u)$ differently in the ε -neighbourhood of $S_e^N(u)$ such that the whole sequence $\tilde{S}^N(u)$ converges everywhere to $\tilde{S}(u)$. It therefore follows that

$$\|S_e^N(u) - S_e(u)\|_{L^\infty} \leq 2\varepsilon,$$

for N sufficiently large. The result then follows arguing along the lines of the proof of Theorem 6.1.2. \square

Remark 6.3.3. *In the case that no shocks are present, one may take $\varepsilon = 0$ and obtain the desirable convergence with rate $\psi(N)$. However, it may not be feasible to know a priori if shocks appear in certain areas, and the shock strengths may be too small that the measurement devices cannot detect them.*

6.4 Further discussions

In this section, we discuss an alternative way of applying the Bayesian framework for inverse problems and modify the framework such that it works in a more general space of functions, the Frechet space. We present an application for inverse problems in scalar conservation laws with direct measurement data.

6.4.1 A modified framework

We develop in this section a slightly different framework that helps to approximate the posterior by means of potential functions, for general data (shock-free or not). This works especially when we can have an amount of data that is as large as we wish.

Assume that we want to recover u from a set of data $v_J = (v_{(1)}, v_{(2)}, \dots, v_{(J)})$, with $J \in \mathbb{N}$ big enough. Let $G(u), G^N(u) : X \rightarrow \mathbb{R}^J$ be the measurement maps and assume that the measurement data are now given by

$$y = \frac{1}{J} \begin{pmatrix} v_{(1)} \\ v_{(2)} \\ \dots \\ v_{(J)} \end{pmatrix} + \xi = \frac{1}{J} G(u) + \xi, \quad (6.28)$$

and similarly for $G^N(u)$. We assume that the noise is Gaussian distributed

$$\xi \sim \mathcal{N}(0, \Gamma),$$

where Γ is the covariance operator. We make the following assumption regarding G and G^N (see the next section for the motivation).

Assumption 6.4.1. *The operators G, G^N are bounded. In addition, there exists a function $\psi : \mathbb{N} \rightarrow \mathbb{R}^+$ which is decreasing to 0 as N goes to ∞ , and a constant C , independent of N , such that*

$$|G_j(u) - G_j^N(u)| \leq C\psi(N), \quad (6.29)$$

for every $j = 1, \dots, J$ except a few $j \in S$ where S is a fixed (finite) set with $|S| < J$ and is independent of J .

This assumption allows the measurement operators to be even discontinuous, yet bounded. We have the following result.

Theorem 6.4.2. *Let X be a Frechet space whose topology is induced by the metric d_X . Assume that μ and μ^N are defined respectively by (6.11)-(6.13) and (6.12)-(6.14), where*

μ_0 is a probability measure on X . Let $G(u)$ and $G^N(u)$ satisfy Assumption 6.4.1. Then the measures μ and μ^N are close with respect to d : there is a constant C such that

$$d(\mu, \mu^N) \leq C \left(\frac{1}{J} + \frac{\psi(N)}{J} \right),$$

where d is Hellinger or Wasserstein distance.

Proof. We consider here d to be the Wasserstein distance. The case of Hellinger distance can be proved similarly. Throughout the proof C is a constant may change from line to line. Since the measurement data is finite, the potential Φ and Φ^N are given as

$$\Phi(u) = \frac{1}{2}|y - \frac{1}{J}G(u)|_{\Gamma}^2, \quad \Phi^N(u) = \frac{1}{2}|y - \frac{1}{J}G^N(u)|_{\Gamma}^2,$$

where $|\cdot|_{\Gamma} = |\Gamma^{-1/2} \cdot|$ is the weighted norm on \mathbb{R}^J . Hence,

$$\begin{aligned} |\Phi(u) - \Phi^N(u)| &\leq \frac{1}{2J}|G(u) - G^N(u)|_{\Gamma}|2y - \frac{1}{J}G(u) - \frac{1}{J}G^N(u)|_{\Gamma} \\ &\leq \frac{1}{2J}(2|y|_{\Gamma} + \max |G_j| + \max |G_j^N|)|G(u) - G^N(u)|_{\Gamma} \\ &\leq \frac{1}{J}C|G(u) - G^N(u)|_{\Gamma}, \end{aligned}$$

thanks to the boundedness of G and G^N . We then have

$$\int_X (1 + d_X(u, 0))|\Phi(u) - \Phi^N(u)|\mu_0(du) \leq \int_X \frac{1}{J}C(1 + d_X(u, 0))|G(u) - G^N(u)|_{\Gamma}\mu_0(du). \quad (6.30)$$

Without loss of generality, assume that Γ has the diagonal form of the eigenvalues $(\gamma_1^2, \dots, \gamma_J^2)$, where $\gamma_i > 0$. Thus,

$$\begin{aligned} |G(u) - G^N(u)|_{\Gamma} &= \sqrt{\sum_{i=1}^J \frac{1}{\gamma_i^2} |G_i(u) - G_i^N(u)|^2} \\ &\leq \sum_{i=1}^J \frac{1}{\gamma_i} |G_i(u) - G_i^N(u)| \\ &\leq \sum_{i \in S} \frac{1}{\gamma_i} |G_i(u) - G_i^N(u)| + \sum_{i \notin S} \frac{1}{\gamma_i} |G_i(u) - G_i^N(u)|. \end{aligned} \quad (6.31)$$

We now see that, since $|S| < \infty$ and G_i, G_i^N are bounded, we have

$$\sum_{i \in S} \frac{1}{\gamma_i} |G_i(u) - G_i^N(u)| \leq |S|C \frac{1}{\min_i \gamma_i}, \quad (6.32)$$

By virtue of (6.29), it is clear that

$$\sum_{i \notin S} \frac{1}{\gamma_i} |G_i(u) - G_i^N(u)| \leq C \frac{1}{\min_i \gamma_i} \psi(N). \quad (6.33)$$

