Hamilton-Jacobi equations

Tommi Brander

Matematiikan pro gradu

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Kevät 2012

Tiivistelmä

Aihe: Pro-gradu tutkielma (Hamiltonin ja Jacobin yhtälöt)
Alkuperäinen nimi: Hamilton-Jacobi equations
Tekijä: Tommi Brander
Päiväys: 23.4.2012

Tämä Pro gradu-tutkielma käsittelee Hamiltonin ja Jacobin yhtälöitä, jotka kuvaavat mekaanisen järjestelmän kehitystä klassisen mekaniikan puitteissa. Hamiltonin ja Jacobin yhtälöitä käytetään myös säätöteoriassa sekä kvanttimekaniikassa. Hamiltonin mekaaniikan kehitti Sir William Rowan Hamilton valon käytöksen mallintamiseen ja Carl Gustav Jacob Jacobi kehitti sitä edelleen.

Tutkielmassa annamme ehdot, joiden nojalla Hopfin ja Laxin kaava antaa ratkaisun Hamiltonin ja Jacobin yhtälöihin liittyvään alkuarvo-ongelmaan. Sen jälkeen määritämme sopivan heikon ratkaisun käsitteen ja näytämme heikkojen ratkaisujen olevan yksikäsitteisiä tietyillä ehdoilla. Lähestymme Hamiltonin ja Jacobin alkuarvo-ongelmaa asettamalla variaatio-ongelman, jonka Hopfin ja Laxin kaava ratkaisee. Osoitamme, että Hopfin ja Laxin kaavan antama ratkaisuehdokas on Lipschitz-jatkuva ja toteuttaa dynaamisen ohjelmoinnin periaatteen, joka kytkee sen optimaalisen säädön teoriaan. Sen jälkeen näytämme, että Hopfin ja Laxin kaavan antama funktio todella ratkaisee Hamiltonin ja Jacobin yhtälön alkuarvo-ongelman.

Tärkeä työkalu Hopfin ja Laxin kaavan käsittelyssä on Legendren muunnos, joka muuntaa funktion sen konveksiksi duaaliksi. Näytämme, että konvekseille ja tarpeeksi nopeasti kasvaville funktioille Legendren muunnos sovellettuna kahteen kertaan antaa alkuperäisen funktion takaisin. Tutkielmassa tutkitaan Hamiltonin ja Lagrangen funktioita, jotka täyttävät nämä ehdot.

Lopuksi määrittelemme, mitä tarkoitamme heikolla ratkaisulla Hamiltonin ja Jacobin yhtälön alkuarvo-ongelmaan. Määritelmässä käytämme semikonkaaveja funktioita. Osoitamme, että alkuehtojen semikonkaavius tai Hamiltonin funktion vahva konveksisuus takaavat heikkojen ratkaisuiden semikonkaaviuden, ja että semikonkaaveja ratkaisuja voi olla vain yksi, kunhan alkuarvo-ongelma täyttää sopivat säännöllisyysehdot.

Contents

1 Introduction

In this master's thesis I will present an existence and a uniqueness result for the Hamilton-Jacobi equation and focus on two tools used in the proofs: Legendre transformation and the Hopf-Lax formula. The thesis closely follows section 3.3 in monograph [8].

The Hamilton-Jacobi partial differential equation is of the form

$$H(x, D_x u(x, \alpha, t), t) + \partial_t u(x, \alpha, t) = K(\alpha, t)$$

[18] where x is a vector, t the time variable, α a real-valued parameter, and H and K given maps. The function u is unknown and $D_x u$ and $\partial_t u$ are its derivatives with regards to space and time, respectively. Hamiltonian mechanics were introduced by Sir William Rowan Hamilton, originally to model the behaviour of light, and were later developed further by Carl Gustav Jacob Jacobi [18].

Contemporarily the Hamilton-Jacobi equation is used in optimal control theory [3], quantum theory and mechanics [18]. In mechanics the equation is used for finding invariants or approximate invariants [18].

For methods of solving the equation, see e.g. sections 3.3 and 10 of the monograph [8] and section 46 of the monograph [1].

We will consider a simplification of the Hamilton-Jacobi equation. We assume some regularity and that the Hamiltonian H only depends on the gradient of u, the time derivative of u only depends on the space variable x, and that K is identically zero. We consider initial value problem with initial value g. That is, for given maps $g, H \colon \mathbb{R}^n \to \mathbb{R}$ the problem we will investigate is

$$\partial_t u + H(Du) = 0 \text{ in } \mathbb{R}^n \times]0, \infty[\tag{1.0.1}$$

$$u = g \text{ on } \mathbb{R}^n \times \{0\}. \tag{1.0.2}$$

1.1 Hopf-Lax formula

To solve the initial value problem (1.0.1)–(1.0.2) we minimise the action, getting the formula

$$u(x,t) = \inf \left\{ \int_0^t L(\dot{w}(s))\, ds + g(w(0)) \right\} \tag{1.1.1}$$

with infimum taken over all smooth functions w with $w(t) = x$. The formula can be derived via optimal control [8]. Hopf-Lax formula gives one explicit solution to the variational problem (1.1.1):

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

Hopf-Lax formula was originally stated by Peter D. Lax in [15] and by Eberhard Hopf for a special case in [11].

The formulae are equivalent, since both consider the initial conditions $y = w(0)$, and the path between two points is determined by minimising the energy, so we only need to minimise over the initial value. Minimising over the initial value is also consistent with an assumption of classical mechanics—that initial velocity and position are sufficient to describe a closed system, which is also called Newton's principle of determinancy [1]. For more detail on the equivalence see the proof of Theorem 4.2.1.

The Hopf-Lax formula admits the dynamic programming principle (in analogy with Theorem 1 in section 10.3. in the monograph [8])

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

which can be understood as dividing the optimisation problem into distinct components.

The function u defined by the Hopf-Lax formula is Lipschitz continuous (given certain restrictions on g and H). Hopf-Lax formula indeed gives a solution that is, by Rademacher's theorem (Theorem 2.2.5), almost everywhere differentiable.

1.2 Uniqueness

To study the uniqueness of solutions, we introduce the notion of semiconcavity. We show the solution u given by the Hopf-Lax formula is semiconcave, if the initial value g is semiconcave of if the Hamiltonian H is strongly convex. Under some regularity assumptions we show the uniqueness of semiconcave Lipschitz continuous solutions to the Hamilton-Jacobi equation.

1.3 Outline

The first section is the introduction. In the second section notation and various tools are introduced; reader might do well to only skim the second section and return to it when necessary or when a result is referenced. The third section contains physical background and applications of the Hamilton-Jacobi equation, as well as connections between the Legendre transformation and two sets of ordinary differential equations used in physics: Euler-Lagrange equations and Hamilton's equations.

In the fourth section we introduce the Legendre transformation and the Hopf-Lax formula. We use Legendre transformation to show certain properties of the Hopf-Lax formula, and then use these tools to establish regularity of solutions to the Hamilton-Jacobi equation. This regularity justifies the notion of solutions, as we show in the section.

In the fifth section we show that uniqueness does not hold for the solutions of Hamilton-Jacobi equation in the current generality, so we introduce additional regularity, semiconcavity, to guarantee uniqueness.

In the sixth section we mention some further developments of the theory.

2 Notation and preliminaries

In this section we introduce the notation and definitions used throughout the thesis, and state theorems and lemmata that will be used later.

2.1 Vectors and matrices

Let $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ be vectors in $\mathbb{R}^n$. By $a \cdot b$ we denote the inner product of vectors, i.e. $a \cdot b = \sum_{i=1}^{n} a_i b_i$. By $|a|$ we denote the norm in $\mathbb{R}^n$. We have the Cauchy-Schwarz inequality.

Theorem 2.1.1 (Cauchy-Schwarz inequality). *Let a and b be vectors in $\mathbb{R}^n$. Then*

$$|a \cdot b| \leq |a||b|.$$

Let M be a matrix. We denote its transpose by M^T.

Definition 2.1.2 (Positive semidefinite matrix)**.** *Let A and B be symmetric $n \times n$ square matrices. We say A is positive semidefinite, denoted by $A \geq 0$, if and only if for all $x \in \mathbb{R}^n$*

$$x^T Ax \geq 0.$$

We also write $A \geq B$ to mean $A - B \geq 0$.

The following lemma can be proven by diagonalising the matrices.

Lemma 2.1.3. *Suppose A and B are symmetric $n \times n$-matrices, with $0 \leq A \leq M_1 I$ and $B \leq M_2 I$, where M_1 and M_2 are positive constants. Then*

$$\sum_{i,j=1}^{n} a_{ij} b_{ij} \leq n M_1 M_2.$$

2.2 Differentiation and integration

In this thesis, by the integral we mean the Lebesgue integral.

Theorem 2.2.1 (Dominated convergence theorem)**.** *Let the functions $(f_k)_{k\in\mathbb{N}}$ be measurable and suppose there exists integrable g such that for all k $f_k \leq g$ holds almost everywhere. Suppose further that almost everywhere f_k have the limit f as k increases. Then f is measurable and*

$$\lim_{k\to\infty} \int f_k = \int f,$$

where the integrals are taken over the domain where f, f_k and g are measurable and the inequalities and limits hold.

Proof can be found in e.g. [8], Theorem 5 in appendix E.

For measurable $A \subset \mathbb{R}^n$ with finite but positive Lebesgue measure $|A|$ we write the mean integral of an integrable function $f\colon A \to \mathbb{R}$ as

$$\fint_A f\,dx = \frac{1}{|A|} \int_A f\,dx.$$

Let $U \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ be open (and nonempty). We denote by $p_i, i \in \{1, \ldots, n\}$ the first n variables, by $x_i, i \in \{n+1, \ldots, 2n\}$ the second n variables, and by t the variable numbered $2n+1$. That is, $(p, x, t) \in U$ with $p, x \in \mathbb{R}^n$ and $t \in \mathbb{R}$.

For $F\colon U \to \mathbb{R}$ differentiable at $(p, x, t) \in U$ we write its partial derivatives as $\partial_{p_1} F, \ldots, \partial_{p_n} F, \partial_{x_1} F, \ldots, \partial_{x_n} F, \partial_t F$. For the gradient of F with regards to the p-variables we write

$$D_p F = (\partial_{p_1} F, \ldots, \partial_{p_n} F)$$

and $D_x F$ is defined similarly.

