0.1 Introduction

As previously mentioned, in many instances the value function arising from an optimal control problem may fail to be continuously differentiable. If that happens the derivation of the Hamilton-Jacobi equation is no longer valid, but more importantly the notion of classical solution to it does hold anymore. Therefore, we have to weaken the notion of solution in order to get a consistent and unique solution to the dynamic programming equation for non-differentiable value functions. The *viscosity solution* is exactly what we are searching for. It arises from a standard procedure called vanishing viscosity, which allows us to compute the solution of a fully non-linear first order PDE as the limiting solution of quasilinear parabolic PDEs, obtained via infinitesimal perturbations of second order derivatives.

0.1.1 Non-differentiable value functions

Let us consider the calculus of variation problem:

$$\inf_{x\in Lip([0,1];[-1,1])} \int_t^{t_1} 1 + \frac{1}{4}(\dot{x}(s))^2 \, ds, \tag{1}$$

where $Lip(I;U)$ is the collection of Lipschitz continuous functions from I to U. The Hamiltonian related to this problem is:

$$H(t,x,p) = \max_{v\in[-1,1]} \left\{ -v \cdot p - 1 - \frac{1}{4}v^2 \right\}.$$

We can explicitly compute the Hamiltonian and get:

$$H(t,x,p) = p^2 - 1.$$

Then the Hamilton-Jacobi equations read:

$$\begin{cases} \dot{x}^*(s) = -H_p(s, x^*(s), p^*(s)) = 2p^*(s) \\ \dot{p}^*(s) = H_x(s, x^*(s), p^*(s)) = 0, \end{cases}$$

therefore, we get:

$$\dot{x}(s)^* = 2p^*, \ s \in [0,1],$$

for some $p^* \in \mathbb{R}$. We now compute the exit time of $(s, x(s)) = (s, 2(s-t)p^* + x)$ with initial data (t,x). If $p = 0$ then:

$$\tau = 1, \ |x| < 1.$$

If $p > 0$ then $x(s) = 2(s-t)p + x$ is increasing, which implies that the system is going to exit from the right boundary, that is from $x(s) = 1$, and if that happens before time $s = 1$ the exit time will be determined by:

$$2(s-t)p + x = 1 \Rightarrow s = t + \frac{1-x}{2p}.$$

$x(s) = 1$ for $s < 1$ if:

$$2(1-t)p + x \geqslant 1 \Rightarrow p \geqslant t + \frac{1-x}{2p},$$

therefore:

$$\tau = \begin{cases} 1 & p \geqslant t + \frac{1-x}{2p} \\ t + \frac{1-x}{2p} & p > t + \frac{1-x}{2p}. \end{cases}$$

Analogously, if $p < 0$:

$$\tau = \begin{cases} 1 & p \leqslant t - \frac{1+x}{2p} \\ t - \frac{1+x}{2p} & p < t - \frac{1+x}{2p}. \end{cases}$$

We now solve:

$$\inf_{p \in \mathbb{R}} \int_t^\tau 1 + p^2 \, ds = \inf_{p \in \mathbb{R}} (1+p^2)(\tau - t) = \begin{cases} (1+p^2)(1-t), & p = 0 \text{ or } p > 0 \wedge p \geqslant \frac{1-x}{2(1-t)} \text{ or } p < 0 \wedge p \leqslant \frac{-1-x}{2(1-t)} \\ (1+p^2)\frac{1-x}{2p}, & p > 0 \wedge p \geqslant \frac{1-x}{2(1-t)} \\ -(1+p^2)\frac{1+x}{2p}, & p \leqslant \frac{-1-x}{2(1-t)}, \end{cases}$$

which is solved as follows:

$$V(t,x) = \begin{cases} 1-t & |x| \leqslant t \\ 1-t & |x| \geqslant t, \end{cases}$$

which is continuous on the whole space, but clearly no differentiable in $|x| = t$.