From (6.31)–(6.33) we obtain

$$|G(u) - G^N(u)|_\Gamma \leq C(1 + \psi(N)).$$

Combining with (6.30) and using Fernique's theorem we arrive at

$$\begin{aligned} \int_X (1 + d_X(u, 0)) |\Phi(u) - \Phi^N(u)| \mu_0(du) &\leq \int_X \frac{1}{J} C(1 + d_X(u, 0))(1 + \psi(N)) \mu_0(du) \\ &\leq C \left(\frac{1}{J} + \frac{\psi(N)}{J} \right). \end{aligned} \quad (6.34)$$

Now using the inequality $|e^{-x} - e^{-y}| \leq |x - y|$ for all $x, y \geq 0$ and (6.34), we see that

$$\begin{aligned} |Z - Z^N| &\leq \int_X |\exp(-\Phi(u)) - \exp(-\Phi^N(u))| \mu_0(du) \\ &\leq \int_X |\Phi(u) - \Phi^N(u)| \mu_0(du) \\ &\leq \int_X (1 + d_X(u, 0)) |\Phi(u) - \Phi^N(u)| \mu_0(du) \\ &\leq C \left(\frac{1}{J} + \frac{\psi(N)}{J} \right). \end{aligned} \quad (6.35)$$

Now, from the definition of Wasserstein distance,

$$\begin{aligned} d_W(\mu, \mu^N) &= \sup_{\|h\|_{Lip} \leq 1} \left| \int_X h d\mu - \int_X h d\mu^N \right| \\ &= \sup_{\|h\|_{Lip} \leq 1} \left| \int_X h(u) \frac{\exp(-\Phi(u))}{Z} - \frac{\exp(-\Phi^N(u))}{Z^N} \mu_0(du) \right| \\ &\leq I_1 + I_2, \end{aligned} \quad (6.36)$$

where

$$\begin{aligned} I_1 &= \frac{1}{Z} \sup_{\|h\|_{Lip} \leq 1} \int_X |h(u)| |\exp(-\Phi(u)) - \exp(-\Phi^N(u))| \mu_0(du) \\ &\leq C \sup_{\|h\|_{Lip} \leq 1} \int_X |h(u)| |\Phi(u) - \Phi^N(u)| \mu_0(du) \\ &\leq C \int_X (1 + d_X(u, 0)) |\Phi(u) - \Phi^N(u)| \mu_0(du) \\ &\leq C \left(\frac{1}{J} + \frac{\psi(N)}{J} \right), \end{aligned} \quad (6.37)$$

thanks to previous estimates and (6.19); and

$$\begin{aligned} I_2 &= |Z^{-1} - (Z^N)^{-1}| \sup_{\|h\|_{Lip} \leq 1} \int_X |h(u)| \exp(-\Phi^N(u)) \mu_0(du) \\ &\leq C |Z - Z^N| \int_X (1 + d(u, 0)) \mu_0(du) \\ &\leq C \left(\frac{1}{J} + \frac{\psi(N)}{J} \right), \end{aligned} \quad (6.38)$$

thanks to Fernique's theorem and (6.35). Therefore, by gathering (6.36)–(6.38), we arrive at

$$d_W(\mu, \mu^N) \leq C \left(\frac{1}{J} + \frac{\psi(N)}{J} \right).$$

This completes the proof. \square

Remark 6.4.3. • *This result shows that if J is fixed but reasonably large, we can keep the two posteriors reasonably close. This is an attempt to get some convergent-like estimate for the posterior when the solution of the forward problem develops discontinuities. The task is challenging and it seems to be the best we can get if the data is collected directly from the solution.*

- *The Fernique theorem used in the proof of Theorem 6.4.2 is in fact a general version of the classical theorem, where the underlining space X is not necessarily a separable Banach space but a locally convex topological space. See [Bog98] for a proof of this general version.*
- *One can also prove that, by processing similarly, the approximation converges in Hellinger distance with the same rate*

$$d_{\text{Hell}}(\mu, \mu^N) \leq C \left(\frac{1}{J} + \frac{\psi(N)}{J} \right).$$

- *We note also that, for the Wasserstein distance it is not expected to have the same convergence of expectation as for Hellinger's.*

6.4.2 Applications

In this section we illustrate some situations of conservation laws which satisfies the conditions required for the approximation we have discussed in the last subsection. Before moving to a direct example, we discuss briefly the nature of shocks of conservation laws.

Despite the simple form of the scalar conservation law, the understanding of its shock set for general initial data and flux is still limited. In [Liu81], Liu shows that the set of points of discontinuity of a bounded variation entropy solution of (6.1)-(6.2) which is constructed by the random choice method of Glimm [Gli65] is the countable union of Lipschitz arcs. Consequently, the set of shocks is countable. Beyond that, there is no information on the global structure of the shock set [Daf85].

However, when the initial data and the flux is sufficiently smooth, the shock set of the solution could behave quite “nicely”. More precisely, for *generic* initial data, the solution of the scalar conservation law has a finite set of discontinuities. This is due to the following regularity result by Schaeffer (see [Sch73], also Theorem 2.5.3).

Lemma 6.4.4 (Schaeffer). *Denote $\mathcal{S}(\mathbb{R})$ to be the Schwartz space. Assume that f is smooth and uniformly convex. Then there exists a set $\mathfrak{F} \subset \mathcal{S}(\mathbb{R})$ which is open and dense in $\mathcal{S}(\mathbb{R})$ such that for every $v_0 \in \mathfrak{F}$, the entropy solution v of the scalar conservation law*

(6.1)-(6.2) is piecewise smooth: the number of shocks of v is finite and v is smooth in the complement of the shock set.

