View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Helsingin yliopiston digitaalinen arkisto

University of Helsinki
Faculty of Science
Department of Mathematics and Statistics

Master's Thesis

Hamilton-Jacobi Equations and Scalar Conservation Laws

Author:
Teemu Säilynoja

Instructor:
Professor Xiao Zhong

May 26, 2017

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department
Matemaattis-luonnontieteellinen tiedekunta	Matematiikan ja tilastotieteen laitos

Tekijä — Författare — Author

Teemu Säilynoja

Työn nimi — Arbetets titel — Title

Hamilton-Jacobi Equations and Scalar Conservation Laws

Oppiaine — Läroämne — Subject

Matematiikka

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro gradu -tutkielma	Toukokuu 2017	73 s.

Tiivistelmä — Referat — Abstract

Työssä esitellään moniulotteisten osittaisdifferentiaaliyhtälöiden alkuarvo-ongelmia. Erityisesti keskitytään Hamilton-Jacobi -yhtälöihin, sekä yhden tilamuuttujan säilyvyyslakeihin. Näitä yhtälöitä kohdataan usein mallinnettaessa fysikaalisten systeemien käyttäytymistä matemaattisesti, eikä yhtälöille ole välttämättä ratkaisua perinteisessä mielessä.

Työn tavoitteena onkin esitellä ensin karakteristisena menetelmänä tunnettu ratkaisukeino, jossa osittaisdifferentiaaliyhtälö muunnetaan systeemiksi tavallisia differentiaaliyhtälöitä. Näiden tavallisten differentiaaliyhtälöiden ratkaisujen avulla voidaan määrittää alkuperäisen ongelman ratkaisu, ainakin tutkittavan alueen reunan läheisyydessä.

Seuraavaksi nostetaan esille karakteristisen menetelmän puutteita. Kuten säilyvyyslakeihin liittyvistä esimerkeistä havaitaan, menetelmä ei pysty tarjoamaan joillekin alkuarvotehtäville jatkuvaa, eikä siten differentioituvaa ratkaisua. Joissakin tapauksissa menetelmän antama tieto ratkaisusta ei puolestaan riitä kattamaan koko ratkaisun haluttua määrittelyjoukkoa.

On kuitenkin mahdollista määrittää alkuarvo-ongelman integraaliratkaisu, rajoitettu funktio, jolla on sileällä kompaktikantajaisella testifunktiolla kerrottaessa tutkittavan osittaisdifferentiaaliyhtälön ratkaisua muistuttavia ominaisuuksia. Annettuun ongelmaan ei kuitenkaan välttämättä ole aina yksikäsitteistä integraaliratkaisua, eikä osa integraaliratkaisun määritelmän toteuttavista funktioista tarjoa fysiikan ongelmista johdettuihin yhtälöihin mielekästä ratkaisua. Nopeasti havaitaan, että osa näistä epätoivotuista ratkaisuista saadaan karsittua vaatimalla integraaliratkaisulta tiettyjen entropiaehtojen täyttämistä.

Haluamme kuitenkin tarjota riittävät ehdot yksikäsitteisen, entropiaehdot toteuttavan ratkaisun löytämiseksi. Tämän tavoitteen saavuttamiseksi työssä esitellään variaatiolaskennan perusteita ja Hopf-Lax -kaavan johtaminen Hamilton-Jacobi -yhtälöistä. Hopf-Lax -kaavan antamien integraaliratkaisujen avulla Hamilton-Jacobi -yhtälöille voidaan määritellä niin kutsuttu heikko ratkaisu.

Lopuksi Hamilton-Jacobi -yhtälöiden heikkojen ratkaisujen määrittämiseksi esiteltyä teoriaa voidaan käyttää antamaan säilyvyyslakien alkuarvotehtäville entropiaehto, joka takaa ratkaisujen yksikäsitteisyyden nollamittaista joukkoa lukuunottamatta koko halutussa määrittelyjoukossa. Hamilton-Jacobi -yhtälöiden ja Hopf-Lax -kaavan avulla voidaan johtaa Lax-Oleinik -kaava, joka antaa sopivissa tapauksissa alkuarvo-ongelman entropiaratkaisun.

Avainsanat — Nyckelord — Keywords

Karakteristinen menetelmä, Säilyvyyslaki, Hamilton-Jacobi -yhtälö, Hopf-Lax -kaava, Lax-Oleinik -kaava, Heikot ratkaisut, Integraaliratkaisut, Entropiaratkaisut

Säilytyspaikka — Förvaringsställe — Where deposited

Kumpulan tiedekirjasto

Muita tietoja — Övriga uppgifter — Additional information

Contents

Introduction **2**

1 Method of Characteristics **4**
1.1 Characteristic ODEs 4
1.2 Existence of Local Solutions 6
1.2.1 Flattening the Boundary 6
1.2.2 Inverse and Implicit Function Theorems 7
1.2.3 Local Existence Theorem 11

2 Weak solutions **17**
2.1 Quasilinear Partial Differential Equations 17
2.2 Scalar Conservation Laws 20
2.3 Integral Solutions 20
2.3.1 Rankine-Hugoniot Jump Conditions 21
2.3.2 Entropy Conditions 25

3 Hamilton-Jacobi Equations **28**
3.1 Derivation of Hamilton's Equations 29
3.2 Introduction to the Calculus of Variations 30
3.3 Legendre Transform 34
3.4 Hopf-Lax Formula 36
3.5 Weak Solutions 45
3.6 Mollifiers 50
3.7 Uniqueness of Weak Solutions 54

4 Conservation Laws Revisited **59**
4.1 Lax-Oleinik Formula 59
4.2 Entropy Condition 64

Introduction

In this thesis, we consider boundary value problems for first-order partial differential equations (PDEs) of the form

$$\begin{cases} F(Du(x), u(x), x) = 0 & \text{for } x \in \Omega; \\ u(x) = g(x) & \text{for } x \in \Gamma \subset \partial\Omega. \end{cases}$$

Above Ω is assumed to be a bounded domain in $\mathbb{R}^n$ with a smooth boundary $\partial\Omega$. Here

$$F\colon \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R} \text{ and } g\colon \mathbb{R}^n \to \mathbb{R}$$

are given smooth functions and the unknown solution $u\colon \overline{\Omega} \to \mathbb{R}$ is expected to be continuously differentiable in Ω. The gradient of u is denoted by Du.

We focus closely on the properties of two specific types of PDEs, scalar conservation laws and Hamilton-Jacobi equations. In general these equations do not have continuous solutions in the whole domain. Therefore we consider the weak solutions to these equations.

In Section 1, we introduce the method of characteristics to solve boundary value problems for nonlinear first-order PDE. The method is used to transform the n-dimensional problem into a system of up to $2n + 1$ first-order ordinary differential equations (ODEs) that characterize the behaviour of the solution near the boundary $\partial\Omega$.

First we derive this system of characteristic ODE and then prove the existence of unique local solution of the original initial value problem.

In Section 2, we consider scalar conservation laws, quasilinear first-order PDEs of the form

$$u_t(x, t) + [F(u(x, t))]_x = 0,$$

where $F\colon \mathbb{R} \to \mathbb{R}$ is a given function and $u\colon \mathbb{R} \times [0, \infty) \to \mathbb{R}$ is unknown. By u_t, we denote the partial derivative of u with respect to t. Similarly, $[F(u(x, t))]_x$ denotes the partial derivative $\frac{\partial}{\partial x} F(u(x, t))$.

These types of PDEs often arise from problems modelling the time behaviour of physical systems of objects, liquids or gases.

There does not, in general, exist a smooth solution u for all times $t > 0$. Instead, it is possible to define so called integral solutions for the problem. These solutions form a family of bounded functions that satisfy the PDE almost everywhere in $\mathbb{R} \times (0, \infty)$.

Not all of the integral solutions are physically meaningful or possible. As we seek to have a unique solution with properties that are useful in modelling physical systems, we introduce an entropy condition to exclude the unwanted solutions. Later we show that by further improving the entropy condition, there exist at most one entropy solution to the problem, provided that function F is a convex C^2 function and g is essentially bounded.

In Section 3, we introduce the Hamilton-Jacobi equations

$$\begin{cases} u_t + H(D_x u, x) = 0 & \text{in } \mathbb{R}^n \times (0, \infty); \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

These equations have a connection to the scalar conservation laws and we build theory for defining weak solutions of Hamilton-Jacobi equations and prove the uniqueness of these solutions.

In section 4, we return to the initial value problem for scalar conservation laws. As there turns out to be many similarities to the initial value problems for the Hamilton-Jacobi Equations, we use the theory constructed in Section 3 to derive the Lax-Oleinik formula and define entropy solutions to initial value problems for the scalar conservation laws. We proceed to prove that, up to a set of measure zero, there exists unique entropy solution, provided $F:\mathbb{R}\to\mathbb{R}$ is a convex C^2 function and g is essentially bounded. If additionally F is strongly convex, this solution is given by the Lax-Oleinik formula.

1 Method of Characteristics

Consider the following first-order partial differential equation

$$F(Du, u, x) = 0 \text{ in } \Omega, \tag{1.1}$$

where $F\colon \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is a given smooth function. Here $\Omega \subset \mathbb{R}^n$ is a bounded domain with a smooth boundary $\partial\Omega$. See Section 1.2 for the definition. Suppose further that the solution $u\colon \overline{\Omega} \to \mathbb{R}$ is twice continuously differentiable in Ω, i.e. $u \in C^2(\Omega)$, and that u satisfies the boundary condition

$$u = g \text{ on } \Gamma, \tag{1.2}$$

where $\Gamma \subseteq \partial\Omega$ and $g\colon \Gamma \to \mathbb{R}$ is a given smooth function.

To evaluate u at a point $x \in \Omega$, we form a path from $x^0 \in \Gamma$ to x. As $u(x^0)$ is given by $g(x^0)$, we use this known value to evaluate u upon the path and to obtain the value $u(x)$.

Let $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, $p = (p_1, \cdots, p_n) \in \mathbb{R}^n$ and $z \in \mathbb{R}$. Let $F = F(p, z, x)$ and denote the partial derivative of F with respect to p_k by F_{p_k}. We write

$$\begin{cases} D_pF(p, z, x) = (F_{p_1}(p, z, x), F_{p_2}(p, z, x), \cdots, F_{p_n}(p, z, x)); \\ D_zF(p, z, x) = F_z(p, z, x); \\ D_xF(p, z, x) = (F_{x_1}(p, z, x), F_{x_2}(p, z, x), \cdots, F_{x_n}(p, z, x)). \end{cases}$$

In deriving the characteristic ODEs, we refer to section 3.2. in [3]. In the proofs of the Inverse Function Theorem and the Implicit Function Theorem presented in section 1.2.2, we refer to sections 13.3 and 13.4 in [1] and to theorems 9.24 and 9.28 in [8]

1.1 Characteristic ODEs

Let us consider a path $\mathbf{x}(s) \subset \Omega$, where $\mathbf{x}(s) = (x^1(s), x^2(s), \cdots, x^n(s))$ and $s \in I$ with I being a subinterval in $\mathbb{R}$. Suppose $u \in C^2(\Omega)$ is a solution to (1.1) and define

$$\begin{cases} z(s) = u(\mathbf{x}(s)) = u(x^1(s), x^2(s), \cdots, x^n(s)); \\ \mathbf{p}(s) = (p^1(s), p^2(s), \cdots, p^n(s)) \\ \qquad = (u_{x_1}(\mathbf{x}(s)), u_{x_2}(\mathbf{x}(s)), \cdots, u_{x_n}(\mathbf{x}(s))). \end{cases} \tag{1.3}$$

We aim to form a system of first order ODEs for $\mathbf{p}(s)$, $z(s)$ and $\mathbf{x}(s)$ to solve the boundary value problem (1.1), (1.2). These are called the characteristic ODEs of (1.1). We use the following notation for the regular derivative:

$$\dot{\mathbf{p}}(s) := \frac{d}{ds}\mathbf{p}(s).$$

Differentiating $p^i(s)$, where $i \in \{1, 2, \cdots, n\}$, we arrive at the following equation involving second order partial derivatives of u:

$$\dot{p}^i(s) = \frac{d}{ds} u_{x_i}(\mathbf{x}(s)) = \sum_{k=1}^{n} u_{x_i x_k}(\mathbf{x}(s)) \dot{x}^k(s). \tag{1.4}$$

We will choose $\mathbf{x}(s)$ in a way such that the second order derivatives in (1.4) do not appear in the characteristic ODE. To this end, we differentiate equation (1.1) with respect to x_i to obtain:

$$\sum_{k=1}^{n} F_{p_k}(Du, u, x) u_{x_k x_i} + F_z(Du, u, x) u_{x_i} + F_{x_i}(Du, u, x) = 0.$$

This gives us that

$$\sum_{k=1}^{n} F_{p_k}(Du, u, x) u_{x_k x_i} = -F_z(Du, u, x) u_{x_i} - F_{x_i}(Du, u, x). \tag{1.5}$$

To form the characteristic ODEs for $\mathbf{p}(s)$, we choose $\mathbf{x}(s)$ so that

$$\dot{x}^k(s) := F_{p_k}(\mathbf{p}(s), z(s), \mathbf{x}(s)). \tag{1.6}$$

It follows from (1.4), (1.5) and (1.6) that

$$\begin{aligned} \dot{p}^i(s) &= \sum_{k=1}^{n} u_{x_i x_k}(\mathbf{x}(s)) F_{p_k}(\mathbf{p}(s), z(s), \mathbf{x}(s)) \\ &= -F_z(\mathbf{p}(s), z(s), \mathbf{x}(s)) p^i(s) - F_{x_i}(Du, u, x). \end{aligned} \tag{1.7}$$

To give the characteristic ODE for $z(s)$ we calculate $\dot{z}(s)$:

$$\dot{z}(s) = \frac{d}{ds} u(\mathbf{x}(s)) = Du(\mathbf{x}(s)) \cdot \dot{\mathbf{x}}(s) = \mathbf{p}(s) \cdot D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)). \tag{1.8}$$

Let us collect equations (1.6), (1.8) and (1.7) all together. We arrive at the following system of $2n + 1$ characteristic ODEs:

$$\begin{cases} \dot{\mathbf{x}}(s) = D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)); & \text{(1.9a)} \\ \dot{z}(s) = D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)) \cdot \mathbf{p}(s); & \text{(1.9b)} \\ \dot{\mathbf{p}}(s) = -D_x F(\mathbf{p}(s), z(s), \mathbf{x}(s)) - D_z F(\mathbf{p}(s), z(s), \mathbf{x}(s)) \mathbf{p}(s). & \text{(1.9c)} \end{cases}$$

The functions $\mathbf{p}(s)$, $z(s)$ and $\mathbf{x}(s)$ solving the system of characteristic ODEs are called the *characteristics* and $\mathbf{x}(s)$ is called the *projected characteristic* onto Ω.

In the above we proved the following theorem.

Theorem 1.1. *Let $u \in C^2(\Omega)$ be a solution to* (1.1)*. Suppose $\mathbf{x}(s)$ is a solution to* (1.9a)*. Then for $s \in \mathbb{R}$ such that $\mathbf{x}(s) \in \Omega$, functions $z(s)$ and $\mathbf{p}(s)$ which are defined as in* (1.3)*, solve equations* (1.9b) *and* (1.9c) *respectively.*

1.2 Existence of Local Solutions

1.2.1 Flattening the Boundary

Next we show that, when considering a sufficiently small neighbourhood $U \subset \mathbb{R}^n$ of any fixed $x^0 \in \partial\Omega$, the boundary of Ω can be assumed to lie on a hyperplane.

Definition 1.2. The boundary $\partial\Omega$ of a bounded domain $\Omega \subset \mathbb{R}^n$ is called smooth, if for every $x^0 \in \partial\Omega$ there exists a radius $r > 0$, a rotation ϕ in $\mathbb{R}^n$ and a smooth function $\gamma\colon \mathbb{R}^{n-1} \to \mathbb{R}$ such that

$$B(x^0, r) \cap \Omega = \left\{x \in B(x^0, r)\colon \phi^n(x) > \gamma\left(\phi^1(x), \cdots, \phi^{n-1}(x)\right)\right\}.$$

We define a mapping $\Phi\colon \mathbb{R}^n \to \mathbb{R}^n$ that flattens the boundary of Ω near x^0. Let

$$\begin{cases} \Phi^k(x) = \phi^k(x) =: y_k & k \in \{1, \cdots, n-1\}; \\ \Phi^n(x) = \phi^n(x) - \gamma\left(\phi^1(x), \cdots, \phi^{n-1}(x)\right) =: y_n. & \end{cases}$$

We obtain the inverse mapping $\Psi = \Phi^{-1}$ by setting

$$\begin{cases} \Psi^k(y) = \psi^k(y) = x_k & k \in \{1, \cdots, n-1\}; \\ \Psi^n(y) = \psi^n(y_n + \gamma(y_1, \cdots, y_{n-1})) = x_n, & \end{cases}$$

where $\psi = \phi^{-1}$. We note that the Jacobians, as defined in Section 1.2.2, of both Ψ and Φ are equal to one.

We can use these mappings to define a change of variable for our PDE. Set $\widetilde{\Omega} \coloneqq \Phi(\Omega)$ and define $\tilde{u}\colon \widetilde{\Omega} \to \mathbb{R}$ by $\tilde{u}(y) = u\left(\Psi(y)\right)$ for all $y \in \widetilde{\Omega}$. Let us now write PDE (1.1) and the boundary value condition (1.2) in a form involving $\tilde{u}$ and $y = \Phi(x)$.

$$\begin{aligned} 0 &= F(D_x u(x), u(x), x) \\ &= F\big(D_y\tilde{u}\left(\Phi(x)\right) \cdot D_x\Phi(x), \tilde{u}\left(\Phi(x)\right), x\big) \\ &= F\big(D_y\tilde{u}(y) \cdot D_x\Phi(\Psi(y)), \tilde{u}(y), \Psi(y)\big) \\ &=: \tilde{F}\big(D_y\tilde{u}(y), \tilde{u}(y), y\big). \end{aligned}$$

In the above, $\widetilde{F}\colon \mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n \to \mathbb{R}$ is a smooth function. Let $\tilde{g}(y) := g(\Psi(y))$ for $y \in \widetilde{\Gamma} = \Phi(\Gamma)$. Now $\tilde{u}$ satisfies the following boundary value problem:

$$\begin{cases} \tilde{F}(D\tilde{u}, \tilde{u}, y) = 0 & \text{in } \tilde{\Omega}; \\ \tilde{u} = \tilde{g} & \text{on } \tilde{\Gamma}. \end{cases} \tag{1.10}$$

The boundary value problem (1.10) is of the same form as (1.1). Furthermore, by Definition 1.2, it holds for every $y \in B(\Phi(x^0), r) \cap \partial\widetilde{\Omega}$ that $y_n = \Phi^n(x) = 0$. Therefore we can assume without loss of generality, that for every fixed $x^0 \in \partial\Omega$ there exists a neighbourhood $V \subset \partial\Omega$ of x^0 that lies on a hyperplane.

1.2.2 Inverse and Implicit Function Theorems

To prove the existence of local solutions to (1.1) with the boundary condition (1.2), we need the Inverse Function Theorem and the Implicit Function Theorem.

We use the Implicit Function Theorem to show that for suitable points $x^0 \in \Gamma$, there exists a neighbourhood $V \subset \Gamma$ of x^0 where we can give initial values for the characteristic ODEs.

The Inverse Function Theorem is used to show that there exists a neighbourhood $U \subset \mathbb{R}^n$ of V, such that each $x \in U$ lies on a unique projected characteristic with an initial value given on V.

Definition 1.3. The Jacobian determinant of a mapping $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^n$ is

$$J\mathbf{f}(x) = \det(D\mathbf{f})(x),$$

where the $D\mathbf{f}(x)$ is the differential matrix of f at $x \in \mathbb{R}^n$

$$D\mathbf{f}(x) = \begin{bmatrix} f^1_{x_1}(x) & \cdots & f^1_{x_n}(x) \\ \vdots & \ddots & \vdots \\ f^n_{x_1}(x) & \cdots & f^n_{x_n}(x) \end{bmatrix}.$$

Theorem 1.4. (Inverse Function Theorem). *Let $\Omega \subset \mathbb{R}^n$ be open and $\mathbf{f} \in C^1(\Omega; \mathbb{R}^n)$. Suppose that for some $x^0 \in \Omega$, $J\mathbf{f}(x^0) \neq 0$. Then there exist neighbourhoods $V \subset \Omega$ of x^0 and $U \subset \mathbf{f}(\Omega)$ of $\mathbf{f}(x^0)$, such that $\mathbf{f} : V \to U$ is one-to-one and the inverse function $\mathbf{f}^{-1}$ belongs to $C^1(U; V)$.*

Proof. Let us first find the neighbourhoods V of x^0 and U of $\mathbf{f}(x^0)$ where $\mathbf{f}$ is one-to-one and therefore has an inverse function. Since $\mathbf{f} \in C^1(\Omega; R^n)$, the Jacobian $J\mathbf{f}$ is a continuous function in Ω. By assumption, $J\mathbf{f}(x^0) \neq 0$ and thus there is a ball $B(x^0, r) \subset \Omega$, where $r > 0$, such that $J\mathbf{f}(x) \neq 0$ for every $x \in B(x^0, r)$. Denote $V := B(x^0, r)$.

As f^k is continuously differentiable for each $k \in \{1, 2, \cdots, n\}$, the Fundamental Theorem of Calculus gives for every $x, \tilde{x} \in V$ that

$$\begin{aligned} f^k(\tilde{x}) - f^k(x) &= \int_0^1 \frac{d}{dt} f^k(x + t(\tilde{x} - x))\, dt \\ &= \int_0^1 Df^k(x + t(\tilde{x} - x)) \cdot (\tilde{x} - x)\, dt. \end{aligned} \tag{1.11}$$

By the Mean Value Theorem, there is a $t_k \in (0, 1)$ such that

$$\begin{aligned} \int_0^1 Df^k(x + t(\tilde{x} - x)) \cdot (\tilde{x} - x)\, dt &= Df^k(x + t_k(\tilde{x} - x)) \cdot (\tilde{x} - x) \\ &= \sum_{i=1}^{n} f^k_{x_i}(x + t_k(\tilde{x} - x))(\tilde{x}_i - x_i). \end{aligned} \tag{1.12}$$

Suppose $\mathbf{f}(x) = \mathbf{f}(\tilde{x})$ for some $x, \tilde{x} \in V$. Then $f^k(\tilde{x}) - f^k(x) = 0$ for each $k \in \{1, \ldots, n\}$ and, by equations (1.11) and (1.12)

$$0 = f^k(\tilde{x}) - f^k(x) = \sum_{i=1}^{n} (\tilde{x}_i - x_i) f^k_{x_i}(x + t_k(\tilde{x} - x)). \tag{1.13}$$

As (1.13) holds for every $k \in \{1, 2, \cdots, n\}$, and $J\mathbf{f} \neq 0$ in V, the following system

$$\sum_{i=1}^{n} f^k_{k_i}(x + t_k(\tilde{x} - x)) y_i = 0 \qquad k \in \{1, 2, \cdots, n\},$$

has only one solution, $y_i = 0$ for $i \in \{1, 2, \cdots, n\}$. This implies that $x_i = \tilde{x}_i$ for every $i \in \{1, 2, \cdots, n\}$. Thus $x = \tilde{x}$.

