Viscosity solution of the Hamilton–Jacobi equation by a limiting minimax method

Qiaoling Wei

Université Paris 7, UFR de Mathématiques, Bâtiment Sophie Germain 5, rue Thomas Mann, 75205 Paris Cedex 13, France

E-mail: weiqiaoling@math.jussieu.fr

Received 18 July 2013, revised 3 November 2013
Accepted for publication 15 November 2013
Published 16 December 2013

Abstract

For non-convex Hamiltonians, the viscosity solution and the more geometric minimax solution of the Hamilton–Jacobi equation do not coincide in general. They are nevertheless related: we show that iterating the minimax procedure during shorter and shorter time intervals one recovers the viscosity solution.

Mathematics Subject Classification: 35D40, 35F21, 37J05, 49L25

(Some figures may appear in colour only in the online journal)

1. Introduction

Consider the Cauchy problem for the Hamilton–Jacobi equation

\[
\begin{align*}
\partial_t u(t, x) + H(t, x, \partial_x u) &= 0, \\
\partial_t u(0, x) &= v(x).
\end{align*}
\]

By the classical characteristics method (section 3.1), one can to some extent solve this problem when \(v\) is \(C^2\): solving Hamilton’s equations \(\dot{x} = \partial_p H(t, x, p)\) and \(\dot{p} = -\partial_x H(t, x, p)\), one gets the characteristic lines \((t, x(t), p(t))\), from which the solution \(u(t, x)\) of the Cauchy problem can be obtained as follows: setting \(u_t(x) = u(t, x)\), the graph of \(d u_t\) is the section at time \(t\) of the union \(L\) of the characteristic lines passing through \(\{(0, d v(x))\}\), whereas \(\partial_t u\) is provided by the equation.

When \(H\) is not linear with respect to \(p\), this procedure does not yield a global solution of the problem in general, as the geometric solution \(L\) is not the graph of a function \(p(t, x)\). In other words the wavefront \(F\), obtained in \((t, x, u)\) space by solving the equation \(d u = -H(t, x, p) d t + p d x\) restricted to \((t, x, p) \in L\) —and taking into account the values of \(v\) and not only \(d v\)—is not the graph of a function (which otherwise would be a solution of the problem): the projected characteristics \((t, x(t))\) may have a variation of slopes and cross after some time.
In some applications, e.g. to geometrical optics, the wavefront $\mathcal{F}$ provides a solution of the physical problem at hand. In other cases, however, one does look for a single-valued solution $u(t, x)$, which can not be $C^2$ since, otherwise, its graph would be $F$. Therefore, some new idea is needed to obtain a weak solution in the sense of distributions.

The most natural idea, assuming that the projection of $F$ into $(t, x)$ space is onto, is to get such a solution as a section of the wavefront, obtained by selecting a single $u$ over each $(t, x)$. When the Hamiltonian is convex enough with respect to $p$ (and $v$ is not too wild at infinity), such a 'graph selector' consists in choosing the lowest point of $\mathcal{F}$ above each $(t, x)$, i.e. taking for $u(t, x)$ the smallest $u$ with $(t, x, u) \in \mathcal{F}$. Analytically, this is expressed by formulæ due to Hopf–Lax [Hop65] for Hamiltonians of the form $H(p)$ and to Lax–Oleinik in the general case [Fat]; these formulæ, which extend to non-differentiable initial data $v$, are widely used in weak KAM theory; they imply that the min selector provides a Lipschitz weak solution of the problem.

This min solution turns out to be the viscosity solution, which owes its name to the fact that it was first introduced as the limit when $\varepsilon \to 0^+$ of the solution of the same Cauchy problem for the equation $\partial_\tau u(t, x) + H(t, x, \partial_x u) = \varepsilon \Delta_x u(t, x)$; the much more tractable general definition\(^1\) was given in 1983 by Crandall, Evans and Lions for general nonlinear first order partial differential equations [Lio82, CEL84].

In the non-convex case, it has been known for quite a long time that the viscosity solution, when it exists, may not be a section of the wavefront (see for example [Che74]). On the other hand, following a remark of Sikorav, Chaperon introduced in 1991 [Cha91] weak solutions whose graph is a section of the wavefront, obtained by a 'minmax' procedure which generalizes the minimum considered in the convex case and relies on the existence of suitable generating families for the geometric solution—in addition to [Che74], see [Vit96, VO95, BC11] for examples showing that minmax and viscosity solutions may differ.

The aim of this paper is to relate such minmax solutions to the viscosity solutions, which possess very general existence, uniqueness and stability properties in a large class of ‘good’ cases but are not so easy to understand geometrically in general, one of the motivations for this work.

Among the properties of viscosity solutions not shared by minmax solutions—see appendix B—is the ‘semi-group’ (Markov) property. Hence, it is natural to try and get this semi-group as a limit of (discrete) semi-groups obtained by dividing a given time interval into small pieces and taking the minimax step by step (‘iterated minimax procedure’). Our main purpose is to show that when the small time intervals go to zero, one gets a limit solution which is indeed the viscosity solution (see theorem 3.19). This answers a question of Chaperon.

2. Generating families and minmax selector

Consider the cotangent bundle $T^* M$ of a connected manifold $M$ of dimension $m$, endowed with its canonical symplectic form $\omega = dp \wedge dq$. A Lagrangian submanifold $L$ is a submanifold of dimension $m$ such that $\omega$ vanishes on $L$.

**Definition 2.1.** A generating family for a Lagrangian submanifold $L \subset T^* M$ is a $C^2$ function $S : M \times \mathbb{R}^k \to \mathbb{R}$ such that $0$ is a regular value of the map $(x, \eta) \mapsto \partial S(x, \eta)/\partial \eta$ and

$$L = \left\{ \left( x, \frac{\partial S}{\partial x} (x, \eta) \right) : \frac{\partial S}{\partial \eta} = 0 \right\};$$

\(^1\) For conservation laws, more particular characterizations called entropy conditions had been obtained by Oleinik in dimension one [Lax73] and by Kruzkov in arbitrary dimensions [Kru70].
more precisely, the condition that 0 is a regular value implies that the critical locus \( \Sigma_S := \{(x, \eta) | \partial_x S = 0\} \) is a submanifold and that the map
\[
i_S : \Sigma_S \to T^* M, \quad (x, \eta) \mapsto (x, \partial_x S(x, \eta))
\]
is an immersion; we require that \( i_S \) be an embedding and, of course, \( i_S(\Sigma_S) = L \).

A function \( S \) on \( M \times \mathbb{R}^k \) need not have critical points. However, it does have critical points if we prescribe some behaviour at infinity as in the following definition:

**Definition 2.2.** A generating family \( S : M \times \mathbb{R}^k \to \mathbb{R} \) of \( L \) is quadratic at infinity if there exists a non-degenerate quadratic form \( Q \) such that, for any compact subset \( K \subset M \), the differential \( |\partial_x (S(x, \eta) - Q(\eta))| \) is bounded on \( K \times \mathbb{R}^k \).

We will abbreviate ‘generating family quadratic at infinity’ as GFQI. Remark that, the \( C^2 \) GFQI such defined can be made into ‘exactly quadratic at infinity’: for every compact subset \( K \subset M \), \( (S - Q)|_{K \times \mathbb{R}^k} \) has compact support modulo a fibrewise diffeomorphism. See for example [Thé96], or [Wei13] proposition 1.21.

Consider the sub-level sets \( S_a^x := \{ \eta : S(x, \eta) \leq a \} \), the homotopy type of \( (S_a^x, S^{-a}^x) \) does not depend on \( a \) and coincides with the homotopy type of \( (Q^a, Q^{-a}) \) when \( a \) is large enough, we may write it as \( (S_\infty^x, S^{-\infty}_x) \). If the Morse index of \( Q \) is \( k_\infty \), then
\[
H_i(S_\infty^x, S^{-\infty}_x; \mathbb{Z}_2) = H_i(Q_\infty, Q^{-\infty}; \mathbb{Z}_2) \simeq \begin{cases} \mathbb{Z}_2, & i = k_\infty, \\ 0, & \text{otherwise} \end{cases}
\]

**Definition 2.3.** The minmax function is defined as
\[
R_S(x) := \inf_{[\sigma] \mapsto A} \max_{\eta \in [\sigma]} S(x, \eta),
\]
where \( A \) is a generator of the homology group \( H_{k_\infty}(S_\infty^x, S^{-\infty}_x; \mathbb{Z}_2) \) and \([\sigma]\) denotes the image of the relative singular homology cycle \( \sigma \). Such a cycle \( \sigma \) will be called a descending cycle.

The function \( R_S \) is determined (up to the addition of a constant) by \( L \) and does not depend on the particular choice of its GFQI \( S \); indeed, by a theorem of Viterbo and Théret [Thé99], for a given Lagrangian submanifold \( L \), \( S \) is unique up to the following three operations.

(a) Fibrewise diffeomorphism: \( \hat{S}(x, \eta) := S(x, \varphi(x, \eta)) \), where \( (x, \eta) \mapsto (x, \varphi(x, \eta)) \) is a fibrewise diffeomorphism.
(b) Addition of a constant: \( \tilde{S}(x, \eta) := S(x, \eta) + C \).
(c) Stabilization: \( \tilde{S}(x, \eta, \xi) := S(x, \eta) + q(\xi) \), where \( q \) is a non-degenerate quadratic form.

It is well known that the minmax \( R_S(x) \) is a critical value of \( S(x, \cdot) \) for every \( x \), a property which we shall establish in a more general Lipschitz framework (proposition 2.22).

In the rest of the paper, we will take the manifold \( M \) to be \( \mathbb{R}^d \), in which case the generating families are constructed explicitly. The case where \( M = \mathbb{T}^d \) can be treated via the same construction. For a general manifold, one can embed it into some \( \mathbb{R}^d \) and use the trick of Chekanov [Che96, Bru91] to get generating families from those in \( \mathbb{R}^d \).

The contents of this section are organized as follows: first, we present an explicit formula for generating families, due to Chaperon and Sikorav; then the notion of ‘quadratic at infinity’ is discussed; finally, we will arrive at a generalization of the GFQIs to Lipschitz cases.

\footnote{Hence, the graph of \( R_S \) is a section of the ‘wavefront’ obtained in \((x, u)\)-space (up to vertical translation) by solving the equation \( du = p dx \) restricted to \( L \).}
2.1. Construction of generating functions and families

In the following, we equip \( \mathbb{R}^k \) with the Euclidean \( \ell^2 \) norm \( \| \cdot \| \), and matrices in \( \mathbb{R}^k \) with the associated operator norm. We denote by \( \text{Lip}(f) \) the Lipschitz constant of a function \( f \) and by \( \pi : T^*\mathbb{R}^d \to \mathbb{R}^d \) the canonical projection \( \pi(x, y) = x \).

We denote by \( H : [0, T] \times T^*\mathbb{R}^d \to \mathbb{R} \) a \( C^2 \) Hamiltonian satisfying

\[
\c_H := \sup \{ D^2 H(x, y) | < \infty \}
\]

and by \( X_H \), the associated time-depending Hamiltonian vector field\(^3\). By the general theory of differential equations, as \( c_H = \max \text{Lip}(DH) = \max \text{Lip}(X_H) \), the Hamiltonian transformation \( \psi^{H,t}_s \) obtained by integrating \( X_H \) from \( s = t \) to \( s = t \) is a well-defined diffeomorphism for all \((s, t) \in [0, T] \). For simplicity, we sometimes write \( \psi^{H,t}_s = (X^t_s, Y^t_s) := \psi^{H,t}_s \) without mentioning \( H \).

We will be mostly interested in the special case where \( H \) has compact support, and consider the Lagrangian submanifolds of \( T^*\mathbb{R}^d \) which are Hamiltonianly isotopic to the zero section:

\[
\mathcal{L} := \{ L = \varphi(dv), \quad v \in C^2 \cap C^{\text{Lip}}(\mathbb{R}^d), \varphi \in \text{Ham}_c(T^*\mathbb{R}^d) \};
\]

here \( C^{\text{Lip}}(\mathbb{R}^d) \) denotes the space of globally Lipschitz functions and

\[
\text{Ham}_c(T^*\mathbb{R}^d) = \{ \varphi = \varphi_H, H \in C^2([0, 1] \times T^*\mathbb{R}^d) \}
\]

where \( \varphi_H = \psi^{0,1}_H \) is the endpoint of the isotopy (‘Hamiltonian flow’ defined by \( H \).)