With this regularity result at hand, we are able to apply the approximation procedure given in the last subsection, employing the new setting (6.28). This gives a convergence result for the approximation of the inverse problem of finding initial data in $\mathcal{S}(\mathbb{R})$. Recall that the Schwartz space $\mathcal{S}(\mathbb{R})$ is defined as follows

$$\mathcal{S}(\mathbb{R}) = \{f \in C^\infty(\mathbb{R}, \mathbb{R}) : \|f\|_{\alpha, \beta} := \sup_{x \in \mathbb{R}} |x^\alpha (D^\beta f)(x)| < \infty, \text{ for all } \alpha, \beta \in \mathbb{N}\}.$$

The family of semi-norms $\{\|\cdot\|_{\alpha, \beta}, \alpha, \beta \in \mathbb{N}\}$ is countable, hence $\mathcal{S}(\mathbb{R})$ is metrizable. We define the metric d_X on $\mathcal{S}(\mathbb{R})$ by

$$d_X(u, v) = \sum_{\alpha + \beta =: k=0}^{\infty} 2^{-k} \frac{\|u - v\|_{\alpha, \beta}}{1 + \|u - v\|_{\alpha, \beta}}, \quad u, v \in \mathcal{S}(\mathbb{R}).$$

Assume now that v_u is the solution of (6.1)-(6.2) with initial data $v_0 = u \in X = \mathcal{S}(\mathbb{R})$ and v_u^N is the solution of the approximation PDE of (6.1)-(6.2) such that

$$|v_u^N - v_u| \leq C\psi(N), \quad (6.39)$$

at every continuity point of v_u and v_u^N , where $\psi(N) \rightarrow 0$ as $N \rightarrow \infty$. Let $\mathfrak{F} \in \mathcal{S}$ be the set given in Lemma 6.4.4. Denote

$$S = \{i : v_{(i)} \text{ or } v_{(i)}^N \text{ are collecting from discontinuities of } v_u, v_u^N \text{ for } u \in \mathfrak{F}\},$$

hence (6.39) may not be satisfied in S . Thanks to Lemma 6.4.4, the number of shocks is essentially finite, hence J can be chosen large enough such that $|S| < J$. Let μ_0 be a probability measure on $X = \mathcal{S}$ whose support lies outside $X \setminus \mathfrak{F}$. Hence, $\mu_0(X \setminus \mathfrak{F}) = 0$ and then

$$\int_X (1 + d_X(u, 0)) |\Phi(u) - \Phi^N(u)| \mu_0(du) = \int_{\mathfrak{F}} (1 + d_X(u, 0)) |\Phi(u) - \Phi^N(u)| \mu_0(du)$$

Therefore, all assumptions of Theorem 6.4.2 still hold. We come to the following result.

Corollary 6.4.5. *Assume that μ is the solution of the Bayesian inverse problem (6.3) of finding $u = v_0$ in (6.1)-(6.2), which is defined by (6.11)-(6.13) where μ_0 is a probability measure on the metric space (\mathcal{S}, d_X) whose support lies outside $\mathcal{S} \setminus \mathfrak{F}$. Let μ^N be defined by (6.12)-(6.14). Then μ^N is an approximation of μ in the sense that: there is a constant C such that*

$$d(\mu, \mu^N) \leq C \left(\frac{1}{J} + \frac{\psi(N)}{J} \right),$$

where d is Hellinger or Wasserstein distance.

In many situations in practice, we may only obtain the convergence without knowing the rate. In other words, we cannot transfer the rate of approximation of the forward PDE into the rate of approximation of the corresponding inverse problem, but the continuity is still obtained.

6.5 Summary and notes

Due to the nature of shocks and irreversibility, inverse problems of the physical phenomena involving these features are challenging. There have been only a few attempts in studying inverse problems for hyperbolic conservation laws, despite their wide range applications. Some of the work addressing some particular aspects of these inverse problems may be found in [JS99, CJS03, KT05, CZ11, BD13, HPR14, BCD17, CP20, LZ20].

This chapter investigates the Bayesian inverse problems for scalar conservation laws in a systematic way. Two types of general inverse problems of recovering the initial field and flux function are considered. Based on the study of the particle trajectory theory for the forward problems in Chapter 5, it is proved that the posterior for each inverse problem is well-defined and stable with respect to data. The approximations of the posterior in terms of the forward map are also addressed thoroughly. The presentation here is an expansion of part of our work [DD20].

The discussions in Section 6.4 (not included in [DD20]), where we talk about the Bayesian framework in Frechet space, are barely touching the surface. Questions regarding how to construct a probability prior on X are not yet studied.

There are many other open and interesting problems related to the Bayesian inverse problems that we consider in this chapter. Some of them are listed in Chapter 7.

Chapter 7

Conclusion and Outlook

7.1 Conclusion

In this thesis, we employ a Bayesian approach to study inverse problems in scalar conservation laws. Inverse problems involving hyperbolic conservation laws are challenging due to the development of shock waves and their irreversibility. These characteristic properties of scalar conservation laws are featured in Chapter 2, where we summarised the most significant aspects of scalar conservation laws in a concise and informative way. Background on inverse problems, including regularization methods and the Bayesian approach to inverse problems, are outlined in Chapter 3.

The heart of the thesis lies in Chapter 5 and Chapter 6. Chapter 5 is devoted to the study of the particle trajectories for scalar conservation laws. Motivated from a model of traffic flow, we consider some ordinary differential equations where their right-hand side depends on solutions of scalar conservation laws. Despite the presence of shocks, by using Filippov theory for differential equations, which was outlined in Chapter 4, it is showed that the particle trajectories are well-posed. Moreover, we prove that this particle trajectory theory is compatible with the front tracking approximation method and the vanishing viscosity method in the sense that the approximate particle trajectories given by either of these methods converge uniformly to the particle trajectories corresponding to the entropy solution of scalar conservation laws. For certain flux functions, illustrated by traffic flow, we are able to obtain the convergence rate for the approximate particle trajectories with respect to changes in the initial field or flux function. This task is done by combining the front tracking method with Filippov theory.