Hence $\mathbf{f}\colon V \to \mathbf{f}(V)$ is one-to-one and continuous, and therefore has an inverse function $\mathbf{f}^{-1}\colon \mathbf{f}(V) \to V$ that is also one-to-one and continuous. The set $U \coloneqq \mathbf{f}(V) \subset \mathbb{R}^n$ is open, as it is the preimage of the open set V under the continuous function $\mathbf{f}^{-1}$.

Let us conclude the proof by showing that $\mathbf{f}^{-1} \in C^1(U; V)$. Define $\mathbf{g} \coloneqq \mathbf{f}^{-1}$ in U and let $y \in U$. Let $x = \mathbf{g}(y)$ and $x^h = \mathbf{g}(y + he_i)$, where $h \in \mathbb{R}$ and e_i is the ith coordinate vector of $\mathbb{R}^n$. Since U is open, $y + he_i \in U$ and $x^h \in V$ when $|h|$ is small enough. Now

$$\mathbf{f}(x^h) - \mathbf{f}(x) = he_i,$$

and for $k \in \{1, 2, \cdots, n\}$ we have

$$f^k(x^h) - f^k(x) = h\delta_{k,i}, \tag{1.14}$$

where $\delta_{k,i}$ is the Kronecker delta function, that is,

$$\delta_{k,i} = \begin{cases} 1, & \text{if } k = i; \\ 0, & \text{if } k \neq i. \end{cases}$$

By the Mean Value Theorem, there exists a $t_k \in (0, 1)$ satisfying

$$\mathbf{f}^k(x^h) - \mathbf{f}^k(x) = \sum_{j=1}^{n} (x_j^h - x_j) f_{x_j}^k \left(x + t_k(x^h - x)\right). \tag{1.15}$$

Substituting $x = \mathbf{g}(y)$ and $x^h = \mathbf{g}(y + he_i)$, (1.14) and (1.15) give the following system

$$\delta_{k,i} = \sum_{j=1}^{n} \frac{g^j(y + he_i) - g^j(y)}{h} f_{x_j}^k \left(x + t_k(x^h - x)\right), \qquad k \in \{1, 2, \cdots, n\}.$$

As $J\mathbf{f} \neq 0$ in V, this system has a unique solution for

$$\frac{g^k(y + he_i) - g^k(y)}{h}, \quad k \in \{1, 2, \cdots, n\}.$$

By Cramer's rule, this solution is of the form

$$\frac{g^k(y + he_i) - g^k(y)}{h} = \frac{\det(A_{k,i})}{J\mathbf{f}(t_k x^h + (1 - t_k)x)},$$

where $A_{k,i}$ is the matrix obtained by substituting $\delta_{k,i}$ for $f_{x_k}^i$ in $D\mathbf{f}$.

By the continuity of $\mathbf{g}$, $x^h \to x$ as $h \to 0$. Since $J\mathbf{f} \neq 0$ in V, the limit

$$g_{y_i}^k = \lim_{h \to 0} \frac{g^k(y + he_i) - g^k(y)}{h} = \lim_{h \to 0} \frac{\det(A_{k,i})}{J\mathbf{f}(x + t_k(x^h - x))}$$

exists for each $k \in \{1, 2, \cdots, n\}$. Hence we have that $\mathbf{g} \in C^1(U; V)$. $\square$

Let $V \subset \mathbb{R}^{n+m}$ be an open set and $\mathbf{f} \in C^1(V; \mathbb{R}^n)$. Then we use the following notation for the blocks of the differential matrix of $\mathbf{f}$ corresponding to $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$:

$$D\mathbf{f} = \begin{bmatrix} f_{x_1}^1 & \cdots & f_{x_n}^1 & f_{y_1}^1 & \cdots & f_{y_m}^1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f_{x_1}^n & \cdots & f_{x_n}^n & f_{y_1}^n & \cdots & f_{y_m}^n \end{bmatrix} = (D_x\mathbf{f}, D_y\mathbf{f}).$$

Similarly, we write $J_x\mathbf{f}(x, y) \coloneqq \det D_x\mathbf{f}(x, y)$.

Theorem 1.5. (Implicit Function Theorem). *Let $U \in \mathbb{R}^{n+m}$ be open and $\mathbf{f} \in C^1(U;\mathbb{R}^n)$. Suppose that $J_t\mathbf{f}(t^0, x^0) \neq 0$ and $\mathbf{f}(t^0, x^0) = 0$ for some $(t^0, x^0) \in U$. Then there exists a neighbourhood $V \subset \mathbb{R}^m$ of x^0 and a unique function $\mathbf{g} \in C^1(V;\mathbb{R}^n)$, for which $\mathbf{g}(x^0) = t^0$ and $\mathbf{f}(\mathbf{g}(x), x) = 0$ for every $x \in V$.*

Proof. Let us define $\mathbf{F}\colon \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ as

$$\mathbf{F}(t, x) = (\mathbf{f}(t, x), x). \tag{1.16}$$

Now $\mathbf{F}(t^0, x^0) = (0, x^0)$ and

$$J\mathbf{F}(t^0, x^0) = \det(D\mathbf{F}(t^0, x^0)) = \det \begin{bmatrix} D_t\mathbf{f} & D_x\mathbf{f} \\ 0 & I_{m\times m} \end{bmatrix} = \det D_t\mathbf{f}(t^0, x^0) \neq 0.$$

By the Inverse Function Theorem, Theorem 1.4, there exist neighbourhoods $\widetilde{V} \subset \mathbb{R}^{n+m}$ of (t^0, x^0) and $\widetilde{U} \subset \mathbb{R}^{n+m}$ of $(0, x^0)$, such that $\mathbf{F}\colon \widetilde{V} \to \widetilde{U}$ is one-to-one and there is a one-to-one inverse function $\mathbf{F}^{-1}\colon \widetilde{U} \to \widetilde{V}$.

Let $(\tau, \xi) \in \widetilde{U}$ and denote $\mathbf{F}^{-1} := \mathbf{G}$. As $\mathbf{F}$ is one-to-one, the equation

$$\mathbf{F}(t, x) = (\tau, \xi) \tag{1.17}$$

has a unique solution for $(t, x) \in \widetilde{V}$ given by

$$\begin{cases} t = \mathbf{h}(\tau, \xi); \\ x = \xi, \end{cases} \tag{1.18}$$

where $\mathbf{h} = (\mathbf{G}^1, \mathbf{G}^2, \cdots, \mathbf{G}^n) \in C^1(\widetilde{U};\mathbb{R}^n)$.

By (1.16) and (1.18), $\mathbf{f}(\mathbf{h}(\tau, x), x) = \tau$ for every $(\tau, x) \in \widetilde{U}$. Let $V \subset \mathbb{R}^m$ be a neighbourhood of x^0, for which $(0, x) \in \widetilde{U}$ for every $x \in V$. By defining $\mathbf{g}\colon V \to \mathbb{R}^n$ as $\mathbf{g}(x) = \mathbf{h}(0, x)$, we have $\mathbf{g}(x^0) = t^0$ and

$$\mathbf{f}(\mathbf{g}(x), x) = 0, \text{ for every } x \in V.$$

Lastly, recall that $\mathbf{F}^{-1}$ is one to one from $\widetilde{U}$ to $\widetilde{V}$. Therefore, for every $x \in V$ it holds that $\mathbf{F}^{-1}(0, x) = (t, x)$ for some unique $t \in \mathbb{R}^n$. By definition $\mathbf{g}(x) = \mathbf{h}(0, x) = t$ and thus $\mathbf{g}$ is the unique function satisfying

$$\mathbf{f}(\mathbf{g}(x), x) = 0$$

for every $x \in V$. □

1.2.3 Local Existence Theorem

We begin to construct a local solution to the original boundary value problem, (1.1) and (1.2), near a point $x^0 \in \Gamma$. As shown in Section 1.2.1, we may assume without loss of generality, that for the neighbourhood $V \subset \Gamma$ of x^0, the boundary of Ω lies on any hyperplane. Therefore we assume hereafter that for the neighbourhoods $V \subset \Gamma$ in this section, $y_n = 0$ for every $y \in V$.

Let us denote the initial conditions of the characteristic ODEs as

$$\mathbf{x}(0) = x^0,\ z(0) = z^0 \text{ and } \mathbf{p}(0) = p^0.$$

It follows from the boundary value condition (1.2), that is, $u = g$ on Γ, that

$$z^0 = z(0) = u(\mathbf{x}(0)) = g(\mathbf{x}(0)) = g(x^0). \tag{1.19}$$

To determine p^0, we use condition (1.2) again, and see that for $y \in \Gamma$ near x^0 it holds that

$$u(y) = u(y_1, \cdots, y_{n-1}, 0) = g(y_1, \cdots, y_{n-1}, 0).$$

Since function g is smooth, we have

$$p_k^0 = p^k(0) = u_{x_k}(\mathbf{x}(0)) = u_{x_k}(x^0) = g_{x_k}(x^0), \tag{1.20}$$

for $k \in \{1, \cdots, n-1\}$. We define p^0 as follows.

$$\begin{cases} p_k^0 = g_{x_k}(x^0) & \text{for } k \in \{1, \cdots, n-1\}, \\ F(p^0, z^0, x^0) = 0. \end{cases} \tag{1.21}$$

Conditions (1.19) and (1.21) are called the *compatibility conditions* and the triple (p^0, z^0, x^0) that satisfies these conditions is called *admissible.*

Our goal is to evaluate the solution u at each point $x \in U \cap \Omega$, where $U \subset \mathbb{R}^n$ is some neighbourhood of $x^0 \in \Gamma$. To this end, we need to give initial conditions to each $y \in V$ for some neighbourhood $V \subset \Gamma$ of x^0.

As before, the initial value conditions for $\mathbf{x}(s)$ and $z(s)$ are given for each $y \in \Gamma$ by

$$\mathbf{x}(0) = y \text{ and } z(0) = g(y). \tag{1.22}$$

The next theorem gives a condition for the existence of a neighbourhood $V \subset \Gamma$ of x^0, on which suitable initial value conditions for $\mathbf{p}(s)$ can be given.

Theorem 1.6. *Let $x^0 \in \Gamma$ and suppose (p^0, z^0, x^0) is an admissible triple for the boundary value problem given by* (1.1) *and* (1.2). *Let $\mathbf{x}(s)$ and $z(s)$ satisfy the initial value conditions* (1.22). *Suppose further that*

$$F_{p_n}(p^0, z^0, x^0) \neq 0. \tag{1.23}$$

Then there exists a neighbourhood $V \subset \Gamma$ of x^0 and a unique $\mathbf{q} \in C^1(V; R^n)$ for which $(\mathbf{q}(y), z(y), y)$ is admissible for every $y \in V$ and $\mathbf{q}(x^0) = p^0$.

The triple (p^0, z^0, x^0) is called *noncharacteristic* if it satisfies (1.23). This condition ensures that the projected characteristic emanating from x^0 is not parallel to Γ and therefore enters Ω at x^0.

Proof. Assume the triple $(p^0, z^0, x^0) \in \mathbb{R}^n \times \mathbb{R} \times \Gamma$ to be admissible and noncharacteristic. Let $p \in \mathbb{R}^n$ and $y \in \Gamma$ and define $\mathbf{G}\colon \mathbb{R}^n \times \Gamma \to \mathbb{R}^n$ by

$$\begin{cases} G^k(p, y) = p - g_{x_k}(y), & k \in \{1, \cdots, n-1\}; \\ G^n(p, y) = F(p, g(y), y). \end{cases}$$

As functions g and F in (1.1) and (1.2) are smooth, $\mathbf{G}$ is smooth. By assumptions, $\mathbf{G}(p^0, x^0) = 0$. We also have

$$\begin{aligned} D_p\mathbf{G}(p, y) &= \begin{bmatrix} G^1_{p_1}(p, y) & \cdots & G^1_{p_n}(p, y) \\ \vdots & \ddots & \vdots \\ G^n_{p_1}(p, y) & \cdots & G^n_{p_n}(p, y) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ F_{p_1}(p, g(y), y) & & \cdots & & F_{p_n}(p, g(y), y) \end{bmatrix}. \end{aligned}$$

As the triple (p^0, z^0, x^0) is noncharacteristic by assumption, we have

$$J_p\mathbf{G}(p^0, x^0) = F_{p_n}(p^0, g(x^0), x^0) \neq 0.$$

Therefore, by the Implicit Function Theorem, Theorem 1.5, there exists a neighbourhood $V \subset \Gamma$ of x^0 and a unique function $\mathbf{q} \in C^1(V; \mathbb{R}^n)$, for which $\mathbf{q}(x^0) = p^0$ and $\mathbf{G}(\mathbf{q}(y), y) = 0$ for every $y \in V$.

By the definition of $\mathbf{G}$, we have for every $y \in V$ that

$$\begin{cases} q^k(y) = g_{x_k}(y) & \text{for } k \in \{1, \cdots, n-1\}; \\ F(\mathbf{q}(y), g(y), y) = 0. \end{cases} \tag{1.24}$$

Therefore, the triple $(\mathbf{q}(y), g(y), y)$ is admissible for every $y \in V$. □

Let $x^0 \in \Gamma$ and suppose the triple (p^0, z^0, x^0) is noncharacteristic. Take $V \subset \Gamma$ to be a neighbourhood of x^0 as in Theorem 1.6. In addition to the parameter $s \in \mathbb{R}$, the solutions to the characteristic ODEs depend on the initial point $y \in V$. As assumed earlier, $y_n = 0$ for $y \in V$. We define the characteristics as

$$\begin{cases} \mathbf{x}(y,s) \coloneqq \mathbf{x}(y_1, \cdots, y_{n-1}, s); \\ z(y,s) \coloneqq z(y_1, \cdots, y_{n-1}, s); \\ \mathbf{p}(y,s) \coloneqq \mathbf{p}(y_1, \cdots, y_{n-1}, s), \end{cases} \tag{1.25}$$

where $\mathbf{x}(y, \cdot)$, $z(y, \cdot)$ and $\mathbf{p}(y, \cdot)$ are solutions to the characteristic ODEs with the initial conditions given by

$$\mathbf{x}(y, 0) = y, \; z(y, 0) = g(y) \text{ and } \mathbf{p}(y, 0) = \mathbf{q}(y). \tag{1.26}$$

To emphasise, $\mathbf{x}(y, s)$, $z(y, s)$ and $\mathbf{p}(y, s)$ are here viewed as functions from $\widetilde{V} \times I \subset \mathbb{R}^n$ to $\mathbb{R}^n$, where

$$\widetilde{V} = \{\tilde{y} \in \mathbb{R}^{n-1} \mid (\tilde{y}_1, \tilde{y}_2, \cdots, \tilde{y}_{n-1}, 0) \in V\}.$$

Next we show that for every $x \in \Omega$ near x^0 there exists a unique characteristic curve with the projected characteristic passing through x.

Theorem 1.7. *Let $x^0 \in \Gamma$ and suppose (p^0, z^0, x^0) is noncharacteristic. Then there exist neighbourhoods $V \subset \Gamma$ and $U \subset \mathbb{R}^n$ of x^0 and an interval $I \subset \mathbb{R}$ containing 0, such that for each $x \in U$ there are unique $y \in V$ and $s \in I$ for which it holds that*

$$x = \mathbf{x}(y, s).$$

Proof. By the initial conditions it holds that

$$\mathbf{x}(x^0, 0) = x^0.$$

Therefore the claim holds by the Inverse Function Theorem, Theorem 1.4, if

$$J\mathbf{x}(x^0, 0) \neq 0.$$

As it holds for every $y \in \Gamma$ that

$$\mathbf{x}(y, 0) = \mathbf{x}(y_1, y_2, \cdots, y_{n-1}, 0) = (y_1, y_2, \cdots, y_{n-1}, 0),$$

we have for $i \in \{1, \cdots, n-1\}$ that

$$x^k_{y_i}(x^0, 0) = \begin{cases} \delta_{i,k} & k \in \{1, 2, \cdots, n-1\}; \\ 0 & k = n. \end{cases} \tag{1.27}$$

The characteristic ODEs (1.9a) corresponding to $\mathbf{x}(s)$ provide that

$$\dot{x}(x^0, 0) = D_p F(p^0, z^0, x^0). \tag{1.28}$$

Equations (1.27) and (1.28) together give us that

$$D\mathbf{x}(x^0, 0) = \begin{bmatrix} 1 & & 0 & F_{p_1}(p^0, z^0, x^0) \\ & \ddots & & \vdots \\ 0 & & 1 & \\ 0 & \cdots & 0 & F_{p_n}(p^0, z^0, x^0) \end{bmatrix}.$$

Since the triple (p^0, z^0, x^0) is noncharacteristic, $F_{p_n}(p^0, z^0, x^0) \neq 0$ and therefore

$$J\mathbf{x}(x^0, 0) = F_{p_n}(p^0, z^0, x^0) \neq 0.$$

By the Inverse Function Theorem, Theorem 1.4, there are neighbourhoods $\widetilde{V} \subset \mathbb{R}^{n-1}$ of $(x_1^0, x_2^0, \cdots, x_{n-1}^0)$, $U \subset \mathbb{R}^n$ of x^0 and $I \subset \mathbb{R}$ of 0, for which there exists a unique inverse function $\mathbf{x}^{-1} \colon U \to \widetilde{V} \times I$ that is one-to-one.

Let $V = \widetilde{V} \times \{0\}$. Now $V \subset \Gamma$ is a neighbourhood of x^0. As $\mathbf{x}^{-1}$ is one-to-one, there are unique $y \in V$ and $s \in I$ for each $x \in U$, such that $(y, s) = \mathbf{x}^{-1}(x)$ and thus

$$\mathbf{x}(y, s) = x.$$

The proof is finished. □

Let again $x^0 \in \Gamma$ and (p^0, z^0, x^0) be noncharacteristic for the boundary value problem (1.1) and (1.2). Let $y \in V$, $s \in I$ and $x \in U$ as in Theorem 1.7. Denote the solutions to

$$\mathbf{x}(y, s) = x$$

by $y = \mathbf{y}(x)$ and $s = s(x)$. We use (1.25) to define

$$\begin{cases} u(x) := z(\mathbf{y}(x), s(x)); \\ \mathbf{p}(x) := \mathbf{p}(\mathbf{y}(x), s(x)), \end{cases} \tag{1.29}$$

for $x \in U$.

We arrive at the Local Existence Theorem that ties together the solutions to the characteristic ODEs and the solution to the original boundary value problem.

Theorem 1.8. (Local Existense Theorem). *The function u defined in* (1.29) *is a C^2 solution to the boundary value problem*

$$\begin{cases} F(Du, u, x) = 0 & \text{in } U; \\ u = g & \text{on } U \cap \Gamma. \end{cases} \tag{1.30}$$

Proof. Fix a point $y \in U \cap \Gamma$. Consider $\mathbf{p}(y, \cdot)$, $z(y, \cdot)$ and $\mathbf{x}(y, \cdot)$ that solve the characteristic ODEs with the initial value conditions for $y \in U \cap \Gamma$ given as

$$\mathbf{x}(y, 0) = y, \; z(y, 0) = g(y) \text{ and } \mathbf{p}(y, 0) = \mathbf{q}(y),$$

where $\mathbf{q}$ is given by Theorem 1.6. Next we have the following two claims. *Claim 1.* For every $s \in \mathbb{R}$ it holds that when $y \in U \cap \Gamma$,

$$F(\mathbf{p}(y, s), z(y, s), \mathbf{x}(y, s)) = 0. \tag{1.31}$$

This, together with Theorem 1.7, gives us

$$F(\mathbf{p}(x), u(x), x) = 0 \quad \text{ for } x \in U.$$

Claim 2. It holds for $\mathbf{p}(x)$ and $u(x)$ defined in (1.29), that

$$\mathbf{p}(x) = Du(x) \quad \text{ for } x \in U.$$

We first prove claim 1. By Theorem 1.6, it holds for $y \in U \cap \Gamma$ that

$$F(\mathbf{p}(y, 0), z(y, 0), \mathbf{x}(y, 0)) = F(\mathbf{q}(y), g(y), y) = 0. \tag{1.32}$$

Let us denote $F(\mathbf{p}(y, s), z(y, s), \mathbf{x}(y, s))$ in the following equation by F, and differentiate (1.31) with regard to s:

$$\frac{\partial F}{\partial s} = D_p F \cdot \dot{\mathbf{p}} + D_z F \dot{z} + D_x F \cdot \dot{\mathbf{x}} = 0.$$

By the characteristic ODEs we then have

$$\begin{aligned} \frac{\partial F}{\partial s} &= D_p F \cdot (-D_x F - D_z F \mathbf{p}) + D_z F (D_p F \cdot \mathbf{p}) + D_x F \cdot (D_p F) \\ &= 0. \end{aligned} \tag{1.33}$$

By equations (1.32) and (1.33) it holds for every $y \in U \cap \Gamma$ and $s \in I$ that

$$F(\mathbf{p}(y, s), z(y, s), \mathbf{x}(y, s)) = 0.$$

By Theorem 1.8, for every $x \in U$ there are unique $y \in U \cap \Gamma$ and $s \in I$ satisfying $x = \mathbf{x}(y, s)$. Therefore it holds for every $x \in U$ that

$$F(\mathbf{p}(x), u(x), x) = 0. \tag{1.34}$$

This proves Claim 1.