**Definition 2.4.** A diffeomorphism \( \varphi : T^*\mathbb{R}^d \to T^*\mathbb{R}^d \) admits a *generating function* \( \phi \), if \( \phi : T^*\mathbb{R}^d \to \mathbb{R} \) is of class \( C^2 \), such that \((x, y), (X, Y) \in \text{Graph}(\varphi)\) if and only if

\[
x = X + \partial_x \phi(x, y),
\]

\[
Y = y + \partial_y \phi(x, y).
\]

This definition can be interpreted as follows: the isomorphism

\[
I : T^*\mathbb{R}^d \times T^*\mathbb{R}^d \to T^*(T^*\mathbb{R}^d),
(x, y, X, Y) \mapsto (X, y, Y - y, x - X)
\]

is symplectic if \( T^*\mathbb{R}^d \) is equipped with the standard symplectic form \( \omega = dx \wedge dy \) and \( T^*\mathbb{R}^d \times T^*\mathbb{R}^d \) with the symplectic form \( -\omega \otimes \omega = dX \wedge dY \wedge dx \wedge dy \); this symplectic isomorphism \( I \) maps the diagonal of the space \( T^*\mathbb{R}^d \times T^*\mathbb{R}^d \) to the zero section of the cotangent space \( T^*(T^*\mathbb{R}^d) \) and \( \text{Graph}(\varphi) \) to \( \text{Graph}(\phi) \).

Hence, if it exists, the generating function \( \phi \) is unique up to the addition of a constant.

**Lemma 2.5 ([Cha90]).** If \( \delta_H := c_H^{-1} \ln 2 \), then, for \( |s - t| < \delta_H \), the map

\[
\alpha^{H,t}_s : (x, y) \mapsto (X^t_s(x, y), y)
\]

is a diffeomorphism. As a consequence, \( \varphi^{H,t}_s \) admits the generating function

\[
\phi^{H,t}_s(X, y) = \int_{s}^{t} \left( (Y^t_s - y)X^t_s - H(\tau, X^t_s, Y^t_s) \right) d\tau
\]

where \((X^t_s(X, y), Y^t_s(X, y)) = \varphi^{H,t}_s \circ (\alpha^{H,t}_s)^{-1}(X, y) \) and the dot denotes the derivative with respect to \( \tau \).

\(^3\) We use the convention of sign that \( X_H = (\partial_p H, -\partial_q H) \).
Lemma 2.6. For the generating function $\phi'_t$ defined by (2.2), we have
\[ \frac{\partial}{\partial t} \phi'_t(X, y) = H(s, x, y), \quad \frac{\partial}{\partial t} \phi'_t(X, y) = -H(t, X, Y). \]

Proof. Differentiating (2.2) on both sides, we have
\[
\frac{\partial}{\partial t} \phi'_t(X, y) = H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + (Y^t_s - y) \frac{d}{d\tau} Y^t_s \frac{d}{ds} X^t_s - \frac{d}{d\tau} X^t_s \frac{d}{ds} Y^t_s \right) d\tau.
\]
\[
= H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + Y^t_s \frac{d}{d\tau} X^t_s \right) d\tau - \int_t^s Y^t_s \frac{d}{d\tau} X^t_s d\tau
\]
\[
= H(s, x, y)
\]
where we have used $\frac{\partial}{\partial t} H_t(X_t, Y_t) = -\frac{\partial}{\partial t} H(t, X, Y)$.

Proof. Differentiating (2.2) on both sides, we have
\[
\frac{\partial}{\partial t} \phi'_t(X, y) = H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + (Y^t_s - y) \frac{d}{d\tau} Y^t_s \frac{d}{ds} X^t_s - \frac{d}{d\tau} X^t_s \frac{d}{ds} Y^t_s \right) d\tau
\]
\[
= H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + Y^t_s \frac{d}{d\tau} X^t_s \right) d\tau - \int_t^s Y^t_s \frac{d}{d\tau} X^t_s d\tau
\]
\[
= H(s, x, y)
\]
where we have used $\frac{\partial}{\partial t} H_t(X_t, Y_t) = -\frac{\partial}{\partial t} H(t, X, Y)$.

Proof. Differentiating (2.2) on both sides, we have
\[
\frac{\partial}{\partial t} \phi'_t(X, y) = H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + (Y^t_s - y) \frac{d}{d\tau} Y^t_s \frac{d}{ds} X^t_s - \frac{d}{d\tau} X^t_s \frac{d}{ds} Y^t_s \right) d\tau
\]
\[
= H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + Y^t_s \frac{d}{d\tau} X^t_s \right) d\tau - \int_t^s Y^t_s \frac{d}{d\tau} X^t_s d\tau
\]
\[
= H(s, x, y)
\]
where we have used $\frac{\partial}{\partial t} H_t(X_t, Y_t) = -\frac{\partial}{\partial t} H(t, X, Y)$.

Proof. Differentiating (2.2) on both sides, we have
\[
\frac{\partial}{\partial t} \phi'_t(X, y) = H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + (Y^t_s - y) \frac{d}{d\tau} Y^t_s \frac{d}{ds} X^t_s - \frac{d}{d\tau} X^t_s \frac{d}{ds} Y^t_s \right) d\tau
\]
\[
= H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + Y^t_s \frac{d}{d\tau} X^t_s \right) d\tau - \int_t^s Y^t_s \frac{d}{d\tau} X^t_s d\tau
\]
\[
= H(s, x, y)
\]
where we have used $\frac{\partial}{\partial t} H_t(X_t, Y_t) = -\frac{\partial}{\partial t} H(t, X, Y)$.

Proof. Differentiating (2.2) on both sides, we have
\[
\frac{\partial}{\partial t} \phi'_t(X, y) = H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + (Y^t_s - y) \frac{d}{d\tau} Y^t_s \frac{d}{ds} X^t_s - \frac{d}{d\tau} X^t_s \frac{d}{ds} Y^t_s \right) d\tau
\]
\[
= H(s, x, y) + \int_t^s \left( \frac{d}{ds} Y^t_s \frac{d}{d\tau} X^t_s + Y^t_s \frac{d}{d\tau} X^t_s \right) d\tau - \int_t^s Y^t_s \frac{d}{d\tau} X^t_s d\tau
\]
\[
= H(s, x, y)
\]
where we have used $\frac{\partial}{\partial t} H_t(X_t, Y_t) = -\frac{\partial}{\partial t} H(t, X, Y)$.
Proof.

(i) As the Hamiltonian flow is a ‘two-parameter groupoid’, we have that
\[ \varphi_{s,t}^H = \varphi_{t,0}^H = \varphi_{tN,0}^H \circ \cdots \circ \varphi_{0,t}^H; \]
hence, if \(|t_{i+1} - t_i| < \delta_H\) for all \(i\), it follows from the composition formula in proposition 2.7 that formula 2.4 does define a generating family for \(\varphi_{s,t}^H (dv)\);
(ii) is clear;
(iii) is proved by inspection (and very important).

\[ \square \]

2.2. Generating families quadratic at infinity

Lemma 2.9. If \(H\) has compact support, the generating families constructed in Corollary 2.8 are quadratic at infinity when the \(C^2\) function \(v\) is Lipschitzian.

Proof. Each \(\varphi_{t_i,t_{i+1}}^H\) has compact support and therefore bounded derivatives. Hence \(S(x, \eta) = \psi(x, \eta) + Q(\eta)\) is quadratic at infinity with \(Q(\eta) := -x_{N-1}y_{N-1} + \sum_{0 \leq i < N-1} (x_{i+1} - x_i) y_i\) and \(\psi(x, \eta) = v(x_0) + x y_{N-1} + \sum_{0 \leq i < N} \varphi_{t_i,t_{i+1}}^H (x_{i+1}, y_i)\) having bounded derivatives with respect to \(\eta\) for \(x\) on compact subsets of \(\mathbb{R}^d\).

Denote \(\pi\) the projection \(T^*\mathbb{R}^d \to \mathbb{R}^d\). A necessary condition for \(L\) to admit a GFQI is that, for any compact subset \(K\) of \(\mathbb{R}^d\), the intersection \(L \cap \pi^{-1}(K)\) be compact and non-empty: indeed, a function on \(\mathbb{R}^k\) whose differential, up to a bounded map, equals that of a non-degenerate quadratic form off a compact set must have critical points.

It follows that there does not always exist a GFQI for \(L = \varphi_{0,\pi/4}^H (dv)\) if \(H\) is not compactly supported, even when it satisfies (2.1) and \(v\) has as little growth at infinity as possible:

Example 2.10. If the Hamiltonian \(H \in C^2(\mathbb{R} \times T^*\mathbb{R})\) is given by \(H(t, x, y) = x^2 + y^2\) for \(|y| \leq 1\), \(\frac{\partial h}{\partial y}, \frac{\partial^2 h}{\partial y^2} \leq 1/2\) and \(h\) vanishes off a compact set of \(\mathbb{R}^2\). If \(v(x) = \frac{1}{3} x^3\), then for \(t = 1\),

\[ \pi \circ \varphi_{0,1}^H (dv) \subset \{x + x^2, |x| \leq 1\} \cup \left[\frac{1}{2}, +\infty\right] \cup \left[-\infty, -\frac{1}{2}\right]. \]

One sees that the image under the projection \(\pi\) of \(L = \varphi_{0,1}^H (dv)\) is not \(\mathbb{R}\).

It is also essential to require that \(v\) has little growth at infinity.

Example 2.11. It is essential that \(v\) be Lipschitzian: indeed, if \(d = 1\), \(H(t, x, y) = h(x, y) = \frac{1}{2} y^2\) for \(|y| \leq 1\), \(\frac{\partial h}{\partial y}, \frac{\partial^2 h}{\partial y^2} \leq 1/2\) and \(h\) vanishes off a compact set of \(\mathbb{R}^2\). If \(v(x) = \frac{1}{3} x^3\), then for \(t = 1\),

\[ \pi \circ \varphi_{0,1}^H (dv) \subset \{x + x^2, |x| \leq 1\} \cup \left[\frac{1}{2}, +\infty\right] \cup \left[-\infty, -\frac{1}{2}\right]. \]

One sees that the image under the projection \(\pi\) of \(L = \varphi_{0,1}^H (dv)\) is not \(\mathbb{R}\).

As the main ingredient in the construction of generating families is the Hamiltonian flow, what matters essentially over a given compact subset of \(\mathbb{R}^d\) is the region swept by the Hamiltonian flow; this is the idea of what is called the property of finite propagation speed in [CV08, appendix A]:

22
Proposition 2.12. Let \([s, t] \subset [0, T]\) and \(L = \psi^s_t(d\nu)\). If for any compact subset \(K \subset \mathbb{R}^d\), the set

\[
\mathcal{U}_K := \bigcup_{\tau \in [s, t]} \left\{ \tau \times \left\{ \psi^s_\tau \left( \psi^s_t(\pi^{-1}(K)) \cap d\nu \right) \right\} \right\}
\]

is non-empty and compact, then \(L\) admits GFQI’s in the sense that each \(L|_K := L \cap \pi^{-1}(K)\) has a GFQI.

Proof. For any \(K\), let \(\tilde{H} = \chi H\), where \(\chi\) is a compactly supported smooth function on \([0, T] \times T^*\mathbb{R}^d\) equal to 1 in a neighbourhood of \(\mathcal{U}_K\). Then formula (2.4) with \(H := \tilde{H}\) gives a GFQI \(S_{\tilde{H}}\) for \(L|_K := \psi^s_t(\pi^{-1}(K) \cap d\nu)\).

Remark 2.13. One can also truncate \(v\), as the effective region for \(v\) is \(\pi \left( \psi^{s,t}_H(\pi^{-1}(K)) \right)\). This may help to localize the minmax.

Condition (2.1) is not required here, provided \(H\) is \(C^2\) and such that \(\psi^{s,t}_H\) is defined for all \(s, t \in [0, T]\).