For subsets of $\mathbb{R}^n \times \mathbb{R}$ we use similar notation, but write the gradient operator with regards to first variables simply as D.

Notation $\dot{x}$ means the derivative of x with regards to the time variable, which is usually t or s.

Divergence of a differentiable vector-valued function $f = (f_1, \ldots, f_n)\colon A \to \mathbb{R}^n$, with A an open subset of $\mathbb{R}^n$, we write as

$$\operatorname{div} f(x) = \sum_{j=1}^{n} \partial_{x_j} f_j(x).$$

The Laplacian of a twice differentiable function $g\colon A \subset \mathbb{R}^n \to \mathbb{R}$ is defined as divergence of the gradient of g:

$$\Delta g = \operatorname{div} Dg.$$

Lemma 2.2.2 (Product rule for divergence)**.** *Let $A \subset \mathbb{R}^n$ be an open set. For differentiable $f\colon A \to \mathbb{R}^n$ and differentiable $g\colon A \to \mathbb{R}$ it holds that*

$$\operatorname{div}(gf) = f \cdot Dg + g \operatorname{div} f.$$

Proof. The proof is a simple calculation:

$$\operatorname{div}(gf) = \sum_{i=1}^{n} \partial_{x_i}(gf_i) = \sum_{i=1}^{n} (f_i \partial_{x_i} g + g \partial_{x_i} f_i) = f \cdot Dg + g \operatorname{div} f.$$

□

The divergence theorem is also known by the names of Ostrogradsky, Gauss or Green.

Theorem 2.2.3 (Divergence theorem)**.** *Suppose $B \subset \mathbb{R}^n$ is a ball with boundary ∂B and normal vector ν pointing outward from B. Further suppose that $f\colon B \to \mathbb{R}^n$ is continuously differentiable. Then*

$$\int_B \operatorname{div} f \, dx = \int_{\partial B} f \cdot \nu \, dS.$$

An elementary proof can be found in e.g. [12], chapter 0.d.

Lemma 2.2.4. *Suppose $A \subset \mathbb{R}^n$ is open and $f\colon A \to \mathbb{R}$ is Lipschitz with constant C. Then $|Df| \leq C$ whenever f is differentiable.*

Proof. Let $x \in A$ be a point where the function f is differentiable. Then $(Df)(x)$ exists and we suppose $(Df)(x) \neq 0$. We write $D = (Df)(x)$, $D_1 = D/|D|$, and calculate

$$\frac{|f(x + hD_1) - f(x)|}{h} \leq \frac{Ch|D_1|}{h} = C,$$

from which follows $C \geq |Df|$. □

We will use Rademacher's theorem to provide regularity for solutions of the Hamilton-Jacobi equation.

Theorem 2.2.5 (Rademacher's theorem)**.** *Let $A \subset \mathbb{R}^n$ be open and $f\colon A \to \mathbb{R}^m$ be Lipschitz. Then f is differentiable almost everywhere in A.*

For proof see e.g. Theorems 4–6 in section 5.8 in [8].

The following theorem is originally by Grönwall [10].

Theorem 2.2.6 (Differential form of Grönwall's inequality)**.** *Let T be positive and $\eta\colon [0,T] \to [0,\infty[$ absolutely continuous such that there are integrable functions $\phi, \psi\colon [0,T] \to [0,\infty[$ satisfying for almost all t*

$$\eta'(t) \leq \phi(t)\eta(t) + \psi(t).$$

Then for all t

$$\eta(t) \leq \left(\eta(0) + \int_0^t \psi(s)\, ds\right) \exp\left(\int_0^t \phi(s)\, ds\right).$$

The proof can be found in e.g. [8], inequality j in appendix B.

2.3 Properties of functions

We introduce superlinearity and several notions of convexity and concavity. We also define the Lipschitz continuous maps.

Definition 2.3.1. *Let (X, d_X) and (Y, d_Y) be metric spaces. Then the mapping $f\colon X \to Y$ is Lipschitz with constant L if and only if for all $x_1, x_2 \in X$ we have $d_Y(f(x_1), f(x_2))) \leq L d_X(x_1, x_2)$.*

We write $\mathrm{Lip}(f)$ *for the optimal Lipschitz constant of a Lipschitz mapping f.*

We define superlinear (also known as coercive) functions.

Definition 2.3.2 (Superlinearity)**.** *We say that the mapping $f\colon \mathbb{R}^n \to \mathbb{R}$ is superlinear if and only if*

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

We next define convex and concave functions. Later we will also define strongly convex (Definition 2.3.8) and semiconcave (Definition 2.3.10) functions.

Definition 2.3.3 (Convexity and concavity)**.** *Let $A \subset \mathbb{R}^n$ be a convex set. We call a function $f\colon A \to \mathbb{R}$ convex if and only if for all $x, y \in A$ and for all $\lambda \in]0,1[$*

$$\lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda) y).$$

We call a function f concave if and only if the function $-f$ is convex.

The following result holds in finite-dimensional spaces.

Theorem 2.3.4. *[Continuity of convex functions] A convex or concave function $f\colon A \to \mathbb{R}$ is continuous on a convex open set $A \subset \mathbb{R}^n$.*

We refer to Theorem 10.1 of [17] for a proof.

The following theorem is originally by Jensen [13].

Theorem 2.3.5 (Jensen's inequality)**.** *Let $f\colon \mathbb{R} \to \mathbb{R}$ be convex, $A \subset \mathbb{R}^n$ open and bounded, and $g \in L^1(A; \mathbb{R})$. Then*

$$f\left(⨍_A g(x)\,dx\right) \leq ⨍_A f(g(x))\,dx.$$

For proof see e.g. Theorem 2 in appendix B of [8].

Lemma 2.3.6. *[Supporting hyperplane of a convex function] A convex function $f\colon A \to \mathbb{R}$ is supported by a hyperplane at every point, which means that for all $x \in A$ there is $r \in \mathbb{R}^n$ so that for all $y \in A$*

$$f(y) \geq f(x) + r \cdot (y - x).$$

Proof proceeds by convex analysis: since f is convex, its epigraph is a convex set, and therefore has a supporting hyperplane at every boundary point. For details see for example [5, chapter 2.5.2].

The second derivative of a convex function is positive:

Lemma 2.3.7. *For convex function f we have $D^2 f \geq 0$ whenever the second derivative is defined. Conversely, if f is twice differentiable and $D^2 f \geq 0$, then f is convex.*

For proof, see e.g. Theorem 4.5 of [17].

Strongly convex functions are convex. Sometimes they are called uniformly convex.

Definition 2.3.8 (Strong convexity)**.** *Let $A \subset \mathbb{R}^n$ be a convex set. We call a function $f\colon A \to \mathbb{R}$ strongly convex with constant $m > 0$ if and only if for all $x, y \in A$ and for all $\lambda \in]0,1[$*

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y) - \frac{m}{2} \lambda (1-\lambda) |x-y|^2.$$

Norm squared is an example of a strongly convex function.

Example 2.3.9. *The function $x \mapsto |x|^2$ is strongly convex. To check this, first let x and y be vectors in $\mathbb{R}^n$ and let λ be a positive number less than one. Since*

$$|\lambda x + (1-\lambda)y|^2 - \lambda|x|^2 - (1-\lambda)|y|^2 = -\lambda(1-\lambda)|x-y|^2,$$

norm squared is strongly convex with constant 2.

Semiconcave function is concave by removing some fixed quadratic term.

Definition 2.3.10 (Semiconcavity)**.** *Let $A \subset \mathbb{R}^n$ be a convex set. We call a function $f\colon A \to \mathbb{R}$ semiconcave with constant $C \geq 0$ if and only if for all $x, z \in A$*

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

In some cases another definition of semiconcavity is easier to verify.

Lemma 2.3.11. *Let A be as above and $f\colon A \to \mathbb{R}$ a function. The function f is semiconcave with constant C if and only if the function*

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

is concave.

Proof can be found in [6] as Proposition 1.1.3.

2.4 Mollifiers and convolution

Mollifiers are smooth functions with special properties, used to construct sequences of smooth functions approximating nonsmooth function, via convolution.

Definition 2.4.1 (Convolution)**.** *Let $A \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}$ and $g\colon \mathbb{R}^n \to \mathbb{R}$ be measurable functions with f locally integrable and g having compact support. Their convolution is defined as $(f * g)(x) = \int_A f(y)g(x-y)\,dy$ and is a function on $\mathbb{R}^n$.*

Let n be a positive integer, $A \subset \mathbb{R}^n$ an open set and $\varepsilon > 0$. Write $A_\varepsilon = \{x \in A \mid \operatorname{dist}(x, \partial A) > \varepsilon\}$, where ∂ indicates boundary and dist distance.

Definition 2.4.2 (Standard mollifier)**.** *We define $\eta\colon \mathbb{R}^n \to \mathbb{R}$ by*

$$\eta(x) = Cexp\left(-\frac{1}{1-|x|^2}\right)$$

for $|x| < 1$, and 0 elsewhere, with C chosen so that $\int_{\mathbb{R}^n} \eta(x)\,dx = 1$.

We define $\eta_\varepsilon\colon \mathbb{R}^n \to \mathbb{R}$ by

$$\eta_\varepsilon(x) = \varepsilon^{-n}\eta\left(x/\varepsilon\right)$$

and call them standard mollifiers.

Smoothening or mollification of a function is convolution of the function with a mollifier.

Definition 2.4.3 (Mollification)**.** *Assume $u\colon U \to \mathbb{R}$ is locally integrable. Its mollification $u^\varepsilon\colon U_\varepsilon \to \mathbb{R}$ is the convolution $\eta_\varepsilon * u$.*

Convolution commutes by change of variables, so we have $u^\varepsilon = u * \eta_\varepsilon$.

Lemma 2.4.4 (Properties of mollifiers)**.** *The following hold for locally integrable mapping u:*

- *The function u^ε is infinitely differentiable in U_ε.*
- *Almost everywhere $u^\varepsilon \to u$ as $\epsilon \to 0$.*

Proof can be found in [8], Theorem 6 in appendix C.