Second, we prove Claim 2. Let us calculate the components of $Du(x)$.

$$\begin{aligned} u_{x_k}(x) &= \frac{\partial z}{\partial x_k}(\mathbf{y}(x), s(x)) \\ &= \sum_{i=1}^{n-1} z_{y_i}(\mathbf{y}(x), s(x)) y^i_{x_k}(x) + \dot{z}(\mathbf{y}(x), s(x)) s_{x_k}(x). \end{aligned} \tag{1.35}$$

From the characteristic ODEs, we have

$$\begin{aligned} \dot{z}(y,s) &= D_pF(\mathbf{p}(y,s), z(y,s)), \mathbf{x}(y,s)) \cdot \mathbf{p}(y,s) \\ &= \dot{\mathbf{x}}(y,s) \cdot \mathbf{p}(y,s). \end{aligned}$$

The chain rule gives for $i \in \{1, 2, \cdots, n-1\}$ that

$$\begin{aligned} \dot{z}_{y_i}(y,s) &= \frac{\partial}{\partial y_i} (\dot{\mathbf{x}}(y,s) \cdot \mathbf{p}(y,s)) \\ &= \dot{\mathbf{x}}_{y_i}(y,s) \cdot \mathbf{p}(y,s) + \dot{\mathbf{x}}(y,s) \cdot \mathbf{p}_{y_i}(y,s). \end{aligned} \tag{1.36}$$

From equation (1.31) we have $F_{y_i}(\mathbf{p}, z, \mathbf{x}) = 0$. On the other hand, the characteristic ODEs for $\mathbf{x}$ in (1.9a) give us that

$$\begin{aligned} F_{y_i}(\mathbf{p}, z, \mathbf{x}) &= D_pF(\mathbf{p}, z, \mathbf{x}) \cdot \mathbf{p}_{y_i} + D_zF(\mathbf{p}, z, \mathbf{x}) z_{y_i} + D_xF(\mathbf{p}, z, \mathbf{x}) \cdot \mathbf{x}_{y_i} \\ &= \dot{\mathbf{x}} \cdot \mathbf{p}_{y_i} + D_zF(\mathbf{p}, z, \mathbf{x}) z_{y_i} + D_xF(\mathbf{p}, z, \mathbf{x}) \cdot \mathbf{x}_{y_i}. \end{aligned}$$

Since $F_{y_i}(\mathbf{p}, z, \mathbf{x}) = 0$, the above equation becomes

$$\dot{\mathbf{x}} \cdot \mathbf{p}_{y_i} = -D_zF z_{y_i} - D_xF \cdot \mathbf{x}_{y_i}.$$

By substituting the identity above into (1.36), we obtain the following ODE for z_{y_i}:

$$\begin{aligned} \dot{z}_{y_i} &= \dot{\mathbf{x}}_{y_i} \cdot \mathbf{p} - D_zF z_{y_i} - D_xF \cdot \mathbf{x}_{y_i} \\ &= \dot{\mathbf{x}}_{y_i} \cdot \mathbf{p} - D_zF z_{y_i} + (\dot{\mathbf{p}} + D_zF\mathbf{p}) \cdot \mathbf{x}_{y_i} \\ &= \dot{\mathbf{x}}_{y_i} \cdot \mathbf{p} + \dot{\mathbf{p}} \cdot \mathbf{x}_{y_i} - D_zF z_{y_i} + D_zF\mathbf{p} \cdot \mathbf{x}_{y_i} \\ &= \frac{\partial}{\partial s} (\mathbf{p} \cdot \mathbf{x}_{y_i}) - D_zF(z_{y_i} - \mathbf{p} \cdot \mathbf{x}_{y_i}). \end{aligned} \tag{1.37}$$

Note that when $s = 0$ it holds by the compatibility conditions (1.19), (1.21) and Theorem 1.6 that

$$z_{y_i}(y,0) = g_{y_i}(y) = \mathbf{q}^i(y) = \mathbf{p}^i(y,0) \cdot \mathbf{x}_{y_i}(y,0).$$

Therefore $z_{y_i}(y,s) = \mathbf{p}(y,s) \cdot \mathbf{x}_{y_i}(y,s)$ is the solution to the ODE (1.37) with the initial condition $z_{y_i}(y,0) = \mathbf{p}(y,0) \cdot \mathbf{x}_{y_i}(y,0)$.

Let us now calculate $u_{x_k}(x)$:

$$
\begin{aligned}
u_{x_k}(x) &= \sum_{i=1}^{n-1} z_{y_i}(y,s)y^i_{x_k} + \dot{z}(y,s)s_{x_k} \\
&= \sum_{i=1}^{n-1} (\mathbf{p}(y,s)\cdot \mathbf{x}_{y_i}(y,s))y^i_{x_k} + (\dot{\mathbf{x}}(y,s)\cdot \mathbf{p}(y,s))s_{x_k} \\
&= \sum_{i=1}^{n-1}\sum_{j=1}^{n} p^j(y,s)x^j_{y_i}(y,s)y^i_{x_k} + \sum_{j=1}^{n} \dot{x}^j(y,s)p^j(y,s)s_{x_k} \\
&= \sum_{j=1}^{n} p^j(y,s)\left[\dot{x}^j(y,s)s_{x_k} + \sum_{i=1}^{n-1} x^j_{y_i}(y,s)y^i_{x_k}\right] \\
&= \sum_{j=1}^{n} p^j(y,s)x^j_{x_k}(y,s) \\
&= \sum_{j=1}^{n} p^j(x)\delta_{j,k} = p^k(x)
\end{aligned}
$$

Therefore $\mathbf{p}(x) = Du(x)$ for every $x \in U$. Since $\mathbf{p} \in C^1(U;\mathbb{R}^n)$ and $Du = \mathbf{p}$ in U, we have that $u \in C^2(U)$. By equation (1.34), u solves the PDE in U. As $u(x) = z(\mathbf{y}(x), s(x))$, the initial condition

$$z(y,0) = g(y) \text{ for } y \in U \cap \Gamma$$

ensures that u satisfies the boundary condition. Therefore u is a solution to the boundary value problem (1.30). □

We have shown that the method of characteristics can be used to find unique local solutions to boundary value problems. Next we show how this method is useful for studying the behaviour of the solutions.

2 Weak solutions

2.1 Quasilinear Partial Differential Equations

In this section we focus on scalar conservation laws, which are a type of quasilinear partial differential equations. A PDE of first order is said to be quasilinear, if it is of the form

$$F(Du(x), u(x), x) = \mathbf{a}(x, u(x)) \cdot Du(x) + b(x, u(x)) = 0. \tag{2.1}$$

The functions $\mathbf{a}\colon \mathbb{R}^{n+1} \to \mathbb{R}^n$ and $b\colon \mathbb{R}^{n+1} \to \mathbb{R}$ are assumed to be continuous and the unknown $u\colon \Omega \to \mathbb{R}$ is assumed to be continuously differentiable in some bounded domain $\Omega \subset \mathbb{R}^n$.

Let $p, x \in \mathbb{R}^n$ and $z \in \mathbb{R}$. Now

$$F(p, z, x) = \mathbf{a}(x, z) \cdot p + b(x, z),$$

and the method of characteristics gives the following characteristic ODEs:

$$\begin{cases} \dot{\mathbf{x}}(s) &= D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)) = \mathbf{a}(\mathbf{x}(s), z(s)); \\ \dot{z}(s) &= D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)) \cdot \mathbf{p}(s) = \mathbf{a}(\mathbf{x}(s), z(s)) \cdot \mathbf{p}(s); \\ \dot{\mathbf{p}}(s) &= -D_x F(\mathbf{p}(s), z(s), \mathbf{x}(s)) - D_z F(\mathbf{p}(s), z(s), \mathbf{x}(s))\mathbf{p}(s). \end{cases}$$

By equation (2.1), we have

$$\mathbf{a}(\mathbf{x}(s), z(s)) \cdot \mathbf{p}(s) + b(\mathbf{x}(s), z(s)) = 0.$$

Therefore we have

$$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{a}(\mathbf{x}(s), z(s)); & (2.2a) \\ \dot{z}(s) = -b(\mathbf{x}(s), z(s)). & (2.2b) \end{cases}$$

As the characteristic ODEs of $\mathbf{x}(s)$ and $z(s)$ are independent of $\mathbf{p}(s)$, we have a system of $n+1$ ODEs which can be solved independently of $\mathbf{p}(s)$. The equations for $\mathbf{p}(s)$ are therefore not needed to give the solution to boundary value problems for quasilinear PDE.

Suppose the boundary of Ω is smooth and we are given a boundary condition

$$u = g \quad \text{on } \Gamma \subset \partial\Omega.$$

As we may again suppose Γ to be flat close to any point $x^0 \in \Gamma$, the nonchracteristic condition (1.23) takes the form

$$F_{p_n}(p^0, z^0, x^0) = \mathbf{a}_n(x^0, g(x^0)) \neq 0.$$

Again, only the initial conditions x^0 and z^0 need to be considered. Without flattening the boundary of Ω, this condition is

$$\mathbf{a}(x^0, g(x^0)) \cdot \nu(x^0) \neq 0, \tag{2.3}$$

where $\nu(x^0)$ is the outward pointing unit normal vector of Γ at point x^0.

Example 2.1. (Inviscid Burgers' equation). Let us consider a quasilinear PDE of first order, known as inviscid Burgers' equation:

$$u_t(x,t) + u(x,t)u_x(x,t) = 0. \tag{2.4}$$

Here $u\colon \mathbb{R}\times[0,\infty) \to \mathbb{R}$ is assumed to be continuously differentiable. Let $p = (p_x, p_t) \in \mathbb{R}^2$, $y = (x,t) \in \mathbb{R}^2$ and $z \in \mathbb{R}$. Now we have

$$F(p,z,y) = p_t + zp_x = (z,1)\cdot p.$$

The characteristic ODEs (2.2a) and (2.2b) are now

$$\begin{cases} \dot{\mathbf{y}}(s) = (z(s),1); & (2.5a)\\ \dot{z}(s) = 0, & (2.5b)\end{cases}$$

where $\mathbf{y}(s) \coloneqq (x(s),t(s))$. As we can see, we do not need to introduce the characteristic ODEs for $\mathbf{p}(s)$.

Suppose that we are given the initial value condition

$$u(x,0) = g(x) \quad \text{for } x\in\mathbb{R}, \tag{2.6}$$

with $g\colon \mathbb{R}\to\mathbb{R}$ bounded. This gives the following initial conditions for the characteristic ODEs,

$$\begin{cases} \mathbf{y}(0) = (x(0),t(0)) = (x^0,0);\\ z(0) = g(x^0).\end{cases}$$

Since $F_p(p,z,y) = (z,1) \neq 0$, the noncharacteristic condition (2.3) holds for every $x\in\mathbb{R}$. Solving the initial value problems for the characteristic ODEs gives us

$$\begin{cases} \mathbf{y}(t) = (x(t),t) = (g(x^0)t + x^0, t);\\ z(t) = g(x^0).\end{cases}$$

The solution $u(x,t)$ takes the constant value $u(x,t) = g(x^0)$ along the projected characteristic $\mathbf{y}(t)$, that is, on the half-line

$$(x,t) = (g(x^0)t + x^0, t), \quad t \geq 0.$$

Suppose that $x^0, x^1 \in \mathbb{R}$ are two distinct points with $g(x^0) \neq g(x^1)$. Now $u(x,t) = g(x^0)$ for $(x,t) = (g(x^0)t + x^0, t)$ and $u(x,t) = g(x^1)$ for $(x,t) = (g(x^1)t + x^1, t)$. Suppose it holds for some $t_1 > 0$ that

$$t_1 = \frac{x^0 - x^1}{g(x^1) - g(x^0)}.$$

Then we have

$$g(x^0)t_1 + x^0 = g(x^1)t_1 + x^1,$$

which is a contradiction, since the solution takes two different values at $(g(x^0)t_1 + x^0, t_1)$. We arrive at the conclusion that the initial value problem (2.4) and (2.6) does not in general have a continuous solution defined for all $t > 0$, as it is possible for two projected characteristics to intersect, without giving a unique value for the solution.

2.2 Scalar Conservation Laws

In the light of Example 2.1, we need different notions of solutions. In this section we define integral solutions to initial value problems for scalar conservation laws with one space variable:

$$\begin{cases} u_t + [F(u)]_x = 0 & \text{in } \mathbb{R} \times (0, \infty); \\ u(x, 0) = g(x) & \text{for } x \in \mathbb{R}. \end{cases} \tag{2.7}$$

Functions $F\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ are given, with F being differentiable and g bounded. The unknown $u\colon \mathbb{R} \times [0, \infty) \to \mathbb{R}$ is a function of one space variable and one time variable.

In defining integral solutions and deriving the Rankine-Hugoniot jump condition, we follow the theory introduced in section 3.4. in [3] and examples shown on the lecture handling conservation laws in [7]. The initial value problems for inviscid Burgers' equation provides clear insight into the integral solutions with shock waves and rarefaction waves.

2.3 Integral Solutions

We introduce solutions that are not necessarily differentiable at every point $(x, t) \in \mathbb{R} \times (0, \infty)$. To be more precise, we want u to solve (2.7) almost everywhere in $\mathbb{R} \times [0, \infty)$, with respect to the Lebesgue measure in $\mathbb{R}^2$.

We assume temporarily that u is differentiable in $\mathbb{R} \times (0, \infty)$ and solves the initial value problem (2.7). Let $v\colon \mathbb{R} \times [0, \infty) \to \mathbb{R}$ be a smooth function with a compact support in $\mathbb{R} \times [0, \infty)$. By equation (2.7), we have

$$v(u_t + [F(u)]_x) = 0.$$

Integrating both sides of the identity over $\mathbb{R} \times [0, \infty)$, we obtain

$$0 = \int_0^\infty \int_{-\infty}^\infty v(x,t) u_t(x,t)\, dx\, dt + \int_0^\infty \int_{-\infty}^\infty v(x,t)\, [F(u(x,t))]_x\, dx\, dt.$$

As the support of v is compact, we have by integration by parts that

$$\begin{aligned}
0 &= -\int_0^\infty \int_{-\infty}^\infty v_t(x,t)u(x,t)\,dx\,dt - \int_{-\infty}^\infty v(x,0)u(x,0)\,dx \\
&\quad - \int_0^\infty \int_{-\infty}^\infty v_x(x,t)F(u(x,t))\,dx\,dt \\
&= -\int_0^\infty \int_{-\infty}^\infty v_t(x,t)u(x,t) + v_x(x,t)F(u(x,t))\,dx\,dt - \int_{-\infty}^\infty v(x,0)u(x,0)\,dx.
\end{aligned}$$

Since $u(x,0) = g(x)$, we obtain

$$0 = \int_0^\infty \int_{-\infty}^\infty v_t(x,t)u(x,t) + v_x(x,t)F(u(x,t))\,dx\,dt + \int_{-\infty}^\infty v(x,0)g(x)\,dx. \tag{2.8}$$

Hence we have arrived at an identity that only requires u to be an an essentially bounded function, denoted as $u \in L^\infty(\mathbb{R}\times(0,\infty))$. That is, there exists a bound $M > 0$ such that the set $\{(x,t) \in (\mathbb{R}\times(0,\infty)) \mid |u(x)| > M\}$ is of Lebesgue measure zero. We use (2.8) to define integral solutions to the initial value problem (2.7).

Definition 2.2. Function $u \in L^\infty(\mathbb{R}\times(0,\infty))$ is called an *integral solution* to the initial value problem (2.7), if it satisfies identity (2.8) for every smooth test function $v\colon \mathbb{R}\times(0,\infty) \to \mathbb{R}$ with a compact support in $\mathbb{R}\times[0,\infty)$.

2.3.1 Rankine-Hugoniot Jump Conditions

Let us further investigate the properties of integral solutions. Suppose u is an integral solution to (2.7). Suppose further that inside some bounded domain $V \subset \mathbb{R}\times(0,\infty)$ solution u is differentiable on both sides of a C^1 curve S. Denote the components of $V\setminus S$ as V_l and V_r. To use integration by parts, we assume that u and its first order partial derivatives are uniformly continuous in both V_l and V_r.

Let $v\colon \mathbb{R}^2 \to \mathbb{R}$ be a smooth function with a compact support in V. Since u is an integral solution of (2.7), identity (2.8) holds for u and v. As v has a compact support in V, it holds that $v(x,0) = 0$ for all $x \in \mathbb{R}$. Therefore (2.8) becomes

$$\iint_V v_t u + v_x F(u)\,dx\,dt = 0. \tag{2.9}$$

We can divide the integration domain into V_l and V_r. As u is differentiable in these domains, we may use integration by parts to obtain

$$\begin{aligned} 0 &= \iint\limits_{V_l} v_t u + v_x F(u)\, dx\, dt + \iint\limits_{V_r} v_t u + v_x F(u)\, dx\, dt \\ &= -\iint\limits_{V_l} v\left(u_t + [F(u)]_x\right)\, dx\, dt + \int\limits_S v\left(u_l \nu^2 + F(u_l)\nu^1\right)\, ds \\ &\quad - \iint\limits_{V_r} v\left(u_t + [F(u)]_x\right)\, dx\, dt - \int\limits_S v\left(u_r \nu^2 + F(u_r)\nu^1\right)\, ds, \end{aligned} \tag{2.10}$$

where $\nu = (\nu^1, \nu^2)$ is the unit normal vector of S pointing into V_r. By u_l and u_r we denote the limits of u on S from V_l and V_r respectively.

Suppose that the support of v is a subset of V_l. Then (2.10) is reduced to

$$0 = -\iint\limits_{V_l} v\left(u_t + [F(u)]_x\right)\, dx\, dt. \tag{2.11}$$

Since (2.11) holds for every smooth function with a compact support in V_l, we have that

$$u_t + [F(u)]_x = 0 \quad \text{in } V_l.$$

Similarly, we also have

$$u_t + [F(u)]_x = 0 \quad \text{in } V_r.$$

Therefore equation (2.10) is reduced to

$$\begin{aligned} 0 &= \int\limits_S v\left(u_l \nu^2 + F(u_l)\nu^1\right)\, ds - \int\limits_S v\left(u_r \nu^2 + F(u_r)\nu^1\right)\, ds \\ &= \int\limits_S v\left[(u_l - u_r)\nu^2 + (F(u_l) - F(u_r))\nu^1\right]\, ds. \end{aligned}$$

Again, this identity must hold for every smooth v with a compact support in V. Therefore it holds for u on S that

$$(F(u_l) - F(u_r))\nu^1 + (u_l - u_r)\nu^2 = 0.$$

This can be written as

$$\frac{F(u_l) - F(u_r)}{u_l - u_r} = -\frac{\nu^2}{\nu^1}. \tag{2.12}$$

Let us assume that the curve S can be parametrized with some differentiable function $s\colon \mathbb{R} \to \mathbb{R}$ as

$$S = \{(x,t) \in \mathbb{R}\times[0,\infty) \mid x = s(t),\, t > 0\}.$$

As the unit normal vector ν of S is perpendicular to $(\dot{s}(t), 1)$ for every $t > 0$, it holds that

$$\dot{s}(t) = -\frac{\nu^2(s(t),t)}{\nu^1(s(t),t)}.$$

Now we can rewrite identity (2.12) as

$$\frac{F(u_l) - F(u_r)}{u_l - u_r} = \dot{s}.$$

The above identity imposed on an integral solution u of (2.7) is called the *Rankine-Hugoniot jump condition* and is useful as it characterizes the behaviour of an integral solution on a C^1 curve of discontinuity.

In the above we proved the following theorem.

Theorem 2.3. (Rankine-Hugoniot Jump Condition). *Let $u\colon \mathbb{R} \to [0,\infty)$ be an integral solution to* (2.7)*. Let $s \in C^1(\mathbb{R})$ parametrize a curve S by*

$$S = \{(x,t) \in \mathbb{R}\times[0,\infty) \mid x = s(t), t > 0\}.$$

Suppose $V \subset \mathbb{R}\times(0,\infty)$ is a bounded domain that S divides into two components V_l and V_r. Then, if u and its partial derivatives are uniformly continuous in V_l and V_r, it holds that

$$F(u_l) - F(u_r) = \dot{s}(u_l - u_r) \quad \text{on } S.$$

Here u_l and u_r are limits of u from V_l and V_r, respectively.

Example 2.4. Consider the initial value problem for inviscid Burgers' equation

$$\begin{cases} u_t + uu_x = 0 & \text{in } \mathbb{R}\times(0,\infty); \\ u = g & \text{on } \mathbb{R}\times\{t=0\}, \end{cases} \tag{2.13}$$

with $g\colon \mathbb{R} \to \mathbb{R}$ defined as

$$g(x) = \begin{cases} 1, & \text{if } x < 0; \\ 0, & \text{if } x \geq 0. \end{cases}$$

Note that inviscid Burgers' equation is indeed a scalar conservation law and can be obtained from (2.7) by setting

$$F(u) = \frac{1}{2}u^2.$$

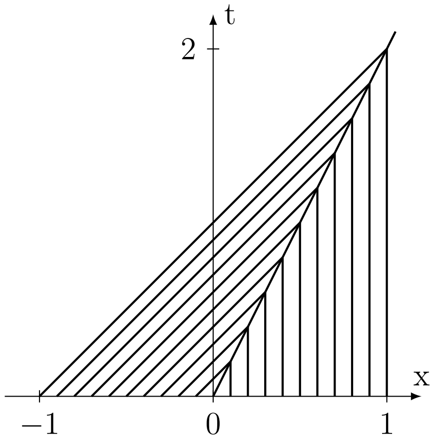

Figure 1: The projected characteristics and curve of discontinuity S for the integral solution of Example 2.4.

Let $x^0 \in \mathbb{R}$. By the calculations in Example 2.1, a continuous solution to (2.13) takes the constant value $u(\mathbf{y}(x^0,t)) = g(x^0)$ on the projected characteristic $\mathbf{y}(x^0,t) = (g(x^0)t + x^0, t)$. Since it holds that

$$\mathbf{y}(x^0,t) = \begin{cases} (t + x^0, t) & \text{for } x^0 < 0; \\ (x^0, t) & \text{for } x^0 \geq 0, \end{cases}$$

every projected characteristic starting from a point on the negative real axis intersects with the projected characteristic starting from the origin.

Therefore there does not exist a solution that would be continuous in the whole half plane $\mathbb{R}\times[0,\infty)$. We apply the Rankine-Hugoniot jump condition to find an integral solution to (2.13). Based on the initial condition, we seek to parameterize a curve of discontinuity S, with $u_l = 1$ and $u_r = 0$. Thus it holds that $F(u_l) = \frac{1}{2}$ and $F(u_r) = 0$. By the Rankine-Hugoniot jump condition, the function $s\colon \mathbb{R} \to \mathbb{R}$ parameterizing S satisfies

$$\dot{s}(t) = \frac{\frac{1}{2} - 0}{1 - 0} = \frac{1}{2} \quad \text{for } t > 0.$$

Since the discontinuity starts at the origin, we set $s(t) = \frac{t}{2}$ and define the integral solution u as

$$u(x,t) = \begin{cases} 1 & \text{for } x < \frac{t}{2}; \\ 0 & \text{for } x \geq \frac{t}{2}. \end{cases}$$

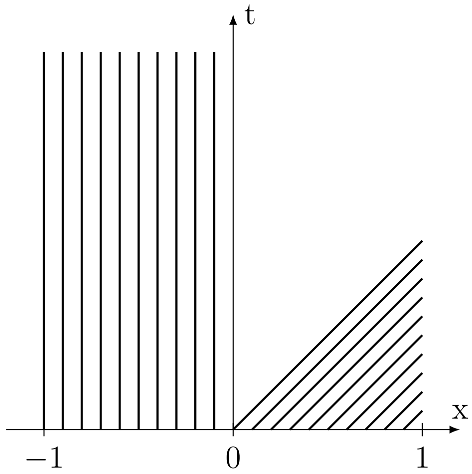

Figure 2: Trajectories of the projected characteristics given by the method of characteristics in Example 2.5.