Lemma 2.14. If two families \(S\) and \(S'\) are quadratic at infinity with \(|S - S'|_{C^0} < \infty\), then the associated minmax functions satisfy

\[|R_S(x) - R_{S'}(x)| \leq |S - S'|_{C^0}.\]

Proof. If \(S \leq S'\), then by definition \(R_S(x) \leq R_{S'}(x)\). Hence, in general, the inequality \(S \leq S' + |S - S'|_{C^0}\) yields \(R_S(x) \leq R_{S'}(x) + |S - S'|_{C^0}\). We conclude by exchanging \(S\) and \(S'\).

Proposition 2.15. Under the hypotheses of proposition 2.12 and with the notation of its proof, the Lagrangian submanifold \(L\) determines a minmax function, given by

\[R(x) = \inf_{x \in K} \max_{\pi(L)} S_{\tilde{H}}(x, \eta), \quad \text{if } x \in K \subset \mathbb{R}^d\]

and independent of the truncation \(\tilde{H}\) and the subdivision of \([s, t]\) used to define \(S_{\tilde{H}}\).

Proof. Let \(\tilde{H}\) and \(H')\) be two truncations for \(H\) on \(\mathcal{U}_K\) as in the proof of proposition 2.12. Let \(H^\mu = \mu H + (1 - \mu)H'\), \(\mu \in [0, 1]\); as the constant \(c_{H_\mu}\) of (2.1) is uniformly bounded, one can find a subdivision \(s = t_0 < t_1 \cdots < t_N = t\) satisfying \(|t_i - t_{i+1}| < \delta_{H_\mu}\) for all \(\mu\) (see lemma 2.5); if \(S_\mu\) denotes the corresponding GFQI of \(L|_K = L \cap \pi^{-1}(K)\) for \(0 \leq \mu \leq 1\) then, by lemma 2.14, as \(S_\mu\) depends continuously on \(\mu\), so does the minmax \(R_{S_\mu}(x)\) for \(x \in \pi(L)\).

On the other hand, \(R_{S_\mu}(x)\) is a critical value of the map \(\eta \mapsto S_\mu(x, \eta)\), and, by (2.6), the set of all such critical values is independent of \(\mu\) and the subdivision, and depends only on \(\mathcal{U}_K\); as it has measure zero by Sard’s theorem, \(R_{S_\mu}(x)\) is constant for \(\mu \in [0, 1]\).

The fact that the critical value \(R_{S_\mu}(x)\) itself does not depend on the subdivision is established in lemma 3.4.

Example 2.16. If the base manifold is \(M = \mathbb{T}^d\), taking its universal covering \(\mathbb{R}^d\), we can consider \(v : \mathbb{R}^d \to \mathbb{R}\) a periodic function and \(H : \mathbb{R} \times T^*\mathbb{R}^d \to \mathbb{R}\) periodic in \(x\). Then in order that \(L = \psi^s_t(d\nu)\) admits a GFQI, it is enough to require that the flow \(\psi^s_t\) is well-defined for \(\tau \in [s, t]\). Indeed, since \(d\nu\) is compact, \(\bigcup_{\tau \in [s, t]} (\tau \times \psi^s_t(d\nu))\) is compact, hence the condition of finite propagation speed is satisfied automatically.
Example 2.17. The following hypotheses yield the finite propagation speed property:

\[ |\partial_t H| \leq C_H(1 + |x|), \quad |\partial_y H| \leq C_H(1 + |y|). \]

It is a classical condition for the existence and uniqueness of viscosity solutions in \( \mathbb{R}^d \), see [CL87].

2.3. Generalized generating families and minmax in the Lipschitz setting

Already if \( d = 1 \), \( H(t, x, y) = \frac{1}{2}y^2 \) and \( u(x) = \arctan x \), the Lagrangian submanifold \( \psi^0_H(du) = \{(x + t/(1 + x^2), 1/(1 + x^2)): x \in \mathbb{R}\} \) is not the graph of a function for \( t > 0 \) large enough, and the minimal of its generating family \( S_t(x; x_0, y_0) = \arctan x_0 + \frac{1}{2}y_0^2 + (x - x_0)y_0 \) is not a \( C^1 \) function, though it is locally Lipschitzian (see proposition 2.23 herafter).

Hence, in order to iterate the minmax procedure, one is led to defining the minmax when the Cauchy datum is a Lipschitzian function. We will use Clarke’s generalization of the derivatives of \( C^1 \) functions in the Lipschitz setting [Cla83], see appendix A.

Proposition 2.18. Under the hypothesis (2.1) and with the notation of corollary 2.8, if \( u \) is only locally Lipschitzian, the family \( S \) given by (2.4) generates \( L = \psi^s_H(\partial u) \) in the sense that

\[ L = \{(x, \partial_y S(x; \eta))|0 \in \partial_y S(x; \eta)\}, \quad (2.7) \]

where \( \partial \) denote Clarke’s generalized derivative and \( \partial u := \{(x, p), p \in \partial u(x)\} \).

Proof. The equation 0 \( \in \partial \eta S(x; \eta) \) means that \( y_0 \in \partial u(x_0) \) and \( y_{i+1} = y_i + \partial_i \phi_H^{s, i+1}(x_{i+1}, y_i) \), \( x_i = x_{i+1} + \partial_j \phi_H^{s, i+1}(x_{i+1}, y_i) \) for \( 0 \leq i < N \), where \( x := x_N \) and \( \eta := (x_i, y_i) \in \mathbb{R}^{i+1} \).

However, this definition of a generating family is not invariant by fibrewise diffeomorphism, even by the following very simple (and useful) one:

\[ \chi; (x_i)_{0 \leq i < N}, (y_i)_{0 \leq i < N} \] \[ \mapsto \chi; (x_{i+1} - x_i, y_i)_{0 \leq i < N} := \chi; (x_i, y_i)_{0 \leq i < N}; \]

indeed, it transforms the family \( S \) given by (2.4) into

\[ S'(x; (\xi_i, y_i)_{0 \leq i < N}) := u(x - \sum_{0 \leq i < N} \xi_i) + \sum_{0 \leq j < N} \phi_{H}^{s, i+1}(x - \sum_{i < j < N} \xi_j, y_j) + \sum_{0 \leq i < N} \xi_i y_i, \]

for which \( \partial \chi S'(x; (\xi_i, y_i)_{0 \leq i < N}) \) is not a point, but the subset

\[ \partial u(x - \sum_{0 \leq i < N} \xi_i) + \sum_{0 \leq j < N} \partial \phi_{H}^{s, i+1}(x - \sum_{i < j < N} \xi_j, y_j). \]

As often, this difficulty is overcome by finding the right definition:\

Definition 2.19. A Lipschitz family \( S : \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R} \) is called a generating family for \( L \subset T^*\mathbb{R}^d \) when

\[ L = \{(x, y) \in T^*\mathbb{R}^d | \exists \eta \in \mathbb{R}^k : (y, 0) \in \partial S(x, \eta)\}. \]

Lemma 2.20. This definition of a generating family is invariant by fibrewise \( C^1 \) diffeomorphisms.

But this example exhibits one of the features of the Clarke derivative: the relation \( (y, 0) \in \partial S'(x, \eta) \) is definitely not equivalent to \( y \in \partial \chi S'(x, \eta), 0 \in \partial_y S'(x, \eta) \) and 0 \( \in \partial_y S'(x, \eta) \).
Proof. If $\Phi(x, \eta') = (x, \phi(x, \eta'))$ is a fibrewise diffeomorphism of $\mathbb{R}^d \times \mathbb{R}^k$, and $S' := S \circ \Phi$, then the chain rule (see appendix A, lemma A.7) yields

$$\partial S'(x, \eta') = \left\{ \left( y + \frac{\partial}{\partial x} \phi(x, \eta'), \frac{\partial}{\partial \eta} \phi(x, \eta') \right) \right\} \in \partial S(x, \phi(x, \eta')),$$

as $\eta' \mapsto \phi(x, \eta')$ is a diffeomorphism, it does follow that the two conditions

$$\exists \eta \in \mathbb{R}^k : (y, 0) \in \partial S(x, \eta) \quad \text{and} \quad \exists \eta' \in \mathbb{R}^k : (y, 0) \in \partial S'(x, \eta')$$

are equivalent. \hfill \Box

We are now ready to consider GFQI's for the elements of $\mathbb{R}^d$. Proposition 2.21. If $H : [0, T] \times T^* \mathbb{R}^d \to \mathbb{R}$ is $C^2$ and has compact support, then for each $v \in C^{1,0}(\mathbb{R}^d)$, the generating family of $L = \phi^H_v(\partial v) \in \mathcal{L}$ given by (2.4), namely

$$S(x; \eta) = v(x_0) + \sum_{0 \leq i < N} \phi^{i+1}_H(x_{i+1}, y_i) + \sum_{0 \leq i < N} (x_{i+1} - x_i) y_i,$$

where $x_N := x, \eta := (x_i, y_i)_{0 \leq i < N}$, is 'quadratic at infinity' in the following sense: let

$$Q(\eta) := -x_{N-1} y_{N-1} + \sum_{0 \leq i < N-1} (x_{i+1} - x_i) y_i,$$

the Lipschitz constant of each $S(x; \cdot) - Q : (T^* \mathbb{R}^d)^N \to \mathbb{R}$ is bounded, uniformly with respect to $x$ on each compact subset of $\mathbb{R}^d$.

Hence, for each compact $K \subset \mathbb{R}^d$, if $\theta \in C^\infty(\mathbb{R}^d, [0, 1])$ equals 1 in a neighbourhood of 0, there exists a positive constant $\alpha_K$ such that the function

$$S_K(x; \eta) = \psi_K(x; \eta) + Q(\eta), \quad \text{where} \ \psi_K(x; \eta) := \theta(\eta/\alpha_K) \left( S(x, \eta) - Q(\eta) \right), x \in K.$$

(2.8)

is a GFQI of $L_K := L \cap \pi^{-1}(K)$ in that sense.

Proof. Denote $\psi(x, \eta) = S(x, \eta) - Q(\eta)$, and $Q(\eta) = \frac{1}{2} \eta^T B \eta$. For a fixed compact subset $K$, let $c = \max_{x \in K} \text{Lip}(\psi(x, \cdot))$, and assume that $|D \theta| \leq 1$. By lemma A.8,

$$\partial_\eta S_K(x, \eta) = \partial_\eta \left( \theta \left( \frac{\eta}{\alpha_K} \right) \psi(x, \eta) + Q(\eta) \right) \leq \frac{1}{\alpha_K} D \theta \left( \frac{\eta}{\alpha_K} \right) \psi(x, \eta) + \theta \left( \frac{\eta}{\alpha_K} \right) \partial_\eta \psi(x, \eta) + DQ(\eta).$$

We have

$$|\psi(x, \eta)| \leq |\psi(x, 0)| + |\psi(x, \eta) - \psi(x, 0)| \leq b + c|\eta|$$

where $b := \max_{x \in K} |\psi(x, 0)|$. Hence,

$$\left| \frac{1}{\alpha_K} D \theta \left( \frac{\eta}{\alpha_K} \right) \psi(x, \eta) + \theta \left( \frac{\eta}{\alpha_K} \right) \partial_\eta \psi(x, \eta) \right| \leq \frac{1}{\alpha_K} (b + c|\eta|) + c \leq \frac{1}{2} |B^{-1}|^{-1}|\eta| < |DQ(\eta)|,$$

when $|\eta| \geq b_K$ for some $b_K$ with $\alpha_K, b_K$ large enough. In addition, we can choose $\alpha_K, b_K$ such that for $|\eta| \leq b_K$, $\theta(\eta/\alpha_K) = 1$. Thus $S_K = \pi$ for $|\eta| \leq b_K$ and there are no critical points of $S, S_K$ outside $[|\eta| \leq b_K]$, from which $L_K = \{(x, \partial_\eta S_K(x, \eta)) | 0 \in \partial_\eta S_K(x, \eta) \}$. \hfill \Box

In the following, unless otherwise specified, we consider families $S$ of the form (2.4) or (2.5) and families $S_K$ of the form (2.8). The advantage is that $S$ generates $L$ in the simple
sense \((2.7)\), which helps to express the properties of minmax \(R_S(x)\) in a clear way, similar to the \(C^2\) case.

To study the minmax function \(R_S\) for such \(S\), we use the extension of classical results in critical point theory to locally Lipschitz functions described in appendix \(A\).