As the second main part of the thesis, recorded in Chapter 6, we study some Bayesian inverse problems for scalar conservation laws and establish several well-posedness and ap-

proximation results. Specifically, we consider two types of inverse problems: the inverse problem of recovering the upstream field and the inverse problem of finding the flux function, both from observations of appropriate functions of the entropy solutions of scalar conservation laws. Based on the theory of particle trajectories developed in the first part of the thesis and the Bayesian inversion theory developed by Stuart *et.al.*, we prove that the statistical solutions to these inverse problems are well-posed and stable with respect to changes in the forward model. Rates of convergence of the posterior approximations are also given for certain inverse problems.

Finally, for the data given from measuring directly the entropy solution, we propose some different ways to approximate the posterior without assuming the strong continuity of the forward model. Some of those are such as measuring the data in L^1 balls, collecting shock-free data, or adjusting the Bayesian framework.

7.2 Outlook and future work

We list here some interesting research questions that are closely related to the work of this thesis and that we would like to pursue in the near future.

7.2.1 The consistency problem

An important problem arising after the work of this thesis is the consistency (a statistical property) of the inverse problems. We want to know if the underlying truth can be recovered provided that we are able to collect as many data points as we wish. To this end, we need some stability structures of the inverse maps $z \mapsto \rho_0$ and $z \mapsto f$. The L^1 contraction property of the forward map provides a good picture and may help to understand these structures. This is our work in progress [DD21].

7.2.2 The non-convex flux case

One of our assumptions on f is that it is either strictly convex or strictly concave. However, there are models of scalar conservation laws in which neither of the above is the case, such as the Buckley-Leverett equation with applications in oil recovery industry. When f has inflection points, the structure of solutions becomes more intricate. Also, the famous Oleinik decay estimate fails, resulting in possibly the non-uniqueness of the particle trajectories. In [Mar04], the author considered the flux function with a single inflection point and obtained a well-posedness result for the Filippov solution. The stability properties

of the trajectories when either the initial field or the flux function is perturbed are both open for future studies.

7.2.3 Rate of convergence for vanishing viscosity approximation

We establish in this thesis the convergence of the approximate trajectory arising from the vanishing viscosity method. However, having a rate for that convergence is still an open problem. To be able to go further, one may need to have a deeper understanding of the vanishing viscosity approximation and how it relates to the structural properties of the entropy solution.

7.2.4 Numerical experiments

The work that has been carried out in this thesis is mostly theoretical. It is important to perform numerical experiments to illustrate these results. This may require some special techniques developed for hyperbolic conservation laws.

Moreover, the study of the statistical solutions for inverse problems in order to extract information from them is not easy and typically involves numerical methods, such as MCMC method. I plan to carry this task in my future work.

7.2.5 Other related questions

- Finding the optimal rates for the convergence of the approximate trajectories (it is not currently known if $\epsilon^{1/2}$ is optimal).
- Extending to the multi-dimensional problems (when $x \in \mathbb{R}^d$, for $d > 1$).

I also plan to apply the Bayesian inversion framework to other PDE models, such as problems from electrical impedance tomography (the Calderón problem).

Bibliography

- [AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*, volume 254. Clarendon Press Oxford, 2000. [18](#)
- [AGN12] Fabio Ancona, Olivier Glass, and Khai T. Nguyen. Lower compactness estimates for scalar balance laws. *Communications on Pure and Applied Mathematics*, 65(9):1303–1329, 2012. [23](#)
- [Amb04] Luigi Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Inventiones Mathematicae*, 158(2):227–260, 2004. [66](#)
- [AN19] Kweku Abraham and Richard Nickl. On statistical caldéron problems. *Mathematical Statistics and Learning (to appear)*, *arXiv preprint arXiv:1906.03486*, 2019. [37](#)
- [AVK66] Aleksandr Aleksandrovich Andronov, Aleksandr Adol’fovich Vitt, and Semen Emmanuilovich Khaikin. *Theory of Oscillators: Adiwes International Series in Physics*, volume 4. Elsevier, 1966. [38](#)
- [Bak67] Anatolii Borisovich Bakushinskii. A general method of constructing regularizing algorithms for a linear incorrect equation in hilbert space. *Zh. Vychisl. Mat. Mat. Fiz.*, 7:672–677, 1967. [37](#)
- [Bal12] Guillaume Bal. Introduction to inverse problems. *Lecture Notes-Department of Statistics and Mathematics, University of Chicago*, 2012. [30](#), [37](#)
- [Bat15] Harry Bateman. Some recent researches on the motion of fluids. *Monthly Weather Review*, 43(4):163–170, 1915. [10](#)
- [BB05] Stefano Bianchini and Alberto Bressan. Vanishing viscosity solutions of non-linear hyperbolic systems. *Annals of Mathematics*, pages 223–342, 2005. [24](#)