Lemma 2.4.5. *Let $A \subset \mathbb{R}^n$ be an open set and $f\colon A \to \mathbb{R}$ smooth. Then $D^\alpha(f*\eta_\varepsilon) = D^\alpha(f)*\eta_\varepsilon$ for any multi-index α with f at least $|\alpha|$ times continuously differentiable.*

Proof. It suffices to to prove the lemma for the case $|\alpha| = 1$. The other cases can be proved by induction. Since

$$\begin{aligned}
\partial_k(f*\eta_\varepsilon)(x) = \partial_k \int f(x-z)\eta_\varepsilon(z)\,dz &= \int \partial_k(f(x-z)\eta_\varepsilon(z))\,dz \\
&= \int \eta_\varepsilon(z)\partial_k(f(x-z))\,dz = \int \eta_\varepsilon(z)\,(\partial_k f)\,(x-z)\,dz \\
&= (\eta_\varepsilon * \partial_k f)\,(x),
\end{aligned}$$

we are done. □

A mollified Lipschitz function is still Lipschitz with the same constant and its derivatives also converge.

Lemma 2.4.6. *Let u be Lipschitz continuous with constant C. Then $|Du^\varepsilon| \le C$ and almost everywhere $Du^\varepsilon \to Du$ as $\epsilon \to 0$.*

Proof. For the first part we calculate for arbitrary $x_1, x_2 \in \mathbb{R}^n$

$$\begin{aligned}
|u^\varepsilon(x_1) - u^\varepsilon(x_2)| &= \left| \int \eta_\varepsilon(y)u(x_1-y)\,dy - \int \eta_\varepsilon(y)u(x_2-y)\,dy \right| \\
&= \left| \int \eta_\varepsilon(y)(u(x_1-y) - u(x_2-y))\,dy \right| \\
&\le \int \eta_\varepsilon(y)|u(x_1-y) - u(x_2-y)|\,dy \\
&\le C|x_1-x_2| \int \eta_\varepsilon(y)\,dy = C|x_1-x_2|,
\end{aligned}$$

where the integrals are taken over $B(0,\varepsilon)$.

For the second part, we first mention that by Rademacher's theorem (Theorem 2.2.5) Du exists almost everywhere. Further, by Lemma 2.4.4 $u^\varepsilon \to u$ almost everywhere and Du^ε is defined everywhere, so in particular all of these hold almost everywhere. By Lemma 2.4.5 for any $k \in \{1,\dots,n\}$ $(\partial_k u)^\varepsilon = \partial_k(u^\varepsilon)$, so by Lemma 2.4.4 we only need to show that $\partial_k u$ is locally integrable. By Lemma 2.2.4 it is bounded and by definition it is a limit of a sequence of measurable functions and therefore measurable. Therefore it is locally integrable. □

Lemma 2.4.7. *Let $f\colon \mathbb{R}^n \times [0,\infty[\to \mathbb{R}$ be continuous and almost everywhere differentiable. Suppose the mapping $x \mapsto f(x,t)$ is concave for all positive t. Then for all $t > 0$ the mapping $x \mapsto f^\varepsilon(x,t)$ is concave.*

Proof. Let $(a,t),(b,t) \in \mathbb{R}^n \times]0,\infty[$. By the concavity assumption we know that the mapping

$$\alpha \mapsto Df(\alpha a + (1-\alpha)b, t)\cdot(b-a)$$

is decreasing for $\alpha \in]0,1[$ since the function restricted to a line is concave and has decreasing one-dimensional derivative. We note that the function is differentiable almost everywhere, so in particular it is possible that the gradient does not even exist. A calculation for $x \neq y$ where integrals are taken over $B(0,\varepsilon)$ proves the theorem:

$$
\begin{aligned}
Df^{\varepsilon}(x,t)\cdot(y-x) - Df^{\varepsilon}(y,t)\cdot(y-x) &= (Df^{\varepsilon}(x,t) - Df^{\varepsilon}(y,t))\cdot(y-x) \\
=&\Big(D\int f(x-z,t-h)\eta_{\varepsilon}(z,h)\,d(z,h) \\
&- D\int f(y-z,t-h)\eta_{\varepsilon}(z,h)\,d(z,h)\Big)\cdot\ (y-x) \\
=&\Big(\int Df(x-z,t-h)\eta_{\varepsilon}(z,h)\,d(z,h) \\
&- \int Df(y-z,t-h)\eta_{\varepsilon}(z,h)\,d(z,h)\Big)\cdot(y-x) \\
=&\int (Df(x-z,t-h) - Df(y-z,t-h))\cdot(y-x)\eta_{\varepsilon}(z,h)\,d(z,h),
\end{aligned}
$$

which is non-positive, since the integrand is non-positive. This implies that $x \mapsto f^{\varepsilon}(x,t)$ is concave. We used Lemma 2.4.5 in the calculation. □

3 Physical background

Euler-Lagrange equations, Hamilton's equations and the Legendre transformation all concern the Hamiltonian $H = H(p,x)$ and the Lagrangian $L = L(q,x)$, where p is generalised momentum, x is position and q holds the position for $\dot{x}$, velocity. The Lagrangian can be interpreted as the difference between kinetic energy and potential energy, while the Hamiltonian is the total energy of the system, which means the sum of kinetic and potential energy. We refer to chapter 3 of [1] for more details.

Euler-Lagrange equations

$$-\frac{d}{ds}(D_qL(\dot{x}(s),x(s))) + D_xL(\dot{x}(s),x(s)) = 0$$

give necessary and sufficient conditions for function x to be an extreme value of a given functional. In our case the extreme values represent the motions of certain mechanical system (section 13 in [1]). Hamilton's equations

$$\begin{cases} \dot{x} &= D_pH(p,x) \\ \dot{p} &= -D_xH(p,x) \end{cases} \tag{3.0.1}$$

are one reformulation of classical mechanics and they are equivalent to Euler-Lagrange equations, see section 15 of [1].

Let us now derive the Hamilton's ordinary differential equations. Assume the Lagrangian $L\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ to be twice continuously differentiable. For all $x_0, y \in \mathbb{R}^n$ and $t > 0$ define the admissible class

$$\mathbb{A} = \{w \in C^2([0,t];\mathbb{R}^n) | w(0) = y, w(t) = x_0\}$$

and the action functional $I\colon \mathbb{A} \to \mathbb{R}$,

$$I(w) = \int_0^t L(\dot{w}(s),w(s))ds. \tag{3.0.2}$$

Theorem 3.0.8 (Euler-Lagrange equations)**.** *Suppose that function $x \in \mathbb{A}$ minimises the action functional I. Then the function x solves the Euler-Lagrange equations*

$$-\frac{d}{ds}(D_q L(\dot{x}(s), x(s))) + D_x L(\dot{x}(s), x(s)) = 0. \tag{3.0.3}$$

Proof. Let $v\colon [0,t] \to \mathbb{R}^n$, $v(0) = v(t) = 0$ be twice continuously differentiable function. Since $x \in \mathbb{A}$, $x + \tau v$ is also in the admissible class $\mathbb{A}$ for all real numbers τ. Set $i\colon \mathbb{R} \to \mathbb{R}$, $i(\tau) = I(x + \tau v)$. Now $i(0) = I(x)$ is a minimum of the action functional (3.0.2), and supposing it is differentiable at 0, $i'(0) = 0$.

Let us compute $i'(0)$ by using the fundamental theorem of calculus. Since

$$\begin{aligned}
i'(\tau) =& \frac{d}{d\tau} I(x + \tau v) = \frac{d}{d\tau} \int_0^t L(\dot{x}(s) + \tau \dot{v}(s), x(s) + \tau v(s))\, ds \\
=& \int_0^t \frac{d}{d\tau} L(\dot{x}(s) + \tau \dot{v}(s), x(s) + \tau v(s))\, ds \\
=& \int_0^t \sum_{i=1}^n (\partial_{q_i} L)(\dot{x}(s) + \tau \dot{v}(s), x(s) + \tau v(s)) \dot{v}_i(s) \\
& + (\partial_{x_i} L)(\dot{x}(s) + \tau \dot{v}(s), x(s) + \tau v(s)) v_i(s)\, ds,
\end{aligned}$$

we have for $\tau = 0$

$$\begin{aligned}
0 = i'(0) &= \sum_{i=1}^n \int_0^t (\partial_{q_i} L)(\dot{x}(s), x(s)) \dot{v}_i(s) + (\partial_{x_i} L)(\dot{x}(s), x(s)) v_i(s)\, ds \qquad (3.0.4) \\
&= \sum_{i=1}^n \int_0^t -\frac{d}{ds} \left((\partial_{q_i} L)(\dot{x}(s), x(s)) \right) v_i(s) + (\partial_{x_i} L)(\dot{x}(s), x(s)) v_i(s)\, ds \\
&= \sum_{i=1}^n \int_0^t v_i(s) \left((\partial_{x_i} L)(\dot{x}(s), x(s)) - \frac{d}{ds} \left((\partial_{q_i} L)(\dot{x}(s), x(s)) \right) \right) ds.
\end{aligned}$$

Now suppose the claim is false; that is, for some $i \in \{1, ..., n\}$ and for some $s \in [0, t]$

$$-\frac{d}{ds}(\partial_{q_i} L(\dot{x}(s), x(s))) + \partial_{x_i} L(\dot{x}(s), x(s)) \neq 0. \tag{3.0.5}$$

By continuity of the expression (3.0.5) there is a small interval $J \subset [0, t]$, $s \in J$, in which the expression is strictly positive or strictly negative. By selecting a mollifier v with support in J (Definition 2.4.2), we get a contradiction with equation (3.0.4).

We supposed that the function i is differentiable at 0, and now prove it. By dominated convergence theorem 2.2.1 we only need to show that the difference quotients $\frac{1}{h}(L(\dot{x} + h\dot{v}, x + hv)) - L(\dot{x}, x))$ have a common integrable dominant for small h.