2.3.2 Entropy Conditions

Example 2.5. Let us continue with another initial value problem for inviscid Burgers' equation. Define the initial value g_1 as

$$g_1(x) = \begin{cases} 0 & \text{if } x < 0; \\ 1 & \text{if } x \geq 0. \end{cases}$$

Now the projected characteristics do not intersect. In fact, the projected characteristics are

$$\mathbf{y}(x^0, t) = \begin{cases} (x^0, t) & \text{for } x^0 < 0; \\ (t + x^0, t) & \text{for } x^0 \geq 0. \end{cases}$$

With the method of characteristics, we can uniquely define the solution for those $(x, t) \in \mathbb{R} \times [0, \infty)$ satisfying $x < 0$, or $x \geq t$, as

$$u(x, t) = \begin{cases} 0 & x < 0; \\ 1 & x \geq t. \end{cases}$$

As demonstrated in Figure 2, the method of characteristics yields no information about determining the solution when $0 < x < t$.

Let us again construct an integral solution by the use of the Rankine-Hugoniot jump condition. As before, we will define an integral solution with a single curve of discontinuity.

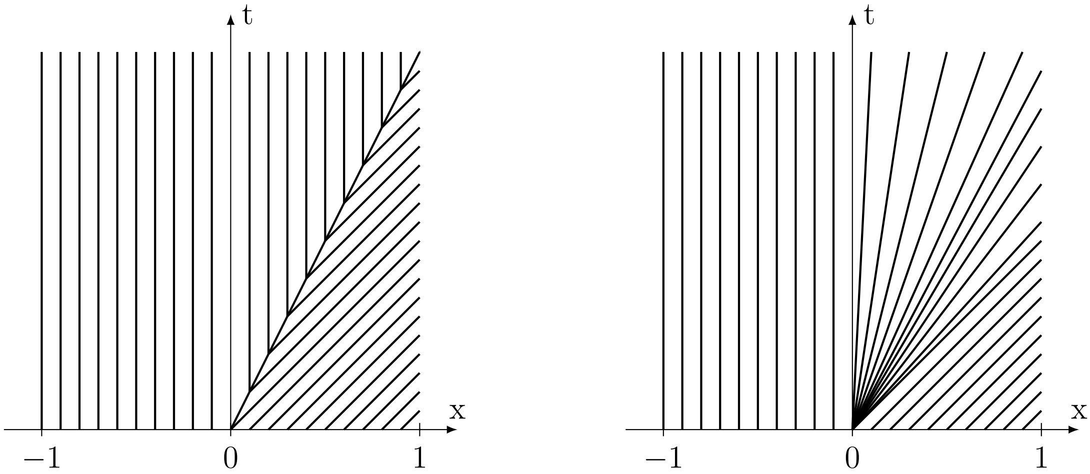

Figure 3: Visualizations of the two integral solutions given for example 2.5.

By the Rankine-Hugoniot jump condition the function s parameterizing the curve of discontinuity must satisfy

$$\dot{s}(t) = \frac{F(u_l) - F(u_r)}{u_l - u_r} = \frac{0 - \frac{1}{2}}{0 - 1} = \frac{1}{2}.$$

As the discontinuity again starts at the origin, we set $s(t) = \frac{t}{2}$ and define

$$u(x,t) = \begin{cases} 0 & x < \frac{t}{2}; \\ 1 & x \geq \frac{t}{2}, \end{cases}$$

for $(x,t) \in \mathbb{R} \times [0, \infty)$.

Instead of giving a discontinuous integral solution, we can also construct a continuous integral solution called a *rarefaction wave*. Define

$$u(x,t) = \begin{cases} 0 & x < 0; \\ \frac{x}{t} & 0 \leq x < t; \\ 1 & x \geq t. \end{cases}$$

Now there are two curves,

$$S_1 = \{(x,t) \in \mathbb{R} \times (0,\infty) \mid x = 0,\, t > 0\}$$

and

$$S_2 = \{(x,t) \in \mathbb{R} \times (0,\infty) \mid x = t,\, t > 0\},$$

where the solution is not differentiable. These curves can be represented parametrically with functions $s_1(t) = 0$ and $s_2(t) = x$ respectively. It holds that both

$$\frac{F(u_{l_1}) - F(u_{r_1})}{u_{l_1} - u_{r_1}} = \frac{\frac{1}{2}(u_{l_1} - u_{r_1})^2}{u_{l_1} - u_{r_1}} = \frac{u_{l_1} + u_{r_1}}{2} = 0 = \dot{s}_1(t)$$

and

$$\frac{F(u_{l_2}) - F(u_{r_2})}{u_{l_2} - u_{r_2}} = \frac{u_{l_2} + u_{r_2}}{2} = \frac{1+1}{2} = 1 = \dot{s}_2(t).$$

Hence u meets the Rankine-Hugoniot jump condition on both S_1 and S_2. It also holds for $0 < x < t$, that

$$\begin{aligned} u_t(x,t) + [F(u(x,t))]_x &= u_t(x,t) + u(x,t)u_x(x,t) \\ &= -\frac{x}{t^2} + \frac{x}{t^2} = 0. \end{aligned}$$

Therefore u satisfies equation (2.13) in $\mathbb{R} \times [0, \infty)$ with the exception of the curves S_1 and S_2. As u meets the Rankine-Hugoniot jump condition on these curves, it is therefore an integral solution to inviscid Burgers' equation with the initial value condition given by function g_1.

As demonstrated by the previous example, an integral solution is not necessarily unique. Of the two solutions given in Example 2.5, the later can be seen as a more desirable one, since it does not contain discontinuities and could therefore, for example, give a better representation of a physical wave.

To eliminate some of the physically less interesting solutions, we consider the two initial value problems given above. In Example 2.4 the projected characteristics intersect and therefore having a solution with discontinuities is necessary. On the other hand, in Example 2.5 the projected characteristics do not intersect and we were able to give a continuous integral solution.

In general, it holds for scalar conservation laws that the solution takes a constant value $g(x^0)$ on projected characteristics

$$\mathbf{y}(t) = (F'(g(x^0))t + x^0, t) \quad t \geq 0.$$

Therefore we would only like to accept discontinuous solutions, if the projected characteristics on the left side of the curve of discontinuities move faster than the projected characteristics on the right side. This holds, if

$$F'(u_l) > \dot{s} > F'(u_r). \tag{2.14}$$

The inequalities above are called the *entropy condition*, or the Lax entropy condition. A curve of discontinuity that meets both the Rankine-Hugoniot jump condition and the Lax entropy condition, is called a *shock curve* and we call an integral solution, that only has discontinuities on shock curves, an *admissible solution*.

As we aim to give conditions to ensure the existence of unique weak solutions to initial value problems for scalar conservation laws, it turns out that the entropy condition plays an important role. In Section 4 we give a different form of the entropy condition and achieve uniqueness of entropy solutions to

$$\begin{cases} u_t + [F(u)]_x = 0 & \text{in } \mathbb{R} \times (0, \infty); \\ u = g & \text{on } \mathbb{R} \times \{t = 0\}, \end{cases}$$

when $g \in L^\infty(\mathbb{R})$ and F is a convex C^2 function.

In the next section we give an introduction to Calculus of Variations. We aim to use this theory in Section 4 to define the Lax-Oleinik formula for scalar conservation laws. This formula then is used to prove the existence and uniqueness of entropy solutions.

3 Hamilton-Jacobi Equations

Suppose that u is a solution to the initial value problem for a scalar conservation law

$$\begin{cases} u_t + [F(u)]_x = 0 & \text{in } \mathbb{R} \times (0, \infty); \\ u = g & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases}$$

Suppose further that $u = w_x$ for some smooth $w\colon \mathbb{R} \times [0, \infty) \to \mathbb{R}$. Now w is a solution to the following initial value problem:

$$\begin{cases} w_{xt} + [F(w_x)]_x = 0 & \text{in } \mathbb{R} \times (0, \infty); \\ w_x = g & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases} \tag{3.1}$$

Therefore, (3.1) is satisfied, if w is the solution to

$$\begin{cases} w_t + F(w_x) = 0 & \text{in } \mathbb{R} \times (0, \infty); \\ w = h & \text{on } \mathbb{R} \times \{t = 0\}. \end{cases} \tag{3.2}$$

In the above $h\colon \mathbb{R} \to \mathbb{R}$ is defined as

$$h(x) = \int_0^x g(y)\, dy.$$

Problem (3.2) is an initial value problem for a 1-dimensional version of the *Hamilton-Jacobi equations*:

$$\begin{cases} u_t + H(D_x u, x) = 0 & \text{in } \mathbb{R}^n \times (0, \infty); \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{3.3}$$

In the above, function $g\colon \mathbb{R}^n \to \mathbb{R}$ and Hamiltonian $H\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ are given and $u\colon \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is the unknown. Note that equation (3.2), is a special case of (3.3). In the following, we consider the general equation (3.3) and in Section 3.3 we return to (3.2).

In this section we study the existence and uniqueness of weak solutions to the Hamilton-Jacobi equations. In the theory presented, we refer to sections 3.3. and 3.4. in [3].

3.1 Derivation of Hamilton's Equations

We use the method of characteristics to consider the Hamilton-Jacobi equations. Let $z \in \mathbb{R}$ and define $q, y \in \mathbb{R}^{n+1}$ as

$$\begin{cases} q := (p, p_t) := (p_1, p_2, \cdots, p_n, p_t); \\ y := (x, t) := (x_1, x_2, \cdots, x_n, t). \end{cases}$$

To apply the method of characteristic, we define $F\colon \mathbb{R}^{n+1} \times \mathbb{R} \times \mathbb{R}^{n+1} \to \mathbb{R}$ as

$$F(q, z, y) = p_t + H(p, x).$$

Now the Hamilton-Jacobi equations are given in $\mathbb{R}^n \times (0,\infty)$ by F as follows:

$$F(Du(x,t), u(x,t), (x,t)) = 0.$$

The gradients of F appearing in the characteristic ODEs are

$$\begin{cases} D_qF(q, z, y) = (D_pH(p, x), 1); \\ D_zF(q, z, y) = 0; \\ D_yF(q, z, y) = (D_xH(p, x), 0). \end{cases}$$

Therefore we have the following characteristic ODEs:

$$\begin{cases} \dot{\mathbf{y}}(s) = (D_pH(\mathbf{p}(s), \mathbf{x}(s)), 1); & (3.4a) \\ \dot{z}(s) = (D_pH(\mathbf{p}(s), \mathbf{x}(s)), 1) \cdot \mathbf{q}(s); & (3.4b) \\ \dot{\mathbf{q}}(s) = -(D_xH(\mathbf{p}(s), \mathbf{x}(s)), 0). & (3.4c) \end{cases}$$

As $\dot{\mathbf{y}}^{n+1}(s) = 1$ in (3.4a), we set $s = t$ to reduce the ODEs above into a system of $2n+1$ characteristic ODEs:

$$\begin{cases} \dot{\mathbf{x}}(t) = D_pH(\mathbf{p}(t), \mathbf{x}(t)); & (3.5a) \\ \dot{z}(t) = D_pH(\mathbf{p}(t), \mathbf{x}(t)) \cdot \mathbf{p}; & (3.5b) \\ \dot{\mathbf{p}}(t) = -D_xH(\mathbf{p}(t), \mathbf{x}(t)). & (3.5c) \end{cases}$$

Equations (3.5a) and (3.5c) together form a closed system of $2n$ ODEs. The ODE for $z(t)$ can be solved by substituting $\mathbf{x}(t)$ and $\mathbf{p}(t)$ into (3.5b) and integrating the obtained equation with respect to t. Therefore the system of characteristic ODEs for the Hamilton-Jacobi equation can be reduced to

$$\begin{cases} \dot{\mathbf{x}} = D_p H(\mathbf{p}, \mathbf{x}); \\ \dot{\mathbf{p}} = -D_x H(\mathbf{p}, \mathbf{x}). \end{cases} \tag{3.6}$$

These equations are called *Hamilton's equations* and are used in Hamiltonian mechanics. To further study and solve these equations, we use calculus of variations.

3.2 Introduction to the Calculus of Variations

A fundamental problem in the calculus of variations is to find a minimizer for the action integral

$$I[\mathbf{w}] = \int_0^t L(\dot{\mathbf{w}}(s), \mathbf{w}(s))\, ds \tag{3.7}$$

in the class of admissible paths

$$\mathcal{K} = \{\mathbf{w} \in C^2([0,t]; \mathbb{R}^n) \mid \mathbf{w}(0) = y, \mathbf{w}(t) = x\}. \tag{3.8}$$

Above the mapping $(q, x) \mapsto L(q, x)$ from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$ is called the *Lagrangian* and is here assumed to be a given C^2 function.

In problems modelling physical systems, the admissible class $\mathcal{K}$ could for example represent the class of possible paths an object can take from point x to y in a given time step t. The action integral could in this case be an integral over time of the difference between kinetic and potential energy of the object on the path. The path taken by a physical system is the minimizer, or at least a critical point of the action integral I.

Theorem 3.1. *Let* $\mathbf{x} \in \mathcal{K}$ *be a path minimizing the action integral* (3.7). *Then,* $\mathbf{x}$ *is a solution to the* Euler-Lagrange equations*:*

$$-\frac{d}{ds} D_q L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) + D_x L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) = 0, \tag{3.9}$$

when $0 < s < t$.

Proof. Suppose $\mathbf{x} \in \mathcal{K}$ is a minimizer of the action integral (3.7) and take $\mathbf{q} \in C^\infty([0,t];\mathbb{R}^n)$ satisfying $\mathbf{q}(0) = \mathbf{q}(t) = 0$. Then the paths defined as

$$\varphi_\tau(t) \coloneqq \mathbf{x}(t) + \tau\mathbf{q}(t),$$

are in $\mathcal{K}$ for every $\tau \in \mathbb{R}$. Let $A\colon \mathbb{R} \to \mathbb{R}$ be defined as

$$A(\tau) \coloneqq I[\varphi_\tau].$$

As $\mathbf{x}$ is a minimizer of I, A has a minimum at $\tau = 0$. Thus it holds that

$$A'(0) = 0.$$

Let $q, x \in \mathbb{R}^n$. We write $D_qL(q,x)$ and $D_xL(q,x)$ for the gradients of L with respect to the first and second n-dimensional variables q and x respectively. Let us next calculate $A'(\tau)$ from the action integral. By definition of A, we have

$$A'(\tau) = \frac{d}{d\tau}\int_0^t L(\dot\varphi_\tau(s), \varphi_\tau(s))\, ds$$

As L, φ_τ and $\dot\varphi_\tau$ are differentiable functions, we have

$$\frac{d}{d\tau}\int_0^t L(\dot\varphi_\tau(s), \varphi_\tau(s))\, ds = \int_0^t \frac{d}{d\tau} L(\dot\varphi_\tau(s), \varphi_\tau(s))\, ds.$$

By the chain rule,

$$\begin{aligned}\int_0^t \frac{d}{d\tau} L(\dot\varphi_\tau(s), \varphi_\tau(s))\, ds &= \int_0^t D_qL(\dot\varphi_\tau(s), \varphi_\tau(s)) \cdot \frac{d}{d\tau}\dot\varphi_\tau(s)\, ds \\ &+ \int_0^t D_xL(\dot\varphi_\tau(s), \varphi_\tau(s)) \cdot \frac{d}{d\tau}\varphi_\tau(s)\, ds.\end{aligned}$$

Recall that $\varphi_\tau(s) = \mathbf{x}(s) + \tau\mathbf{q}(s)$. Therefore, we have

$$\begin{aligned}A'(\tau) = &\int_0^t D_qL(\dot{\mathbf{x}}(s) + \tau\dot{\mathbf{q}}(s), \mathbf{x}(s) + \tau\mathbf{q}(s)) \cdot \dot{\mathbf{q}}(s)\, ds \\ &+ \int_0^t D_xL(\dot{\mathbf{x}}(s) + \tau\dot{\mathbf{q}}(s), \mathbf{x}(s) + \tau\mathbf{q}(s)) \cdot \mathbf{q}(s)\, ds.\end{aligned}$$

Let $\tau = 0$. Since $A'(0) = 0$, we have

$$0 = \int_0^t D_q L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) \cdot \dot{\mathbf{q}}(s)\, ds + \int_0^t D_x L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) \cdot \mathbf{q}(s)\, ds.$$

Note that $\mathbf{q}(0) = \mathbf{q}(t) = 0$. By integration by parts, we obtain that

$$0 = \int_0^t \left(-\frac{d}{ds} D_q L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) + D_x L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) \right) \cdot \mathbf{q}(s)\, ds.$$

Since the identity above holds for every smooth $\mathbf{q}$ with $\mathbf{q}(0) = \mathbf{q}(t) = 0$, we conclude that

$$-\frac{d}{ds} D_q L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) + D_x L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) = 0,$$

for $s \in (0, t)$. This concludes the proof. □

Let $x, p \in \mathbb{R}^n$. We define the Hamiltonian $H \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ associated to the Lagrangian L as

$$H(p, x) = p \cdot \mathbf{q}(p, x) - L(\mathbf{q}(p, x), x), \tag{3.10}$$

where we assume $\mathbf{q}(p, x)$ to be the unique differentiable solution to

$$p = D_q L(\mathbf{q}, x). \tag{3.11}$$

Theorem 3.2. *Suppose $\mathbf{x} \colon [0, t] \to \mathbb{R}^n$ is a minimizer of the action integral* (3.7)*, among the class of admissible paths* (3.8)*. Define*

$$\mathbf{p}(s) := D_q L(\dot{\mathbf{x}}(s), \mathbf{x}(s)) \quad \textit{for } s \in (0, t). \tag{3.12}$$

Then $\mathbf{x}(s)$ and $\mathbf{p}(s)$ are solutions to Hamilton's equations

$$\begin{cases} \dot{\mathbf{x}}(s) = D_p H(\mathbf{p}(s), \mathbf{x}(s)); \\ \dot{\mathbf{p}}(s) = -D_x H(\mathbf{p}(s), \mathbf{x}(s)), \end{cases}$$

for $s \in (0, t)$.

Proof. As the Hamiltonian is defined by (3.10), it holds for every $p, x \in \mathbb{R}^n$ that

$$\begin{aligned} D_p H(p, x) &= \mathbf{q}(p, x) + p \cdot D_p \mathbf{q}(p, x) - D_p L(\mathbf{q}(p, x), x) \\ &= \mathbf{q}(p, x) + p \cdot D_p \mathbf{q}(p, x) - D_q L(\mathbf{q}(p, x), x) \cdot D_p \mathbf{q}(p, x) \\ &= \mathbf{q}(p, x) + (p - D_q L(\mathbf{q}(p, x), x)) \cdot D_p \mathbf{q}(p, x). \end{aligned}$$

By (3.11), it therefore holds that

$$D_pH(p,x) = \mathbf{q}(p,x). \tag{3.13}$$

On the other hand, $\mathbf{q}(\mathbf{p}(s), \mathbf{x}(s)) = \dot{\mathbf{x}}(s)$ for $s \in (0,t)$, as

$$D_qL(\dot{\mathbf{x}}(s), \mathbf{x}(s)) = D_qL(\mathbf{q}(\mathbf{p}(s), \mathbf{x}(s)), \mathbf{x}(s))$$

by (3.12) and (3.11). Hence we have for $s \in (0,t)$, that

$$D_pH(\mathbf{p}(s), \mathbf{x}(s)) = \dot{\mathbf{x}}(s).$$

Thus we have arrived at the first of Hamilton's equations. By the definition of the Hamiltonian in (3.10), we have for every $x, p \in \mathbb{R}^n$ that

$$\begin{aligned} D_xH(p,x) &= p \cdot D_x\mathbf{q}(p,x) - D_x\big(L(\mathbf{q}(p,x),x)\big) \\ &= p \cdot D_x\mathbf{q}(p,x) - D_qL(\mathbf{q}(p,x),x) \cdot D_x\mathbf{q}(p,x) - D_xL(\mathbf{q}(p,x),x) \\ &= (p - D_qL(\mathbf{q}(p,x),x)) \cdot D_x\mathbf{q}(p,x) - D_xL(\mathbf{q}(p,x),x). \end{aligned}$$

By (3.12) and (3.11), we have for $\mathbf{p}(s)$ and $\mathbf{x}(s)$ that

$$D_xH(\mathbf{p}(s), \mathbf{x}(s)) = -D_xL(\mathbf{q}(\mathbf{p}(s), \mathbf{x}(s)), \mathbf{x}(s)) = -D_xL(\dot{\mathbf{x}}(s), \mathbf{x}(s)).$$

As $\mathbf{x}(s)$ is a solution to the Euler-Lagrange equations presented in Theorem 3.1, we have

$$D_xL(\dot{\mathbf{x}}(s), \mathbf{x}(s)) = \frac{d}{ds}D_qL(\dot{\mathbf{x}}(s), \mathbf{x}(s)) = \dot{\mathbf{p}}(s).$$

Therefore, for $s \in (0,t)$,

$$\dot{\mathbf{p}}(s) = -D_xH(\mathbf{p}(s), \mathbf{x}(s)).$$

This proves the theorem. $\square$

In the above, we have shown the connection between the minimizer of the action integral and the solutions of Hamilton's equations. In Section 3.4 we use this connection to solve the Hamilton-Jacobi equations by finding a minimizer to the action integral.

3.3 Legendre Transform

We continue to study the connection between the Lagrangian and the Hamiltonian associated to it.

We assume the Lagrangian to be a continuous mapping from $\mathbb{R}^n$ to $\mathbb{R}$. Further, we assume that the Lagrangian is convex and satisfies the following growth condition:

$$\lim_{|q|\to\infty} \frac{L(q)}{|q|} = \infty. \tag{3.14}$$

Definition 3.3. The *Legendre transform* of $L\colon \mathbb{R}^n \to \mathbb{R}$ is

$$L^*(p) = \sup_{q\in\mathbb{R}^n} \{p \cdot q - L(q)\} \quad \text{for } p \in \mathbb{R}^n.$$

By condition (3.14) it holds for every $p \in \mathbb{R}^n$, that

$$\lim_{|q|\to\infty} p \cdot q - L(q) = -\infty.$$

As L is continuous, by the extreme value theorem, there exists at least one $q^* \in \mathbb{R}^n$ for each $p \in \mathbb{R}^n$ such that

$$L^*(p) = p \cdot q^* - L(q^*). \tag{3.15}$$

If L is differentiable, q^* is a solution, albeit not necessarily the unique one, to $p = D_q L(q^*)$. Provided $q^* = \mathbf{q}(p)$ is unique for $p \in \mathbb{R}^n$ and this solution is differentiable, the Legendre transform gives the Hamiltonian associated with the Lagrangian L as defined in (3.10). Consequently, we write $H = L^*$.

Since originally we assumed the Hamiltonian, rather than the Lagrangian, to be the function given in the Hamilton-Jacobi equations, let us clarify the connection between these two functions and show how one can be obtained when the other is known.

Theorem 3.4. *Let $L\colon \mathbb{R}^n \to \mathbb{R}$ be convex and satisfy the growth condition* (3.14)*. Then $H = L^*$ is convex and for $p \in \mathbb{R}^n$,*

$$\lim_{p\to\infty} \frac{H(p)}{|p|} = \infty. \tag{3.16}$$

Moreover, $L = H^*$.