0.1.2 Vanishing viscosity

We now euristichally expose a technique called vanishing viscosity, which is widely used in calculus of variations problems and will show us one of the origins of the viscosity solution notion. Let us consider the initial value problem:

$$\begin{cases} u_t + H(u, Du) = 0, & \mathbb{R}^n \times (0, +\infty) \\ u = g, & \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{2}$$

The method of characteristics shows that there cannot be a smooth solution of the above problem over the whole positive real line. Indeed, a weaker notion of solution is needed. One approach is to use Hopf-Lax solution concept. We are not interested in it, instead we start by perturbing the system as:

$$\begin{cases} u_t^\epsilon + H(u^\epsilon, Du^\epsilon) - \epsilon \Delta u^\epsilon = 0, & \mathbb{R}^n \times (0, +\infty) \\ u^\epsilon = g, & \mathbb{R}^n \times \{t = 0\}, \end{cases} \tag{3}$$

so that the fully non-linear system in 2 becomes a semilinear one, which turn out to have a smooth solution. Then, we take $\epsilon \to 0$. We expect the solution u^ϵ to lose the bounds on the derivatives, as they strongly depend on the regularization effect of $\epsilon \Lambda$. Turns out that many times Ascoli-Arzela theorem's hypotheses are satisfied, that is $(u^\epsilon)_\epsilon$ is uniformly bounded and equicontinuous, then we have local uniform convergence along a subsequence u^{ϵ_j}. We now use the limit $u \xleftarrow{j \to +\infty} u^{\epsilon_j}$ as a solution. We now it to be continuous but we lack information about its derivatives. We will then verify these information using test functions. Unlike the classical variational weak solution concept, where integration by part play the central role, we will use the maximum principle to translate the derivates of u onto the test functions.

Let us take $v \in C^\infty(\mathbb{R}^n \times (0, +\infty))$ and suppose that $u - v$ has a strict local maximum at (x_0, t_0), then:

$$(u - v)(x_0, t_0) > (u - v)(x, t),$$

for all (t, x) sufficiently close to (x_0, t_0). It can be shown that it implies that there exists $J > 0$ such that for all $j > J$ there exists $(x_{\epsilon_j}, t_{\epsilon_j})$ such that:

$$(u^\epsilon - v)(x_{\epsilon_j}, t_{\epsilon_j}) \geqslant (u^\epsilon - v)(x, t),$$

for (x, t) sufficiently close to $(x_{\epsilon_j}, t_{\epsilon_j})$ and such that:

$$(x_{\epsilon_j}, t_{\epsilon_j}) \xrightarrow{j \to +\infty} (x_0, t_0).$$

Because $(u^\epsilon - v)$ has a local maximum at $(x_{\epsilon_j}, t_{\epsilon_j})$:

$$Du^{\epsilon_j}(x_{\epsilon_j}, t_{\epsilon_j}) = Dv(x_{\epsilon_j}, t_{\epsilon_j}),\ u_t^{\epsilon_j}(x_{\epsilon_j}, t_{\epsilon_j}) = v(x_{\epsilon_j}, t_{\epsilon_j}),$$

and:

$$-\Lambda u^{\epsilon_j}(x_{\epsilon_j}, t_{\epsilon_j}) \geqslant -\Lambda v(x_{\epsilon_j}, t_{\epsilon_j}).$$

Therefore, we get:

$$v_t(x_{\epsilon_j}, t_{\epsilon_j}) + H(Dv(x_{\epsilon_j}, t_{\epsilon_j}), x_{\epsilon_j}) \leqslant \Lambda v(x_{\epsilon_j}, t_{\epsilon_j}) \xrightarrow{j \to +\infty} 0$$

Analogous computations can be done for local minimum of $u - v$, obtaining the opposite inequality above.

We can now grasp the intuition behind the following definition.