- [BB18] Martin Benning and Martin Burger. Modern regularization methods for inverse problems. *Acta Numerica*, 27:1–111, 2018. [37](#)
- [BCD17] Raimund Bürger, Julio Careaga, and Stefan Diehl. Entropy solutions of a scalar conservation law modeling sedimentation in vessels with varying cross-sectional area. *SIAM Journal on Applied Mathematics*, 77(2):789–811, 2017. [84](#)
- [BD13] Raimund Bürger and Stefan Diehl. Convexity-preserving flux identification for scalar conservation laws modelling sedimentation. *Inverse Problems*, 29(4):045008, 2013. [84](#)
- [BGS07] Sylvie Benzoni-Gavage and Denis Serre. *Multi-dimensional Hyperbolic Partial Differential Equations: First-order Systems and Applications*. Oxford University Press, 2007. [24](#)
- [BL99] Alberto Bressan and Philippe G LeFloch. Structural stability and regularity of entropy solutions to hyperbolic systems of conservation laws. *Indiana Univ. Math. J.*, 48(1):43–84, 1999. [3](#), [51](#), [52](#), [61](#), [77](#), [78](#)
- [BM17] Stefano Bianchini and Elio Marconi. On the structure of L^∞ -entropy solutions to scalar conservation laws in one-space dimension. *Archive for Rational Mechanics and Analysis*, 226(1):441–493, 2017. [4](#), [17](#)
- [Bog98] Vladimir I. Bogachev. *Gaussian Measures*. Mathematical Surveys and Monographs. AMS, 1998. [82](#)
- [Bre92] Alberto Bressan. Global solutions of systems of conservation laws by wave-front tracking. *Journal of Mathematical Analysis and Applications*, 170(2):414–432, 1992. [24](#)
- [Bre00] Alberto Bressan. *Hyperbolic Systems of Conservation Laws: The One-dimensional Cauchy Problem*, volume 20. Oxford University Press, 2000. [2](#), [20](#), [21](#), [24](#)
- [Bre03] Alberto Bressan. An ill posed Cauchy problem for a hyperbolic system in two space dimensions. *Rendiconti del Seminario Matematico della Università di Padova*, 110:103–117, 2003. [66](#)
- [Bre10] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer Science & Business Media, 2010. [22](#), [28](#)

- [BS98] Alberto Bressan and Wen Shen. Uniqueness for discontinuous ODE and conservation laws. *Nonlinear Analysis*, 34(5):637–652, 1998. [66](#)
- [BT20] Bernard Brogliato and Aneel Tanwani. Dynamical systems coupled with monotone set-valued operators: Formalisms, applications, well-posedness, and stability. *SIAM Review*, 62(1):3–129, 2020. [46](#)
- [BTGMS13] Tan Bui-Thanh, Omar Ghattas, James Martin, and Georg Stadler. A computational framework for infinite-dimensional Bayesian inverse problems, Part I: The linearized case, with application to global seismic inversion. *SIAM Journal on Scientific Computing*, 35(6):A2494–A2523, 2013. [37](#)
- [Bur48] Johannes Martinus Burgers. A mathematical model illustrating the theory of turbulence. In *Advances in Applied Mechanics*, volume 1, pages 171–199. Elsevier, 1948. [10](#)
- [CDRS09] Simon L Cotter, Massoumeh Dashti, James C Robinson, and Andrew M Stuart. Bayesian inverse problems for functions and applications to fluid mechanics. *Inverse Problems*, 25(11):115008, 2009. [35](#), [37](#)
- [CDS10] Simon L Cotter, Massoumeh Dashti, and Andrew M Stuart. Approximation of Bayesian inverse problems for PDEs. *SIAM J. Numer. Anal.*, 48(1):322–345, 2010. [35](#), [37](#)
- [CEL84] Michael G Crandall, Lawrence C Evans, and P-L Lions. Some properties of viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 282(2):487–502, 1984. [22](#)
- [CJS03] Aníbal Coronel, François James, and Mauricio Sepúlveda. Numerical identification of parameters for a model of sedimentation processes. *Inverse Problems*, 19(4):951, 2003. [84](#)
- [CL55] Earl A Coddington and Norman Levinson. *Theory of Ordinary Differential Equations*. Tata McGraw-Hill Education, 1955. [46](#)
- [CL71] Michael G Crandall and Thomas M Liggett. Generation of semi-groups of nonlinear transformations on general Banach spaces. *American Journal of Mathematics*, 93(2):265–298, 1971. [24](#)

- [CM03] Rinaldo M Colombo and Andrea Marson. A Hölder continuous ODE related to traffic flow. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 133(4):759–772, 2003. [6](#), [49](#), [66](#)
- [COW08] Gianluca Crippa, Felix Otto, and Michael Westdickenberg. Regularizing effect of nonlinearity in multidimensional scalar conservation laws. In *Transport equations and multi-D hyperbolic conservation laws*, pages 77–128. Springer, 2008. [17](#), [23](#)
- [CP20] Rinaldo M Colombo and Vincent Perrollaz. Initial data identification in conservation laws and Hamilton–Jacobi equations. *Journal de Mathématiques Pures et Appliquées*, 138:1–27, 2020. [84](#)
- [Cra72] Michael G Crandall. The semigroup approach to first order quasilinear equations in several space variables. *Israel Journal of Mathematics*, 12(2):108–132, 1972. [24](#)
- [CS17] Laura Caravenna and Laura V. Spinolo. Schaeffer’s regularity theorem for scalar conservation laws does not extend to systems. *Indiana Univ. Math. J.*, 66(1):101–160, 2017. [24](#)
- [CZ11] Carlos Castro and Enrique Zuazua. Flux identification for 1-D scalar conservation laws in the presence of shocks. *Mathematics of computation*, 80(276):2025–2070, 2011. [84](#)
- [Daf72] Constantine M Dafermos. Polygonal approximations of solutions of the initial value problem for a conservation law. *Journal of Mathematical Analysis and Applications*, 38(1):33–41, 1972. [19](#)
- [Daf77] Constantine M Dafermos. Generalized characteristics and the structure of solutions of hyperbolic conservation laws. *Indiana Univ. Math. J.*, 26(6):1097–1119, 1977. [3](#), [6](#), [48](#)
- [Daf85] Constantine M Dafermos. Regularity and large time behaviour of solutions of a conservation law without convexity. *Proc. Roy. Soc. Edinburgh Sect. A*, 99(3-4):201–239, 1985. [24](#), [77](#), [82](#)
- [Daf16] Constantine M Dafermos. *Hyperbolic Conservation Laws in Continuum Physics*, 4th ed. Springer, 2016. [2](#), [3](#), [13](#), [15](#), [17](#), [22](#), [23](#), [24](#), [60](#), [61](#), [65](#)