Since $v \in C^2$, both v and $\dot{v}$ are are bounded on the compact interval, so there is some constant $C > 0$ such that for all $h \in \mathbb{R}$ and for all $s \in [0, t]$

$$|(\dot{x} + h\dot{v}, x + hv) - (\dot{x}, x)| \leq C|h|.$$

Since the Lagrangian L is twice continuously differentiable, its derivative is continuous and so L is Lipschitz on any subset of a compact set, such as

$$\{(\dot{x}(s) + h\dot{v}(s), x(s) + hv(s)) \mid s \in [0, t], |h| \leq M\}$$

for some positive M. Hence for all $s \in [0,t]$

$$|L(\dot{x}(s)+h\dot{v}(s),x(s)+hv(s)))-L(\dot{x}(s),x(s))| \leq C_M|h|,$$

where C_M is a positive constant that depends on the choice of M. Hence for all $s \in [0,t]$ and for $|h| \leq M$ we have $|\frac{1}{h}(L(\dot{x}+h\dot{v},x+hv))-L(\dot{x},x))| \leq C_M$, which is sufficient to guarantee uniform integrability on a compact interval. $\square$

We call the solution $x \in \mathbb{A}$ of the Euler-Lagrange equations (3.0.3) a critical point. Our goal is to derive Hamilton's ordinary differential equations (3.0.1) from the Euler-Lagrange equations.

Let x be a critical point and set $p\colon [0,t] \to \mathbb{R}^n$,

$$p(s) = D_q L(\dot{x}(s),x(s)). \tag{3.0.6}$$

For the development of theory in this section we assume the following lemma without proof:

Lemma 3.0.9. *Let $L\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be in $C^2(\mathbb{R}^n \times \mathbb{R}^n)$. Then there is a unique smooth mapping $q\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ that solves $p = D_q L(q,x)$. We write $q = q(p,x)$.*

Now we can define the Hamiltonian and prove the Hamilton's equations.

Definition 3.0.10 (Hamiltonian)**.** *Given q and L as in the lemma 3.0.9, we define the associated Hamiltonian $H\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as*

$$H(p,x) = p \cdot q(p,x) - L(q(p,x),x). \tag{3.0.7}$$

Theorem 3.0.11 (Hamilton's equations)**.** *Functions x and p satisfy Hamilton's equations*

$$\begin{cases} \dot{x} &= D_p H(p,x) \\ \dot{p} &= -D_x H(p,x). \end{cases} \tag{3.0.8}$$

.

Proof. Lemma 3.0.9 and equation (3.0.6) imply the equality $\dot{x} = q$. Hence for any natural number $i \in \{1,\ldots,n\}$ we have

$$\begin{aligned}
\partial_{p_i} H(p,x) &= \partial_{p_i}\left(p \cdot q(p,x) - L(q(p,x),x)\right) \\
&= \partial_{p_i}\left(\sum_{j=1}^n p_j q_j(p,x)\right) - \partial_{p_i} L(q(p,x),x) \\
&= q_i(p,x) + \sum_j p_j \partial_{p_i} q_j(p,x) - \sum_{j=1}^n (\partial_{q_j} L)\partial_{p_i} q_j(p,x) \\
&= q_i(p,x) + \sum_j \partial_{p_i} q_j(p,x)\left(p_j - \partial_{q_j} L\right) = q_i(p,x) = \dot{x}_i,
\end{aligned}$$

which proves the first set of Hamilton's ordinary differential equations. To prove the second set we calculate for any $i \in \{1,\ldots,n\}$

$$\begin{aligned}
-\partial_{x_i} H(p,x) &= -\sum_{j=1}^n \partial_{x_i}\left(p_j q_j(p,x)\right) + \partial_{x_i} L(q(p,x),x) \\
&= -\sum_{j=1}^n p_j \partial_{x_i} q_j(p,x) + \sum_{j=1}^n \partial_{q_j} L(q(p,x),x)\partial_{x_i} q_j(p,x) + \partial_{x_i} L(q(p,x),x) \\
&= \partial_{x_i} L(q(p,x),x) = \frac{d}{ds}\partial_{q_i} L(q(p,x),x) = \frac{d}{ds} p_i = \dot{p}_i,
\end{aligned}$$

where we used the Euler-Lagrange equations (3.0.3) and equation (3.0.6). $\square$

The Hamiltonian stands for the total energy of the system, so given a closed physical system it should be constant with regards to time. Indeed it is:

Corollary 3.0.12. *The Hamiltonian $H(p,x)$ as a function of time $s \in [0,t]$ is constant.*

Proof. It suffices to show that the time derivative of the Hamiltonian is identically zero.

$$\frac{d}{ds}H(p(s),x(s)) = (D_pH)(p(s),x(s)) \cdot \dot{p}(s) + (D_xH)(p(s),x(s)) \cdot \dot{x}(s) = 0$$

by Theorem 3.0.11. $\square$

4 Weak solution via Hopf-Lax formula

In this section we introduce a notion of weak solution. In our case, weak solution is not necessarily everywhere differentiable, but does solve the equation almost everywhere.

4.1 Legendre transformation

We now assume that the Hamiltonian H does not depend on the x variable—that is, for all $x, y, p \in \mathbb{R}^n$, $H(p,x) = H(p,y)$ and we write $H\colon \mathbb{R}^n \to \mathbb{R}$, $H(p) = H(p,x)$ for any x.

Let $L\colon \mathbb{R}^n \to \mathbb{R}$ be convex and superlinear. Then it is continuous by theorem 2.3.4. Our goal is to formulate the Lagrangian to be independent of x and to show how the Lagrangian and the Hamiltonian relate to each other.

Definition 4.1.1 (Legendre transformation)**.** *Legendre transformation of the Lagrange's function L is $L^*\colon \mathbb{R}^n \to \mathbb{R}$,*

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

The Legendre transformation is well-defined and, further, for every $p \in \mathbb{R}^n$ there is $q \in \mathbb{R}^n$ where the supremum is reached.

Lemma 4.1.2.

$$L^*(p) = \max_{q \in \mathbb{R}^n} \left(p \cdot q - L(q)\right)$$

Proof. Let us fix $p \in \mathbb{R}^n$. The mapping $q \mapsto p \cdot q - L(q) =: a(q)$ is continuous since L is, so it reaches its maximum value on any compact set. Let us see what happens when $|q|$ grows:

$$p \cdot q - L(q) = |q|\left(p \cdot \frac{q}{|q|} - \frac{L(q)}{|q|}\right) \leq |q|\left(|p| - \frac{L(q)}{|q|}\right) \longrightarrow -\infty$$

by superlinearity of L as $|q| \to \infty$.

So for q outside some sufficiently large closed ball $p \cdot q - L(q) \leq -L(0)$ and so the function a reaches its maximum within the large closed ball, as it is compact. $\square$

The simplest example of Legendre transformation is that of the function $x \mapsto x^2/2$.

Example 4.1.3. *We define the Lagrangian $L\colon \mathbb{R}^n \to \mathbb{R}$ by $L(q) = |q|^2/2$. Its Legendre transformation is*

$$L^*(p) = \max_{q\in\mathbb{R}^n} \left(p \cdot q - \frac{1}{2}\ |q|^2 \right).$$

The maximum is achieved when $q = p$. Thus we have

$$L^*(p) = \frac{1}{2}|p|^2.$$

We can generalise this for the Lagrangian $L_n(q) = \frac{1}{n}|q|^n$ for real numbers $n \geq 1$, and further defining $L_\infty(q) = 0$ when $|q| \leq 1$ and infinite otherwise. For finite $n > 1$ we have

$$L_n^*(p) = \left(1 - \frac{1}{n}\right) |p|^{\frac{n}{n-1}}.$$

Thus we have $(L_n^*)^* = L_n$ for $n > 1$. This also holds for $n = 1$ or ∞, given the conventions established above. The next theorem states that this duality holds in general.

We note that the cases $n = 1$ or ∞ are not in the scope of our definition of Legendre transform, since they are not real-valued superlinear functions.

We proved that for each p there is q^* such that value $L^*(p)$ is reached, and that it is the maximum of the mapping $q \mapsto p \cdot q - L(q)$, still denoted by a. Therefore the derivative of the mapping a vanishes there. In particular

$$DL(q^*) = D_q(p \cdot q^*) - Da(q^*) = D_q \sum_{i=1}^{n} p_i q_i^* = p.$$

We remark that q^* may not be unique.

By Definition 3.0.10

$$H(p) = p \cdot q(p) - L(q(p)) = L^*(p)$$

when we select $q\colon \mathbb{R}^n \to \mathbb{R}^n$, $q(p) = q^*$.

Theorem 4.1.4 (Convex duality of the Hamiltonian and the Lagrangian)**.** *For convex superlinear Lagrangian L with $H = L^*$ the Hamiltonian is convex and superlinear. Further, $L = H^*$.*

Proof. We start with the superlinearity of H. We want to show that for all $M \in \mathbb{R}$ there exists a positive α_0 so that for all $\alpha \geq \alpha_0$ and for all $p \in \mathbb{R}^n$ with $|p| = 1$ there is $q \in \mathbb{R}^n$ such that

$$\frac{\alpha p}{|\alpha p|} \cdot q - \frac{L(q)}{|\alpha p|} = p \cdot q - \frac{1}{\alpha} L(q) \geq M.$$

Select $q = 2Mp$. Then what needs to be shown is that

$$2M - \frac{1}{\alpha} L(2Mp) \geq M.$$

Since L is continuous, it is bounded on the closed ball $\overline{B(0, 2M)}$, not taking values in excess of some positive $K \in \mathbb{R}$. For $\alpha \geq K/M$

$$\frac{L(2Mp)}{\alpha} \leq \frac{K}{\alpha} \leq M,$$

so by selecting $\alpha_0 = K/M$ we confirm that the Hamiltonian indeed is superlinear.

To establish convexity of the Hamiltonian, let $\beta \in]0,1[$ and $p_1, p_2 \in \mathbb{R}^n$. Then

$$\begin{aligned}
H(\beta p_1 + (1-\beta)p_2) &= \max_{q\in\mathbb{R}^n} \left((\beta p_1 + (1-\beta)p_2)\cdot q - L(q)\right) \\
&= \max_{q\in\mathbb{R}^n} \left(\beta p_1 \cdot q - \beta L(q) + (1-\beta)p_2\cdot q - (1-\beta)L(q)\right) \\
&\le \beta \max_{q\in\mathbb{R}^n} \left(p_1\cdot q - L(q)\right) + (1-\beta)\max_{q\in\mathbb{R}^n}\left(p_2\cdot q - L(q)\right) \\
&= \beta H(p_1) + (1-\beta)H(p_2),
\end{aligned}$$

which shows the convexity.