Definition 0.1.1. *A viscosity solution of 2 is a function u bounded and uniformly continuous on $\mathbb{R}^n \times [0, T]$ for all $T > 0$ such that for all $v \in C^{+\infty}(\mathbb{R}^n \times (0, +\infty))$:*

$$v_t(x, t) + H(Dv(x, t), x) \leqslant 0$$

for all $(x, t) \in \arg\max\{u - v\}$ and:

$$v_t(x, t) + H(Dv(x, t), x) \geqslant 0$$

for all $(x, t) \in \arg\min\{u - v\}$. Furthermore, $u \equiv g$ for $t = 0$.

0.1.3 Abstract dynamic programming and viscosity solutions

We now present an abstraction of the dynamic programming principle, which will allow us to define the viscosity solutions of the dynamic programming equation. Let Σ be a closed subset of a Banach space and $\mathcal{C}$ a collection of functions on Σ, closed under addition:

$$\phi, \psi \in \mathcal{C} \Rightarrow \phi + \psi \in \mathcal{C}.$$

We consider the family of operators $\{\mathcal{T}_{tr}\}_{t_0 \leqslant t \leqslant r \leqslant t_1}$ such that:

$$\mathcal{T}_{tt}\phi = \phi,\ \forall\, \phi \in \mathcal{C}, \tag{4}$$

$$\mathcal{T}_{tr}\phi \leqslant \mathcal{T}_{ts}\psi \text{ if } \phi \leqslant \mathcal{T}_{rs}\psi, \tag{5}$$

and:

$$\mathcal{T}_{tr}\phi \geqslant \mathcal{T}_{ts}\psi \text{ if } \phi \geqslant \mathcal{T}_{rs}\psi. \tag{6}$$

Conditions 5 and 6 are a weaker version of monotonicity; they imply it together with 4. Moreover, they also imply the semigroup property, provided that $\mathcal{T}_{rt} : \mathcal{C} \to \mathcal{C}$. Under this assumption, the two conditions are equivalent to monotonicity. The semigroup property:

$$\mathcal{T}_{tr}\left(\mathcal{T}_{rs}\psi\right) = \mathcal{T}_{ts}\psi, \ \mathcal{T}_{rs}\psi \in \mathcal{C}, \tag{7}$$

is going to be the dynamic programming principle. Indeed, let us consider the classical optimal control problem defined on a bounded set $O \subset \mathbb{R}^n$, which we set to be $\Sigma = \overline{O}$ and $\mathcal{C} = \mathcal{M}(\Sigma)$, the collection of measurable functions bounded by below. Then as in chapter 1 we aim at minimize a functional, we set this functional to be the operator $\mathcal{T}$. Let us define:

$$\mathcal{T}_{t,r;u}\psi(x) = \int_t^{\tau \wedge r} L(s, x(s), u(s)), \ ds + g(\tau, x(\tau))\chi_{\tau < r} + \psi(x(r))\chi_{\tau \geqslant r}, \tag{8}$$

which gives:

$$\mathcal{T}_{tr}\psi = \inf_{u \in \mathcal{U}(t,x)} \mathcal{T}_{t,r;u}\psi. \tag{9}$$

Under the usual assumption on the running and terminal costs, as well as on the control space U we now that the value function defined in ref4-1-valuefunctreform is measurable and bounded by below. Therefore, $\mathcal{T}_{rt} : \mathcal{C} \to \mathcal{C}$ and we can formulate the semigroup property just by asking $\psi \in \mathcal{C}$. It is clear by its definition that the dynamic programming principle is translated as:

$$\mathcal{T}_{tt_1}\psi(x) = \mathcal{T}_{tr}\left(\mathcal{T}_{rt_1}\psi\right)(x),$$

for $(t,x) \in \overline{Q}$ and $\psi \in \mathcal{C}$. We now derive the abstract dynamic programming equation. The same procedure as in chapter 1 gives:

$$-\frac{1}{h}\left[\mathcal{T}_{tt+h}V(t+h, \cdot)(x) - V(t,x)\right] = 0.$$

What happens if we let $h \to 0$? We ask for the existence of a family of non-linear operators which will play the role of the Hamiltonian. Let $\Sigma' \subset \Sigma$ and $\mathcal{D} \subset C([t_0, t_1] \times \Sigma')$ and $\{\mathcal{G}_t\}_{t \in [t_0, t_1]}$ functions on Σ such that:

$$\lim_{h \to 0} \frac{1}{h}\left[\mathcal{T}_{tt+h}V(t+h, \cdot)(x) - V(t,x)\right] = \frac{\partial}{\partial t}w(t,x) - (\mathcal{G}_t w(t, \cdot))(x) \tag{10}$$

for all $w \in \mathcal{D}, (t,x) \in Q = [t_0, t_1) \times \Sigma'$. The space $\mathcal{D}$ is such that:

$$\forall \ w \in \mathcal{D} \text{ the functions } \frac{\partial w}{\partial t}, \ \mathcal{G}_t w(t, \cdot) \text{ are continuous on } Q \text{ and } w(t, \cdot) \in \mathcal{C} \ \forall \ t \in [t_0, t_1], \tag{11}$$

and $\mathcal{D}$ is a vector space:

$$w, \tilde{w} \in \mathcal{D} \Rightarrow w + \tilde{w} \in \mathcal{D}, \ \lambda w \in \mathcal{D}. \tag{12}$$

The elements of $\mathcal{D}$ are called test functions and $\mathcal{G}_t$ the infinitesimal generator of $\mathcal{T}_{tr}$. Explicit choices of $\mathcal{C}$ and $\mathcal{D}$ will vary from case to case, usually they are chosen to satisfy certain integrability conditions on the functions.

If we require the existence of a test function space and an infinitesimal generator of the semigroup given by the value function, the dynamic programming equation becomes:

$$-\frac{\partial}{\partial t}V(t,x) + (\mathcal{G}_t V(t, \cdot))(x) = 0, \ (t,x) \in Q. \tag{13}$$

Then if a function $V \in \mathcal{D}$ satisfies 13 point wise it is called a classical solution of it. Thanks to this reformulation we will be able to weaken this notion of solution and finally get the viscosity solution of the dynamic programming equation.

We point out that in the canonical deterministic optimal control studied in chapter 1, the infinitesimal generator has the form:

$$(\mathcal{G}_t \phi)(x) = H(t, x, D\phi(x)) = \sup_{v \in U} \{-f(t, x, v) \cdot D\phi(x) - L(t, x, v)\},$$

with test functions space $\mathcal{D} = C^1(Q) \cap \mathcal{M}(\overline{Q})$ and $\Sigma' = O$.

In view of what we saw in previous sections, we give the definition of viscosity solution.

Definition 0.1.2. *Let $W \in C([t_0, t_1] \times \Sigma)$. Then:*

1. *W is a viscosity subsolution of 13 in Q if for every $w \in \mathcal{D}$:*

$$-\frac{\partial}{\partial t} w\left(\bar{t}, \bar{x}\right) + \left(\mathcal{G}_{\bar{t}} w\left(\bar{t}, \cdot\right)\right)(\bar{x}) \leqslant 0, \tag{14}$$

 at every:

$$\left(\bar{t}, \bar{x}\right) \in \arg \max_{(t,x) \in Q} \{(W - w)(t, x)\},$$

 and $W\left(\bar{t}, \bar{x}\right) = w\left(\bar{t}, \bar{x}\right)$.

2. *W is a viscosity supersolution of 13 in Q if for every $w \in \mathcal{D}$:*

$$-\frac{\partial}{\partial t} w\left(\bar{t}, \bar{x}\right) + \left(\mathcal{G}_{\bar{t}} w\left(\bar{t}, \cdot\right)\right)(\bar{x}) \geqslant 0, \tag{15}$$

 at every:

$$\left(\bar{t}, \bar{x}\right) \in \arg \min_{(t,x) \in Q} \{(W - w)(t, x)\},$$

 and $W\left(\bar{t}, \bar{x}\right) = w\left(\bar{t}, \bar{x}\right)$.