**Proposition 2.22.** The minmax \(R_S(x)\) is well-defined and it is a critical value\(^5\) of the map \(\eta \mapsto S(x, \eta)\). For each compact subset \(K \subset \mathbb{R}^d\) and each truncation \(S_K\) of \(S\) of the form \((2.8)\) generating \(L_K\), we have that \(R_S(x) = R_{S_K}(x)\) for \(x \in K\).

**Proof.** By proposition 2.21, \(f(\eta) := S(x, \eta) = \psi(\eta) + Q(\eta)\) with \(\psi\) Lipschitzian and \(Q\) a non-degenerate quadratic form. Hence, \(f\) satisfies the Palais–Smale (PS) condition (appendix \(A\), example \(A.4\)). If \(c = R_S(x)\) were not a critical value, the flow \(\varphi_t\) of theorem \(A.5\) in appendix \(A\) would deform the descending cycles in \(f^{c+\epsilon}\) into descending cycles in \(f^{c-\epsilon}\), hence the contradiction \(c = \inf \max \varphi \leq c - \epsilon\).

To see that \(R_{S_K}|_K = R_{S_K}\), just notice that every descending cycle \(\sigma\) of \(S(x, -)\) or \(S_K(x, -)\), \(x \in K\), can be deformed into a common descending cycle \(\sigma'\) with \(\max S(x, \sigma'(-)) = \max S_K(x, \sigma'(-))\) by using the gradient flow of \(Q\), suitably truncated. \(\square\)

**Proposition 2.23.** The minmax \(R_S(x)\) is a locally Lipschitz function.

**Proof.** Let \(K \subset \mathbb{R}^d\) be compact. By proposition 2.22, we have that \(R_S|_K = R_{S_K}\), where \(S_K : K \times \mathbb{R}^d \to \mathbb{R}\) writes \(S(x, \eta) = \psi_K(x, \eta) + Q(\eta)\) with \(Q\) a non-degenerate quadratic form and \(\psi_K\) a compactly supported Lipschitz function. Given \(x, x' \in K\), for all \(\epsilon > 0\), there exists a descending cycle \(\bar{\sigma}\) such that \(\max_{\eta \in \mathbb{R}^d} S_K(x, \eta) \leq R_S(x) + \epsilon\); if \(\max_{\eta \in \mathbb{R}^d} S_K(x', \eta)\) is reached at \(\bar{\eta}\), then

\[
R_S(x') - R_S(x) \leq S_K(x', \bar{\eta}) - S_K(x, \bar{\eta}) + \epsilon = \psi_K(x', \bar{\eta}) - \psi_K(x, \bar{\eta}) + \epsilon \\
\leq \text{Lip}(|x - x'|) + \epsilon.
\]

If we let \(\epsilon \to 0\) and exchange \(x\) and \(x'\), we obtain

\[
|R_S(x) - R_S(x')| \leq \text{Lip}(\text{Lip}(|x - x'|),
\]

which proves our result. \(\square\)

**Proposition 2.24.** The sets \(C(x) = \{\eta \in \partial \eta S(x, \eta), S(x, \eta) = R_S(x)\}\) are compact\(^6\) and the set-valued map (‘correspondence’) \(x \mapsto C(x)\) is upper semi-continuous: for every convergent sequence \((x_k, \eta_k) \to (x, \eta)\) with \(\eta_k \in C(x_k)\), one has \(\eta \in C(x)\). In other words, the graph \(C = \{(x, \eta) \mid \eta \in C(x)\}\) of the correspondence is closed.

**Proof.** Let \((x_k, \eta_k) \to (x, \eta)\) with \(\eta_k \in C(x_k)\); then, as \(S\) is \(C^1\) in \(x\), \(\partial S = \partial_x S \times \partial_\eta S\). Now \(\partial S : (x, \eta) \mapsto \partial_x S \times \partial_\eta S\) is upper-semi-continuous (appendix \(A\), proposition \(A.2\)), the limit \((\partial_x S(x, \eta), 0)\) of the sequence \((\partial_x S(x_k, \eta_k), 0)\) belongs to \(\partial S(x, \eta)\), hence \(0 \in \partial_\eta S(x, \eta)\); as the continuity of \(S\) and \(R_S\) implies that \(S(x_k, \eta_k) \to S(x, \eta)\) and \(R_S(x_k) \to R_S(x)\), this proves \(\eta \in C(x)\). \(\square\)

---

\(^5\) Appendix \(A\), definition \(A.3\).

\(^6\) Appendix \(A\), example \(A.4\).
Lemma 2.25. Given any $\delta > 0$, there exists an $\epsilon > 0$ such that

$$R_S(x) = \inf_{\sigma \in \Sigma} \max_{\eta \in \sigma(1)} S(x, \eta)$$

where $\Sigma = \{\sigma | \max_{\eta} S(x, \eta) \leq R_S(x) + \epsilon\}$ and $C_\delta(x) = B_\delta(C(x))$ denotes the $\delta$-neighbourhood of the critical set $C(x)$.

Proof. This is a direct consequence of the deformation lemma (appendix A, theorem A.6) for $S_\delta := S(x, \cdot)$: for $\delta > 0$, we have that $\epsilon > 0$ and $V$ such that $\varphi_\delta(S^- \setminus C_\delta(x)) \subseteq S^\cdot$. In particular, we remark that for $\sigma \in \Sigma$, the intersection $\sigma \cap C_\delta(x)$ is non-empty, otherwise, the flow $\varphi_\delta$ may take $\sigma$ to a descending cycle $\sigma' = \varphi_\delta(\sigma)$ such that $\max_{\eta \in \sigma} S_\delta(\eta) \leq R_S(x) - \epsilon$, contradiction with the definition of minmax.

Remark 2.26. When $S$ is $C^2$, the $S_\delta$’s are generically Morse functions: indeed, $S_\delta$ is Morse if and only if $x$ is a regular value of the projection $\pi : L \to M$, $(x, p) \mapsto x$, whose regular values, by Sard’s theorem and the compactness of $\Crit(S_\delta)$, form an open set of full measure. In this case, $S_\delta$ is indeed a deformation retract of $S^\cdot$, for $\epsilon > 0$ small enough, hence $\inf \max$ deserves its name ‘minmax’, that is, there exists a descending cycle $\sigma$ such that $R_S(x) = \max_{\sigma \in \sigma(1)} S(x, \eta) = \max_{\eta \in C(\pi)} S(x, \eta)$.

Proposition 2.27. The generalized derivative of $R_S$ satisfies

$$\partial R_S(x) \subset \co[\partial_\partial, S(x, \eta) | \eta \in C(x)].$$

Proof. First, we claim that, if $R_S$ is differentiable at $\bar{x}$, then

$$dR_S(\bar{x}) \subset \co[\partial_\partial, S(\bar{x}, \eta) | \eta \in C(\bar{x})].$$

Take $\delta$ and $\epsilon$ for $\bar{x}$ as in lemma 2.25. Consider $K = \overline{B_\delta(\bar{x})}$, and $S_K$ obtained in proposition 2.21, one can choose a $\varrho \in (0, 1)$ such that for $x \in B_\delta(\bar{x})$,

$$|S_K(x, \cdot) - S(x, \cdot)|_{C^0} \leq \epsilon / 4.$$

Now let $y \in \mathbb{R}^d$ and $\lambda < 0$ be small so that $x_\lambda := \bar{x} + \lambda y \in B_\delta(\bar{x})$ and $\lambda^2 < \epsilon / 4$. Then by lemma 2.25, for each $x_\lambda$, there is a descending cycle $\sigma_\lambda$ such that

$$\max_{\eta} S(x_\lambda, \eta) \leq R_S(x_\lambda) + \lambda^2,$$

then,

$$\max_{\eta} S(\bar{x}, \eta) \leq \max_{\eta} S(x_\lambda, \eta) + \frac{\epsilon}{4} \leq R_S(x_\lambda) + \frac{\epsilon}{2} \leq R_S(\bar{x}) + \frac{3\epsilon}{4}$$

and

$$R_S(\bar{x}) \leq \max_{\sigma \in \sigma(1)} S(\bar{x}, \eta) = S(\bar{x}, \eta_\lambda), \quad \text{for some } \eta_\lambda \in \sigma_\lambda \cap C_\delta(\bar{x}).$$

Hence we have

$$\lambda^{-1}[R_S(x_\lambda) - R_S(\bar{x})] \leq \lambda^{-1}[S(x_\lambda, \eta_\lambda) - S(\bar{x}, \eta_\lambda)] - \lambda$$

$$= \langle \partial_\partial S(x_\lambda, \eta_\lambda), y \rangle - \lambda,$$

where the last equality is given by the mean value theorem for some $x_\lambda$ in the line segment between $\bar{x}$ and $x_\lambda$.

Take the lim sup of both sides in the above inequality and let $\delta \to 0$, we get

$$\langle dR_S(\bar{x}), y \rangle \leq \max_{\eta \in C(\pi)} \langle \partial_\partial S(\bar{x}, \eta), y \rangle, \quad \forall y \in \mathbb{R}^d.$$
Note that this implies that \( dR_S(\bar{x}) \) belongs to the sub-derivative of the convex function \( f(y) := \max_{\eta \in C(\bar{x})} (\partial_y S(\bar{x}, \eta), y) \) at \( v = 0 \). Thus we get (2.10). In general,

\[
\partial R_S(x) = \text{co} \{ \lim_{x' \to x} dR_S(x') \} \subset \text{co} \{ \partial_x S(x, \eta), \eta \in C(x) \}
\]

by the upper-semi continuity of \( x \mapsto C(x) \) and the continuity of \( \partial_x S \).

The formula (2.9) gives us somehow a generalized graph selector. While for a classical graph selector, we require that for almost every \( x \),

\[
\partial R_S(x) = \partial_x S(x, \eta).
\]

from which \( (x, dR_S(x)) \in L \). Following Chaperon, Sikorav [Cha91, PPS03], this occurs if \( S \) is a GFQI of \( L = \varphi(dv) \in L \) for \( v \in \hat{C} \). Indeed, in this case, \( S_x := S(x, \cdot) \) is an excellent Morse function for almost every \( x \), in which cases \( C(x) \) consists of a single point, hence \( \partial R_S(x) = \partial_x S(x, \eta) \) for a unique \( \eta \), proving that \( R_S \) is a true graph selector for \( L \).

There are still some mysteries for the generalized graph selector. One may ask whether the minmax \( R_S \) is also a true graph selector for \( L \in \hat{L} \). Is it true that, when \( R_S \) is differentiable at \( x \), one has \( (x, dR_S(x)) \in L \), where \( L \in L \) or even \( \hat{L} \)?

3. Viscosity solutions and minmax solutions

We look at the solutions of the (H–J) Cauchy problem, assuming that \( H \in C^2([0, T] \times T^* \mathbb{R}^d) \) and \( v \in C^{1,\text{Lip}}(\mathbb{R}^d) \) satisfy the condition of finite propagation speed. Unless otherwise specified, we assume that \( H \) has compact support (as a function on \([0, T] \times T^* \mathbb{R}^d \) when \( H \) and \( v \) are periodic).

3.1. Geometric solution and its minmax selector

Following the classical geometric method for first order partial differential equations, the Hamilton–Jacobi equation is considered to be a hypersurface in the cotangent bundle \( T^*(\mathbb{R} \times \mathbb{R}^d) \). See [Arn00, Arn88].

More precisely, let

\[
\mathcal{H}(t, x, e, p) := e + H(t, x, p), \quad (t, x, e, p) \in T^*(\mathbb{R} \times \mathbb{R}^d)
\]

and at the moment suppose that the initial function \( v \) is \( C^2 \).