We still want to show that $L = H^*$. By Definition 4.1.1 and the equality $H = L^*$, we have for all $p, q \in \mathbb{R}^n$

$$L(q) + H(p) \ge p\cdot q,$$

from which it follows that

$$L(q) \ge \sup_{p\in\mathbb{R}^n} \left(p\cdot q - H(p)\right) = H^*(q).$$

We next want to show that $L \le H^*$. Let us expand H^*:

$$\begin{aligned}
H^*(q) &= \sup_{p\in\mathbb{R}^n}\left(p\cdot q - H(p)\right) = \sup_{p\in\mathbb{R}^n}\left(p\cdot q - L^*(p)\right) \\
&= \sup_{p\in\mathbb{R}^n}\left(p\cdot q - \sup_{q_0\in\mathbb{R}^n}\left(p\cdot q_0 - L(q_0)\right)\right) \\
&= \sup_{p\in\mathbb{R}^n}\left(p\cdot q + \inf_{q_0\in\mathbb{R}^n}\left(-p\cdot q_0 + L(q_0)\right)\right) \\
&= \sup_{p\in\mathbb{R}^n}\inf_{q_0\in\mathbb{R}^n}\left(p\cdot q - p\cdot q_0 + L(q_0)\right) \\
&= \sup_{p\in\mathbb{R}^n}\inf_{q_0\in\mathbb{R}^n}\left(p\cdot (q-q_0) + L(q_0)\right). \qquad (4.1.1)
\end{aligned}$$

The convexity of the Lagrangian L means that it is supported by a hyperplane (see Lemma 2.3.6), that is, there is $r \in \mathbb{R}^n$ so that for all $q_0 \in \mathbb{R}^n$ we have

$$L(q_0) \ge L(q) + r\cdot(q_0 - q),$$

which together with (4.1.1) gives

$$\begin{aligned}
H^*(q) &\ge \sup_{p\in\mathbb{R}^n}\inf_{q_0\in\mathbb{R}^n}\left(p\cdot(q-q_0) + L(q) + r\cdot(q_0-q)\right) \\
&= \sup_{p\in\mathbb{R}^n}\inf_{q_0\in\mathbb{R}^n}\left((p-r)\cdot(q-q_0) + L(q)\right) \\
&\ge \inf_{q_0\in\mathbb{R}^n}\left(0\cdot(q-q_0) + L(q)\right) = L(q).
\end{aligned}$$

□

Example 4.1.5. *Let $L\colon \mathbb{R}\to\mathbb{R}$ be defined by $L(q) = e^{|q|}$. Now the Lagrangian is convex and superlinear. We calculate the Legendre transformation L^*:*

$$H(p) = L^*(p) = \max_{q\in\mathbb{R}}(pq - L(q)) = \max_{q\in\mathbb{R}}\left(pq - e^{|q|}\right).$$

We have by an easy calculation that

$$H(p) = \begin{cases} p(1 - \log(-p)) & \text{for } p < -1 \\ -1 & \text{for } p \in [-1, 1] \\ p(\log p - 1) & \text{for } p > 1. \end{cases}$$

Though the dual of a convex function is convex, strict convexity is lost in this example.

4.2 Hopf-Lax formula

In this section we solve the initial value problem

$$\partial_t u + H(Du) = 0 \text{ in } \mathbb{R}^n \times]0, \infty[\tag{4.2.1}$$

$$u = g \text{ on } \mathbb{R}^n \times \{0\}. \tag{4.2.2}$$

We assume that $H\colon \mathbb{R}^n \to \mathbb{R}$ is convex and superlinear and that $g\colon \mathbb{R}^n \to \mathbb{R}$ is Lipschitz-continuous. We define the function $u\colon \mathbb{R}^n \times [0, \infty[\to \mathbb{R}$ by

$$u(x,t) = \inf \left\{ \int_0^t L(\dot{w}(s))\, ds + g(y) \middle| w(0) = y, w(t) = x, w \in C^1 \right\}. \tag{4.2.3}$$

This choice follows from optimal control theory, see e.g. section 10.3 of [8].

Theorem 4.2.1 (Hopf-Lax formula)**.** *For u defined by equation* (4.2.3) *and for all $x \in \mathbb{R}^n$ and for all $t > 0$*

$$u(x,t) = \min_{y \in \mathbb{R}^n} \left(tL\left(\frac{x-y}{t}\right) + g(y) \right). \tag{4.2.4}$$

The right-hand side of the equation in the Theorem 4.2.1 is called the Hopf-Lax formula, see Theorem 4 in [8, section 3.3].

Proof. Let us first show that u equals a slightly modified Hopf-Lax formula: one with infimum instead of minimum. We accomplish this by considering a function w that is linear between y and x and checking that it is optimal.

We define $w\colon [0, t] \to \mathbb{R}$ by

$$w(s) = y + \frac{s}{t}(x - y).$$

The function w clearly fulfills all requirements set in equation (4.2.3), and hence

$$\begin{aligned} u(x,t) &= \inf \left\{ \int_0^t L(\dot{w}(s))\, ds + g(y) | w(0) = y, w(t) = x, w \in C^1 \right\} \\ &\le \int_0^t L(\dot{w}(s))\, ds + g(y) = \int_0^t L\left(\frac{x-y}{t}\right) ds + g(y) \\ &= tL\left(\frac{x-y}{t}\right) + g(y) \end{aligned}$$

for arbitrary y, which implies

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

We use Jensen's inequality, Theorem 2.3.5, to prove the other direction, so convexity of the Lagrangian is essential for the proof. Any continuously differentiable w with $w(t) = x$ has

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

and multiplying by t and adding $g(y)$ gives

$$tL\left(\frac{1}{t}\int_0^t \dot{w}(s)\,ds\right) + g(y) \leq \int_0^t L(\dot{w}(s))\,ds + g(y),$$

where

$$\int_0^t \dot{w}(s)\,ds = w(t) - w(0) = x - y$$

for $y = w(0)$ by the fundamental theorem of calculus. Since w and thereby y were arbitrary, we have

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

and hence equality.

We still need to check that the infimum is reached. Let us consider the map

$$y \mapsto tL\left(\frac{x-y}{t}\right) + g(y), \tag{4.2.5}$$

which is continuous because L and g are. We only need to show that the function (4.2.5) takes large values when y is large. The claim follows from the facts that L is superlinear and g is Lipschitz. The proof is similar to that of lemma 4.1.2. Writing C for the Lipschitz constant of g we get

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

Now the claim follows from the superlinearity of the Lagrangian L. □

Example 4.2.2. *Let $g\colon \mathbb{R} \to \mathbb{R}$ be defined as $g(x) = |x|$ and $H\colon \mathbb{R} \to \mathbb{R}$ as $H(x) = x^2/2$. Then the corresponding Lagrangian is, by Example 4.1.3, identical to the Hamiltonian. Hence, by definition, we have*

$$u(x,t) = \inf\left\{\int_0^t \frac{1}{2}|\dot{w}(s)|^2\,ds + |w(0)| \,\middle|\, w(t) = x, w \in C^1\right\},$$

which is hard to calculate in any obvious way. The Hopf-Lax formula makes the calculation easy:

$$u(x,t) = \min_{y\in\mathbb{R}}\left(\frac{t}{2}\left|\frac{x-y}{t}\right|^2 + |y|\right) = \min_{y\in\mathbb{R}}\left(\frac{1}{2t}|x-y|^2 + |y|\right)$$

admits a minimising argument. By differentiating and checking separately the case $y = 0$ we get

$$u(x,t) = \begin{cases} \frac{t}{2} + |x+t| & \text{for } x < -t \\ \frac{1}{2t}x^2 & \text{for } x \in [-t,t] \\ \frac{t}{2} + |x-t| & \text{for } t < x. \end{cases}$$

Lemma 4.2.3. *For all $t > 0$ the mapping $x \mapsto u(x,t)$ is Lipschitz with constant* $\mathrm{Lip}(g)$.

Proof. Let $t > 0$ and $x, \hat{x} \in \mathbb{R}^n$. Recall that g is assumed to be Lipschitz, and write $\mathrm{Lip}(g) = C > 0$. Then for z that is a minimiser in Hopf-Lax formula (4.2.4) for $(\hat{x}, t)$ (see Theorem 4.2.1)

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

By the same argument $u(\hat{x},t) - u(x,t) \le C|x-\hat{x}|$. $\square$

Next we show the dynamic programming principle for the Hopf-Lax formula.

Lemma 4.2.4 (Dynamic programming principle)**.** *For all $x \in \mathbb{R}^n$ and for every $s \in]0,t[$*

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

Proof. Let $y \in \mathbb{R}^n$, $t > 0$, $s \in]0,t[$ and select $z \in \mathbb{R}^n$ such that

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

Since $x - z = x - y + (y-z) = (t-s)\frac{x-y}{t-s} + s\frac{y-z}{s}$, we have by the convexity of the Lagrangian

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

and therefore

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

Since y was arbitrary, the result holds for all y and hence for their minimum.

To check the other direction fix $(x,t) \in \mathbb{R}^n \times]0,\infty[$ and let w be the minimiser in Hopf-Lax formula:

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

Set $y = \frac{s}{t}x + (1-\frac{s}{t})w$, from which follow $\frac{y-w}{s} = \frac{x-w}{t}$ and $\frac{x-y}{t-s} = \frac{(1-s/t)(x-w)}{t-s} = \frac{x-w}{t}$, and calculate

$$\begin{aligned}(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) &= (t-s)L\left(\frac{x-w}{t}\right) + \inf_{z\in\mathbb{R}^n}\left(sL\left(\frac{y-z}{s}\right) + g(z)\right) \\ &\leq (t-s)L\left(\frac{x-w}{t}\right) + sL\left(\frac{y-w}{s}\right) + g(w) \\ &= tL\left(\frac{x-w}{t}\right) + g(w) = u(x,t).\end{aligned}$$

Since $u(x,t)$ is greater than the formula for given y, it is necessarily greater than the minimum over all y. $\square$

By the variational equality (4.2.3) we have for all $x \in \mathbb{R}^n$ the equation $u(x,0) = g(x)$. This is consistent with the Hopf-Lax formula (4.2.1) as shown in the next lemma.