Proof. First we show that $H = L^*$ is convex. Indeed, let $\tau \in (0,1)$ and $p, r \in \mathbb{R}^n$. It holds that

$$\begin{aligned} H(\tau p + (1-\tau)r) &= \sup_{q\in\mathbb{R}^n} \{(\tau p + (1-\tau)r)\cdot q - L(q)\} \\ &= \sup_{q\in\mathbb{R}^n} \{\tau\,(p\cdot q - L(q)) + (1-\tau)\,(r\cdot q - L(q))\} \\ &\le \tau \sup_{q\in\mathbb{R}^n} \{p\cdot q - L(q)\} + (1-\tau)\sup_{q\in\mathbb{R}^n} \{r\cdot q - L(q)\} \\ &= \tau H(p) + (1-\tau)H(r). \end{aligned}$$

Hence H is convex.

Second, we prove (3.16). Let $\tau > 0$ and $p \in \mathbb{R}^n$ with $p \neq 0$. Let $\tilde{q} = \tau \frac{p}{|p|}$. Now it holds that

$$\begin{aligned} H(p) &= \sup_{q\in\mathbb{R}^n} \{p\cdot q - L(q)\} \ge p\cdot\tilde{q} - L(\tilde{q}) \\ &= \tau|p| - L(\tau\frac{p}{|p|}) \ge \tau|p| - \max_{q\in B(0,\tau)} L(q). \end{aligned}$$

As it holds that

$$\lim_{|p|\to\infty} \frac{1}{|p|}\left(|p|\tau - \max_{q\in B(0,\tau)} L(q)\right) = \tau,$$

we have shown that for every $\tau > 0$,

$$\liminf_{|p|\to\infty} \frac{H(p)}{|p|} \ge \tau.$$

Therefore

$$\lim_{|p|\to\infty} \frac{H(p)}{|p|} = \infty.$$

Finally, we prove that $L = H^*$. Let $q \in \mathbb{R}^n$ be fixed. By Definition 3.3 it holds that

$$\begin{aligned} H^*(q) &= \sup_{p\in\mathbb{R}^n} \{q\cdot p - H(p)\} \\ &= \sup_{p\in\mathbb{R}^n} \left\{ q\cdot p - \sup_{r\in\mathbb{R}^n} \{p\cdot r - L(r)\} \right\} \qquad (3.17) \\ &= \sup_{p\in\mathbb{R}^n} \inf_{r\in\mathbb{R}^n} \{p\cdot(q-r) + L(r)\}. \end{aligned}$$

Since L is convex, there is a $s \in \mathbb{R}^n$ such that

$$L(r) + s\cdot(q-r) \ge L(q), \quad \text{for every } r \in \mathbb{R}^n. \qquad (3.18)$$

By (3.18) and (3.17) we have

$$\begin{aligned} H^*(q) &= \sup_{p\in\mathbb{R}^n} \inf_{r\in\mathbb{R}^n} \{p\cdot(q-r)+L(r)\} \\ &\geq \inf_{r\in\mathbb{R}^n} \{s\cdot(q-r)+L(r)\} \\ &\geq L(q). \end{aligned} \tag{3.19}$$

On the other hand, for every $p \in \mathbb{R}^n$,

$$H(p) \geq p\cdot q - L(q).$$

Therefore it holds that

$$L(q) \geq \sup_{p\in\mathbb{R}^n} \{q\cdot p - H(p)\} = H^*(q). \tag{3.20}$$

By (3.19) and (3.20) we have $L(q) = H^*(q)$ for every $q \in \mathbb{R}^n$. □

3.4 Hopf-Lax Formula

Theorem 3.2 shows that for $L \in C^2$ the minimizer $\mathbf{w} \in \mathcal{K}$ of the action integral

$$I[\mathbf{w}] = \int_0^t L(\dot{\mathbf{w}}(s))\,ds$$

is a solution to Hamilton's equations, which form the system of characteristic ODEs for the Hamilton-Jacobi equations. This suggests that there is a direct connection between the Calculus of Variations and the Hamilton-Jacobi equations. To solve the following initial value problem for the Hamilton-Jacobi equations,

$$\begin{cases} u_t + H(D_x u) = 0 & \text{in } \mathbb{R}^n \times (0,\infty); \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}, \end{cases} \tag{3.21}$$

we modify the action integral to take into account the initial function g.

Let $\mathcal{K} := \{\mathbf{w} \in C^1([0,t];\mathbb{R}^n) \mid \mathbf{w}(t) = x, \mathbf{w}(0) = y\}$ and define the action integral for $\mathbf{w} \in \mathcal{K}$ as

$$I[\mathbf{w}(\cdot)] = \int_0^t L(\dot{\mathbf{w}}(s))\,ds + g(y).$$

We want to show that function $u\colon \mathbb{R}^n \times (0,\infty) \to \mathbb{R}$ defined by

$$u(x,t) = \inf \left\{ \int_0^t L(\dot{\mathbf{w}}(s))\,ds + g(y) \right\}, \tag{3.22}$$

where the infimum is taken over $\mathbf{w} \in \mathcal{K}$ and $y \in \mathbb{R}^n$, solves the initial value problem for the Hamilton-Jacobi equation (3.21) almost everywhere in $\mathbb{R}^n \times [0, \infty)$.

We assume hereafter that the Hamiltonian H is a convex C^2 function that satisfies the growth condition condition

$$\lim_{|p|\to\infty} \frac{H(p)}{|p|} = \infty.$$

By Theorem 3.4, the Lagrangian H is associated to is obtained as $L = H^*$, and the Lagrangian also satisfies these conditions.

A more simplified form for u satisfying (3.22) is offered by the *Hopf-Lax formula.* For this formulation we further assume, that the function g giving the initial condition, is Lipschitz continuous in $\mathbb{R}^n$.

Theorem 3.5. (Hopf-Lax Formula). *Let $g\colon \mathbb{R}^n \to \mathbb{R}$ be Lipschitz continuous. Assume that $L\colon \mathbb{R}^n \to \mathbb{R}$ is a continuous convex function satisfying*

$$\lim_{|q|\to\infty} \frac{L(q)}{|q|} = \infty. \tag{3.23}$$

Define $u\colon \mathbb{R}^n \times (0, \infty) \to \mathbb{R}$ as follows.

$$u(x,t) = \inf \left\{ \int_0^t L(\dot{\mathbf{w}}(s))\, ds + g(y) \right\}, \tag{3.24}$$

where the infimum is taken over all $w \in \mathcal{K}$, with

$$\mathcal{K} := \left\{ \mathbf{w} \in C^1([0,t]; \mathbb{R}^n) \mid \mathbf{w}(t) = x,\ \mathbf{w}(0) = y \right\},$$

and over all $y \in \mathbb{R}$. Then

$$u(x,t) = \min_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\}. \tag{3.25}$$

Equation (3.25) is called the Hopf-Lax formula.

Proof. Fix $x, y \in \mathbb{R}^n$ and $t > 0$. Let u be defined as above. Then for any $\mathbf{w} \in \mathcal{K}$ we have that

$$u(x,t) \leq \int_0^t L(\dot{\mathbf{w}}(s))\, ds + g(y). \tag{3.26}$$

Define $\mathbf{w}(s) \coloneqq y + \frac{s}{t}(x-y)$ for $s \in [0,t]$. Now $\mathbf{w} \in \mathcal{K}$ and

$$\begin{aligned} \int_0^t L(\dot{\mathbf{w}}(s))\, ds + g(y) &= \int_0^t L\left(\frac{x-y}{t}\right) ds + g(y) \\ &= tL\left(\frac{x-y}{t}\right) + g(y). \end{aligned} \tag{3.27}$$

By (3.26) and (3.27), we have for every $y \in \mathbb{R}^n$ that

$$u(x,t) \leq tL\left(\frac{x-y}{t}\right) + g(y).$$

Therefore it holds that

$$u(x,t) \leq \inf_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\}. \tag{3.28}$$

By Jensen's inequality, it holds for the convex function L that

$$L\left(\frac{1}{t}\int_0^t \dot{\mathbf{w}}(s)\, ds\right) \leq \frac{1}{t}\int_0^t L\left(\dot{\mathbf{w}}(s)\right)\, ds.$$

Thus we have for all $\mathbf{w} \in \mathcal{K}$, that

$$L\left(\frac{x-y}{t}\right) \leq \frac{1}{t}\int_0^t L\left(\dot{\mathbf{w}}(s)\right)\, ds$$

and further

$$tL\left(\frac{x-y}{t}\right) + g(y) \leq \int_0^t L\left(\dot{\mathbf{w}}(s)\right)\, ds + g(y).$$

By taking the infimum over $y \in \mathbb{R}^n$, we obtain

$$\inf_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \leq u(x,t). \tag{3.29}$$

Thus (3.28) and (3.29) imply that

$$u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\}.$$

To show that the infimum is reached by some $y \in \mathbb{R}^n$, we use the Lipschitz continuity of g. By definition, g is Lipschitz continuous, if there exists a $C > 0$ satisfying

$$|g(x) - g(y)| \leq C|x-y| \quad \text{for every } x, y \in \mathbb{R}^n. \tag{3.30}$$

We denote the smallest constant satisfying this condition by $\|g\|_{Lip}$.

Denote $f(y) := tL\left(\frac{x-y}{t}\right) + g(y)$ for $y \in \mathbb{R}^n$. By the Lipschitz continuity of g, it holds that

$$\begin{aligned} f(y) &\geq tL\left(\frac{x-y}{t}\right) - \|g\|_{Lip}|x-y| + g(x) \\ &= |x-y|\left(\frac{t}{|x-y|}L\left(\frac{x-y}{t}\right) - \|g\|_{Lip}\right) + g(x). \end{aligned}$$

Since $\frac{|x-y|}{t} \to \infty$ as $|y| \to \infty$, we have by (3.23) that

$$\lim_{|y|\to\infty} \frac{t}{|x-y|}L\left(\frac{x-y}{t}\right) = \infty.$$

Hence $f(y) \to \infty$ as $|y| \to \infty$.

Let $y^0 \in \mathbb{R}^n$. Now there exists a $r > 0$, such that

$$f(y) > f(y^0), \text{ if } |y - y^0| > r.$$

Note that f is continuous. By the extreme value theorem, there exists a $y^* \in \mathbb{R}^n$ satisfying

$$f(y^*) = \min\{f(y) \mid |y - y^0| \leq r\}.$$

Therefore,

$$f(y^*) = \inf_{y\in\mathbb{R}^n} f(y) = u(x,t).$$

□

Lemma 3.6. *Suppose that the assumptions of Theorem 3.5 hold. Let $x \in \mathbb{R}^n$ and $t > 0$ and define $u(x,t)$ by the Hopf-Lax formula. Then for all $s \in (0,t)$,*

$$u(x,t) = \min_{y\in\mathbb{R}^n}\left\{(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s)\right\}. \tag{3.31}$$

Proof. Fix $y \in \mathbb{R}^n$ and $s,t \in \mathbb{R}$ satisfying $0 < s < t$. As shown above, there exists a solution $z \in \mathbb{R}^n$ to

$$u(y,s) = sL\left(\frac{y-z}{s}\right) + g(z). \tag{3.32}$$

On the other hand, it holds that

$$u(x,t) \leq tL\left(\frac{x-z}{t}\right) + g(z). \tag{3.33}$$

Since

$$\frac{x-z}{t} = \frac{s}{t}\frac{y-z}{s} + \left(1-\frac{s}{t}\right)\frac{x-y}{t-s}$$

and L is convex, we have

$$L\left(\frac{x-z}{t}\right) \leq \frac{s}{t}L\left(\frac{y-z}{s}\right) + \left(1-\frac{s}{t}\right)L\left(\frac{x-y}{t-s}\right). \tag{3.34}$$

Combining (3.32), (3.33) and (3.34), we obtain that

$$\begin{aligned} u(x,t) &\leq tL\left(\frac{x-z}{t}\right) + g(z) \\ &\leq sL\left(\frac{y-z}{s}\right) + (t-s)L\left(\frac{x-y}{t-s}\right) + g(z) \\ &= (t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s). \end{aligned} \tag{3.35}$$

In Lemma 3.9 we show that u is Lipschitz continuous. By assumptions,

$$\lim_{|q|\to\infty} \frac{L(q)}{|q|} = \infty.$$

Therefore, by the extreme value theorem, there exists a $y \in \mathbb{R}^n$ minimizing

$$(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s).$$

It follows from (3.35), that for every $s \in (0,t)$,

$$u(x,t) \leq \min_{y\in\mathbb{R}^n}\left\{(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s)\right\}. \tag{3.36}$$

In the following, we prove that

$$u(x,t) \geq \min_{y\in\mathbb{R}^n}\left\{(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s)\right\}. \tag{3.37}$$

We choose $z \in \mathbb{R}^n$ such that

$$u(x,t) = tL\left(\frac{x-z}{t}\right) + g(z). \tag{3.38}$$

By definition of u, for any $y \in \mathbb{R}^n$ we have

$$u(y,s) \leq sL\left(\frac{y-z}{s}\right) + g(z). \tag{3.39}$$

Choose $y_0 \in \mathbb{R}^n$ as

$$y_0 = \frac{s}{t}x + (1 - \frac{s}{t})z.$$

Now y_0 satisfies

$$\frac{x - y_0}{t - s} = \frac{y_0 - z}{s} = \frac{x - z}{t}$$

and we have by (3.38) and (3.39) that

$$\begin{aligned} u(x,t) &= tL\left(\frac{x-z}{t}\right) + g(z) \\ &\geq tL\left(\frac{x-z}{t}\right) - sL\left(\frac{x-y_0}{t-s}\right) + u(y_0, s) \\ &= (t-s)L\left(\frac{x-y_0}{t-s}\right) + u(y_0, s). \end{aligned}$$

The above inequality holds for $y_0 \in \mathbb{R}^n$. This proves (3.37). The claimed identity (3.31) follows from (3.36) and (3.37). This finishes the proof. □

Theorem 3.7. *Suppose that the assumptions of Theorem 3.5 hold. Suppose further that the function u given by the Hopf-Lax formula* (3.25) *is differentiable at point $(x, t) \in \mathbb{R}^n \times (0, \infty)$. Then*

$$u_t(x,t) + H(D_x u(x,t)) = 0, \tag{3.40}$$

where $H = L^*$ is the Legendre transform of L.

Proof. Let $q \in \mathbb{R}^n$ and $h > 0$. By Lemma 3.6,

$$\begin{aligned} u(x + hq, t + h) &= \min_{y \in \mathbb{R}^n} \left\{ hL\left(\frac{x + hq - y}{h}\right) + u(y, t) \right\} \\ &\leq hL(q) + u(x, t). \end{aligned}$$

This gives the following estimate:

$$u_t(x,t) + q \cdot D_x u(x,t) = \lim_{h \to 0^+} \frac{u(x + qh, t + h) - u(x,t)}{h} \leq L(q). \tag{3.41}$$

By definition, $H = L^*$, therefore (3.41) implies

$$u_t(x,t) + H(D_x u(x,t)) = u_t(x,t) + \max_{q \in \mathbb{R}^n} \{q \cdot D_x u(x,t) - L(q)\} \leq 0. \tag{3.42}$$

By Theorem 3.5, we can choose $z \in \mathbb{R}^n$ for which

$$u(x,t) = \min_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} = tL\left(\frac{x-z}{t}\right) + g(z).$$

Thus it holds for any $y \in \mathbb{R}^n$ and $s > 0$ that

$$\begin{aligned} u(x,t) - u(y,s) &= tL\left(\frac{x-z}{t}\right) + g(z) - u(y,s) \\ &\geq tL\left(\frac{x-z}{t}\right) + g(z) - sL\left(\frac{y-z}{s}\right) - g(z) \\ &= tL\left(\frac{x-z}{t}\right) - sL\left(\frac{y-z}{s}\right). \end{aligned} \tag{3.43}$$

Choose $y_0 \in \mathbb{R}^n$ as

$$y_0 = \frac{s}{t}(x-z) + z.$$

Now

$$\frac{x-z}{t} = \frac{y_0 - z}{s},$$

and we obtain from (3.43) that

$$u(x,t) - u\left(\frac{s}{t}(x-z) + z, s\right) \geq (t-s)L\left(\frac{x-z}{t}\right).$$

By further setting $h := t - s$, we arrive at the following inequality:

$$\frac{u(x,t) - u(x + \frac{h}{t}(z-x), t-h)}{h} \geq L\left(\frac{x-z}{t}\right).$$

By letting $h \to 0^+$, we have:

$$u_t(x,t) + \frac{x-z}{t} \cdot D_x u(x,t) \geq L\left(\frac{x-z}{t}\right). \tag{3.44}$$

As $H = L^*$, (3.44) gives us

$$\begin{aligned} u_t(x,t) + H(D_x u(x,t)) =& u_t(x,t) + \max_{q \in \mathbb{R}^n} \{q \cdot D_x u(x,t) - L(q)\} \\ \geq& u_t(x,t) + \frac{x-z}{t} \cdot D_x u(x,t) - L\left(\frac{x-z}{t}\right) \\ \geq& 0. \end{aligned} \tag{3.45}$$

By (3.42) and (3.45), we end up with (3.40). This finishes the proof. □

Theorem 3.7 shows that the function u given by the Hopf-Lax formula (3.25) satisfies the Hamilton-Jacobi equations (3.40) at every point $(x,t) \in \mathbb{R}^n \times (0,\infty)$ where u is differentiable. Next we show that u is Lipschitz continuous in $\mathbb{R}^n \times (0,\infty)$. Thus, by the Rademacher's Theorem, u is differentiable at almost every point $(x,t) \in \mathbb{R}^n \times (0,\infty)$.

Theorem 3.8. (Rademacher's Theorem). *Let $f\colon \Omega \to \mathbb{R}^m$ be Lipschitz continuous in the domain $\Omega \subset \mathbb{R}^n$. Then f is differentiable at almost every point in Ω.*

For the proof of Theorem 3.8, we refer to section 5.8. in [3] and Theorem 3.1 in [4].

Lemma 3.9. *The function $u\colon \mathbb{R}^n \times (0,\infty) \to \mathbb{R}$, defined by the Hopf-Lax formula* (3.25)*, is Lipschitz continuous.*

Proof. Let $t > 0$ and $x \in \mathbb{R}^n$. By Theorem 3.5, there exists a $y \in \mathbb{R}^n$ that satisfies

$$u(x,t) = tL\left(\frac{x-y}{t}\right) + g(y).$$

Then, for every $\tilde{x} \in \mathbb{R}^n$

$$\begin{aligned} u(\tilde{x},t) &= \min_{z\in\mathbb{R}^n}\left\{tL\left(\frac{\tilde{x}-z}{t}\right) + g(z)\right\} \\ &\le tL\left(\frac{x-y}{t}\right) + g(\tilde{x}-x+y) \\ &= u(x,t) - g(y) + g(\tilde{x}-x+y) \\ &\le u(x,t) + \|g\|_{Lip}|\tilde{x}-x|. \end{aligned}$$

Therefore

$$u(\tilde{x},t) - u(x,t) \le \|g\|_{Lip}|\tilde{x}-x|.$$

Similarly, we can prove that

$$u(x,t) - u(\tilde{x},t) \le \|g\|_{Lip}|x-\tilde{x}|.$$

Therefore

$$|u(\tilde{x},t) - u(x,t)| \le \|g\|_{Lip}|\tilde{x}-x| \quad \text{for } x,\tilde{x} \in \mathbb{R}^n. \tag{3.46}$$

Therefore $u(\cdot,t)$ is Lipschitz continuous for every fixed $t > 0$. Next we fix $x \in \mathbb{R}^n$ and $0 < s < t$. By Lemma 3.6, we have

$$\begin{aligned} u(x,t) &= \min_{y\in\mathbb{R}^n}\left\{(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s)\right\} \\ &\le (t-s)L\left(0\right) + u(x,s) \\ &= L(0)|t-s| + u(x,s). \end{aligned} \tag{3.47}$$

Denote $R = ||u(\cdot, s)||_{Lip}$. The Lipschitz continuity of $u(\cdot, s)$ gives us the following estimate

$$
\begin{aligned}
u(x,t) &= \min_{y\in\mathbb{R}^n} \left\{ (t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \right\} \\
&\geq \min_{y\in\mathbb{R}^n} \left\{ (t-s)L\left(\frac{x-y}{t-s}\right) - R|x-y| + u(x,s) \right\} \\
&= u(x,s) + (t-s)\min_{y\in\mathbb{R}^n} \left\{ L\left(\frac{x-y}{t-s}\right) - R\frac{|x-y|}{t-s} \right\}.
\end{aligned}
$$

For the second term on the right hand side, we have

$$
\begin{aligned}
\min_{y\in\mathbb{R}^n} \left\{ L\left(\frac{x-y}{t}\right) - R\frac{|x-y|}{t} \right\} &= \max_{z\in\mathbb{R}^n} \left\{ R|z| - L(z) \right\} \\
&= \max_{y\in B(0,R)} \max_{z\in\mathbb{R}^n} \{ y\cdot z - L(z) \} \\
&= \max_{y\in B(0,R)} H(y).
\end{aligned}
$$

Therefore we have that

$$
u(x,t) \geq u(x,s) + (t-s)\max_{y\in B(0,R)} H(y). \tag{3.48}
$$

By inequalities (3.48) and (3.47) we have

$$
|u(x,t) - u(x,s)| \leq \max\left\{ |L(0)|, \max_{y\in B(0,R)} |H(y)| \right\} |t-s|
$$

and therefore $u(x,\cdot)$ is Lipschitz continuous. Because $u(x,t)$ is Lipschitz continuous with respect to x and t, it is Lipschitz continuous in $\mathbb{R}^n \times (0,\infty)$. □

As mentioned above, Lipschitz continuity of u gives differentiability almost everywhere in $\mathbb{R}^n \times (0,\infty)$. Therefore, by Theorem 3.7, u solves the Hamilton-Jacobi equation almost everywhere. What remains to be shown, is that u can be extended to satisfy the initial value condition $u = g$ on $\mathbb{R}^n \times \{t = 0\}$. This is easily achieved through similar calculations as above in the proof of Lemma 3.9.

Lemma 3.10. *Suppose $g\colon \mathbb{R}^n \to \mathbb{R}$ is Lipchitz continuous. Then the function u defined by the Hopf-Lax formula* (3.25) *satisfies*

$$
\lim_{t\to 0^+} u(x,t) = g(x) \quad \textit{for } (x,t) \in \mathbb{R}^n \times (0,\infty).
$$

Proof. By definition of u we have that

$$
\begin{aligned}
u(x,t) =& \min_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \\
\leq& tL(0) + g(x).
\end{aligned}
$$

On the other hand we have

$$
\begin{aligned}
u(x,t) =& \min_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \\
\geq& \min_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(x) - \|g\|_{Lip}|x-y| \right\} \\
=& g(x) - t\max_{z\in\mathbb{R}^n} \{\|g\|_{Lip}|z| - L(z)\} \\
=& g(x) - t \max_{y\in B(0,\|g\|_{Lip})} H(y).
\end{aligned}
$$

These two estimates together give us that

$$
|u(x,t) - g(x)| \leq t\max\left\{ |L(0)|, \max_{y\in B(0,\|g\|_{Lip})} H(y) \right\}.
$$

Thus, $\lim_{t\to 0^+} u(x,t) = g(x)$ for every $x\in\mathbb{R}^n$. □

To summarize, we proved the following theorem.