**Definition 3.1.** Let \( \varphi_H^t \) denote the Hamiltonian flow of \( \mathcal{H} \), which preserves the levels of \( \mathcal{H} \), and let

\[
\Gamma_v = \{(0, x, -H(0, x, dv(x)), dv(x))\};
\]

then, the geometric solution of the Cauchy problem (H–J) is

\[
L_{\mathcal{H}, v} := \bigcup_{t \in [0, T]} \varphi_H^t(\Gamma_v).
\]

It is a Lagrangian submanifold containing the initial isotropic submanifold \( \Gamma_v \) and contained in the hypersurface

\[
\mathcal{H}^{-1}(0) = \{(t, x, e, p)|e + H(t, x, p) = 0\} \subset T^*(\mathbb{R} \times \mathbb{R}^d).
\]

For a convex function \( f \), the sub-derivative at a point \( x \) is the set of \( \xi \) such that \( f(y) - f(x) \geq \langle \xi, y-x \rangle, \forall y \).
As every Lagrangian submanifold $L$ of $T^*(\mathbb{R} \times \mathbb{R}^d)$ contained in $\mathcal{H}^{-1}(0)$ is locally invariant by $\varphi^t_H$, this geometric solution is in some sense maximal.

Writing $T^*(\mathbb{R} \times \mathbb{R}^d)$ as $T^*\mathbb{R} \times T^*\mathbb{R}^d$, we have $X_H = (1, -\partial_tH, X_H)$, and

$$L_{\mathcal{H},u} = \bigcup_{t \in [0,T]} \{(t, -H(t, \varphi^t_H(du)), \varphi^t_H(du))\}$$

where $\varphi^t_H := \varphi^0_H$ is the Hamiltonian isotopy generated by $H$.

**Lemma 3.2.** Formula (2.5) defines a GFQI of $L_{\mathcal{H},v}$.

**Proof.** For simplicity, we may assume that $T \in (0, \delta_H)$, hence that

$$S : [0,T] \times \mathbb{R}^d \times \mathbb{R}^{2d} \to \mathbb{R}, \quad S(t,x,y_0) = v(x_0) + y_0 + \varphi^t_H(x,y_0) - x_0y_0.$$

Let $(x_0, y_0) \in \Sigma_S$, then

$$(\partial_tS(t,x,y_0), \partial_x S(t,x,y_0)) = (\partial_t\varphi^t_H(x,y_0), \partial_x \varphi^t_H(x,y_0)) = (-H(t,x,y(t)), y(t)),$$

where $(x,y(t)) = \varphi^t_H(x_0, y_0)$ with $y_0 = dv(x_0)$.

Hence

$$\{(t,x,\partial_tS(t,x,y_0), \partial_x S(t,x,y_0)) | (x_0, y_0) \in \Sigma_S\} = L_{\mathcal{H},v}.$$

The wave fronts of the geometric solution $L_{\mathcal{H},v}$ are then given by

$$\mathcal{F}^e := \{(x,S(t,x;\eta)) | \frac{\partial S}{\partial \eta}(t,x;\eta) = 0\} \subset J^1(\mathbb{R}^d).$$

If there exists a $C^1$ function $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ such that

$$L = L_{\mathcal{H},v} = \{(t,x,\partial_t u(t,x), \partial_x u(t,x)) \in T^*([0,T] \times \mathbb{R}^d)\}$$

we say that $L$ is a 1-graph in $T^*([0,T] \times \mathbb{R}^d)$. In this case, $u$ is a global solution of the Cauchy problem of (H–J) equation. In general, $L$ may be the graph of the derivatives of a multi-valued function.

An equivalent but more economic way to describe the geometric solution is to identify (as in the introduction) each $\varphi^t_H(\Gamma_0)$ with $[s] \times \varphi^t_H(du)$ by the inverse of the map $(t,x,p) \mapsto (t,x,-H(t,x,p),p)$. In this way, we also call the union

$$L_{\mathcal{H},v} := \bigcup_{t \in [0,T]} \{t \times \varphi^t_H(du) \subset \mathbb{R} \times T^*\mathbb{R}^d$$

the geometric solution.

If we look at the projection of the characteristics, that is, the image of the graph of the solutions $\{\{t, \varphi^t_H(x_0, p_0)\}\}_{t \in [0,T]}$, $(x_0, p_0) \in T^*\mathbb{R}^d$, of Hamilton’s equations under the projection

$$\pi : [0,T] \times T^*\mathbb{R}^d \to \mathbb{R} \times \mathbb{R}^d, \quad (t,x,p) \mapsto (t,x),$$

then $L$ is not a 1-graph when the corresponding characteristics intersect under the projection. Without ambiguity, we will simply say that the characteristics intersect.

For the initial functions $v \in C^{1,\text{Lip}}(\mathbb{R}^d)$, set

$$L_{\mathcal{H},v} = \bigcup_{t \in [0,T]} \{t \times \varphi^t_H(\partial v) := \bigcup_{t \in [0,T]} \{t \times \{\varphi^t_H(x,p) : p \in \partial v(x)\}\},$$

where $\partial$ is Clarke’s generalized derivative. We call them generalized geometric solutions. They are also generated by the GFQI’s given by formula (2.5).
For any time $0 \leq s < t \leq T$, we define the minimax operator\(^8\)
\[ R^s_t : C^{L_p}(\mathbb{R}^d) \to C^{L_p}(\mathbb{R}^d), \quad \tau \in [s, t] \]
for the (H–J) equation as
\[ R^s_t v(x) = \inf_{\eta} S(\tau, x, \eta) \]
where $S : [s, t] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is given by (2.5).

For completeness, without referring to the uniqueness theorem for GFQI’s, we give a proof that the minmax is well-defined independently of the subdivisions.

**Lemma 3.4.** The minimax $R_\delta(x) = \inf_{\eta} S(x, \eta)$ given by (2.4) or (2.5) is independent of the subdivision of time in the construction of $S$.

**Proof.** First assume $t - s < \delta_H$; given $\tau \in (s, t)$, consider the family of subdivisions $\zeta_\mu := \{s \leq s + \mu(\tau - s) < t\}$; then,
\[ S_\mu(x; x_0, y_0, x_1, y_1) = v(x_0) + \phi_H^{s, s + \mu(\tau - s)}(x_1, y_0) + (x_1 - x_0)y_0^0 \]
\[ + \phi_H^{s, s + \mu(\tau - s)}(x_2, y_1) + (x_2 - x_1)y_1, \]
where $x_2 := x$, is the generating family defined by (2.4) and associated to $\zeta_\mu$, $\mu \in [0, 1]$. The function $S_0$ is continuous in $\mu$ and the minimax $R_{S_0}(x)$ is a critical value of the map $\eta \mapsto S_\mu(x; \eta)$ with $\eta := (x_0, x_1, y_0, y_1)$. By (2.6), the set of all such critical values is independent of $\mu$; as it has measure zero by Sard’s theorem, $R_{S_0}$ is constant for $\mu \in [0, 1]$. In particular, letting $x_1 := x_1 - x_0$ and $y_0 = y_0 - y_1$, we get
\[ S_0 = S_0(x; x_0, y_0, x_1, y_1') = v(x_0) + \phi_H^{s, s + \mu(\tau - s)}(x_2, y_1) + (x_2 - x_0)y_1 + x_1'y_1'. \]
It is obtained by adding the quadratic form $x_1'y_1'$ to $S(x; x_0, y_1) = v(x_0) + \phi_H^{s, s + \mu(\tau - s)}(x_2, y_1) + (x_2 - x_0)y_1$.

In general, given any two subdivisions $\zeta', \zeta''$ of $[s, t]$ with $|\zeta'|, |\zeta''| < \delta_H$, denote by $\zeta = \zeta' \cup \zeta'' = \{s = t_0 < \cdots < t_\nu = t\}$ the subdivision obtained by collecting the points in $\zeta'$ and $\zeta''$. If $t_0$ is not contained in $\zeta$, we consider the family of subdivisions
\[ \zeta_\mu(j) = \{t_0 < t_{j-1} \leq t_j + \mu(t_j - t_{j-1}) < t_{j+1} < \cdots < t_\nu\}, \quad \mu \in [0, 1]. \]
The same argument as before shows that the minimax relative to $\zeta_\mu(j)$ and $\zeta_\mu(j)$ are the same. Continuing this procedure, we get that the minimax relative to $\zeta'$ and $\zeta$ are the same, and the same holds for $\zeta''$ and $\zeta$. Therefore the minimax with respect to $\zeta'$ and $\zeta''$ are the same.

**Lemma 3.5.** If $v \in C^2 \cap C^{L_p}(\mathbb{R}^d)$, then $R^0_t v(x)$ verifies the (H–J) equation almost everywhere.

**Proof.** This is a direct consequence of the fact that $S$ is a GFQI of $L_{H,v}$ and the minmax is a graph selector in this case.

In general, for a Lipschitzian initial function, we do not know whether the minmax verifies the equation almost everywhere or not. But in view of the estimation of generalized derivatives in proposition 2.27, and that $(x, R^0_t v(x)) \in J^t$, we call $u(t, x) = R^0_t v(x)$ the minmax solution of the (H–J) Cauchy problem.

---

\(^8\) The inclusion $R^\mu_H \left(C^{L_p}(\mathbb{R}^d)\right) \subset C^{L_p}(\mathbb{R}^d)$ is proven in proposition 3.14 hereafter.

\(^9\) For a subdivision $\zeta = \{t_0 < \cdots < t_n\}$, we let $|\zeta| := \max |t_{i+1} - t_i|$. 

---
Lemma 3.6. If $v$ is $C^2$ with bounded second derivative, then there exists an $\epsilon > 0$ such that for $t \in [0, \epsilon)$, the minmax $R_H^{0,t} v(x)$ is $C^2$.

Proof. We will show that, there exists an $\epsilon > 0$, such that for $t \in (0, \epsilon)$, the characteristics beginning from the graph $dv$ do not intersect. More precisely, the map $f_t : x_0 \mapsto X_t (x_0, dv(x_0))$ is a diffeomorphism. Indeed, for $t$ small enough,

$$\text{Lip}(f_t - Id) \leq \text{Lip}(\alpha_t^0 - Id)(1 + \text{Lip}(dv)) < 1$$

where $\alpha_t^0$ and $c_H$ are defined in lemma 2.5. This in turn means that the projection map $L = \phi_{t, H} : R^d \to \mathbb{R}^d, (x, p) \mapsto x$ is a diffeomorphism, hence $L = \{x, dR_H^{0,t} v(x)\}$, from which we obtain that $R_H^{0,t} v(x)$ is $C^2$. \hfill \Box

3.2. Viscosity solutions

Definition 3.7. A function $u \in C^0 \big((0, T) \times \mathbb{R}^d\big)$ is called a viscosity subsolution (respectively, supersolution) of

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

when it has the following property: for every $\psi \in C^1 \big((0, T) \times \mathbb{R}^d\big)$ and every point $(t, x)$ at which $u - \psi$ attains a local maximum (respectively, minimum), one has

$$\partial_t \psi + H(t, x, \partial_x \psi) \leq 0, \quad (\text{respectively} \geq 0).$$

The function $u$ is a viscosity solution if it is both a viscosity subsolution and supersolution.

We remark that one can replace $C^1$ test functions $\psi$ by $C^\infty$ test functions in the definition. Obviously, a classical $C^1$ solution is a viscosity solution.

Theorem 3.8 ([CL87]). If $v \in C^{\text{Lip}}(\mathbb{R}^d)$ and $H \in C^2([0, T] \times T^* \mathbb{R}^d)$, then there exists a unique viscosity solution of the Cauchy problem of the Hamilton–Jacobi equation. Moreover, this solution is globally Lipschitz.

A notable feature of the viscosity solution, is the Markov property: if $J'_t : C^{\text{Lip}}(\mathbb{R}^d) \to C^{\text{Lip}}(\mathbb{R}^d)$ denotes the viscosity solution operator (for a fixed Hamiltonian) which to $v$ associates the time $t$ of the solution equal to $v$ at time $s$, then the ‘two-parameter groupoid’ property $J'_t = J'_s \circ J'_0$ is satisfied. This follows easily from uniqueness.

The following proposition summarizes a well-studied case when the Hamiltonian is convex in $p$, one can refer to [Jou93, Cha90, Wei13].