Lemma 4.2.5. *Let u defined by the Hopf-Lax formula* (4.2.4). *Let us extend it as $u|_{\mathbb{R}^n\times\{0\}} = g$. Then it becomes a continuous function on $\mathbb{R}^n \times [0,\infty[$.*

Proof. Let us take arbitrary $x \in \mathbb{R}^n$. Let t be positive real number. Then

$$u(x,t) = \min_{z\in\mathbb{R}^n}\left(tL\left(\frac{x-z}{t}\right) + g(z)\right) \leq tL(0) + g(x)$$

from which by letting $t \to 0$ we get $u(x,0) \leq g(x)$.

To check the other direction we select $z_t \in \mathbb{R}^n$ to minimise

$$z \mapsto tL\left(\frac{x-z}{t}\right) + g(z)$$

for $t > 0$. Now z_t approaches x as $t \to 0$. If not, there would be $M > 0$ such that there would be arbitrarily small t for which $|z_t - x| > M$, and then by superlinearity of the Lagrangian L we would have

$$tL\left(\frac{x-z_t}{t}\right) + g(z_t) - g(x) \geq \frac{|x-z_t|}{|x-z_t|}tL\left(\frac{x-z_t}{t}\right) - C|z_t - x|$$

which grows to infinity as t vanishes. But this contradicts $\limsup_{t\to 0} u(x,t) \leq g(x)$. By superlinearity of L, there is $R > 0$ so that for all $|z| > R$ it is true that $L(z) \geq 0$. In particular, L is bounded from below, that is, there is L_{min} such that for all z we have $L(z) \geq L_{min}$. Thus

$$u(x,t) \geq tL\left(\frac{x-y_t}{t}\right) - |g(y_t) - g(x)| + g(x) \geq tL_{min} - C|y_t - x| + g(x),$$

which converges to $g(x)$ as $t \to 0$. $\square$

Corollary 4.2.6. *Lemma 4.2.4 also applies to the case $s = 0$. Thus, we have for all $x \in \mathbb{R}^n$ and for all $s \in [0,t[$*

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

Lemma 4.2.7. *The function u defined by the Hopf-Lax formula* (4.2.4) *is Lipschitz.*

Proof. Let us take $(x,t), (x,\hat{t}) \in \mathbb{R}^n \times [0,\infty[$. Without loss of generality we can assume $\hat{t} < t$. Let z_0 be the minimiser in the Hopf-Lax formula (4.2.4). By Lemma 4.2.4, where we select $s = t - \hat{t}$, we have

$$\begin{aligned}
u(x,t) - u(x,\hat{t}) = &\min_{y\in\mathbb{R}^n}\left(\hat{t}L\left(\frac{x-y}{\hat{t}}\right) + u(y, t-\hat{t})\right)\\
&- \min_{z\in\mathbb{R}^n}\left(\hat{t}L\left(\frac{x-z}{\hat{t}}\right) + g(z)\right)\\
\leq& u(z_0, t-\hat{t}) - g(z_0)\\
=&\min_{y\in\mathbb{R}^n}\left((t-\hat{t})L\left(\frac{z_0-y}{t-\hat{t}}\right) + g(y) - g(z_0)\right)\\
\leq&(t-\hat{t})L(0).
\end{aligned}$$

For the other direction we use Lemma 4.2.3, which gives us control over the behaviour of u as x varies. By lemma 4.2.4 with $s = \hat{t}$ we get

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

so we only need to show that the function defined by formula $h(z) := L(z) - C|z|$ is bounded from below. This follows from the superlinearity of L by the same argument as before.

Now we know that the mappings $x \to u(x,t)$ and $t \to u(x,t)$ are Lipschitz with constants that do not depend on x or t. This finishes the proof of the lemma. □

4.3 Existence of a solution

Now we show the existence of a solution to the initial value problem

$$\partial_t u + H(Du) = 0 \text{ in } \mathbb{R}^n \times]0,\infty[\tag{4.3.1}$$

$$u = g \text{ on } \mathbb{R}^n \times \{0\}. \tag{4.3.2}$$

Lemma 4.3.1. *The function u defined by the Hopf-Lax formula* (4.2.4) *is differentiable almost everywhere on the set $\mathbb{R}^n \times]0,\infty[$.*

Proof. By Lemma 4.2.7, u is Lipschitz continuous and by Rademacher's theorem, Theorem 2.2.5, Lipschitz functions are differentiable almost everywhere. □

Theorem 4.3.2 (Solving the Hamilton-Jacobi initial value problem)**.** *For $x \in \mathbb{R}^n$, $t > 0$ and u differentiable at (x,t) we have*

$$\partial_t u(x,t) + H(Du(x,t)) = 0.$$

Proof. Let $t, h > 0$ and $x, q \in \mathbb{R}^n$. We want to get control of the derivative of u, so by Lemma 4.2.4 we calculate

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

and therefore

$$\frac{1}{h}\left(u(x+hq,t+h)-u(x,t)\right)\ \le L(q),$$

from which by letting $h\to 0$ and relying on the differentiability of u we get

$$Du(x,t)\cdot q+\partial_t u(x,t)\le L(q).$$

Now, since q was arbitrary, by Theorem 4.1.4 we have

$$0\ge u_t(x,t)+\max_{q\in\mathbb{R}^n}\left(Du(x,t)\cdot q-L(q)\right)=\partial_t u(x,t)+H(Du(x,t)).$$

To show the other direction we can use similar techniques. Let x,t and h be as before and $s=t-h$. We still need to control the derivative of u, so let a be a vector in $\mathbb{R}^n$ to be specified later, and let us calculate

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

Ideally we want to remove g from the equation, so we select $y=z$, and we also want $\frac{x-z}{t}=\frac{a-y}{s}$, which is equivalent to $a=\frac{s}{t}x+\left(1-\frac{s}{t}\right)z$, providing us with the estimate

$$u(x,t)-u(a,s)\ge hL\left(\frac{x-z}{t}\right).$$

Since u is differentiable at (x,t), dividing by h gives

$$\frac{u(x,t)-u(a,s)}{h}=\frac{u((x,t)-h(\frac{x-z}{t},1))-u(x,t)}{-h}\ge L\left(\frac{x-z}{t}\right),$$

where letting $h\to 0$ and theorem 4.1.4 provide us with

$$
\begin{aligned}
Du(x,t)\cdot\frac{x-z}{t}+u_t(x,t)\ge L\left(\frac{x-z}{t}\right)&=\max_{p\in\mathbb{R}^n}\left(\frac{x-z}{t}\cdot p-H(p)\right)\\
&\ge\frac{x-z}{t}\cdot Du(x,t)-H(Du(x,t)).
\end{aligned}
$$

This completes the proof. □

Let us check that the Example 4.2.2 satisfies the Hamilton-Jacobi equation outside a set of measure zero.

Example 4.3.3. *We have the equation $\partial_t u+|Du|^2/2=0$ and the mapping $u\colon\mathbb{R}\times[0,\infty[\to\mathbb{R}$ defined as follows:*

$$u(x,t)=\begin{cases}\frac{t}{2}+|x+t| & \text{for } x<-t\\ \frac{1}{2t}x^2 & \text{for } x\in[-t,t]\\ \frac{t}{2}+|x-t| & \text{for } t<x\end{cases}$$

It is easy to check that u is a solution to the Hamilton-Jacobi equation outside the set $\{(x,t)\in\mathbb{R}\times[0,\infty[\quad|\quad t=|x|\}$, which has measure zero.

In summary, we have:

Theorem 4.3.4. *Hopf-Lax formula* (4.2.1) *defines almost everywhere differentiable Lipschitz-continuous solution to the initial value problem* (1.0.1)–(1.0.2).

5 Uniqueness of solutions by semiconcavity

In section 4 we have shown that there exists at least one solution to the initial value problem for the Hamilton-Jacobi equation. In this section we study the uniqueness of solutions.

Example 5.0.5. *Consider the initial value problem*

$$u = 0 \qquad \text{on } \mathbb{R}^n \times \{0\} \tag{5.0.3}$$

$$\partial_t u + |\partial_x u|^2 = 0 \qquad \text{in } \mathbb{R}^n \times]0, \infty[. \tag{5.0.4}$$

It is easy to see that for all $a \geq 0$ the function $u(x,t) = -a^2 t + a|x|$ when $at \geq |x|$ and $u(x,t) = 0$ otherwise, solves this problem.

Hence we need to place more restrictions on g or H to guarantee uniqueness.

5.1 Semiconcavity

We use regularity of initial value g and that of the Hamiltonian H to establish the semiconcavity (Definition 2.3.10) of u with respect to the variable x.

Lemma 5.1.1. *Let g be semiconcave with constant C and u defined by the Hopf-Lax formula* (4.2.1)*. Then for all $t > 0$ the mapping $x \mapsto u(x,t)$ is semiconcave with constant at most C.*

Proof. Let $t > 0$ and $x, z \in \mathbb{R}^n$. Let C be the semiconcavity constant of g. By Theorem 4.2.1 and semiconcavity of g we have for the y that minimises (4.2.4) for $u(x,t)$

$$\begin{aligned}
&u(x+z,t) + u(x-z,t) - 2u(x,t) \\
&\qquad = \min_{a \in \mathbb{R}^n} \left(tL\left(\frac{x+z-a}{t}\right) + g(a) \right) + \min_{b \in \mathbb{R}^n} \left(tL\left(\frac{x-z-b}{t}\right) + g(b) \right) \\
&\qquad - 2tL\left(\frac{x-y}{t}\right) - 2g(y) \\
&\qquad \leq g(z+y) + g(-z+y) - 2g(y) \leq C|z|^2.
\end{aligned}$$

$\square$

Strong convexity of H also guarantees semiconcavity of u. To prove this, we need a lemma that states how the Lagrangian is affected by the strong convexity of the Hamiltonian.