3. *W is a viscosity solution if it is both a subsolution and a supersolution.*

As a first step after the definition of viscosity solutions, we prove its consistency with the classical notion. To do so we require the operator to satisfy a maximum principle, that is if $\mathcal{G}_t$ is a general operator and:

$$\mathcal{D} = \{W \in C([t_0, t_1] \times \Sigma) \mid W_t(t, x), (\mathcal{G}_t W(t, \cdot))(x) \in C(Q)\},$$

then $\mathcal{G}_t$ satisfies the *maximum principle* if:

$$\mathcal{G}_t \phi(\bar{x}) \geqslant \mathcal{G}_t \psi(\bar{x})$$

for every $\bar{x} \in \arg \max\{(\phi - \psi)(x) \mid x \in \Sigma\} \cap \Sigma'$ with $\phi(\bar{x}) = \psi(\bar{x})$. Now, if W is a classical solution of 13, then for a $w \in \mathcal{D}$ and $\left(\bar{t}, \bar{x}\right) \in \arg \max_{(t,x) \in Q} \{(W - w)(t, x)\}$ and $W\left(\bar{t}, \bar{x}\right) = w\left(\bar{t}, \bar{x}\right)$ then:

$$-\frac{\partial}{\partial t} w\left(\bar{t}, \bar{x}\right) + \left(\mathcal{G}_{\bar{t}} w\left(\bar{t}, \cdot\right)\right)(\bar{x}) \leqslant -\frac{\partial}{\partial t} W\left(\bar{t}, \bar{x}\right) + \left(\mathcal{G}_{\bar{t}} W\left(\bar{t}, \cdot\right)\right)(\bar{x}) = 0, \tag{16}$$

since we asked continuity of time derivatives. For the supersolution recall that:

$$\max\{(\phi - \psi)(x) \mid x \in \Sigma\} = \min\{-(\phi - \psi)(x) \mid x \in \Sigma\}.$$

If $\mathcal{G}_t$ is the infinitesimal generator of a two-parameter semigroup the connection between classical and viscosity solutions is even stronger.

Proposition 0.1.1. *Let $W \in \mathcal{D}$. Then W is a classical solution of 13 if and only if it is a viscosity solution of 13 in Q.*

Proof. If W is a viscosity solution, since it is also a test function then **??** and **??** hold for every point $(t, x) \in Q$, which implies that:

$$-\frac{\partial}{\partial t}w(t, x) + (\mathcal{G}_t w(t, \cdot))(x) = 0, \ \forall \ (t, x) \in Q.$$

If W is a classical solution and prove the subsolution property. Let $w \in \mathcal{D}$ and (t, x) as usual. Since $w \geqslant W$:

$$
\begin{aligned}
-\frac{\partial}{\partial t}w(t, x) + (\mathcal{G}_t w(t, \cdot))(x) &= -\lim_{h \to 0}\frac{1}{h}\left[(\mathcal{T}_{tt+h}w(t + h, \cdot))(x) - w(t, x)\right] \\
&\leqslant -\lim_{h \to 0}\frac{1}{h}\left[(\mathcal{T}_{tt+h}W(t + h, \cdot))(x) - W(t, x)\right] \\
&= -\frac{\partial}{\partial t}W(t, x) + (\mathcal{G}_t W(t, \cdot))(x) = 0.
\end{aligned}
$$

The supersolution property is proven similarly. $\qquad\square$

We now prove that the family of linear operators $\mathcal{T}_{tt_1}$ defined in 9 defines viscosity solution of the dynamic programming equation 13.