Proposition 3.9. Assume $H \in C^2([0, T] \times T^* \mathbb{R}^d)$ strictly convex in $p$, equal to $|p|^2$ off a compact set, and $v \in C^{\text{Lip}}(\mathbb{R}^d)$. Then the minmax solution is reduced to a min, and it possesses the ‘semi-group’ property with respect to time, that is

$$R_{0}^t v(x) = R_{s}^t \circ R_{0}^{s} v(x), \quad 0 \leq s \leq t.$$

Proof. Our hypotheses imply that there exists a constant $\epsilon_H > 0$ such that, for $0 < t - s < \epsilon_H$, the Hamiltonian diffeomorphism $\psi^t_{s}$ of $H$ has a ‘classical’ generating function $\psi^t_{s}(X, x)$ in the sense that $(x, y, (X, Y)) \in \text{Graph}(\psi)$ if and only if

$$Y = \partial_X \psi(X, x),$$

$$y = -\partial_Y \psi(X, x).$$

31
Therefore, for any $0 \leq s < t \leq T$, the subset $L = \varphi_t^s(\partial v)$ has the generating family
\[ F^t_s(x; x_i) = v(x_0) + \sum_{0 \leq i \leq j} \psi^{s+1}_t(x_{i+1}, x_i), \tag{3.13} \]
where $x_{j+1} := x$ and $\{s = \tau_0 < \tau_1 < \cdots < \tau_{j+1} = t\}$ is a subdivision of $[s, t]$ such that $|\tau_i - \tau_{i+1}| < \epsilon_H$, $0 \leq i \leq j$. Up to diffeomorphism, $F^t_s$ is quadratic of index 0 at infinity, since $H$ equals $|p|^2$ off a compact subset. Thus the minmax is reduced to a min:
\[ R^t_Sv(x) = \min_{(x_i)} F^t_x(x, (x_i)). \]
Note that $R^t_j$ is independent of the choice of the subdivision (same argument as for lemma 3.4), hence
\[
R^t_i \circ R^t_j v(x) = \min_{(x_i)} \left( R^t_i v(x_0) + \sum_{0 \leq i \leq j} \psi^{s+1}_t(x_{i+1}, x_i) \right)
\]
\[
= \min_{(x_i)} \left( \min_{(x'_i)} \left( v(x'_0) + \psi^j_0(x_0, (x'_i)) \right) + \sum_{0 \leq i \leq j} \psi^{s+1}_t(x_{i+1}, (x_i)) \right)
\]
\[
= \min_{(x_i)} \left( \min_{(x'_i)} v(x'_0) + \psi^j_0(x_0, (x'_i)) \right) + \sum_{0 \leq i \leq j} \psi^{s+1}_t(x_{i+1}, (x_i))
\]
\[
= R^t_j v(x).
\]

**Remark 3.10.** For $S$ and $F$ defined by (2.5) and (3.13) respectively, we have
\[ \inf \max S(x, (x_i)) = \min F(x, (x_i)). \]
Indeed, when $v$ is $C^2$, we can conclude by the uniqueness theorem of GFOI’s since $S$ and $F$ generates the same Lagrangian submanifold $L = \varphi^{(dv)}$; in the general Lipschitz case, we can apply the continuity dependence of the minmax selector on the generating family (lemma 2.14).

**Remark 3.11.** The hypothesis that $H = |p|^2$ at infinity can be generalized to the case where the condition of finite propagation speed is satisfied. The min solution operator is a finite dimensional ‘discretization’ of the Lax–Oleinik semi-group in weak KAM theory, defined by
\[ T^t_s v(x) = \inf_{\gamma(t) = x} \left\{ v(\gamma(s)) + \int_s^t L(t, \gamma(t), \dot{\gamma}(t)) \, dt \right\}, \]
where $L$ is the Legendre transform of $H$ with respect to the $p$ variable, and the inf is taken over all absolutely continuous paths $\gamma : [s, t] \to \mathbb{R}^d$.

**Theorem 3.12 ([Jou93]).** The min solution $R^0_t v(x)$ is the viscosity solution of the Cauchy problem $H–J$.

### 3.3. Iterated minmax and viscosity solution

In contrast to the case of convex Hamiltonians, where the minmax is reduced to a min and provides the viscosity solution, for general non-convex Hamiltonians, the minmax and the viscosity solution may differ: see [Vit96, VO95, BC11] for counterexamples, and also [Che74] for a very nice geometric illustration of the fact that the viscosity solution is not necessarily contained in the geometric solution.

Particularly, in [Vit96], the author pointed out without proof that the minmax does not provide a semi-group as a consequence of not being viscosity. We will make this point clear by showing that the semi-group property is a sufficient condition for the minmax to be viscosity.
Proposition 3.13. Given $v$, the minmax $R_H^0 v(x)$ is the viscosity solution of the Cauchy problem (H–J) if it has the semi-group property with respect to time, that is,

$$R_H^0 v(x) = R_H^t \circ R_H^0 v(x), \quad 0 \leq s < t \leq T.$$ 

Proof. Suppose $R_H^0 v(x) := R_H^0 t v(x)$ possesses the semi-group property, we first show that $R_H^0 v(x)$ is a minmax, that is, $R_H^0 v(x)$ has the viscosity solution of the Cauchy problem (H–J) with respect to time, that is,

$$R_H^0 v(x) \leq R_H^s v(x), \quad 0 \leq s < t \leq T.$$ 

Proof. Suppose $R_H^0 v(x) := R_H^0 t v(x)$ possesses the semi-group property, we first show that $R_H^0 v(x)$ is a minmax, that is, $R_H^0 v(x)$ has the viscosity solution of the Cauchy problem (H–J) with respect to time, that is,

$$R_H^0 v(x) \leq R_H^s v(x), \quad 0 \leq s < t \leq T.$$ 

As a direct consequence, we get theorem 3.12 since the min solutions form a semi-group (proposition 3.9).

We remark that proposition 3.13 does not essentially depend on the variational formulation of the minmax. Indeed, an operator is the viscosity solution operator if it verifies the semigroup property, the monotonicity, and it is a generator of the (H–J) equation, that is it generates a regular solution at least within small time when the initial data is regular enough. See for example [Ber11, proposition 20] and [FS06, theorem 5.1].

To compensate the fact that the minmax is not a semi-group, an idea due to Chaperon is to replace the ‘minmax’ by some ‘iterated minmax’. Roughly speaking, an iterated minmax is obtained by dividing a given time interval into small pieces and taking the minmax step by step. This is a priori a discrete semi-group with respect to the points of the subdivision. We are going to show that, as the steps of the subdivision go to zero, the iterated minmax converges to a genuine semi-group, and therefore to the viscosity solution.

In the following, we denote the Lipschitz constant of a global Lipschitz function $f$ by $\|\partial f\|$ and $|\cdot|_K$ denotes the maximum norm on a compact set $K$.

Proposition 3.14. Assuming $H \in C^2_c([0, T] \times T^* \mathbb{R}^d)$ and $v \in C^{1,\text{Lip}}(\mathbb{R}^d)$, we have the following estimates:

1. $R_H^{s,t}$ defines an operator from $C^{1,\text{Lip}}(\mathbb{R}^d)$ to $C^{1,\text{Lip}}(\mathbb{R}^d)$, and
   $$\|\partial(R_H^{s,t} v)\| \leq \|\partial v\| + \|\partial_s H\| |t-s|.$$ 

2. For any $0 \leq s < t_i \leq T$, $i = 1, 2$,
   $$|R_H^{s,t_1} v(x) - R_H^{s,t_2} v(x)| \leq |t_1 - t_2| \max_{t \in [n_i, n_{i+1}]} |H(t, x, \cdot)|_Y$$
   where $Y = \{y : |y| \leq \|\partial v\| + \|\partial_s H\| \max |t_i - s|\}$. 

33
(3) Let $H^0$ and $H^1$ be two Hamiltonians, then
\[ |R^{s,t}_{H^0} v - R^{s,t}_{H^1} v|_{C^0} \leq |t - s| \max_{x \in C^0} |(H^0 - H^1)(\tau, \cdot, y)|_{C^0} \]
where $Y' = \{ y : |y| \leq \| \partial_y v \| + \max_{\tau, y} \| \partial_y H' \| |t - s| \}$. 

(4) If $v^0, v^1 \in C^{1,0}(\mathbb{R}^d)$ and $K$ is a compact set in $\mathbb{R}^d$, then there exists a bounded subset $\tilde{K} \subset \mathbb{R}^d$ which depends on $K \times [0, T]$ and the constants $\| \partial v' \|$, such that
\[ |R^{s,t}_{H^0} v^0 - R^{s,t}_{H^1} v^1|_{\tilde{K}} \leq |v^0 - v^1|_{\tilde{K}}, \quad 0 \leq s < t \leq T. \]  

**Proof.** The proof is based on proposition 2.27 with some variation on the original variable $x$, which can be either $t \in [0, T]$, $x \in \mathbb{R}^d$ or some parameter $\lambda$ for the generating family constructed as below.

For simplicity, we may first assume that $|t - s| < \delta_H$ so that $S^{s,t}(x, x_0, y_0) = v(x_0) + \phi^{s,t}_{H}(x, y_0) + x_0 - x_0 y_0$.

Let $(x(\tau), y(\tau))$ denote the Hamiltonian flow, and $C(x)$ be the critical set defined in proposition 2.27.

(1) For $(x_0, y_0) \in C(x)$, we have
\[ \partial_x S^{s,t}(x, x_0, y_0) = \partial_x \phi^{s,t}_{H}(x, y_0) + y_0 = y(t) \]
where
\[ y(t) = y_0 - \int_s^t \partial_x H(\tau, x(\tau), y(\tau)) \, d\tau, \quad y_0 \in \partial v(x_0). \]
Hence by (2.9),
\[ \partial R^{s,t}_{H} v(x) \subset \co \{ y(t), y_0 \in \partial v(x_0) \} \]
thus
\[ \| \partial (R^{s,t}_{H} v) \| \leq \| \partial v \| + \| \partial_x H \| |t - s|. \]

(2) For $(x_0, y_0) \in C(x)$, by lemma 2.6, we have
\[ \partial_y S^{s,t}(x, x_0, y_0) = \partial_y \phi^{s,t}_{H}(x, y_0) = -H(t, x, y(t)). \]
Hence
\[ \partial_i R^{s,t}_{H} v(x) \subset \co \{ -H(t, x, y(t)), y_0 \in \partial v(x_0) \} \]
and therefore
\[ |R^{s,t}_{H} v(x) - R^{s,t}_{H} v(x)| \leq |t_1 - t_2| \max_{x \in C^0} |H(t, x, y)| \]
where $Y = \{ y : |y| \leq \| \partial_y v \| + \| \partial_y H \| |t_1 - t_2| \}$.

(3) Let $H^\lambda = (1 - \lambda) H^0 + \lambda H^1$, $\lambda \in [0, 1]$, and let $S^{s,t}_{H^\lambda}$ be the corresponding generating families. Fix $\lambda$, for $(x_0, y_0)$ in the critical set $C^\lambda(x)$ corresponding to $H^\lambda$,
\[ \partial_x S^{s,t}_{H^\lambda}(x, x_0, y_0) = \partial_x \phi^{s,t}_{H}(x, y_0) = \int_s^t (H^0 - H^1)(\tau, x^k(\tau), y^k(\tau)) \, d\tau, \]
where the proof of the second equality is similar to that of lemma 2.6. Hence,
\[ \partial_x R^{s,t}_{H^\lambda} v(x) \subset \co \left\{ \int_s^t (H^0 - H^1)(\tau, x^k(\tau), y^k(\tau)) \, d\tau, y_0 \in \partial v(x_0) \right\} \]
and therefore
\[ |R^{s,t}_{H^0} v(x) - R^{s,t}_{H^1} v(x)| \leq \int_0^1 \int_s^t |H^0 - H^1|(\tau, x^k(\tau), y^k(\tau)) \, d\tau \, d\lambda \leq |t - s| \max_{x \in C^0} |(H^0 - H^1)(\tau, \cdot, y)|_{C^0} \]
where $Y' = \{ |y| : |y| \leq \| \partial_y v \| + \max_{\tau} \| \partial_y H' \| |t - s| \}$. 

Nonlinearity 27 (2014) 17

Q Wei

(4) Let \( v^\lambda = (1 - \lambda) v^0 + \lambda v^1 \), \( \lambda \in [0, 1] \) and \( S^{\lambda,t}_{\lambda,t} \) denotes the corresponding generating families, then \( \partial_\lambda S^{\lambda,t}_{\lambda,t} (x_0, y_0) = v^1(x_0) - v^0(x_0) \),

\[
\partial_\lambda R^{\lambda,t}_t v^1(x) \subseteq \text{co} \{v^1(x_0) - v^0(x_0) : (x_0, y_0) \in \mathcal{C}(x)\}
\]

with \( \mathcal{C}(x) \subseteq \{(x_0, y_0) : |x_0| \leq |x| + T \|\partial_y (H|_{y \in Y})\|, Y := \{y : |y| \leq \|\partial v\| + T \|\partial_v H\|\} \).