Lemma 5.1.2. *Let H be strongly convex with constant θ and let $q_1, q_2 \in \mathbb{R}^n$ be arbitrary. Then*

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

Proof. By the definition 2.3.8 of strong convexity, where we set $\lambda = 1/2$, we have

$$H\left(\frac{p_1+p_2}{2}\right) \leq \frac{1}{2}\left(H(p_1) + H(p_2)\right) - \frac{\theta}{8}|p_1 - p_2|^2.$$

By Theorem 4.1.4 and Lemma 4.1.2 there exist p_1 and p_2 such that for $j \in \{1,2\}$

$$L(q_j) + H(p_j) = q_j \cdot p_j.$$

Let us estimate

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

Therefore to prove the lemma it is sufficient to show that

$$\frac{1}{4}((p_1 \cdot q_1 + p_2 \cdot q_2) - (p_1 \cdot q_2 + p_2 \cdot q_1)) - \frac{\theta}{8}|p_1 - p_2|^2 \leq \frac{1}{8\theta}|q_1 - q_2|^2,$$

which is equivalent to

$$0 \leq \frac{1}{8\theta}|q_1 - q_2|^2 - \frac{1}{4}((p_1 \cdot q_1 + p_2 \cdot q_2) - (p_1 \cdot q_2 + p_2 \cdot q_1)) + \frac{\theta}{8}|p_1 - p_2|^2.$$

The above inequality holds, since

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

by Cauchy-Schwarz inequality 2.1.1. This completes the proof of the lemma. $\square$

Lemma 5.1.3. *For stronly convex H with constant θ, and for u defined by the Hopf-Lax formula* (4.2.4)*, we have, for all $t > 0$, that the mapping $x \mapsto u(x,t)$ is semiconcave with constant $1/t\theta$.*

Proof. By the Lemma 5.1.2 we have

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

which, together with Lemma 4.1.2, implies

$$
\begin{aligned}
u(x+z,t)+u(x-z,t)-2u(x,t) = &\min_{a\in\mathbb{R}^n}\left(tL\left(\frac{x+z-a}{t}\right)+g(a)\right)+\\
&\min_{b\in\mathbb{R}^n}\left(tL\left(\frac{x-z-b}{t}\right)+g(b)\right)-2tL\left(\frac{x-y}{t}\right)-2g(y)\\
\leq &t\left(L\left(\frac{x+z-y}{t}\right)+L\left(\frac{x-z-y}{t}\right)-2L\left(\frac{x-y}{t}\right)\right)\\
\leq &2t\frac{1}{8\theta}\left|\frac{2z}{t}\right|^2=\frac{1}{t\theta}|z|^2,
\end{aligned}
$$

where y minimises (4.2.4) for $u(x,t)$. □

5.2 Weak solutions

The semiconcavity of u is of interest because it is sufficient to guarantee uniqueness of solutions to the initial value problem (1.0.1)–(1.0.2), under extra conditions. Let us define the weak solutions to the initial value problem and then show the uniqueness.

Definition 5.2.1. *We say a Lipschitz-continuous function $u\colon \mathbb{R}^n\times[0,\infty[\to\mathbb{R}$ is a weak solution of the problem* (1.0.1)–(1.0.2) *if and only if all the following hold:*

1. *for all $x\in\mathbb{R}^n$ we have $u(x,0)=g(x)$;*
2. *for almost all $(x,t)\in\mathbb{R}^n\times]0,\infty[$ we have $u_t(x,t)+H(Du(x,t))=0$;*
3. *there is a constant $C\geq 0$ so that for all $x,z\in\mathbb{R}^n$ and for all positive t it holds that*

$$u(x+z,t)-2u(x,t)+u(x-z,t)\leq C\left(1+\frac{1}{t}\right)|z|^2.$$

To show uniqueness we use mollifications of the solutions. Let u^ε be the standard mollification of u in $n+1$ dimensions, see definition 2.4.3. We first prove the following lemma:

Lemma 5.2.2. *Let I be the $n\times n$ identity matrix and u a weak solution. Then there is $C\in\mathbb{R}$ so that for all $x\in\mathbb{R}^n$, $\varepsilon>0$ and $s>2\varepsilon$*

$$D^2u^\varepsilon(x,s)\leq C\left(1+\frac{1}{s}\right)I.$$

Proof. By semiconcavity of u and Lemma 2.3.11, we know that for all $t>0$ the mapping

$$x\mapsto u(x,t)-\frac{C}{2}\left(1+\frac{1}{t}\right)|x|^2$$

is concave. By lemmata 2.3.7 and 2.4.7 we have

$$
\begin{aligned}
0 \leq & D^2\left(\left(\frac{C}{2}\left(1+\frac{1}{s}\right)|x|^2-u(x,s)\right)*\eta_\varepsilon\right)\\
= & D^2\int_{B(0,\varepsilon)}\left(\left(1+\frac{1}{s}\right)\frac{C}{2}|x-z|^2-u(x-z,s-h)\right)\eta_\varepsilon(z,h)\,d(z,h)\\
= & D^2\left(\int_{B(0,\varepsilon)}\left(1+\frac{1}{s}\right)\frac{C}{2}|x-z|^2\eta_\varepsilon(z,h)\,d(z,h)\right.\\
& \left.-\int_{B(0,\varepsilon)}u(x-z,s-h)\eta_\varepsilon(z,h)\,d(z,h)\right)\\
= & \left(C\left(1+\frac{1}{s}\right)D^2\int_{B(0,\varepsilon)}\frac{1}{2}|x-z|^2\eta_\varepsilon(z,h)\,d(z,h)-D^2u^\varepsilon(x,s)\right),
\end{aligned}
$$

from which we obtain by Lemma 2.4.5

$$
\begin{aligned}
0 \leq & C\left(1+\frac{1}{s}\right)\int_{B(0,\varepsilon)}D^2\left(\frac{1}{2}|x-z|^2\right)\eta_\varepsilon(z,h)\,d(z,h)-D^2u^\varepsilon(x,s)\\
= & C\left(1+\frac{1}{s}\right)\int_{B(0,\varepsilon)}I\eta_\varepsilon(z,h)\,d(z,h)-D^2u^\varepsilon(x,s)\\
= & C\left(1+\frac{1}{s}\right)I-D^2u^\varepsilon(x,s).
\end{aligned}
$$

□

5.3 Uniqueness

Now we prove a uniqueness result for the weak solutions of the Hamilton-Jacobi equation.

Theorem 5.3.1 (Uniqueness of weak solutions)**.** *Suppose the initial condition g is Lipschitz and suppose the Hamiltonian $H \in C^2(\mathbb{R}^n)$ is convex and superlinear. Then there is at most one weak solution to the initial value problem* (1.0.1)–(1.0.2).

Proof. Let u and $\tilde{u}$ be weak solutions and $(y,s) \in \mathbb{R}^n\times]0,\infty[$ be a point where both u and $\tilde{u}$ are differentiable. We write the difference as $w = u - \tilde{u}$. Then by the fundamental theorem of calculus

$$
\begin{aligned}
\partial_t w(y,s) = \partial_t u(y,s) - \partial_t\tilde{u}(y,s) &= -H(Du(y,s)) + H(D\tilde{u}(y,s))\\
&= -\int_0^1 \frac{d}{dr}H(rDu(y,s)+(1-r)D\tilde{u}(y,s))\,dr\\
&= -\int_0^1 DH(rDu(y,s)+(1-r)D\tilde{u}(y,s))\cdot(Du(y,s)-D\tilde{u}(y,s))\,dr\\
&= -\int_0^1 DH(rDu(y,s)+(1-r)D\tilde{u}(y,s))\cdot Dw(y,s)\,dr.
\end{aligned}
$$

For all $j \in \{1,\dots,n\}$ we define

$$
b_j(y,s) = \int_0^1 \partial_j H(rDu(y,s)+(1-r)D\tilde{u}(y,s))\,dr \tag{5.3.1}
$$

and $b = (b_1, \ldots, b_n)\colon \mathbb{R}^n \times [0,\infty[\to \mathbb{R}^n$. Thus

$$w_t(y,s) = b(y,s) \cdot Dw(y,s),$$

which implies that almost everywhere

$$\partial_t w + b \cdot Dw = 0. \tag{5.3.2}$$

Let $\phi\colon \mathbb{R} \to [0,\infty[$ be a smooth function and define $v\colon \mathbb{R}^n \times [0,\infty[\to [0,\infty[$ as $v = \phi(w)$. Multiplying equation (5.3.2) by $\phi'(w)$ gives

$$0 = \phi'(w)(\partial_t w + b \cdot Dw) = \partial_t v + b \cdot Dv, \tag{5.3.3}$$

which holds almost everywhere.

We define $b_\varepsilon\colon \mathbb{R}^n \times [0,\infty[\to \mathbb{R}^n$ by

$$(b_\varepsilon)_j(y,s) = \int_0^1 \partial_j H(rDu^\varepsilon(y,s) + (1-r)D\tilde{u}^\varepsilon(y,s))\,dr$$

for all $j \in \{1,\ldots,n\}$. Adding $b_\varepsilon \cdot Dv$ to (5.3.3), we have

$$\partial_t v + b_\varepsilon \cdot Dv = (b_\varepsilon - b) \cdot Dv,$$

which holds almost everywhere. By Lemma 2.2.2 we have

$$\partial_t v + \operatorname{div}(v b_\varepsilon) = v \operatorname{div} b_\varepsilon + (b_\varepsilon - b) \cdot Dv. \tag{5.3.4}$$

Since $H \in C^2(\mathbb{R}^n)$, we have

$$\begin{aligned}
\operatorname{div} b_\varepsilon &= \sum_{j=1}^n \partial_j (b_\varepsilon)_j = \sum_{j=1}^n \partial_j \int_0^1 (\partial_j H)(rDu^\varepsilon + (1-r)D\tilde{u}^\varepsilon)\,dr \qquad (5.3.5)\\
&= \int_0^1 \sum_{j=1}^n \partial_j \left((\partial_j H)(rDu^\varepsilon + (1-r)D\tilde{u}^\varepsilon)\right)\,dr\\
&= \int_0^1 \sum_{j=1}^n (D\partial_j H)(rDu^\varepsilon + (1-r)D\tilde{u}^\varepsilon) \cdot \partial_j(rDu^\varepsilon + (1-r)D\tilde{u}^\varepsilon)\,dr\\
&= \int_0^1 \sum_{j=1}^n \sum_{k=1}^n (\partial_j\partial_k H)(rDu^\varepsilon + (1-r)D\tilde{u}^\varepsilon)(r\partial_j\partial_k u^\varepsilon + (1-r)\partial_j\partial_k \tilde{u}^\varepsilon)\,dr\\
&\le \int_0^1 \sum_{j=1}^n \sum_{k=1}^n C_0 C_1 \left(1 + \frac{1}{s}\right)\,dr \le C\left(1 + \frac{1}{s}\right)
\end{aligned}$$

by Lemma 2.1.3, with $C_0 = \max_{B(0,C_u)} |D^2 H|$, $C_u = \operatorname{Lip}(u)$, C_1 is the constant C in Lemma 5.2.2, and $s > 2\varepsilon$. Here we applied Lemma 2.4.6.