Theorem 0.1.2. *Let* $\{\mathcal{T}_{tr}\}_{t_0 \leqslant t \leqslant r \leqslant t_1}$ *such that* 4,6,5 *and also there exists a vector space* $\mathcal{D}$ *and another family of operator* $\{\mathcal{G}_t\}_{t \in [t_0, t_1]}$ *such that* 11 *and* 10 *hold. Let:*

$$V(t, x) = (\mathcal{T}_{tt_1}\psi)(x).$$

If $V \in C(Q)$ *then it is a viscosity solution of* 13.

Proof. Let us prove the subsolution condition. Let $w \in \mathcal{D}$ and (t, x) maximizer of $V - w$ in $\overline{Q}$ and $V(t, x) = w(t, x)$. Then $V \leqslant w$ on Q, which implies by 6 that we have:

$$(\mathcal{T}_{tr}w(r, \cdot))(x) \geqslant (\mathcal{T}_{tt_1}\psi)(x) = V(t, x) = w(t, x). \tag{17}$$

Now, because of 11 and 10 we can compute:

$$-\frac{\partial}{\partial t}w(t, x) + (\mathcal{G}_t w(t, \cdot))(x) = -\lim_{h \to 0}\frac{1}{h}\left[(\mathcal{T}_{tt+h}w(t + h, \cdot))(x) - w(t, x)\right] \leqslant 0$$

because of 17. The supersolution condition is proven analogously. $\qquad\square$

0.1.4 The Standard Approach

We have talked about *viscosity solutions* for $\mathcal{G}_t$ that are infinitesimal generators of semigroups. Historically the notion of viscosity solution was introduced for partial differential equations, that is when $\mathcal{G}_t$ is a partial differential operator. In this section we will introduce this classical notion and provide some properties, establishing as well the relevant links to the already proven theory.
Let $\mathcal{G}_t$ be a partial differential operator defined As

$$(\mathcal{G}_t\phi)(x) = F(t, x, D\phi(x), D^2\phi(x), \phi(x)), \tag{18}$$

where F is taken to be continuous. We assume that $\sigma' = O \subset \mathbb{R}^n$ open and σ to be its closure. Furthermore, we assume that test functions $\mathcal{C}$ and $\mathcal{D}$ are such that

$$C_p(\overline{O}) \cap \mathcal{M}(\overline{O}) \subset \mathcal{C}, \ C_p^\infty(\overline{Q}) \cap \mathcal{M}(\overline{Q}) \subset \mathcal{D}, \tag{19}$$

where $Q = [t_0, t_1) \times O$ as usual. A useful result is the following.

Lemma 0.1.3. *The operator* $\mathcal{G}_t$ *obeys to the maximum principle if and only if F is elliptic.*

Proof. Recall that F is elliptic If

$$F(t, x, p, A + B, V) \leqslant F(t, x, p, A, V) \tag{20}$$

for all $(t, x, p, V) \in Q \times \mathbb{R}^n \times \mathbb{R}$ and all $A, B \in Sym(\mathbb{R}^{n \times n})$ with $B \geqslant 0$.

Let us assume that F is elliptic, then for generic $\phi, \psi \in C^2(O)$ and $\overline{x} \in \arg\max\{(\phi - \psi)(x) | x \in \overline{O}\} \cap O$ with $\phi(\overline{x}) = \psi(\overline{x})$ we have that $D\phi(\overline{x}) = D\psi(\overline{x})$ and $D^2\phi(\overleftarrow{x}) \leqslant D^2\psi(\overleftarrow{x})$, therefore

$$F(t, \overline{x}, D\phi(\overline{x}), D^2\phi(\overline{x}), \phi(\overline{x})) \geqslant F(t, \overline{x}, D\psi(\overline{x}), D^2\psi(\overline{x}), \psi(\overline{x})),$$

just by using 20 with $B = D^\psi(\overline{x}) - D^2\phi(\overline{x}) \geqslant 0$. If F is not elliptic for some $(t, \overline{x}, p, A, V)$ then the functions

$$\psi(x) = V + p \cdot (x - \overline{x}) + \frac{1}{2}(A + B)(x - \overline{x}) \cdot (x - \overline{x}),$$

$$\phi(x) = V + p \cdot (x - \overline{x}) + \frac{1}{2}A(x - \overline{x}) \cdot (x - \overline{x}),$$

for $x \in \overline{O}$ are such that $\overline{x}$ maximizes the distance between them, they coincide over it, but the value of F in $\overline{x}$ and ϕ, ψ does not satisfy the maximum principle, since 20 does not hold. $\square$

I this context we give the following definition.