- [DD20] Masoumeh Dashti and Duc Lam Duong. Stability of particle trajectories of scalar conservation laws and applications in Bayesian inverse problems. *Submitted*, 2020. [66](#), [84](#)
- [DD21] Masoumeh Dashti and Duc Lam Duong. Consistency of Bayesian inverse problems for scalar conservation laws. *work in progress*, 2021. [86](#)
- [DiP76] Ronald J DiPerna. Global existence of solutions to nonlinear hyperbolic systems of conservation laws. *Journal of Differential Equations*, 20:187–212, 1976. [24](#)
- [DIS17] Matthew M Dunlop, Marco A Iglesias, and Andrew M Stuart. Hierarchical Bayesian level set inversion. *Statistics and Computing*, 27(6):1555–1584, 2017. [37](#)
- [DL89] Ronald J DiPerna and Pierre-Louis Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Inventiones Mathematicae*, 98(3):511–547, 1989. [66](#)
- [DLOW03] Camillo De Lellis, Felix Otto, and Michael Westdickenberg. Structure of entropy solutions for multi-dimensional scalar conservation laws. *Archive for Rational Mechanics and Analysis*, 170(2):137–184, 2003. [17](#)
- [DLR03] Camillo De Lellis and Tristan Riviere. The rectifiability of entropy measures in one space dimension. *Journal de Mathématiques Pures et Appliquées*, 82(10):1343–1367, 2003. [17](#)
- [DLSV13] Masoumeh Dashti, Kody J.H. Law, Andrew M. Stuart, and Jochen Voss. MAP estimators and their consistency in Bayesian nonparametric inverse problems. *Inverse Problems*, 29(9):095017, 2013. [37](#)
- [DMG14] Maria Laura Delle Monache and Paola Goatin. Scalar conservation laws with moving constraints arising in traffic flow modeling: an existence result. *Journal of Differential Equations*, 257(11):4015–4029, 2014. [51](#), [66](#)
- [DS11] Masoumeh Dashti and Andrew M. Stuart. Uncertainty quantification and weak approximation of an elliptic inverse problem. *SIAM Journal on Numerical Analysis*, 49(6):2524–2542, 2011. [33](#)

- [DS16] Masoumeh Dashti and Andrew M. Stuart. The Bayesian Approach to Inverse Problems. In R. Ghanem, D. Higdon, and H. Owhadi, editors, *Handbook of Uncertainty Quantification*. Springer, 2016. [3](#), [32](#), [33](#), [35](#), [36](#), [37](#)
- [EHN96] Heinz Werner Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375. Springer Science & Business Media, 1996. [25](#), [30](#), [37](#)
- [Eva10] Lawrence C. Evans. *Partial Differential Equations, 2nd ed.* Graduate Studies in Mathematics. AMS, 2010. [13](#), [15](#), [21](#), [22](#)
- [Fil88] Aleksei Fedorovich Filippov. *Differential Equations with Discontinuous Righthand Sides*. Mathematics and its Applications. Kluwer Academic Publishers, Norwell, MA, 1988. [3](#), [38](#), [39](#), [40](#), [41](#), [42](#), [43](#), [46](#), [48](#)
- [Fra70] Joel N. Franklin. Well-posed stochastic extensions of ill-posed linear problems. *Journal of Mathematical Analysis and Applications*, 31(3):682–716, 1970. [3](#), [37](#)
- [FSU19] Joel Feldman, Mikko Salo, and Gunther Uhlmann. Calderón problem: An introduction to inverse problems. *Preliminary notes on the book in preparation*, 2019. [30](#)
- [Gla08] Olivier Glass. An extension of Oleinik’s inequality for general 1D scalar conservation laws. *Journal of Hyperbolic Differential Equations*, 5(1):113, 2008. [23](#)
- [Gli65] James Glimm. Solutions in the large for nonlinear hyperbolic systems of equations. *Comm. Pure Appl. Math.*, 18:697–715, 1965. [24](#), [82](#)
- [Gro93] Charles W Groetsch. *Inverse Problems in the Mathematical Sciences*, volume 52. Springer, 1993. [25](#)
- [GS02] Alison L. Gibbs and Francis Edward Su. On choosing and bounding probability metrics. *International Statistical Review*, 70(3):419–435, 2002. [32](#)
- [GTSA17] Nicolas Garcia Trillos and Daniel Sanz-Alonso. The Bayesian formulation and well-posedness of fractional elliptic inverse problems. *Inverse Problems*, 33(6):065006, 2017. [37](#)

- [Had02] Jacques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902. [27](#)
- [HHHK88] Helge Holden, Lars Holden, and Raphael Høegh-Krohn. A numerical method for first order nonlinear scalar conservation laws in one-dimension. *Comput. Math. Appl.*, 15:595–602, 1988. [19](#)
- [Hof83] David Hoff. The sharp form of Oleinik’s entropy condition in several space variables. *Transactions of the American Mathematical Society*, 276(2):707–714, 1983. [23](#)
- [HOH10] Harald Hanche-Olsen and Helge Holden. The Kolmogorov–Riesz compactness theorem. *Expositiones Mathematicae*, 28(4):385–394, 2010. [22](#)
- [Hop50] Eberhard Hopf. The partial differential equation $u_t + uu_x = \mu u_{xx}$. *Communications on Pure and Applied Mathematics*, 3(3):201–230, 1950. [10](#)
- [HPR14] Helge Holden, Fabio Simone Priuli, and Nils Henrik Risebro. On an inverse problem for scalar conservation laws. *Inverse Problems*, 30(3):035015, 2014. [75](#), [84](#)
- [HR15] Helge Holden and Nils H. Risebro. *Front Tracking for Hyperbolic Conservation Laws, 2nd ed.* Applied Mathematical Sciences. Springer, 2015. [2](#), [3](#), [17](#), [21](#)
- [ILS14] Marco A Iglesias, Kui Lin, and Andrew M Stuart. Well-posed Bayesian geometric inverse problems arising in subsurface flow. *Inverse Problems*, 30(11):114001, 2014. [37](#)
- [ILS16] Marco A Iglesias, Yulong Lu, and Andrew M Stuart. A Bayesian level set method for geometric inverse problems. *Interfaces and Free Boundaries*, 18(2):181–217, 2016. [37](#)
- [Isa06] Victor Isakov. *Inverse Problems for Partial Differential Equations, 2nd ed.* Applied Mathematical Sciences. Springer, 2006. [25](#), [30](#)
- [Iva62] Valentin Konstantinovich Ivanov. On linear problems which are not well-posed. *Soviet Math. Dokl.*, 145:270–272, 1962. [37](#)
- [JS99] François James and Mauricio Sepúlveda. Convergence results for the flux identification in a scalar conservation law. *SIAM Journal on Control and Optimization*, 37(3):869–891, 1999. [84](#)