If we take \( \tilde{K} = \{x_0 : |x_0| \leq |x| + T \|\partial_y (H|_{y \in Y})\|\} \), we obtain

\[
|R^{\lambda,t}_t v^0 - R^{\lambda,t}_t v^1|_{\tilde{K}} \leq |v^0 - v^1|_{\tilde{K}}.
\]

In general, the above results follow from the fact that the critical set \( \mathcal{C}(x) \) defines the Hamiltonian flow \((x(\tau), y(\tau))_{\tau \in [0,T]} \) for any \( 0 \leq s < t \leq T \).

\[\square\]

**Remark 3.15.** The estimates in the proposition, more subtle than needed, precisely reveal that finite propagation speed is enough to define the minimax function.

Now given any compact subset \( K \subseteq \mathbb{R}^d \), we consider \((t, x) \in [0, T] \times K \).

Given a subdivision \( \zeta = [0 = t_0 < t_1 < \cdots < t_n = T] \) of \([0, T]\), to each \( s \in [0, T] \), we associate a number \( [s]_\zeta \) as

\[
[s]_\zeta := t_i, \quad \text{if } t_i \leq s < t_{i+1}.
\]

**Definition 3.16.** The iterated minimax solution operator for the (H–J) equation with respect to a subdivision \( \zeta \) is defined as follows: for \( 0 \leq s' < s \leq T \),

\[
R^{s',s}_{H, \zeta} := R^{s',s}_{H, \eta} \circ R^{s',s}_{H, \eta} \circ \cdots \circ R^{s',t_{i-1}}_{H, \eta} \circ R^{t_{i-1},s}_{H, \eta}, \quad \text{where } t_j = [s]_\zeta, t_i = [s']_\zeta, i \leq j.
\]

When the Hamiltonian \( H \) is fixed, we may abbreviate our notation \( R^{s',s}_{H, \eta} \) as \( R^{s',s}_{\eta} \), and the iterated minimax as

\[
R^{s',s}_{\zeta} := R^{t_{j-1}}_{\eta} \circ R^{s',s}_{\eta} \circ \cdots \circ R^{t_{i-1}}_{\eta} \circ R^{t_i,s}_{\eta}, \quad \text{where } t_j = [s]_\zeta, t_i = [s']_\zeta, i \leq j \quad (3.17)
\]

indicated.

Define the length of \( \zeta \) by \( |\zeta| := \max_i |t_i - t_{i+1}| \). Suppose that \( (\zeta_n) \) is a sequence of subdivisions of \([0, T]\) such that \( |\zeta_n| \) tends to zero as \( n \) goes to infinity, and let \( (R^{s,x}_{0,\zeta_n} v(x))_n \) be the corresponding sequence of iterated minimax solutions for an initial function \( v \in C^{1,\alpha}(\mathbb{R}^d) \).

**Lemma 3.17.** The sequence of functions \( u_n(s, x) := R^{s,x}_{0,\zeta_n} v(x) \) is equi-Lipschitz and uniformly bounded for \((s, x) \in [0, T] \times K \).

**Proof.** By proposition 3.14, one can verify that

\[
\|\partial (R^{s,x}_{0,\zeta_n} v)\| \leq \|\partial v\| + T \|\partial_v H\|,
\]

\[
|R^{s,x}_{0,\zeta_n} v - R^{s,x}_{0,\zeta_n} v|_{K} \leq |H|_K |s - t|, \quad s, t \in [0, T].
\]

where \( K := \{(t, x, y) : t \in [0, T], x \in K, |y| \leq \|\partial v\| + T \|\partial_v H\|\} \).

In particular, taking \( t = 0 \), we get

\[
|R^{s,x}_{0,\zeta_n} v|_{K} \leq |v|_K + T |H|_K, \quad s \in [0, T].
\]

\[\square\]

**Proposition 3.18.** For any sequence \((\zeta_n) \) of subdivisions of \([0, T]\) such that \( |\zeta_n| \rightarrow 0 \) as \( n \rightarrow \infty \), and any compact \( K \subseteq \mathbb{R}^d \), the sequence of iterated minimax \( u_n(s, x) := R^{s,x}_{0,\zeta_n} v(x) \) has a subsequence converging uniformly on \([0, T] \times K \) to the viscosity solution of the (H–J) problem.
Proof. By lemma 3.17 and the Arzela–Ascoli theorem, the sequence of functions $u_n$ takes its values in a compact subset of $C^0([0, T) \times \bar{K})$, hence $(u_n)_n$ has a convergent subsequence $(u_n)_k$. Denote by $\bar{R}_0^k u(x)$ its limit. We will prove that

$$\bar{R}_0^k u(x) = \lim_{k \to \infty} \bar{R}_{\epsilon, \tilde{\tau}_k}^\kappa \circ \bar{R}_0^k u(x), \quad \forall 0 \leq s' < s \leq T. \quad (3.18)$$

Let $\tilde{K} \supset K$ as defined in 4) of proposition 3.14, applying the Arzela–Ascoli theorem to $(u_n)_k \subset C^0([0, T) \times \tilde{K})$ and extracting a subsequence of $n_k$ if necessary, we may suppose that $u_{n_k}$ converges uniformly in $[0, T] \times \tilde{K}$.

For simplicity of notation, we omit the subindex $k$ of $n_k$ in the following.

Denote $[s]_n := [s]_{n_k}$. We first remark that

$$\bar{R}_0^k u(x) = \lim_{n \to \infty} \bar{R}_{0, [s]_n}^k u(x), \quad 0 \leq s \leq T. \quad (3.19)$$

Indeed,

$$|\bar{R}_{0, [s]_n}^k u(x) - \bar{R}_{0, [s]_n}^{|s|} u(x)| = |\bar{R}_{0, [s]_n}^k u(x) - \bar{R}_{0, [s]_n}^{|s|} u(x)| \leq |H|_K (s - [s]_n) \leq |H|_K \xi_n \to 0, \quad n \to \infty.$$

For any $\epsilon > 0$, there exists $N$ large enough such that for any $i, j > N$,

$$|\bar{R}_{0, [s]_n}^k u - \bar{R}_{0, [s]_n}^{|s|} u|_K < \epsilon, \quad \forall s \in [0, T].$$

Hence

$$|\bar{R}_{0, [s]_n}^k u - \bar{R}_{0, [s]_n}^{|s|} u|_K \leq |\bar{R}_{0, [s]_n}^k u - \bar{R}_{0, [s]_n}^{|s|} u|_K \leq |\bar{R}_{0, [s]_n}^k u - \bar{R}_{0, [s]_n}^{|s|} u|_K < \epsilon.$$

Let $j$ go to infinity, we get

$$|\bar{R}_{0, [s]_n}^k u - \bar{R}_{0, [s]_n}^{|s|} u|_K < \epsilon, \quad i > N.$$

Thus the limit $\lim_{i \to \infty} \bar{R}_{[s]_n}^k u(x)$ exists and

$$\lim_{i \to \infty} \bar{R}_{[s]_n}^k u(x) = \bar{R}_0^k u(x), \quad x \in K.$$

We conclude (3.18) by verifying the following, which is similar to (3.19),

$$\lim_{i \to \infty} \bar{R}_{[s]_n}^k u(x) = \lim_{i \to \infty} \bar{R}_{[s]_n}^k u(x) \circ \bar{R}_0^k u(x).$$

Now we show that the limit function $\bar{R}_0^k u(x)$ is the viscosity solution of the (H-J) problem. We first show that it is a viscosity subsolution. For any $(t, x)$, suppose $\psi$ a $C^2$ function defined in a neighbourhood of $(t, x)$, having bounded second derivative and such that $\psi(s, y) \equiv \psi_t(x) \geq \bar{R}_0^k u(y)$, with equality at $(t, x)$,

$$\psi_t(x) = \bar{R}_0^k u(x) = \lim_{k \to \infty} \bar{R}_{[s]_n}^k u(x) \leq \lim_{k \to \infty} \bar{R}_{[s]_n}^k \psi_t(x) = \bar{R}_0^k \psi_t(x) \quad (3.20)$$

the last equality holds for $t - \tau$ small enough, where the characteristics originating from $d\psi_t$ do not intersect, hence the iterated minmax is nothing but the 1-step minmax which is the classical $C^2$ solution. We conclude by applying the same argument in proposition 3.13:

$$\bar{R}_0^k \psi_t(x) = \psi_t(x) - \int_{0}^{t} H(s, x, \partial_x \bar{R}_0^k \psi_t(x)) \, ds. \quad (3.21)$$

Subtracting (3.21) into (3.20), moving $\psi_t(x)$ to the right-hand side, dividing both side by $t - \tau$ and letting $\tau \to t$, we get

$$0 \leq -\partial_t \psi_t(x) - H(t, x, \partial_x \psi_t(x))$$
from which we get a subsolution by definition. Similarly, we can prove that \( R^n_H v(x) \) is a viscosity supersolution.

For given \( H \) and \( v \), we say that the limit of iterated minmax solutions exists in \([s, t]\) if for any sequence of subdivision \( \{\zeta_n\}_{n \in \mathbb{N}} \) of \([s, t]\) such that \( |\zeta_n| \to 0 \) as \( n \to \infty \), the related sequence of iterated minmax solutions \( \{R^n_H \zeta_n v(x)\}_{n \in \mathbb{N}} \) converges uniformly on compact subsets to a limit which is independent of the choice of subdivisions, then, without ambiguity, we denote this limit also by \( R^n_H v(x) \).

We can now prove our main theorem

**Theorem 3.19.** Suppose \( H \in C^2_c([0, T] \times T^*\mathbb{R}^d) \) and \( v \in C^{1,\text{Lip}}(\mathbb{R}^d) \), then for the Cauchy problem of the Hamilton–Jacobi equation

\[
\begin{align*}
\partial_t u + H(t, x, \partial_x u) &= 0, & t \in (0, T], \\
u(x, 0) &= v(x), & x \in \mathbb{R}^d
\end{align*}
\]

the limit of iterated minmax solutions exists and coincides with the viscosity solution.

**Proof.** Let \( K \subset \mathbb{R}^d \) and \( \{\zeta_n\}_n \) is any sequence of subdivisions of \([0, T]\) such that \( |\zeta_n| \to 0 \) as \( n \to \infty \). Denote \( u_n(t, x) := R^n_H \zeta_n v(x) \), and \( u(t, x) \) the viscosity solution of the (H-J) problem. If \( u_n \) does not converge uniformly on \([0, T] \times K \), there exists \( \epsilon > 0 \) and a subsequence \( n_k \) such that \( |u_{n_k} - u|_{[0,T] \times K} > \epsilon \). Note that \( \{\zeta_{n_k}\} \) is itself a sequence of subdivisions, this contradicts with proposition 3.18.

**Acknowledgments**

I am grateful to Marc Chaperon, who introduced me to this subject; I owe him the conjecture of iterated minmax and many useful discussions. I also thank Alain Chenciner for his help and discussions during the work. And I wish to thank Patrick Bernard for motivating the final proof of the main theorem, which simplifies the original one. After the first version of this article was posted on arXiv.org, Professor P E Souganidis kindly informed the author about his paper [Sou85] on the approximation schemes to viscosity solutions. It turns out that the minmax selector behaves like a 'generator' defined in [Sou85], and our limiting iterated minmax procedure fits into his general approximation schemes.

**Appendix A. Lipschitz critical point theory**

The notion of Clarke’s generalized derivative for Lipschitz functions has been well adapted in the Hamilton–Jacobi theory, such as the works in [FM07, Sic09, BdS12, Arn10]. We will give a brief review on Lipschitz critical point theory, extracted from appendix A in [Wei13]. Let us consider a real locally Lipschitz function \( f \) on \( X := \mathbb{R}^d \).