Next, let $x_0 \in \mathbb{R}^n$, $t_0 > 0$, C_u the Lipschitz constant of u and $C_{\tilde{u}}$ the Lipschitz constant of $\tilde{u}$, and define the real number

$$R = \max\left\{|DH(p)| \;\middle|\; p \in \mathbb{R}^n \text{ and } |p| \le \max\{C_u, C_{\tilde{u}}\}\right\}$$

and the ball

$$B = B(x_0, R(t_0 - t)).$$

Further, we set the function $f\colon [0,t_0[\to \mathbb{R}$ by

$$f(t) = \int_B v(x,t)\,dx$$

and split its derivative into two parts:

$$\dot{f}(t) = \int_B \partial_t v(x,t)\,dx - R\int_{\partial B} v(x,t)\,dS.$$

The proof of the above equality is easy. Now we go back to equation (5.3.4). By divergence theorem, Theorem 2.2.3,

$$\begin{aligned}
\dot{f}(t) &= \int_B \partial_t v(x,t)\,dx - R\int_{\partial B} v(x,t)\,dS\\
&= \int_B v\operatorname{div} b_\varepsilon + (b_\varepsilon - b)\cdot Dv - \operatorname{div}(vb_\varepsilon)\,dx - R\int_{\partial B} v\,dS\\
&= \int_B v\operatorname{div} b_\varepsilon + (b_\varepsilon - b)\cdot Dv\,dx - \left(\int_B \operatorname{div}(vb_\varepsilon)\,dx + \int_{\partial B} Rv\,dS\right)\\
&= \int_B v\operatorname{div} b_\varepsilon + (b_\varepsilon - b)\cdot Dv\,dx - \left(\int_{\partial B} vb_\varepsilon\cdot\nu\,dS + \int_{\partial B} Rv\,dS\right)\\
&= \int_B v\operatorname{div} b_\varepsilon + (b_\varepsilon - b)\cdot Dv\,dx - \int_{\partial B} (b_\varepsilon\cdot\nu + R)v\,dS.
\end{aligned}$$

By definition of R and b_ε, and by Cauchy-Schwarz inequality (Theorem 2.1.1), we have $b_\varepsilon\cdot\nu + R \geq 0$, and v is positive by definition, so by the previous calculation and equation (5.3.5) for $2\varepsilon < t$ we have the estimate

$$\begin{aligned}
\dot{f} &\leq \int_B v\operatorname{div} b_\varepsilon + (b_\varepsilon - b)\cdot Dv\,dx \leq C\left(1+\frac{1}{t}\right)\int_B v\,dx + \int_B (b_\varepsilon - b)\cdot Dv\,dx\\
&= C\left(1+\frac{1}{t}\right) f + \int_B (b_\varepsilon - b)\cdot Dv\,dx.
\end{aligned}$$

We will show that the integral in the above inequality vanishes as ϵ does. By definitions of b and b_ε and the regularity assumptions we know that $b_\varepsilon \to b$ as $\epsilon \to 0$. To guarantee that the integral vanishes we use the dominated convergence theorem, Theorem 2.2.1. Since u and $\tilde{u}$ are Lipschitz, Lemmata 2.4.6 and 2.2.4 with the continuity of the derivate of H imply that b and b_ε are bounded. Since we integrate (x,t) over a compact set, the gradient of v is bounded, since ϕ is continuously differentiable and w is Lipschitz, and $v = \phi(w)$. So we have by Cauchy-Schwarz inequality, Theorem 2.1.1, that our integrand is bounded, and therefore for almost every $t \in]0,t_0[$

$$\dot{f}(t) \leq C\left(1+\frac{1}{t}\right) f(t).$$

Let ε be positive but less than t. So far, the function ϕ had been arbitrary. Now we set ϕ to be 0 when its argument is at most $\varepsilon(\operatorname{Lip}(u) + \operatorname{Lip}(\tilde{u}))$, and positive elsewhere. Since for all $z \in \mathbb{R}^n$ $u(z,0) = g(z) = \tilde{u}(z,0)$, we have

$$v(z,\varepsilon) = \phi(w(z,\varepsilon)) = \phi(u(z,\varepsilon) - \tilde{u}(z,\varepsilon)) = 0$$

because

$$\begin{aligned}|u(z,\varepsilon)-\tilde{u}(z,\varepsilon)| &\leq |u(z,\varepsilon)-u(z,0)|+|u(z,0)-\tilde{u}(z,0)|+|\tilde{u}(z,0)-\tilde{u}(z,\varepsilon)| \\ &= |u(z,\varepsilon)-u(z,0)|+|\tilde{u}(z,0)-\tilde{u}(z,\varepsilon)| \\ &\leq \varepsilon(\mathrm{Lip}(u)+\mathrm{Lip}(\tilde{u})).\end{aligned}$$

Now by definition of f we have $f(\varepsilon)=0$ and by Grönwall's inequality, Theorem 2.2.6, for all $r\in]\varepsilon,t[$ it holds that

$$f(r)\leq \exp\left(\int_{\varepsilon}^{r} C\left(1+\frac{1}{s}\right)\,ds\right) f(\varepsilon)=0,$$

which by our selection of ϕ indicates that for all x in B

$$|u(x,r)-\tilde{u}(x,r)|\leq \varepsilon(\mathrm{Lip}(u)+\mathrm{Lip}(\tilde{u})),$$

which implies that u and $\tilde{u}$ are identical on the set $B\times\{r\}$. Since $t_0>t>r$, we can't directly reach the point (x_0,t_0), but we do find r arbitrarily close to t_0 and $u-\tilde{u}$ is continuous, so we have $u(x_0,t_0)=\tilde{u}(x_0,t_0)$. $\square$

Combining Theorems 5.3.1, 5.1.1 and 5.1.3, we obtain the following theorem.

Theorem 5.3.2 (Hopf-Lax formula as weak solution)**.** *Let $H\colon\mathbb{R}^n\to\mathbb{R}$ be twice continuously differentiable, superlinear and convex, and let $g\colon\mathbb{R}^n\to\mathbb{R}$ be Lipschitz continuous. Suppose also that g is semiconcave or H strongly convex. Then $u\colon\mathbb{R}^n\times[0,\infty[\to\mathbb{R}$,*

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

for positive t and $u(x,0)=g(x)$ for the initial condition, is the unique weak solution of the initial value problem (1.0.1)–(1.0.2) *for the Hamilton-Jacobi equation.*

Example 4.2.2 provides the unique weak solution to initial value problem for the Hamilton-Jacobi equation with the initial condition $g(x)=|x|$ and the Hamiltonian $H(x)=|x|^2$.

Example 5.3.3. *By Example 2.3.9 the Hamiltonian $x\mapsto|x|^2$ is strongly convex, so by Theorem 5.3.2, Example 4.2.2 provides a unique weak solution. In particular, the solution is semiconcave for all positive t with respect to variable x with constant $C(1+1/t)$, where C does not depend on t.*

6 Generalisations

The current work could be generalised to several directions. Another perspective on Legendre transformation can be found in e.g. [17], sections 12 and 26, and [9].

Hopf-Lax formula and thereby existence of solutions can be extended to Hamiltonian that not only depends on the derivative, but also the function u itself; see e.g. [4].

Hamilton-Jacobi equation admits a more general notion of solution, that is, viscosity solution, which applies to more general Hamilton-Jacobi-Bellman equation by Theorem 2 in chapter 10.3 of [8]. It was first developed by Crandall and Lions in [7]. We refer to [3] for the connections between optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Generalisations to other directions and several applications can be found; see e.g. [2].

We refer to [6] for detailed discussions on the semiconcave functions, their generalisations and also on their applications.

References

[1] V. I. Arnold. *Mathematical Methods of Classical Mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1978.

[2] V. Barbu and G. Da Prato. *Hamilton-Jacobi equations in Hilbert spaces*, volume 86 of *Research Notes in Mathematics*. Pitman Advanced Publishing Program, 1983.

[3] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Systems & Control: Foundations & Applications. Birkhäuser, Boston, 1997.

[4] E. N. Barron, R. Jensen, and W. Liu. Hopf-Lax-type formula for $u_t+H(u, Du) = 0$. *Journal of differential equations*, 126(1):48–61, Mar 1996.

[5] Stephen P. Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, Cambridge, UK, 2004.

[6] Piermarco Cannarsa and Carlo Sinestrani. *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, volume 58 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, New York, USA, 2004.

[7] Michael G. Crandall and Pierre-Louis Lions. Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 277(1):1–42, May 1983.

[8] Lawrence C. Evans. *Partial Differential Equations*. Number 19 in Graduate Studies in Mathematics. American Mathematical Society, Providence, Rhode Island, 1. edition, 1998. Errata at `http://math.berkeley.edu/~evans/`.

[9] W. Fenchel. On conjugate convex functions. *Canadian journal of mathematics*, 1(1):73–77, 1949.

[10] Thomas Hakon Grönwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *The Annals of Mathematics*, 20(4):pp. 292–296, July 1919.

[11] Eberhard Hopf. The partial differential equation $u_t + u_{xx} = \mu u_{xx}$. *Communications on pure and applied mathematics*, 3(3):201–230, September 1950.

[12] Jaakko Hyvönen. Classical potential theory. In *Summer school in potential theory [14]*. University of Joensuu, department of mathematics and physics, Joensuu, 1983.

[13] Johan L. W. V. Jensen. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30:175–193, January 1906. 10.1007/BF02418571.

[14] Ilpo Laine and Olli Martio, editors. *Summer school in potential theory*. Number 5-M in University of Joensuu, department of mathematics and physics report series. University of Joensuu, department of mathematics and physics, Joensuu, 1983.

[15] Peter D. Lax. Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Communications on pure and applied mathematics*, 7(1):159–193, February 1954. Also in book [16].

[16] Peter D. Lax. *Selected papers*, volume 1. Springer, New York, USA, 2005.

[17] R. Tyrrell Rockafellar. *Convex analysis*. Number 28 in Princeton mathematical series. Princeton university press, Princeton, New Jersey, 1970.

[18] Robert L. Warnock. Hamilton-Jacobi equation. *Scholarpedia*, 5(7):8330, 2010.