Theorem 3.11. *Let $g\colon \mathbb{R}^n \to \mathbb{R}$ be Lipschitz continuous and suppose that $H\colon \mathbb{R}^n \to \mathbb{R}$ is a convex C^2 function satisfying the growth condition* (3.16)*. For $(x,t) \in \mathbb{R}^n \times (0,\infty)$, let $u\colon \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$, be defined by the Hopf-Lax formula* (3.25)*, where $L = H^*$. For $(x,t) \in \mathbb{R}^n \times \{t=0\}$, define u as $u(x,0) = \lim_{t\to 0^+} u(x,t)$. Then u is Lipschitz continuous and satisfies*

$$
\begin{cases}
u_t + H(D_x u) = 0 & a.e. \text{ in } \mathbb{R}^n \times (0,\infty) \\
u = g & \text{on } \mathbb{R}^n \times \{t=0\}.
\end{cases}
$$

3.5 Weak Solutions

In this section we investigate the properties of the function given by the Hopf-Lax formula. We then use these properties to define weak solutions to initial value problems for the Hamilton-Jacobi equations.

The next example demonstrates that there does not, in general, exist a unique Lipschitz continuous integral solution to the initial value problems given for the Hamilton-Jacobi equations.

Example 3.12. Let $H(p) = p^2$ for $p \in \mathbb{R}$ and consider the following initial value problem:

$$\begin{cases} u_t + H(u_x) = 0 & \text{in } \mathbb{R}\times(0,\infty); \\ u = 0 & \text{on } \mathbb{R}\times\{t=0\}. \end{cases} \tag{3.49}$$

Now the initial value is given by a Lipschitz continuous function and the Hamiltonian is clearly a convex function satisfying the growth condition. This initial value problem has the trivial solution

$$u_0(x,t) \equiv 0.$$

On the other hand, function u_1 defined as

$$u_1(x,t) = \begin{cases} 0 & \text{for } |x| \geq t; \\ |x| - t & \text{for } |x| < t, \end{cases}$$

is Lipschitsz continuous and differentiable almost everywhere in $\mathbb{R}\times(0,\infty)$. In addition $u_1(x,0) = 0$ and u_1 is a solution to (3.49) for almost every $(x,t) \in \mathbb{R}\times(0,\infty)$.

Definition 3.13. Let $\Omega \subset \mathbb{R}^n$. Function $H\colon \Omega \to \mathbb{R}$ is *strongly convex*, if there exists a constant $C > 0$ satisfying

$$H(tx + (1-t)y) \leq tH(x) + (1-t)H(y) - \frac{1}{2}t(1-t)C|x-y|^2,$$

for all $x, y \in \mathbb{R}^n$ and $t \in [0,1]$.

Remark 3.14. *Function $H \in C^2(\Omega)$ is strongly convex, if and only if there exists a constant $C > 0$ such that*

$$\xi \cdot D^2H(x) \cdot \xi^T \geq C|\xi|^2 \quad \textit{for every } x \in \Omega \textit{ and } \xi \in \mathbb{R}^n.$$

This holds, if the eigenvalues of the Hessian matrix $D^2H(x)$ are limited from below by C for every $x \in \Omega$.

For matrices $A, B \in \mathbb{R}^{n\times n}$, we write $A \geq B$, if it holds for every $\xi \in \mathbb{R}^n$ that

$$\xi \cdot A \cdot \xi^T \geq \xi \cdot B \cdot \xi^T. \tag{3.50}$$

If $A \in \mathbb{R}^{n\times n}$ satisfies $A \geq 0$, it is said to be positive semidefinite.

Lemma 3.15. *Suppose $H \in C^2(\mathbb{R}^n)$ is strongly convex. Then H satisfies the growth condition*

$$\lim_{|x|\to\infty} \frac{H(x)}{|x|} = \infty.$$

Proof. Let $x \in \mathbb{R}^n$. By choosing $y = 0$ in Definition 3.13, we have for any $t \in [0, 1]$ that

$$H(tx) \leq tH(x) + (1 - t)H(0) - \frac{C}{2}t(1 - t)|x|^2,$$

By dividing both sides of the inequality above with t and then letting $t \to \infty$, we have

$$DH(0) \cdot x \leq H(x) - H(0) - \frac{C}{2}|x|^2.$$

Therefore,

$$\frac{H(x)}{|x|} \geq |x| - |DH(0)| + \frac{H(0)}{|x|}.$$

By taking the limit $|x| \to \infty$, we have

$$\lim_{|x|\to\infty} \frac{H(x)}{|x|} \geq \lim_{|x|\to\infty} |x| - |DH(0)| + \frac{H(0)}{|x|} = \infty,$$

which finishes the proof. □

Definition 3.16. Let $\Omega \subset \mathbb{R}^n$. Function $u \colon \mathbb{R}^n \to \mathbb{R}$ is called *semiconcave*, if there exists a constant $C > 0$ such that

$$u(x + z) - 2u(x) + u(x - z) \leq C|z|^2$$

for every $x, z \in \mathbb{R}^n$.

Remark 3.17. *Function $u \colon \Omega \to \mathbb{R}$ is semiconcave, if and only if the function*

$$x \mapsto u(x) - \frac{C}{2}|x|^2$$

is concave for some $C > 0$.

Lemma 3.18. *Suppose $H \colon \mathbb{R}^n \to \mathbb{R}$ is strongly convex with constant $C > 0$ and define u by the Hopf-Lax formula* (3.25)*. Then u satisfies*

$$u(x + z, t) - 2u(x, t) + u(x - z, t) \leq \frac{1}{Ct}|z|^2$$

for all $z, x \in \mathbb{R}^n$ and $t > 0$.

Proof. We begin by proving the following estimate for the Lagrangian associated with H.

$$\frac{1}{2}L(q_1) + \frac{1}{2}L(q_2) \leq L\left(\frac{q_1 + q_2}{2}\right) + \frac{1}{8C}|q_1 - q_2|^2, \tag{3.51}$$

for all $q_1, q_2 \in \mathbb{R}^n$.

Indeed, by Theorem 3.4, $L = H^*$. Let $q_1, p_1 \in \mathbb{R}^n$ and $q_2, p_2 \in \mathbb{R}^n$ be two pairs satisfying (3.15), that is,

$$L(q) = p \cdot q - H(p). \tag{3.52}$$

Now we have

$$\frac{1}{2}L(q_1) + \frac{1}{2}L(q_2) = \frac{1}{2}(p_1 \cdot g_1 + p_2 \cdot q_2) - \frac{1}{2}(H(p_1) + H(p_2)). \tag{3.53}$$

Since H is strongly convex with constant $C > 0$, we have

$$\frac{1}{2}H(p_1) + \frac{1}{2}H(p_2) \geq H\left(\frac{p_1 + p_2}{2}\right) + \frac{1}{8}C|p_1 - p_2|^2. \tag{3.54}$$

By Theorem 3.4, $H(p) = L^*(p)$, where L^* is the Legendre transform of L given in Definition 3.3. Therefore

$$H\left(\frac{p_1 + p_2}{2}\right) \geq \frac{p_1 + p_2}{2} \cdot \frac{q_1 + q_2}{2} - L\left(\frac{q_1 + q_2}{2}\right). \tag{3.55}$$

Now (3.53), (3.54) and (3.55) give the estimate for the Lagrangian

$$\begin{aligned}\frac{1}{2}L(q_1) + \frac{1}{2}L(q_2) \leq & L\left(\frac{q_1 + q_2}{2}\right) + \frac{1}{2}(p_1 \cdot g_1 + p_2 \cdot q_2) \\ & - \frac{1}{4}(p_1 + p_2) \cdot (q_1 + q_2) - \frac{C}{8}|p_1 - p_2|^2.\end{aligned} \tag{3.56}$$

For the last three terms on the right hand side of (3.56) we have the following estimate.

$$\begin{aligned}&\frac{1}{2}(p_1 \cdot g_1 + p_2 \cdot q_2) - \frac{1}{4}(p_1 + p_2) \cdot (q_1 + q_2) - \frac{C}{8}|p_1 - p_2|^2 \\ &= -\frac{1}{8}\left(C|p_1 - p_2|^2 - 2(p_1 - p_2) \cdot (q_1 - q_2)\right) \\ &\leq -\frac{1}{8}\left(C|p_1 - p_2|^2 - 2|p_1 - p_2||q_1 - q_2|\right) \\ &= -\frac{1}{8}\left(\sqrt{C}|p_1 - p_2| - \frac{1}{\sqrt{C}}|q_1 - q_2|\right)^2 + \frac{1}{8C}|q_1 - q_2|^2 \\ &\leq \frac{1}{8C}|q_1 - q_2|^2.\end{aligned}$$

It follows from (3.56) and the above estimate that for every $q_1, q_2 \in \mathbb{R}^n$

$$\frac{1}{2}L(q_1) + \frac{1}{2}L(q_2) \leq L\left(\frac{q_1 + q_2}{2}\right) + \frac{1}{8C}|q_1 - q_2|^2. \tag{3.57}$$

Let $u\colon \mathbb{R}^n \to \mathbb{R}$ be defined by the Hopf-Lax formula. Let $x, z \in \mathbb{R}^n$, $t > 0$. By Theorem 3.5, there exists a $y \in \mathbb{R}^n$ that satisfies

$$u(x,t) = tL\left(\frac{x-y}{t}\right) + g(y).$$

By (3.57), we have

$$\begin{aligned}
&u(x+z,t) - 2u(x,t) + u(x-z,t) \\
\leq &\left[tL\left(\frac{x+z-y}{t}\right) + g(y)\right] - 2\left[tL\left(\frac{x-y}{t}\right) + g(y)\right] \\
&+ \left[tL\left(\frac{x-z-y}{t}\right) + g(y)\right] \\
=&2t\left[\frac{1}{2}L\left(\frac{x-y+z}{t}\right) + \frac{1}{2}L\left(\frac{x-y-z}{t}\right) - L\left(\frac{x-y}{t}\right)\right] \\
\leq&2t\left[\frac{1}{8C}\left|\frac{2z}{t}\right|^2\right] \leq \frac{1}{Ct}|z|^2.
\end{aligned}$$

□

Lemma 3.19. *Suppose the initial function $g\colon \mathbb{R}^n \to \mathbb{R}$ is semiconcave. Then the function u defined by the Hopf-Lax formula* (3.25) *is semiconcave in the x variable.*

Proof. Let $x \in \mathbb{R}^n$ and $t > 0$. By Theorem 3.5 there exists a $y \in \mathbb{R}^n$ such that

$$u(x,t) = tL\left(\frac{x-y}{t}\right) + g(y).$$

By the Hopf-Lax formula and the semiconcavity of g, we have for each $z \in \mathbb{R}^n$ that

$$\begin{aligned}
&u(x+z,t) - 2u(x,t) + u(x-z,t) \\
\leq &\left[tL\left(\frac{x-y}{t}\right) + g(y+z)\right] - 2\left[tL\left(\frac{x-y}{t}\right) + g(y)\right] \\
&+ \left[tL\left(\frac{x-y}{t}\right) + g(y-z)\right] \\
=&g(x+z) - 2g(x) + g(x-z) \leq C|z|^2.
\end{aligned}$$

Thus, if the initial function g is semiconcave, u defined by the Hopf-Lax formula is also semiconcave. □

Next we define weak solutions to initial value problems for the Hamilton-Jacobi equations and show the uniqueness of these weak solutions.

Definition 3.20. Function $u\colon \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is a weak solution to the initial value problem

$$\begin{cases} u_t + H(D_x u) = 0 & \text{in } \mathbb{R}^n \times (0,\infty); \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases}$$

provided it satisfies the following conditions

i) u is Lipschitz continuous in $\mathbb{R}^n \times [0,\infty)$;

ii) $u_t + H(D_x u) = 0$ for a.e. $(x,t) \in \mathbb{R}^n \times (0,\infty)$;

iii) $u(x,0) = g(x)$ for all $x \in \mathbb{R}^n$;

iv) There exists a $C > 0$ such that

$$u(x+z,t) - 2u(x,t) + u(x-z,t) \le C(1 + \frac{1}{t})|z|^2,$$

for all $x, z \in \mathbb{R}^n$ and $t > 0$.

3.6 Mollifiers

In this section, we introduce *mollifiers* and some properties related to mollifications of functions. These properties are used to prove the uniqueness of weak solutions to the initial value problems for the Hamilton-Jacobi equations.

Definition 3.21. Let $\eta \in C^\infty(\mathbb{R}^n)$ be defined as

$$\eta(x) := \begin{cases} C \exp\left(\frac{1}{|x|^2 - 1}\right) & \text{for } |x| < 1; \\ 0 & \text{for } |x| \ge 1, \end{cases}$$

where $C > 0$ is a constant such that $\int_{\mathbb{R}^n} \eta \, dx = 1$. We call η the *standard mollifier*. For each $\varepsilon > 0$, we define η^ε by

$$\eta^\varepsilon(x) := \frac{1}{\varepsilon^n} \eta\left(\frac{x}{\varepsilon}\right).$$

Note that for each $\varepsilon > 0$, $\eta^\varepsilon \in C^\infty(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} \eta^\varepsilon \, dx = 1$. The support of η^ε lies in $\overline{B(0,\varepsilon)}$, and therefore η^ε is a smooth function in $\mathbb{R}^n$ with a compact support.

Mollifiers are extremely useful in approximating locally integrable functions with smooth functions. This approximation is called a *mollification*.

Definition 3.22. Let $u\colon \Omega \to \mathbb{R}$ be locally integrable in the domain $\Omega \subset \mathbb{R}^n$. For $x \in \Omega_\varepsilon := \{x \in \Omega \mid \operatorname{dist}(x, \mathbb{R}^n \setminus \Omega) > \varepsilon\}$, we define the *mollification* of u by

$$u^\varepsilon(x) := (\eta^\varepsilon * u)(x) = \int_\Omega \eta^\varepsilon(x-y)u(y)\,dy = \int_{B(0,\varepsilon)} \eta^\varepsilon(y)u(x-y)\,dy.$$

Lemma 3.23. *Let Ω be a domain in $\mathbb{R}^n$ and suppose u is locally integrable in Ω. Then $u^\varepsilon \in C^\infty$ for all $\varepsilon > 0$.*

Proof. Fix $x \in \Omega_\varepsilon$ and $h > 0$ such that $x + he_k \in \Omega_\varepsilon$. It holds that

$$\begin{aligned}
&\left|\frac{u^\varepsilon(x+he_k) - u^\varepsilon(x)}{h} - (\eta^\varepsilon_{x_k} * u)(x)\right| \\
=&\left|\int_\Omega \left(\frac{\eta^\varepsilon(x+he_k-y) - \eta^\varepsilon(x-y)}{h} - \eta^\varepsilon_{x_k}(x-y)\right) u(y)\,dy\right| \\
=&\frac{1}{\varepsilon^n}\left|\int_\Omega \left(\frac{1}{h}\left[\eta\left(\frac{x+he_k-y}{\varepsilon}\right) - \eta\left(\frac{x-y}{\varepsilon}\right)\right] - \frac{1}{\varepsilon}\eta_{x_k}(x-y)\right) u(y)\,dy\right|.
\end{aligned}$$

Since η^ε is smooth and has a compact support in $\mathbb{R}^n$,

$$\frac{1}{h}\left[\eta\left(\frac{x+he_k-y}{\varepsilon}\right) - \eta\left(\frac{x-y}{\varepsilon}\right)\right] \to \frac{1}{\varepsilon}\eta_{x_k}\left(\frac{x-y}{\varepsilon}\right)$$

uniformly for $y \in \Omega$, as $h \to 0$. Therefore we have

$$\begin{aligned}
&\left|\frac{u^\varepsilon(x+he_k) - u^\varepsilon(x)}{h} - (\eta^\varepsilon_{x_k} * u)(x)\right| \\
\leq &\frac{1}{\varepsilon^n}\int_\Omega \left|\frac{1}{h}\left[\eta\left(\frac{x+he_k-y}{\varepsilon}\right) - \eta\left(\frac{x-y}{\varepsilon}\right)\right] - \frac{1}{\varepsilon}\eta_{x_k}\left(\frac{x-y}{\varepsilon}\right)\right| |u(y)|\,dy \\
\to &\, 0, \quad \text{as } h \to 0.
\end{aligned}$$

Thus u^ε is differentiable, and we have for every $k \in \{1, 2, \cdots, n\}$ that

$$\frac{\partial}{\partial x_k}u^\varepsilon(x) = (\eta^\varepsilon_{x_k} * u)(x).$$

As the convolution of u with any continuous function is also continuous, the continuity of $\frac{\partial}{\partial x_k}u^\varepsilon$ follows. The derivatives of higher degree can similarly be transferred to the smooth mollifier and thus the smoothness of u^ε follows. □

Lemma 3.24. *Let Ω be a domain in $\mathbb{R}^n$ and suppose $u\colon \Omega \to \mathbb{R}$ is locally integrable. Then $u^\varepsilon \to u$ as $\varepsilon \to 0$ almost everywhere in Ω.*

Proof. Fix $x \in \Omega_\varepsilon$. As $\int_{\mathbb{R}^n} \eta^\varepsilon(x)\,dx = 1$ for every $\varepsilon > 0$, we have

$$\begin{aligned}|u(x) - u^\varepsilon(x)| &= \left| u(x) - \int_{B(0,\varepsilon)} u(x-y)\eta^\varepsilon(y)\,dy \right| \\ &= \left| \int_{B(0,\varepsilon)} (u(x) - u(x-y))\eta^\varepsilon(y)\,dy \right|.\end{aligned}$$

Now

$$\begin{aligned}\left| \int_{B(0,\varepsilon)} (u(x) - u(x-y))\eta^\varepsilon(y)\,dy \right| &\le \int\limits_{B(0,\varepsilon)} |u(x) - u(x-y)|\eta^\varepsilon(y)\,dy \\ &= \int\limits_{B(x,\varepsilon)} |u(x) - u(y)|\eta^\varepsilon(x-y)\,dy \\ &= \frac{C}{\varepsilon^n} \int\limits_{B(x,\varepsilon)} |u(x) - u(y)|\eta\left(\frac{x-y}{\varepsilon}\right)\,dy.\end{aligned}$$

Since $\eta\left(\frac{x-y}{\varepsilon}\right) \le 1$, we have

$$|u(x) - u^\varepsilon(x)| \le C \fint\limits_{B(x,\varepsilon)} |u(x) - u(y)|\,dy.$$

Above $\fint$ is the integral average over the ball $B(x,\varepsilon)$. By assumptions, u is locally integrable in Ω. Therefore, by Lebesgue's differentiation theorem

$$\lim_{\varepsilon \to 0} \fint\limits_{B(x,\varepsilon)} |u(x) - u(y)|\,dy \to 0 \quad \text{for a.e. } x \in \Omega.$$

Therefore, $u^\varepsilon \to u$ as $\varepsilon \to 0$ almost everywhere in Ω. □

Lemma 3.25. *Let $\Omega \subset \mathbb{R}^n$ be an open domain and suppose $u\colon \Omega \to \mathbb{R}$ is Lipschitz continuous. Then $|Du^\varepsilon| \le \|u\|_{Lip}$.*

Proof. Fix $x, y \in \mathbb{R}^n$ and consider the distance of $u^\varepsilon(x)$ form $u^\varepsilon(y)$:

$$\begin{aligned}|u^\varepsilon(x) - u^\varepsilon(y)| &= \left| \int_{B(0,\varepsilon)} u(x-z)\eta^\varepsilon(z)\,dz - \int_{B(0,\varepsilon)} u(y-z)\eta^\varepsilon(z)\,dz \right| \\ &= \left| \int_{B(0,\varepsilon)} [u(x-z) - u(y-z)]\eta^\varepsilon(z)\,dz \right| \\ &\le \int\limits_{B(0,\varepsilon)} \|u\|_{Lip}|x-y|\eta^\varepsilon(z)\,dz = \|u\|_{Lip}|x-y|.\end{aligned}$$

Hence $|Du^\varepsilon| \le \|u\|_{Lip}$. $\square$

Lemma 3.26. *Let $u\colon \mathbb{R}^n \to \mathbb{R}$ be locally integrable. Suppose there exists a constant $C > 0$ such that*

$$u(x+z) - 2u(x) + u(x-z) \le C|z|^2 \tag{3.58}$$

for all $x, z \in \mathbb{R}^n$. Then

$$D^2u^\varepsilon(x) \le CI_{n\times n}$$

for every $\varepsilon > 0$, and $x \in \mathbb{R}^n$.

Proof. Define $h\colon \mathbb{R}^n \to \mathbb{R}$ as

$$h(x) \coloneqq u(x) - \frac{C}{2}|x|^2.$$

Now it holds for every $x, z \in \mathbb{R}^n$ that

$$h(x+z) + h(x-z) = u(x+z) + u(x-z) - \frac{C}{2}\left(|x-z|^2 + |x+z|^2\right).$$

By assumption (3.58), we have

$$\begin{aligned} h(x+z) + h(x-z) &\le 2u(x) - \frac{C}{2}\left(|x+z|^2 + |x-z|^2 - 2|z|^2\right) \\ &= 2\left(u(x) - \frac{C}{2}|x|^2\right) = 2h(x). \end{aligned}$$

Therefore h is concave. Let $\varepsilon > 0$ and h^ε be the mollification of h. We have

$$\begin{aligned} h^\varepsilon(x+z) + h^\varepsilon(x-z) &= \int_{B(0,\varepsilon)} \left(h(x+z-y) + h(x-z-y)\right)\eta^\varepsilon(y)\,dy \\ &\le \int_{B(0,\varepsilon)} \left(2h(x-y)\right)\eta^\varepsilon(y)\,dy \\ &= 2h^\varepsilon(x). \end{aligned}$$

Hence h^ε is concave. By Lemma 3.23, h^ε is smooth. Therefore the Hessian matrix $D^2h^\varepsilon(x)$ is negative semidefinite for each $x \in \mathbb{R}^n$, that is, $D^2h^\varepsilon(x) \le 0$.

We write $h^\varepsilon(x)$ as

$$\begin{aligned}
h^\varepsilon(x) &= \int\limits_{B(0,\varepsilon)} h(x-y)\eta^\varepsilon(y)\,dy \\
&= \int\limits_{B(0,\varepsilon)} \left(u(x-y) - \frac{C}{2}|x-y|^2\right)\eta^\varepsilon(y)\,dy \\
&= u^\varepsilon(x) - \int\limits_{B(0,\varepsilon)} \frac{C}{2}|x-y|^2\eta^\varepsilon(y)\,dy \\
&= u^\varepsilon(x) - C\int\limits_{B(0,\varepsilon)} \frac{1}{2}|x-y|^2\eta^\varepsilon(y)\,dy.
\end{aligned} \tag{3.59}$$

Thus

$$D^2h^\varepsilon(x) = D^2u^\varepsilon(x) - CI_{n\times n}.$$

As $D^2h^\varepsilon(x)$ is negative semidefinite for every $x \in \mathbb{R}^n$, we have

$$D^2u^\varepsilon \leq CI_{n\times n}.$$

This finishes the proof. □

3.7 Uniqueness of Weak Solutions

In this section, we prove the uniqueness of weak solutions for Hamilton-Jacobi equations. We also show that the Hopf-Lax formula gives the weak solution to initial value problems for the Hamilton-Jacobi equations, when the initial value function g is Lipschitz continuous, $H \in C^2(\mathbb{R}^n)$ is convex and additionally, either g is semiconcave and H satisfies the growth condition, or H is strongly convex.