The **Clarke generalized derivative** \( \partial f(a) \) of \( f \) at \( a \in X \) is the convex subset \( \partial f(a) \) of \( X^* = T^*X \) defined as follows: by Rademacher’s theorem, the set \( \text{dom}(df) \) of differentiability points of \( f \) is dense in \( X \); if \( df := \{(x, df(x)) : x \in \text{dom}(df)\} \), we let

\[
\partial f(a) := \text{co}\{y \in X^* : (a, y) \in df\},
\]

where co stands for the convex hull; in other words, \( \partial f(a) \) is the convex hull of the set of limits of convergent sequences \( df(x_n) \) with \( \lim x_n = a \). As \( |df(x)| \) is bounded by the local Lipschitz constant of \( f \) for \( x \) close to \( a \), every sequence \( df(x_n) \) with \( \lim x_n = a \) is bounded and therefore has a convergent subsequence, implying

\[
\forall a \in X \quad \partial f(a) \neq \emptyset;
\]
moreover, $\partial f(a)$ is compact, being the convex hull of a compact subset. The subset
\[ \partial f := \{(x, y) | y \in \partial f(x), x \in X\} \]
is a generalized version of the *enlarged pseudograph* defined for semi-concave functions in [Arn11], where the pseudograph is $df$. In simple one-dimensional cases, it is obtained by adding a vertical segment to $df$ where $f$ is not differentiable:

**Remark A.1.** The set $\partial f(x)$ consists of a single point if and only if $f$ is $C^1$ at $x$ with respect to the set where it is differentiable.

**Proposition A.2.** The set-valued function $x \mapsto \partial f(x)$ is upper-semi-continuous: for every convergent sequence $(x_n, y_n) \to (x, y)$ with $y_n \in \partial f(x_n)$, one has $y \in \partial f(x)$. In other words, $\partial f$ is closed.

**Definition A.3.** A point $x \in X$ is called a critical point of $f$ if $0 \in \partial f(x)$; the number $f(x)$ is then called a critical value of $f$. By proposition A.2, the critical set $\text{Crit}(f)$ of $f$, consisting of its critical points, is closed in $X$.

Setting $\lambda(x) := \min_{w \in \partial f(x)} |w|_X$, we say that $f$ satisfies the *Palais-Smale condition* (PS) if every sequence $(x_n)$ along which $f(x_n)$ is bounded, and such that $\lambda(x_n)$ goes to 0, possesses a convergent subsequence—whose limit is a critical point of $f$ by proposition A.2, as there is a sequence $y_n \in \partial f(x_n)$ converging to 0.

**Example A.4.** The PS condition is satisfied when $\text{Lip}(f - Q) < \infty$ for some nondegenerate quadratic form $Q$ on $X$; moreover, in that case, $\text{Crit}(f)$ is compact.

**Proof.** Indeed, if $\psi := f - Q$, each subset $\partial f(x) = \partial \psi(x) + dQ(x)$ consists of vectors whose norm is at least $|dQ(x)| - \text{Lip}(\psi)$, hence $\lambda(x) \geq |dQ(x)| - \text{Lip}(\psi)$, which tends to $+\infty$ when $|x| \to \infty$; therefore, there exists $R > 0$ such that every sequence $(x_n)$ with $\lim \lambda(x_n) = 0$ satisfies $|x_n| \leq R$ for all large enough $n$, implying both the PS condition and the compactness of $\text{Crit}(f)$. \hfill $\Box$

**Theorem A.5 (Deformation lemma I).** Suppose $f$ satisfies the PS condition and let $f^c := \{x | f(x) \leq c\}$ for each $c \in \mathbb{R}$. If $c$ is not a critical value of $f$, then there exist $\epsilon > 0$ and a bounded smooth vector field $V$ on $X$ equal to 0 off $f^{c+2\epsilon} \setminus f^{-2\epsilon}$, and whose flow $\phi^1_V$ satisfies $\phi^1_V(f^{c+\epsilon}) \subset f^{-\epsilon}$.

**Theorem A.6 (Deformation lemma II).** Suppose $f$ satisfies the PS condition. If $c \in \mathbb{R}$ is a critical value of $f$ and $N$ any neighbourhood of $K_c := \text{Crit}(f) \cap f^{-1}(c)$, then there exist $\epsilon > 0$ and a bounded smooth vector field $V$ on $X$ equal to 0 off $f^{c+2\epsilon} \setminus f^{-2\epsilon}$, whose flow $\phi^1_V$ satisfies $\phi^1_V(f^{c+\epsilon} \setminus N) \subset f^{-\epsilon}$.

**Lemma A.7 (Chain rule).** If $f : X \to \mathbb{R}$ is a Lipschitz function, $F : X \to X$ a $C^1$ diffeomorphism, then
\[ \partial(f \circ F)(x) = \partial f(F(x)) \circ dF(x) := \{dF(x)(\xi), \xi \in \partial f(F(x))\}. \]

**Lemma A.8.** If $f, g : X \to \mathbb{R}$ are Lipschitz functions, then
\[ \partial(fg)(x) \subset f(x)\partial g(x) + g(x)\partial f(x). \]
Appendix B. An example

Let us look at a simple example where the minmax and the viscosity solution differ to see how the iterated minmax converges to the latter. Consider the (H–J) problem for a conservation law in one space dimension, i.e. $H$ depends only on $p$. In this case, the minmax$^{10}$ is given by

$$R_t v(x) = \inf \max S_t(x; x_0, y_0) := \inf \max (v(x_0) + x y_0 - t H(y_0) - x_0 y_0)$$

and the geometric solution and wave front are given by

$$L_t = \phi_t(\partial v) = \{(x, y_0) | x = x_0 + t H'(y_0), y_0 \in \partial v(x_0)\},$$
$$F^t = F^t(v) = \{(x, S_t(x; x_0, y_0)) | x = x_0 + t H'(y_0), y_0 \in \partial v(x_0)\}.$$

For an example, we take $H(p) = -p^3 + p^2 + p$ and the initial function $v(x)$ globally Lipschitz and in a neighbourhood of $x = 0$,

$$v(x) = \begin{cases} -x(x - 1), & x \leq 0, \\ x(x - 1), & x \geq 0. \end{cases}$$

See figure B1, where $p_{\pm 0} = \phi'(0 \pm)$. For $t > s > 0$ small, the geometric solution and wave fronts are depicted as in figures B2–B5. If $A$ is a subset of $\mathbb{R}$, we denote

$$F^t_A := F^t|_{x_0 \in A} = \{(x, S_t(x; x_0, y_0)) \in F^t, x_0 \in A\}.$$

$^{10}$ By the finite propagation speed property, the minmax exists when $v$ is globally Lipschitz and $H(p)$ is $C^1$. 

---

**Figure B1.** $v$ and $H$.

**Figure B2.** $F^t(v)$.

**Figure B3.** $\phi'(\partial v)$.
The presence at each step of these new pieces implies that the minmax does not form a semi-group and that the viscosity solution is not contained in the geometric solution \( \mathcal{F}(v) \), see figure 6: the part not belonging to the geometric solution comes, through the limiting process, from the singularities of the derivative which appear at each step of the iteration.

References

[Arn88] Arnol’d V I 1988 Geometrical Methods in the Theory of Ordinary Differential Equations (Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] vol 250) 2nd edn (New York: Springer)

[Arn00] Arnold V I 2000 Lectures on Partial Differential Equations (New York: Springer)

[Arn10] Arnaud M-C 2010 On a theorem due to Birkhoff Geom. Funct. Anal. 20 1307–16

[Arn11] Arnaud M-C 2011 Pseudographs and the Lax–Oleinik semi-group: a geometric and dynamical interpretation Nonlinearity 24 71–8

[BC11] Bernard O and Cardin F 2011 On \( C^0 \)-variational solutions for Hamilton–Jacobi equations Discrete Contin. Dyn. Syst. 31 385–406

[BdS12] Bernard P and dos Santos J O 2012 A geometric definition of the Mañé–Mather set and a theorem of Marie–Claude Arnaud Math. Proc. Cambr. Phil. Soc. 152 167–78
[Ber11] Bernard P 2011 The Lax–Oleinik semi-group: a Hamiltonian point of view CANPDE crash-course, Edinburgh

[Bru91] Brunella M 1991 On a theorem of Sikorav Enseign. Math. (2) 37 83–7

[CEL84] Crandall M G, Evans L C and Lions P-L 1984 Some properties of viscosity solutions of Hamilton–Jacobi equations Trans. Am. Math. Soc. 282 487–502

[Cha90] Chaperon M 1990 Familles génératrices Cours donné à l’école d’été Erasmus de Samos www.math.jussieu.fr/~chaperon/prepublications.html

[Cha91] Chaperon M 1991 Lois de conservation et géométrie symplectique C. R. Acad. Sci. Paris Sér. I Math. 312 345–8

[Che74] Chenciner A 1974 Aspects géométriques de l’études des chocs dans les lois de conservation Problèmes d’évolution non linéaires, Séminaire de Nice

[Che96] Chekanov Yu V 1996 Critical points of quasifunctions, and generating families of Legendrian manifolds Funkt. Anal. Prilozhen. 30 56–69 96

[CL87] Crandall M G and Lions P-L 1987 Remarks on the existence and uniqueness of unbounded viscosity solutions of Hamilton–Jacobi equations Illinois J. Math. 31 665–88

[CEL83] Clarke F H 1983 Optimization and Nonsmooth Analysis (Canadian Mathematical Society Series of Monographs and Advanced Texts) (New York: Wiley)

[CV08] Cardin F and Viterbo C 2008 Commuting Hamiltonians and Hamilton–Jacobi multi-time equations Duke Math. J. 144 235–84

[Fat] Fathi A 2012 Weak KAM Theorem in Lagrangian Dynamics (Cambridge: Cambridge University Press) [FM07] Fathi A and Maderna E 2007 Weak KAM theorem on non compact manifolds NoDEA Nonlinear Diff. Eqns Appl. 14 1–27

[FS06] Fleming W H and Soner H M 2006 Controlled Markov Processes and Viscosity Solutions vol 25 (New York: Springer)

[Hop65] Hopf E 1965 Generalized solutions of non-linear equations of first order J. Math. Mech. 14 951–73

[Joe93] Joukovskaja T 1993 Singularité de minmax et solutions faibles d’équations aux dérivées partielles PhD Thesis Université de Paris 7

[Kru70] Kruzkov S N 1970 First order quasilinear equations with several independent variables Mat. Sb. (N.S.) 81 228–5

[Lax73] Lax P D 1973 Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves (Philadelphia, PA: Society for Industrial and Applied Mathematics) Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics No 11

[Lao82] Lions P-L 1982 Generalized Solutions of Hamilton–Jacobi Equations (Research Notes in Mathematics vol 69) (Advanced Publishing Program) (Boston, MA: Pitman)

[PPS03] Paternain G P, Polterovich L and Siburg K F 2003 Boundary rigidity for Lagrangian submanifolds, non-removable intersections, and Aubry–Mather theory Mosc. Math. J. 3 593–619 745 Dedicated to Vladimir I. Arnold on the occasion of his 65th birthday

[Sic09] Siconolfi A 2009 Hamilton–Jacobi equations and weak KAM theory Encyclopedia of Complexity and Systems Science pp 4540–61

[Sık87] Sikorav J-C 1987 Problèmes d’intersections et de points fixes en géométrie hamiltonienne Comment. Math. Helv. 62 345–8

[Sou85] Souganidis P E 1985 Approximation schemes for viscosity solutions of Hamilton–Jacobi equations J. Diff. Eqns 59 1–43

[Thé96] Théret D 1996 Utilisation des fonctions génératrices en géométrie symplectique globale PhD Thesis Université de Paris 7

[Thé99] Théret D 1999 A complete proof of Viterbo’s uniqueness theorem on generating functions Topology Appl. 96 249–66

[Vit96] Viterbo C 1996 Solutions d’équations d’Hamilton–Jacobi et géométrie symplectique. Séminaire sur les équations aux dérivées partielles, Palaiseau

[VO95] Viterbo C and Otiolongh A 1995 Variational solutions of Hamilton–Jacobi equations http://math.polytechnique.fr/~viterbo

[Wei13] Wei Q L 2013 Solutions de viscosité des équations de Hamilton–Jacobi et minmax itérés PhD Thesis Université de Paris 7