- [JYPG18] Junxiong Jia, Shigang Yue, Jigen Peng, and Jinghuai Gao. Infinite-dimensional bayesian approach for inverse scattering problems of a fractional helmholtz equation. *Journal of Functional Analysis*, 275(9):2299–2332, 2018. [37](#)
- [Kir11] Andreas Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems*, volume 120. Springer Science & Business Media, 2011. [25](#), [30](#), [37](#)
- [Kru70] Stanislav Nikolaevich Kruzkov. First order quasilinear equations in several independent variables. *Math. USSR Sbornik*, 10:217–243, 1970. [16](#), [18](#), [22](#)
- [KS05] Jari Kaipio and Erkki Somersalo. *Statistical and Computational Inverse problems*. Applied Mathematical Sciences. Springer, 2005. [3](#), [37](#)
- [KS07] Jari Kaipio and Erkki Somersalo. Statistical inverse problems: discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics*, 198(2):493–504, 2007. [37](#)
- [KT05] Hyeonbae Kang and Kazumi Tanuma. Inverse problems for scalar conservation laws. *Inverse Problems*, 21(3):1047, 2005. [20](#), [84](#)
- [Las07] Sari Lasanen. Measurements and infinite-dimensional statistical inverse theory. In *PAMM: Proceedings in Applied Mathematics and Mechanics*, volume 7, pages 1080101–1080102. Wiley Online Library, 2007. [37](#)
- [Lat20] Jonas Latz. On the well-posedness of bayesian inverse problems. *SIAM/ASA Journal on Uncertainty Quantification (accepted)*, *arXiv preprint arXiv:1902.10257*, 2020. [35](#), [37](#)
- [Lax57] Peter D Lax. Hyperbolic systems of conservation laws II. *Communications on Pure and Applied Mathematics*, 10(4):537–566, 1957. [16](#)
- [LeF02] Philippe G LeFloch. *Hyperbolic Systems of Conservation Laws*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2002. [24](#)
- [Leg11] Nicholas Leger. L^2 stability estimates for shock solutions of scalar conservation laws using the relative entropy method. *Archive for Rational Mechanics and Analysis*, 199(3):761–778, 2011. [3](#), [49](#)
- [Leo17] Giovanni Leoni. *A first course in Sobolev spaces, 2nd ed.* Graduate Studies in Mathematics. AMS, 2017. [18](#)

- [Liu81] Tai-Ping Liu. Admissible solutions of hyperbolic conservation laws. *Mem. Amer. Math. Soc.*, 240, 1981. [82](#)
- [LMP11] Corrado Lattanzio, Amelio Maurizi, and Benedetto Piccoli. Moving bottlenecks in car traffic flow: a PDE-ODE coupled model. *SIAM Journal on Mathematical Analysis*, 43(1):50–67, 2011. [66](#)
- [LPS89] Markku S Lehtinen, Lassi Paivarinta, and Erkki Somersalo. Linear inverse problems for generalised random variables. *Inverse Problems*, 5(4):599, 1989. [37](#)
- [LPT94] Pierre-Louis Lions, Benoit Perthame, and Eitan Tadmor. A kinetic formulation of multidimensional scalar conservation laws and related equations. *Journal of the American Mathematical Society*, 7(1):169–191, 1994. [24](#)
- [LS04] Matti Lassas and Samuli Siltanen. Can one use total variation prior for edge-preserving bayesian inversion? *Inverse Problems*, 20(5):1537, 2004. [37](#)
- [LSS09] Matti Lassas, Eero Saksman, and Samuli Siltanen. Discretization-invariant bayesian inversion and besov space priors. *Inverse Problems and Imaging*, 3(1):87–122, 2009. [37](#)
- [Luc86] Bradley J. Lucier. A moving mesh numerical method for hyperbolic conservation laws. *Mathematics of Computation*, 46(173):59–69, 1986. [21](#)
- [LW55] M. James Lighthill and Gerald B. Whitham. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 229(1178):317–345, 1955. [2](#), [10](#)
- [LZ20] Thibault Liard and Enrique Zuazua. Initial data identification for the one-dimensional Burgers equation, preprint. 2020. [84](#)
- [Man84] Avi Mandelbaum. Linear estimators and measurable linear transformations on a hilbert space. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65(3):385–397, 1984. [37](#)
- [Mar04] Andrea Marson. Nonconvex conservation laws and ordinary differential equations. *J. London Mathematical Society*, 69:428–440, 2004. [66](#), [86](#)