Theorem 3.27. *Suppose $H \in C^2(\mathbb{R}^n)$ is convex and $g\colon \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous. Then the initial value problem*

$$\begin{cases} u_t + H(D_xu) = 0 & \text{in } \mathbb{R}^n \times (0,\infty); \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\} \end{cases} \tag{3.60}$$

has at most one weak solution in $\mathbb{R}^n \times [0,\infty)$.

Proof. Suppose that u and $\tilde{u}$ are weak solutions to the initial value problem (3.60) and define $w := u - \tilde{u}$. We show that $w \equiv 0$ in $\mathbb{R}^n \times [0, \infty)$.

Fix $(x, t) \in \mathbb{R}^n \times (0, \infty)$ and suppose u and $\tilde{u}$ are differentiable at (x, t). Then they satisfy the PDE in (3.60) at (x, t) and therefore,

$$w_t(x,t) = -H(D_x u(x,t)) + H(D_x \tilde{u}(x,t)).$$

By using the fundamental theory of calculus, we have

$$\begin{aligned} w_t(x,t) &= -\int_0^1 \frac{d}{dr} H(rD_x u(x,t) + (1-r)D_x\tilde{u}(x,t))\,dr \\ &= -\int_0^1 DH(rD_x u(x,t) + (1-r)D_x\tilde{u}(x,t))\,dr \cdot (D_x u(x,t) - D_x\tilde{u}(x,t)) \\ &=: -\mathbf{b}(x,t) \cdot D_x w(x,t). \end{aligned} \tag{3.61}$$

Let $\phi\colon \mathbb{R} \to [0, \infty)$ be a smooth function, that we will choose later, and define $v\colon \mathbb{R}^n \times (0,\infty) \to \mathbb{R}$ as $v(x,t) := \phi(w(x,t))$. By equation (3.61),

$$\phi'(w(x,t))w_t(x,t) + \mathbf{b}(x,t) \cdot \phi'(w(x,t))D_x w(x,t) = 0.$$

This implies that for almost every $(x,t) \in \mathbb{R}^n \times (0,\infty)$

$$v_t(x,t) + \mathbf{b}(x,t) \cdot D_x v(x,t) = 0. \tag{3.62}$$

Fix $\varepsilon > 0$ and let u^ε and $\tilde{u}^\varepsilon$ be the mollifications of u and $\tilde{u}$ with respect to the x variable, that is

$$u^\varepsilon(x,t) = \int_{\mathbb{R}^n} \eta^\varepsilon(x-y)u(y,t)\,dy$$

and

$$\tilde{u}^\varepsilon(x,t) = \int_{\mathbb{R}^n} \eta^\varepsilon(x-y)\tilde{u}(y,t)\,dy.$$

Define $\mathbf{b}^\varepsilon\colon \mathbb{R}^n \times (0,\infty) \to \mathbb{R}^n$ as

$$\mathbf{b}^\varepsilon(x,t) = \int_0^1 DH(rD_x u^\varepsilon(x,t) + (1-r)D_x\tilde{u}^\varepsilon(x,t))\,dr.$$

By equation (3.62),

$$v_t + \mathbf{b}^\varepsilon \cdot D_x v = (\mathbf{b}^\varepsilon - \mathbf{b}) \cdot D_x v.$$

Since
$$\operatorname{div}_x(\mathbf{b}^\varepsilon v) = \mathbf{b}^\varepsilon \cdot D_x v + \operatorname{div}_x(\mathbf{b}^\varepsilon) v,$$
we have
$$v_t + \operatorname{div}_x(\mathbf{b}^\varepsilon v) = \operatorname{div}_x(\mathbf{b}^\varepsilon) v + (\mathbf{b}^\varepsilon - \mathbf{b}) \cdot D_x v. \tag{3.63}$$
Again, identity (3.63) holds almost everywhere in $\mathbb{R}^n \times (0, \infty)$. As H is a C^2 function, we may calculate $\operatorname{div}_x(\mathbf{b}^\varepsilon)$ directly.

$$\begin{aligned}\operatorname{div}_x(\mathbf{b}^\varepsilon) &= \sum_{k=1}^{n} \frac{\partial}{\partial x_k} b_k^\varepsilon \\ &= \int_0^1 \sum_{k=1}^{n} \sum_{j=1}^{n} H_{p_k p_j}(r D_x u^\varepsilon + (1-r) D_x \tilde{u}^\varepsilon) \left(r u^\varepsilon_{x_k x_j} + (1-r) \tilde{u}^\varepsilon_{x_k x_j} \right) dr.\end{aligned}$$

Since u and $\tilde{u}$ are weak solutions, they are Lipschitz continuous, by *i)* of Definition 3.20. Therefore, by Lemma 3.25, it holds that
$$|r D_x u^\varepsilon + (1-r) D_x \tilde{u}^\varepsilon| \leq \max\{\|u\|_{Lip}, \|\tilde{u}\|_{Lip}\} =: C_0. \tag{3.64}$$
By *iv)* of Definition 3.20, there exists a constant C_1 satisfying
$$u(x+z,t) - 2u(x,t) + u(x-z,t) \leq C_1 \left(1 + \frac{1}{t}\right) |z|^2$$
and
$$\tilde{u}(x+z,t) - 2\tilde{u}(x,t) + \tilde{u}(x-z,t) \leq C_1 \left(1 + \frac{1}{t}\right) |z|^2$$
for every $x, z \in \mathbb{R}^n$ and $t > 0$. Therefore, by Lemma 3.26, it holds for every $t > 0$ and $x \in \mathbb{R}^n$ that
$$\max\{D_x^2 u^\varepsilon(x,t), D_x^2 \tilde{u}^\varepsilon(x,t)\} \leq C_1 \left(1 + \frac{1}{t}\right) I_{n\times n}. \tag{3.65}$$
As H is convex, $D^2 H \geq 0$, which gives us together with (3.64) and (3.65) that
$$\operatorname{div}(\mathbf{b}^\varepsilon) \leq \int_0^1 \max_{x \in B(0,C_0)} \{D^2 H(x)\} C_1 \left(1 + \frac{1}{t}\right) dr. \tag{3.66}$$
Therefore there is a constant $C > 0$ for which
$$\operatorname{div}(\mathbf{b}^\varepsilon) \leq C \left(1 + \frac{1}{t}\right) \quad \text{a.e. in } \mathbb{R}^n \times (0, \infty). \tag{3.67}$$

Fix a point $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$. We define for $t \in [0, t_0)$

$$B(t) \coloneqq B(x_0, R(t_0 - t)) \subset \mathbb{R}^n,$$

where $R \coloneqq \max\{|DH(p)| \mid |p| \leq \max\{\|u\|_{Lip}, \|\tilde{u}\|_{Lip}\}\}$.

In addition, define $e\colon [0, t_0) \to \mathbb{R}$ by

$$e(t) \coloneqq \int\limits_{B(t)} v(x,t)\,dx.$$

By equation (3.63), it holds that

$$v_t = \operatorname{div}_x(\mathbf{b}^\varepsilon)v + (\mathbf{b}^\varepsilon - \mathbf{b}) \cdot D_x v - \operatorname{div}_x(\mathbf{b}^\varepsilon v). \tag{3.68}$$

Thus

$$\begin{aligned}
\dot{e}(t) &= \frac{d}{dt} \int\limits_{B(t)} v(x,t)\,dx = \int\limits_{B(t)} v_t\,dx - R \int\limits_{\partial B(t)} v\,dS \\
&= \int\limits_{B(t)} \operatorname{div}_x(\mathbf{b}^\varepsilon)v + (\mathbf{b}^\varepsilon - \mathbf{b}) \cdot D_x v - \operatorname{div}_x(\mathbf{b}^\varepsilon v)\,dx - R \int\limits_{\partial B(t)} v\,dS \\
&= \int\limits_{B(t)} \operatorname{div}_x(\mathbf{b}^\varepsilon)v + (\mathbf{b}^\varepsilon - \mathbf{b}) \cdot D_x v\,dx - \int\limits_{\partial B(t)} (\mathbf{b}^\varepsilon \cdot \nu + R)v\,dS \\
&\leq \int\limits_{B(t)} C\left(1 + \frac{1}{t}\right) v + (\mathbf{b}^\varepsilon - \mathbf{b}) \cdot D_x v\,dx - \int\limits_{\partial B(t)} (\mathbf{b}^\varepsilon \cdot \nu + R)v\,dS
\end{aligned}$$

In the above, ν is the outward pointing unitary normal vector of $\partial B(t)$. By definition, $R \geq |\mathbf{b}^\varepsilon|$ and thus, the last integrand is positive. As we consider those points, on which u and $\tilde{u}$ are differentiable, $D_x u^\varepsilon \to D_x u$, and $D_x \tilde{u}^\varepsilon \to D_x \tilde{u}$, as $\varepsilon \to 0$. Therefore it holds that $\mathbf{b}^\varepsilon \to \mathbf{b}$ as $\varepsilon \to 0$. Hence it holds for almost every $t \in (0, t_0)$ that

$$\dot{e}(t) \leq C\left(1 + \frac{1}{t}\right) \int\limits_{B} v\,dx = C\left(1 + \frac{1}{t}\right) e(t). \tag{3.69}$$

Fix $\delta \in (0, t_0)$ and choose ϕ to be a smooth function such that $\phi(z) = 0$, if $|z| \leq \delta(\|u\|_{Lip} + \|\tilde{u}\|_{Lip})$ and $\phi(z) > 0$ otherwise. By the initial value condition, $u(x,0) - \tilde{u}(x,0) = 0$ for every $x \in \mathbb{R}^n$, and therefore

$$v(x,s) = \phi(w(x,s)) = \phi(u(x,s) - \tilde{u}(x,s)) = 0 \quad \text{for } s \leq \delta.$$

Thus,

$$e(\delta) = \int_{B(\delta)} v(x,\delta)\,dx = \int_{B(x_0,R(t_0-\delta))} v(x,\delta) = 0.$$

From (3.69), by Grönwall's Lemma, it holds for $r \in (\delta, t_0)$ that

$$e(r) \le e(\delta)\exp\left(\int_{\delta}^{r} C\left(1+\frac{1}{s}\right)ds\right) = 0. \tag{3.70}$$

We refer to Lemma 2.7.2. in [6] for the proof of Grönwall's Lemma. It follows from (3.70) that for $r \in (0, t_0)$,

$$v(x,r) = \phi(w(x,r)) = 0.$$

That is, by definition,

$$|u(x,r) - \tilde{u}(x,r)| \le \delta(\|u\|_{Lip} + \|\tilde{u}\|_{Lip}) \quad \text{for } r \in (0,t_0).$$

Since the above holds with any $\delta \in (0, t_0)$, we have shown that

$$u(x,r) = \tilde{u}(x,r)$$

for all $x \in B(x_0, R(t_0 - r))$. As the weak solutions are Lipschitz continuous, also $u(x_0,t_0) = \tilde{u}(x_0,t_0)$. As this result holds for any $(x_0,t_0) \in \mathbb{R}^n \times (0,\infty)$ the weak solution is unique. □

Together, theorems 3.11 and 3.27 and lemmas 3.19 and 3.18 give the following theorem.

Theorem 3.28. (Hopf-Lax formula as weak solution) *Suppose $g\colon \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous and $H\colon \mathbb{R}^n \to \mathbb{R}$ is a C^2 convex function satisfying the growth condition*

$$\lim_{|p|\to\infty} \frac{H(p)}{\|p\|} = \infty.$$

If either H is strongly convex, or g is semiconcave, the function given by the Hopf-Lax formula,

$$u(x,t) = \min_{y\in\mathbb{R}^n}\left\{tH^*\left(\frac{x-y}{t}\right) + g(y)\right\},$$

is the unique weak solution to

$$\begin{cases} u_t + H(D_x u) = 0 & \text{in } \mathbb{R}^n \times (0,\infty); \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\} \end{cases} \tag{3.71}$$

4 Conservation Laws Revisited

Let us return to the scalar conservation laws. In section 2, we defined an integral solution to the initial value problem

$$\begin{cases} u_t + [F(u)]_x = 0 & \text{in } \mathbb{R}\times(0,\infty); \\ u = g & \text{on } \mathbb{R}\times\{t=0\}. \end{cases} \tag{4.1}$$

As discussed in the beginning of Section 3, by defining $u = w_x$ we can look for w that solves

$$\begin{cases} w_t + F(w_x) = 0 & \text{in } \mathbb{R}\times(0,\infty); \\ w(x,0) = \int\limits_0^x g(y)\,dy & \text{for } x \in \mathbb{R}. \end{cases} \tag{4.2}$$

Suppose that $F \in C^2(\mathbb{R})$ is convex and

$$\lim_{|x|\to\infty} \frac{F(x)}{|x|} = \infty.$$

By Theorem 3.11, the function w given by the Hopf-Lax formula (3.25) is Lipschitz continuous and solves the initial value problem (4.2) almost everywhere. If either F is strongly convex, or $\int_0^x g(y)\,dy$ is semiconcave, w is the unique weak solution to (4.2), by Theorem 3.28.

Thus the function

$$u(x,t) = \frac{\partial}{\partial x}\left[\min_{y\in\mathbb{R}}\left\{tF^*\left(\frac{x-y}{t}\right) + \int\limits_0^y g(z)\,dz\right\}\right], \tag{4.3}$$

that we show in Theorem 4.2 to be defined almost everywhere in $\mathbb{R}\times(0,\infty)$, is a natural candidate for a weak solution to the initial value problem (4.1).

4.1 Lax-Oleinik Formula

In this section, we introduce Lax-Oleinik formula that gives the function defined in (4.3). Hereafter, we assume that $F\colon \mathbb{R} \to \mathbb{R}$ belongs to $C^2(\mathbb{R})$ and is strongly convex. By Lemma 3.15, F satisfies the growth condition at infinity. For strong convexity, see Definition 3.13. Since F is strongly convex, F' is strictly increasing and it has an inverse function. We denote $G := (F')^{-1}$. For the Legendre transform of F, see Definition 3.3, we write $L := F^*$.

Lemma 4.1. *Let $F\colon \mathbb{R} \to \mathbb{R}$ be in $C^2(\mathbb{R})$ and strongly convex. Suppose $F(0) = 0$. Then $L = F^*$ is strongly convex and nonnegative.*

Proof. Let $q \in \mathbb{R}$. The Legendre transform of F is, according to Definition 3.3,

$$L(q) = \max_{p\in\mathbb{R}} \{pq - F(p)\}.$$

Since $F \in C^2(\mathbb{R})$, the maximum in the above equation is reached when $F'(p) = q$, that is, at $p = G(q)$. Therefore, for every $q \in \mathbb{R}$,

$$L(q) = qG(q) - F(G(q)). \tag{4.4}$$

By differentiation, we have

$$L'(q) = G(q) + qG'(q) - F'(G(q))G'(q) = G(q).$$

As F is strongly convex, functions F' and G are increasing. Since we have

$$L''(q) = G'(q) > 0,$$

function L is strongly convex. Now L' is strictly increasing and therefore obtains each value at most once. Suppose that $0 = L'(q_0) = G(q_0)$ for some $q_0 \in \mathbb{R}$. Since L is strongly convex, q_0 is the global minimum of L. By identity (4.4) and the assumption that $F(0) = 0$, it holds that $L(q_0) = 0$. Therefore L is nonnegative. □

Theorem 4.2. (Lax-Oleinik formula). *Assume that $F \in C^2(\mathbb{R})$ is strongly convex and $F(0) = 0$. In addition, let $g \in L^\infty(\mathbb{R})$ and define $h\colon \mathbb{R} \to \mathbb{R}$ as*

$$h(y) = \int_0^y g(z)\,dz.$$

Then for each $t > 0$ the following statements hold.

i) For almost every $x \in \mathbb{R}$, there is a unique $y(x,t)$ satisfying

$$\min_{y\in\mathbb{R}} \left\{ tL\left(\frac{x-y}{t}\right) + h(y) \right\} = tL\left(\frac{x-y(x,t)}{t}\right) + h(y(x,t)).$$

ii) $x \mapsto y(x,t)$ is nondecreasing.

iii) For almost every $x \in \mathbb{R}$, the function u defined in (4.3) is given by the Lax-Oleinik formula:

$$u(x,t) = G\left(\frac{x-y(x,t)}{t}\right). \tag{4.5}$$

Proof. Fix $t > 0$ and consider $x_1, x_2 \in \mathbb{R}$, where $x_1 < x_2$. As g is bounded, there exists a $M > 0$, for which $|g(x)| < M$ for all $x \in \mathbb{R}$. Therefore

$$|h(x_2) - h(x_1)| = \left| \int_{x_1}^{x_2} g(y)\, dy \right| \leq \int_{x_1}^{x_2} |g(y)|\, dy \leq M|x_2 - x_1|.$$

Hence h is Lipschitz continuous. By assumption, F is strongly convex, and by Lemma 4.1, L is strongly convex. As in the proof of Theorem 3.5, there exists at least one point $y_1 \in \mathbb{R}$, for which

$$\min_{y \in \mathbb{R}} \left\{ tL\left(\frac{x_1 - y}{t}\right) + h(y) \right\} = tL\left(\frac{x_1 - y_1}{t}\right) + h(y_1). \tag{4.6}$$

To calculate the minimum

$$\min_{\mathbf{y} \in \mathbb{R}} \left\{ tL\left(\frac{x_2 - y}{t}\right) + h(y) \right\}, \tag{4.7}$$

we show first that

$$tL\left(\frac{x_2 - y_1}{t}\right) + h(y_1) < tL\left(\frac{x_2 - y}{t}\right) + h(y) \quad \text{for all } y < y_1.$$

Let $y < y_1$ and

$$\tau = \frac{y_1 - y}{x_2 - x_1 + y_1 - y}.$$

Now $\tau \in (0,1)$, and we have

$$\begin{cases} x_2 - y_1 & = \tau(x_1 - y_1) + (1-\tau)(x_2 - y) \\ x_1 - y & = (1-\tau)(x_1 - y_1) + \tau(x_2 - y). \end{cases}$$

As L is strongly convex, it holds that

$$L\left(\frac{x_2 - y_1}{t}\right) < \tau L\left(\frac{x_1 - y_1}{t}\right) + (1-\tau) L\left(\frac{x_2 - y}{t}\right) \tag{4.8}$$

and that

$$L\left(\frac{x_1 - y}{t}\right) < (1-\tau) L\left(\frac{x_1 - y_1}{t}\right) + \tau L\left(\frac{x_2 - y}{t}\right). \tag{4.9}$$

By adding the corresponding sides of (4.8) and (4.9) together and multiplying the resulting inequality sidewise by t, we arrive at the following inequality:

$$tL\left(\frac{x_2 - y_1}{t}\right) + tL\left(\frac{x_1 - y}{t}\right) < tL\left(\frac{x_1 - y_1}{t}\right) + tL\left(\frac{x_2 - y}{t}\right). \tag{4.10}$$

As identity (4.6) implies, it holds that

$$tL\left(\frac{x_1-y_1}{t}\right)+h(y_1)\le tL\left(\frac{x_1-y}{t}\right)+h(y). \tag{4.11}$$

Inequalities (4.10) and (4.11) together give us:

$$tL\left(\frac{x_2-y_1}{t}\right)+h(y_1)\le tL\left(\frac{x_2-y}{t}\right)+h(y).$$

Therefore, if $x_2 > x_1$ and y_1 is the minimizer of (4.7) corresponding to x_1, $y_2 \ge y_1$, where y_2 is the minimizer corresponding to x_2.

By defining $y(x,t)$ as the smallest value among the minimizers for each $x \in \mathbb{R}$, the mapping $x \mapsto y(x,t)$ is nondecreasing for every $t > 0$. As the set of discontinuities of $y(x,t)$ is at most countably infinite, $y(x,t)$ is continuous at almost every $x \in \mathbb{R}$. At these x, the minimizer $y = y(x,t)$ is unique. Therefore the first two statements hold.

By Lemma 3.9, the function given by the Hopf-Lax formula is Lipschitz continuous and thus differentiable almost everywhere. As shown above, the mapping $x \mapsto y(x,t)$ is monotone. As shown Theorem 3.4.3 of [5], monotone functions are almost everywhere differentiable. Therefore, given $t > 0$, the function in (4.3) is defined for almost every $x \in \mathbb{R}$.

$$\begin{aligned} u(x,t) &= \frac{\partial}{\partial x}\left[tL\left(\frac{x-y(x,t)}{t}\right)+\int_0^{y(x,t)} g(z)\,dz\right] \\ &= L'\left(\frac{x-y(x,t)}{t}\right)(1-y_x(x,t))+g(y(x,t))y_x(x,t). \end{aligned} \tag{4.12}$$

Since $y(x,t)$ is the minimizer of $tL\left(\frac{x-y(x,t)}{t}\right)+h(y(x,t))$, the differentiable function $y \mapsto tL\left(\frac{x-y}{t}\right)+h(y)$ has a minimum at $y(x,t)$ and therefore,

$$-L'\left(\frac{x-y(x,t)}{t}\right)+g(y(x,t))=0.$$

In the proof of Lemma 4.1, we demonstrated that $L' = G$. With these observations, (4.12) gives us

$$u(x,t)=L'\left(\frac{x-y(x,t)}{t}\right)=G\left(\frac{x-y(x,t)}{t}\right).$$

Hence we have shown that given $t > 0$, u is obtained by the Lax-Oleinik formula (4.5) for almost every $x \in \mathbb{R}$. □

Theorem 4.3. *Assume that $F \in C^2(\mathbb{R})$ is strongly convex and $g \in L^\infty(\mathbb{R})$. Let $u\colon \mathbb{R}\times[0,\infty) \to \mathbb{R}$ be defined for $t > 0$ by the Lax-Oleinik formula* (4.5) *and for $t = 0$ as the limit of $u(x,t)$ as $t \to 0$. Then u is an integral solution to the initial value problem*

$$\begin{cases} u_t + [F(u)]_x = 0 & \text{in } \mathbb{R}\times(0,\infty); \\ u = g & \text{on } \mathbb{R}\times\{t=0\}. \end{cases}$$

Proof. Let $w(x,t)$ be given for $x \in \mathbb{R}$ and $t > 0$ by the Hopf-Lax formula

$$w(x,t) = \min_{y\in\mathbb{R}} \left\{ tL\left(\frac{x-y}{t}\right) + h(y) \right\},$$

where $L = F^*$ and $h(x) = \int_0^x g(y)\,dy$. By Lemma 3.9, w is Lipschitz continuous and therefore, by Theorem 3.8, differentiable almost everywhere in $\mathbb{R}\times(0,\infty)$. By Theorem 3.11, w satisfies

$$\begin{cases} w_t + F(w_x) = 0 & \text{a.e. in } \mathbb{R}\times(0,\infty); \\ w = h & \text{on } \mathbb{R}\times\{t=0\}. \end{cases}$$

Let $v\colon \mathbb{R}\times[0,\infty) \to \mathbb{R}$ be smooth with compact support in $\mathbb{R}\times[0,\infty)$. Then v_x is smooth, compactly supported and satisfies

$$\int_0^\infty \int_{-\infty}^\infty w_t v_x + F(w_x) v_x \,dx\,dt = 0.$$

As w is Lipschitz continuous, we have by integration by parts

$$\begin{aligned} \int_0^\infty \int_{-\infty}^\infty w_t(x,t) v_x(x,t)\,dx\,dt &= -\int_0^\infty \int_{-\infty}^\infty w(x,t) v_{xt}(x,t)\,dx\,dt + \int_{-\infty}^\infty h(x) v_x(x,0)\,dx \\ &= \int_0^\infty \int_{-\infty}^\infty w_x(x,t) v_t(x,t)\,dx\,dt + \int_{-\infty}^\infty g(x) v(x,0)\,dx. \end{aligned}$$

Hence w_x satisfies condition (2.8), that is,

$$0 = \int_0^\infty \int_{-\infty}^\infty w_x(x,t) v_t(x,t) + F(w_x(x,t)) v_x(x,t)\,dx\,dt + \int_{-\infty}^\infty g(x) v_x(x,0)\,dx.$$

By Theorem 4.2, the function u given by the Lax-Oleinik formula satisfies $u = w_x$ almost everywhere. Therefore it holds for u that

$$0 = \int_0^\infty \int_{-\infty}^\infty u(x,t)v_t(x,t) + F(u(x,t))v_x(x,t)\,dx\,dt + \int_{-\infty}^\infty g(x)v_x(x,0)\,dx.$$

Hence u is an integral solution to the given initial value problem. □

4.2 Entropy Condition

To ensure uniqueness of integral solutions to the initial value problems for scalar conservation laws, we introduce a property of the function given by the Lax-Oleinik formula. This property plays a similar role as the semiconcavity condition in the weak solutions to the Hamilton-Jacobi Equations.