- [Ole57] Olga A. Oleinik. Discontinuous solutions of non-linear differential equations. *Uspekhi Matematicheskikh Nauk*, 12(3):3–73, 1957. [23](#)
- [Ole59] Olga A. Oleinik. Uniqueness and stability of the generalized solution of the Cauchy problem for a quasi-linear equation. *Uspekhi Matematicheskikh Nauk*, 14(2):165–170, 1959. [16](#)
- [PC19] Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019. [32](#), [71](#)
- [Ric56] Paul I Richards. Shock waves on the highway. *Operations Research*, 4(1):42–51, 1956. [2](#), [10](#)
- [Ris93] Nils Henrik Risebro. A front-tracking alternative to the random choice method. *Proceedings of the American Mathematical Society*, 117(4):1125–1139, 1993. [24](#)
- [Sch73] David G. Schaeffer. A regularity theorem for conservation laws. *Advances in Mathematics*, 11:368–386, 1973. [24](#), [77](#), [82](#)
- [Ser99] Denis Serre. *Systems of Conservation Laws, Vols. 1 and 2*. Cambridge University Press, 1999. [24](#)
- [Sil19] Luis Silvestre. Oscillation properties of scalar conservation laws. *Communications on Pure and Applied Mathematics*, 72(6):1321–1348, 2019. [17](#)
- [Smo83] Joel Smoller. *Shock Waves and Reaction—Diffusion Equations*. Springer Science, 1983. [1](#), [13](#), [21](#), [24](#)
- [Spr20] Björn Sprungk. On the local lipschitz stability of bayesian inverse problems. *Inverse Problems (accepted)*, 2020. [37](#), [72](#)
- [ST18] Andrew M. Stuart and Aretha Teckentrup. Posterior consistency for Gaussian process approximations of Bayesian posterior distributions. *Mathematics of Computation*, 87(310):721–753, 2018. [37](#)
- [Ste11] David E Stewart. *Dynamics with Inequalities: Impacts and Hard Constraints*, volume 59. SIAM, 2011. [46](#)
- [Stu10] Andrew M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numerica*, 19:451–559, 2010. [3](#), [4](#), [35](#), [37](#)

- [Sul17] Tim J. Sullivan. Well-posed Bayesian inverse problems and heavy-tailed stable quasi-Banach space priors. *Inverse Problems and Imaging*, 11(5):857–874, 2017. [35](#), [37](#)
- [SW20] Andrew M Stuart and Marie-Therese Wolfram. Inverse optimal transport. *SIAM Journal on Applied Mathematics*, 80(1):599–619, 2020. [37](#)
- [TA77] Andrey N Tikhonov and Vasiliy Y Arsenin. *Solutions of ill-posed problems*. New York, 1977. [37](#)
- [Tar05] Albert Tarantola. *Inverse problem theory and methods for model parameter estimation*, volume 89. siam, 2005. [25](#)
- [Tes12] Gerald Teschl. *Ordinary Differential Equations and Dynamical Systems*, volume 140. American Mathematical Soc., 2012. [38](#)
- [Tik43] Andrei Nikolaevich Tikhonov. On the stability of inverse problems. In *Dokl. Akad. Nauk SSSR*, volume 39, pages 195–198, 1943. [37](#)
- [Tik63] Andrei Nikolaevich Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.*, 4:1035–1038, 1963. [37](#)
- [Tik66] Andrei Nikolaevich Tikhonov. On the stability of the functional optimization problem. *USSR Computational Mathematics and Mathematical Physics*, 6(4):28–33, 1966. [37](#)
- [TT93] Eitan Tadmor and Tamir Tassa. On the piecewise smoothness of entropy solutions to scalar conservation laws. *Comm. Partial Differential Equations*, 18(9-10):1631–1652, 1993. [24](#)
- [TT07] Eitan Tadmor and Terence Tao. Velocity averaging, kinetic formulations, and regularizing effects in quasi-linear PDEs. *Communications on Pure and Applied Mathematics*, 60(10):1488–1521, 2007. [24](#)
- [Vil08] Cedric Villani. *Optimal Transport: Old and New*. Springer, 2008. [71](#)
- [Vog02] Curtis R Vogel. *Computational Methods for Inverse Problems*, volume 23. Siam, 2002. [25](#), [30](#)
- [Whi75] G.B. Whitham. *Linear and Nonlinear Waves*. John Wiley, New York, 1975. [20](#)

Index

- $BV(\Omega)$, 18
- $BV_{\text{loc}}(\Omega)$, 18
- L^1 -contraction property, 19
- $TV(u)$, 18
- δ -neighbourhood, 50
- δ -solution, 42, 50
- admissible shock, 16
- basic conditions, 42
- Bayes' theorem, 31, 33
- Bayesian approach, 31
 - approximation, 35
 - well-posedness, 32
- Burgers equation, 10, 57
 - quasilinear, 10
- Carathéodory
 - conditions, 40
 - solution, 40
- Cauchy problem, 9
- differential inclusion, 41
- doubling of variables, 22
- entropy condition
 - Kruzkov, 16
- entropy production, 17
- entropy solution, 15
- entropy-entropy flux pair, 15
 - Kruzkov, 16
- equicontinuity in average, 22
- evaluation operator, 26
- Filippov solution, 41
- finiteness of shocks, 23
- forward map, 26
- front tracking method, 19
- functions of bounded variations, 17
 - locally, 18
- generic property, 23
- genuinely nonlinear, 47
- Hellinger distance, 32
- ill-posed problem, 28
- irreversibility, 17
- least squares method, 29
- normalization constant, 31, 32
- observable noise, 31
- Oleinik decay estimate, 23
- Radon-Nikodym theorem, 32
- random choice method, 24
- Rankine–Hugoniot jump condition, 12
- regularising effect, 23
- regularization
 - basics, 29
 - parameter, 30
 - penalty term, 30
 - strategy, 29
 - Tikhonov, 30

Riemann problem, [19](#), [20](#)

shock admissibility condition

 Lax, [16](#)

 Oleinik, [16](#)

shock speed, [13](#)

shock waves, [12](#)

traffic flow, [10](#)

 LWR model, [10](#)

upstream field, [67](#)

vanishing viscosity method, [21](#)

Wasserstein distance, [71](#)

weak solution, [12](#)