Lemma 4.4. *Let $F \in C^2(\mathbb{R})$ be strongly convex and $g \in L^\infty(\mathbb{R})$. Define $u\colon \mathbb{R}\times(0,\infty) \to \mathbb{R}$ by the Lax-Oleinik formula* (4.5)*. Then there exists a positive constant C satisfying the* entropy condition

$$u(x+z,t) - u(x,t) \le \frac{C}{t}z \tag{4.13}$$

for every $t > 0$, $x \in \mathbb{R}$ and $z > 0$.

Proof. Fix $x \in \mathbb{R}$ and $t > 0$. As F is strongly convex, by Lemma 3.15, it satisfies the growth condition at infinity. Hence there exists a $M > 0$ such that the minimum in (4.3) is reached by some $y \in \mathbb{R}$ for which

$$\left|\frac{x-y}{t}\right| < M.$$

As $G = (F')^{-1}$ is continuously differentiable, it is bounded on the interval $[-M, M]$. Thus there is a constant $C > 0$ satisfying

$$|G(p) - G(q)| \le C|p-q|$$

for every $p, q \in [-M, M]$.

Since both G and $y(\cdot, t)$ are nondecreasing, as shown in the proofs of

Theorem 4.1 and Theorem 4.2, it holds for every $z > 0$ that

$$\begin{aligned} u(x,t) &= G\left(\frac{x - y(x,t)}{x}\right) \\ &\geq G\left(\frac{x - y(x+z,t)}{x}\right) \\ &\geq G\left(\frac{x + z - y(x+z,t)}{x}\right) - \frac{C}{t}z \\ &= u(x+z,t) - \frac{C}{t}z. \end{aligned}$$

This finishes the proof. □

Note that, as $y(x,t)$ is continuous for almost every $x \in \mathbb{R}$, the left and right limits of u exist for almost every $x \in \mathbb{R}$ at any given time $t > 0$. Therefore Lemma 4.4 implies that the integral solution given by the Lax-Oleinik formula satisfies the Lax entropy condition given in (2.14), that is,

$$F'(u_l) > \dot{s} > F'(u_r).$$

Let us conclude by defining entropy solutions and proving that these solutions are unique.

Definition 4.5. A bounded function $u\colon \mathbb{R} \times [0,\infty) \to \mathbb{R}$ is an *entropy solution* to the initial value problem

$$\begin{cases} u_t + [F(u)]_x = 0 & \text{in } \mathbb{R}\times(0,\infty); \\ u = g & \text{on } \mathbb{R}\times\{t=0\}, \end{cases} \tag{4.14}$$

provided that it is an integral solution to (4.14), that is, for all smooth test functions $v\colon \mathbb{R}\times[0,\infty) \to \mathbb{R}$ with compact support in $\mathbb{R}\times[0,\infty)$,

$$0 = \int_0^\infty \int_{-\infty}^\infty v_t(x,t)u(x,t) + v_x(x,t)F(u(x,t))\,dx\,dt + \int_{-\infty}^\infty v(x,0)g(x)\,dx, \tag{4.15}$$

and that there exists a constant $C > 0$ such that

$$u(x+z,t) - u(x,t) \leq C\left(1 + \frac{1}{t}\right)z \tag{4.16}$$

for every $z > 0$ and almost every $(x,t) \in \mathbb{R}\times(0,\infty)$.

Theorem 4.6. *Suppose that $F \in C^2(\mathbb{R})$ is convex and that $g \in L^\infty(\mathbb{R})$. Then, up to a set of measure zero in $\mathbb{R}\times(0,\infty)$, there exists at most one entropy solution to the initial value problem* (4.14)*.*

Proof. Assume that u and $\tilde{u}$ are two entropy solutions to (4.14). Define $w\colon \mathbb{R}\times[0,\infty)\to\mathbb{R}$ as $w(x,t) = u(x,t) - \tilde{u}(x,t)$. By the fundamental theorem of calculus, for any $(x,t)\in\mathbb{R}\times(0,\infty)$,

$$\begin{aligned} F(u(x,t)) - F(\tilde{u}(x,t)) &= \int_0^1 \frac{d}{dr} F(ru(x,t) - (1-r)\tilde{u}(x,t))\,dr \\ &= \left(\int_0^1 F'(ru(x,t) + (1-r)\tilde{u}(x,t))\,dr\right)(u(x,t) - \tilde{u}(x,t)) \\ &=: b(x,t)w(x,t). \end{aligned} \tag{4.17}$$

We approximate w by smooth mollifications. Let $\eta\in C^\infty(\mathbb{R}^2)$ be the standard mollifier, that is,

$$\eta(x,t) = \begin{cases} C\exp\left(\frac{1}{x^2+t^2-1}\right) & \text{for } x^2+t^2<1; \\ 0 & \text{for } x^2+t^2\geq 1, \end{cases}$$

where $C>0$ is a constant such that $\int_{\mathbb{R}^2}\eta\,dx\,dt = 1$. For $\varepsilon>0$, we define η^ε as

$$\eta^\varepsilon(x) := \frac{1}{\varepsilon^2}\eta\left(\frac{x}{\varepsilon},\frac{t}{\varepsilon}\right).$$

Fix $\varepsilon>0$ and let u^ε and $\tilde{u}^\varepsilon$ be mollifications of u and $\tilde{u}$ respectively. We define

$$b^\varepsilon(x,t) := \int_0^1 F'(ru^\varepsilon(x,t) + (1-r)\tilde{u}^\varepsilon(x,t))\,dr. \tag{4.18}$$

Let $v\colon\mathbb{R}\times[0,\infty)\to\mathbb{R}$ be a smooth test function with compact support in $\mathbb{R}\times[0,\infty)$. Since u and $\tilde{u}$ satisfy the initial value condition, by (4.15) and

(4.17) we have

$$
\begin{aligned}
0 &= \int_0^\infty \int_{-\infty}^\infty v_t(u - \tilde{u}) + v_x(F(u) - F(\tilde{u}))\, dx\, dt \\
&= \int_0^\infty \int_{-\infty}^\infty w(v_t + v_x b)\, dx\, dt \qquad (4.19) \\
&= \int_0^\infty \int_{-\infty}^\infty w(v_t + b^\varepsilon v_x)\, dx\, dt + \int_0^\infty \int_{-\infty}^\infty w(b - b^\varepsilon) v_x\, dx\, dt.
\end{aligned}
$$

Fix $T > 0$ and let $\phi\colon \mathbb{R}\times[0,\infty) \to \mathbb{R}$ be a smooth function with a compact support in $\mathbb{R}\times(0,T)$. To prove that $w = 0$ almost everywhere, we show that

$$\int_0^\infty \int_{-\infty}^\infty w\phi\, dx\, dt = 0.$$

Let v^ε be the solution to the following terminal value problem for the linear transport equation

$$
\begin{cases}
v_t^\varepsilon + b^\varepsilon v_x^\varepsilon = \phi & \text{in } \mathbb{R} \times (0,T); \\
v^\varepsilon(x,T) = 0 & \text{for } x \in \mathbb{R}.
\end{cases}
\qquad (4.20)
$$

We solve (4.20) by the method of characteristics. Fix $x \in \mathbb{R}$ and $t \in [0,T]$. Let x^ε be the solution to the following characteristic ODE for (4.20),

$$
\begin{cases}
\dot{x}^\varepsilon(s) = b^\varepsilon(x^\varepsilon(s), s) & \text{for } s \geq t; \\
x^\varepsilon(t) = x.
\end{cases}
\qquad (4.21)
$$

As u^ε and $\tilde{u}^\varepsilon$ are smooth and $F \in C^2(\mathbb{R})$, $b^\varepsilon(y,s)$ is uniformly Lipschitz continuous in y and continuous in s. Therefore, by the Picard-Lindelöf Theorem, the solution x^ε is unique for $s \in [t,T]$. For the proof of the Picard-Lindelöf Theorem, we refer to Theorem 4.4. in [2].

By the fundamental theorem of calculus and the PDE in (4.20), we have

$$
\begin{aligned}
v^\varepsilon(x,t) &= \int_T^t \frac{d}{ds} v^\varepsilon(x^\varepsilon(s),s)\, ds \\
&= -\int_t^T v_t^\varepsilon(x^\varepsilon(s),s) + b^\varepsilon(x^\varepsilon(s),s)\, v_x^\varepsilon(x^\varepsilon(s),s)\, ds \qquad (4.22)\\
&= -\int_t^T \phi(x^\varepsilon(s),s)\, ds.
\end{aligned}
$$

Since (4.21) has a unique solution x^ε for every $(x,t) \in \mathbb{R}\times[0,T]$, $v^\varepsilon(x,t)$ is the unique solution to (4.20). As b^ε is bounded, the length of the path $x^\varepsilon(s)$ for $s \in [t,T]$ is bounded. Therefore, as ϕ has compact support in $\mathbb{R}\times(0,T)$, the support of v^ε is compact in $\mathbb{R}\times[0,T)$.

By (4.19) we have

$$\int_0^\infty \int_{-\infty}^\infty w\phi\, dx\, dt = \int_0^\infty \int_{-\infty}^\infty w[b^\varepsilon - b] v_x^\varepsilon\, dx\, dt. \qquad (4.23)$$

Next, we show that, as $\varepsilon \to 0$,

$$\int_0^\infty \int_{-\infty}^\infty w[b^\varepsilon - b] v_x^\varepsilon\, dx\, dt \to 0.$$

Let us first show that for some $C > 0$,

$$\int_{-\infty}^\infty |v_x^\varepsilon(x,t)| dx \le C, \qquad (4.24)$$

when $t \in (0,\tau)$ and τ is small enough.

Since ϕ has compact support in $\mathbb{R}\times(0,T)$, we can choose $\tau > 0$ satisfying $\phi(x,t) = 0$ for $(x,t) \in \mathbb{R}\times(0,\tau)$. Fix $t \in [0,\tau)$ and let $\{x_0, x_1, \cdots, x_N\} \subset \mathbb{R}$ be such that $x_0 < x_1 < \cdots < x_N$. For $k \in \{0,1,\cdots,N\}$, let x_k^ε be the solution to

$$\begin{cases} \dot{x}_k^\varepsilon(s) = b^\varepsilon(x_k^\varepsilon(s),s) & \text{for } s \in (t,\tau); \\ x_k^\varepsilon(t) = x_k. \end{cases}$$

As these solutions are unique, the paths $(x_k^\varepsilon(s), s)$, where $s \in [t, \tau)$, do not intersect. Therefore, for every $s \in (t, \tau)$,

$$x_0^\varepsilon(s) < x_1^\varepsilon(s) < \cdots < x_N^\varepsilon(s).$$

Since $\phi(x_k^\varepsilon(s), s) = 0$ for $s \in (t, \tau)$, by (4.22), v^ε is constant on each path $(x_k^\varepsilon(s), s)$. Therefore, we have

$$\sum_{k=1}^{N} |v^\varepsilon(x_k, t) - v^\varepsilon(x_{k-1}, t)| = \sum_{k=1}^{N} |v^\varepsilon(x_k^\varepsilon(\tau), \tau)) - v^\varepsilon(x_{k-1}^\varepsilon(\tau), \tau)|.$$

As the variation of v^ε with regard to x is

$$V_x(v^\varepsilon(\cdot, t)) = \sup \sum_{k=1}^{N} |v^\varepsilon(x_k, t) - v^\varepsilon(x_{k-1}, t)| = \int_{-\infty}^{\infty} |v_x^\varepsilon(x, t)| dx,$$

where the supremum is taken over all partitions $\{x_0, x_1, \cdots, x_N\} \subset \mathbb{R}$, we have

$$\int_{-\infty}^{\infty} |v_x^\varepsilon(x, t)| dx \leq \int_{-\infty}^{\infty} |v_x^\varepsilon(x, \tau)| dx.$$

Since v^ε is continuously differentiable and has a compact support, it is absolutely continuous and thus has bounded variation. Here, we refer to Theorem 3.73. of [5], where it is shown that every absolutely continuous function of a closed interval of the real line is of bounded variance. Therefore, there exists a $C > 0$ satisfying

$$\int_{-\infty}^{\infty} |v_x^\varepsilon(x, t)| dx \leq C$$

for every $t \in [0, \tau)$.

Next, we show that for each $\tau > 0$, there is a constant $C_\tau > 0$ satisfying

$$|v_x^\varepsilon(x, t)| \leq C_\tau \tag{4.25}$$

for each $(x, t) \in \mathbb{R} \times (\tau, T)$.

Fix $\tau > 0$. Since both u^ε and $\tilde{u}^\varepsilon$ are smooth mollifications of entropy solutions, inequality (4.16) gives the following upper estimate for the partial derivatives u_x^ε and $\tilde{u}_x^\varepsilon$,

$$u_x^\varepsilon(x, t), \tilde{u}_x^\varepsilon(x, t) \leq C \left(1 + \frac{1}{t}\right).$$

This together with the convexity of $F \in C^2(\mathbb{R})$ and the boundedness of entropy solutions implies that there exists some constant $C > 0$ satisfying

$$b_x^\varepsilon(x,t) = \int_0^1 F''(ru^\varepsilon(x,t) + (1-r)\tilde{u}^\varepsilon(x,t))(ru_x^\varepsilon(x,t) + (1-r)\tilde{u}_x^\varepsilon(x,t))\,dr \\ \leq \frac{C}{\tau} \quad \text{for } (x,t) \in \mathbb{R}\times(\tau,T). \tag{4.26}$$

Differentiating (4.20) with respect to x, we obtain

$$v_{tx}^\varepsilon + b^\varepsilon v_{xx} + b_x^\varepsilon v_x^\varepsilon = \phi_x.$$

Let C be the constant in (4.26) and set $a(x,t) := e^{\lambda t} v_x^\varepsilon(x,t)$, where $\lambda = \frac{C}{\tau} + 1$.
Now

$$\begin{aligned} a_t + b^\varepsilon a_x &= \lambda a + e^{\lambda t}(v_{xt}^\varepsilon + b^\varepsilon v_{xx}^\varepsilon) \\ &= \lambda a + e^{\lambda t}(\phi_x - b_x^\varepsilon v_x^\varepsilon) \\ &= (\lambda - b_x^\varepsilon)a + e^{\lambda t}\phi_x. \end{aligned} \tag{4.27}$$

Since v^ε is continuous and has compact support, a reaches a nonnegative maximum at some point $(x_0, t_0) \in \mathbb{R} \times [\tau, T]$.

1. If $t_0 = T$, $a(x_0, t_0) = 0$. This holds as $v^\varepsilon(x, T) = 0$ for every $x \in \mathbb{R}$.

2. If $t_0 \in [\tau, T)$, $a_t(x_0, t_0) \leq 0$ and $a_x(x_0, t_0) = 0$. In this case, we have

$$a_t(x_0, t_0) + b^\varepsilon(x_0, t_0)a_x(x_0, t_0) \leq 0.$$

By identity (4.27), we have

$$(\lambda - b_x^\varepsilon(x_0, t_0))a(x_0, t_0) + e^{\lambda t_0}\phi_x(x_0, t_0) \leq 0.$$

As $b_x^\varepsilon \leq \frac{C}{\tau}$ by (4.26) and $\lambda = 1 + \frac{C}{\tau}$, we have

$$a(x_0, t_0) \leq -e^{\lambda t_0}\phi_x(x_0, t_0) \leq e^{\lambda T} \sup |\phi_x|.$$

By assuming that a reaches a nonpositive minimum at $(x_1, t_1) \in \mathbb{R}\times[\tau, T]$, we have

$$a(x_1, t_1) \geq -e^{\lambda T} \sup |\phi_x|.$$

Hence, we have for every $(x,t) \in \mathbb{R}\times(\tau,T)$, that

$$|a(x,t)| = |e^{\lambda t} v_x^\varepsilon(x,t)| \leq e^{\lambda T} \sup |\phi_x|.$$

This implies that

$$|v_x^\varepsilon(x,t)| \le e^{\lambda(T-\tau)} \sup |\phi_x|$$

for $(x,t) \in \mathbb{R}\times(\tau,T)$, which proves (4.25).

We return to study the right hand side of (4.23). As v_x^ε is compactly supported in $\mathbb{R}\times(0,T)$, we have

$$\begin{aligned}
\int_0^\infty \int_{-\infty}^\infty w[b^\varepsilon - b]v_x^\varepsilon \, dx\, dt &= \int_0^T \int_{-\infty}^\infty w[b^\varepsilon - b]v_x^\varepsilon \, dx\, dt \\
&= \int_0^\tau \int_{-\infty}^\infty w[b^\varepsilon - b]v_x^\varepsilon \, dx\, dt + \int_\tau^T \int_{-\infty}^\infty w[b^\varepsilon - b]v_x^\varepsilon \, dx\, dt \\
&=: A_\tau^\varepsilon + B_\tau^\varepsilon.
\end{aligned}$$

As u and $\tilde{u}$ are bounded, w is bounded. By Lemma 3.24, as $\varepsilon \to 0$,

$$u^\varepsilon \to u \text{ and } \tilde{u}^\varepsilon \to \tilde{u} \text{ a.e. in } \mathbb{R}\times(0,\infty).$$

Therefore, as $\varepsilon \to 0$,

$$b^\varepsilon \to b \text{ a.e. in } \mathbb{R}\times(0,\infty).$$

Since v_x^ε has a compact support in $\mathbb{R}\times(0,\infty)$, we have for almost every $(x,t) \in \mathbb{R}\times(0,\infty)$, that

$$|w[b^\varepsilon - b]v_x^\varepsilon| \le C_1 |v_x^\varepsilon|. \tag{4.28}$$

Furthermore, by (4.24), we have for small enough τ,

$$|A_\tau^\varepsilon| \le \int_0^\tau \int_{-\infty}^\infty |w[b^\varepsilon - b]v_x^\varepsilon| \, dx\, dt \le \tau C_2. \tag{4.29}$$

By (4.25), $|v_x^\varepsilon|$ is bounded in $\mathbb{R}\times(\tau,T)$ for every $\tau > 0$. Since $|v_x^\varepsilon|$ also has compact support in $\mathbb{R}\times(\tau,T)$, by (4.28) and the dominated convergence theorem, $B_\tau^\varepsilon \to 0$ as $\varepsilon \to 0$. By letting also $\tau \to 0$, we see that

$$\int_0^\infty \int_{-\infty}^\infty w\phi \, dx\, dt = 0.$$

As the above holds for every smooth function ϕ with compact support in $\mathbb{R}\times(0,\infty)$, $w = 0$ almost everywhere in $\mathbb{R}\times(0,\infty)$. This proves that the entropy solution is unique for almost every $(x,t) \in \mathbb{R}\times(0,\infty)$. □

In the Theorem 4.2 4.3 Lemma 4.4 and Theorem 4.6 theorems and lemmas above, we have proved the following theorem.

Theorem 4.7. *Suppose* $F \in C^2(\mathbb{R})$ *is strongly convex and* $g \in L^\infty(\mathbb{R})$. *Then, up to a set of Lebesgue measure zero, the function given by the Lax-Oleinik formula* (4.5) *is the unique entropy solution to*

$$\begin{cases} u_t + [F(u)]_x = 0 & \textit{in } \mathbb{R} \times (0, \infty); \\ u = g & \textit{on } \mathbb{R} \times \{t = 0\}. \end{cases}$$

References

[1] Apostol, T. M., *Mathematical Analysis*, Second Edition, California Institute of Technology, 1974.

[2] Gyllenberg, M., Lamberg, L., Ola, P., Piiroinen, P. & Häsä, J., *Tavalliset Differentiaaliyhtälöt*, lecture notes, University of Helsinki, 2016. Available at `http://wiki.helsinki.fi/pages/viewpage.action?pageId=197658128`

[3] Evans, L. C., *Partial Differential Equations*, Graduate Studies in Mathematics, American Mathematical Society, 1998.

[4] Heinonen, J., *Lectures on Lipschitz Analysis*, Lectures at the 14th Jyväskylä Summer School, University of Jyväskylä, 2004.

[5] Holopainen, I., *Reaalianalyysi I*, lecture notes, University of Helsinki, 2017. Available at `http://wiki.helsinki.fi/pages/viewpage.action?pageId=197658326`

[6] Juutinen, P., *Differentiaaliyhtälöt*, Lecture notes, University of Jyväskylä, 2008. Available at `http://users.jyu.fi/~peanju/dyluennot.pdf`

[7] Levandosky, J., *Partial Differential Equations of Applied Mathematics*, Stanford University, 2002. Available at `http://web.stanford.edu/class/math220a/index.html`

[8] Rudin, W. *Principles of Mathematical Analysis*, International Series in Pure and Applied Mathematics, University of Rochester, 1953.