Definition 0.1.3. *Given a function* $W \in C(\overline{O})$ *we define the equation*

$$-\frac{\partial}{\partial t}W(t, x) + F(t, x, D_x W(t, x), D_x^2 W(t, x), W(t, x)) = 0 \tag{21}$$

(a) *W is a viscosity subsolution of* 21 *in Q if for each $w \in C^\infty(Q)$ then*

$$-\frac{\partial}{\partial t}w(t, x) + F(t, x, D_x w(t, x), D_x^2 w(t, x), w(t, x)) \leqslant 0 \tag{22}$$

at every $(t, x) \in Q$ which locally maximizes $W - w$ yielding 0.

(b) *W is a viscosity supersolution of* 21 *in Q if for each $w \in C^\infty(Q)$ then*

$$-\frac{\partial}{\partial t}w(t, x) + F(t, x, D_x w(t, x), D_x^2 w(t, x), w(t, x)) \geqslant 0 \tag{23}$$

at every $(t, x) \in Q$ which locally minimizes $W - w$ yielding 0.

(c) *W is a viscosity solution if it is both a viscosity subsolution and supersolution.*

We will prove that definitions 0.1.2 and 0.1.3 are equivalent for partial differential operators of the form 18, provided that the function has sub-polynomial growth and is measurable.

An handy way of verifying both the definitions is the following.

Lemma 0.1.4. *If the conditions in the definitions* 0.1.2, 0.1.3 *is satisfied only at strict extrema then it is satisfied at every extrema.*

Proof. We prove it for the minimum of definition 0.1.3, the other assertions are proved similarly. Let (t, x) be a non strict minimum of $W - w$ such that $W(t, x) = w(t, x)$. Let $\epsilon > 0$ and deifne $w^\epsilon = w + \epsilon \xi$ where

$$\xi(s, y) = e^{-(|s-t|^2 + |y-x|^2)} - 1$$

defined over $\overline{Q}$. Since $\mathcal{D}$ is a vector space and $\epsilon \xi \in \mathcal{D}$ by 19, then $w^\epsilon \in \mathcal{D}$. Then for every $\epsilon > 0$ the function $W - w^\epsilon$ has a strict minimum at (t, x) with $W(t, x) = w^\epsilon(t, x)$ which implies that

$$-\frac{\partial}{\partial t}w^\epsilon(t,x) + F(t,x,D_x w^\epsilon(t,x), D_x^2 w^\epsilon(t,x), w^\epsilon(t,x)) \leqslant 0.$$

By continuity of F and smoothness of w^ϵ we let $\epsilon \to 0$ and get the wanted inequality. $\qquad \square$

Finally, we prove the equivalence of these definitions.

Theorem 0.1.5. *Let all the previous assumptions and $W \in C_p(\overline{Q}) \cap \mathcal{M}(\overline{Q})$ and $\mathcal{D} \subset C^{1,2}(Q)$. Then W is a subsolution (supersolution) of 21 as defined by 0.1.2 if and only if it is a subsolution (supersolution) as defined by 0.1.3.*

Proof. Let W be a supersolution as in 0.1.2. We just prove the supersolution property, since if we have W subsolution we can see it as a supersolution. Let W a subsolution and (t,x) a strict maximum of $W - w$ and $W(t,x) = w(t,x)$, then we can find an open set A such that $w > W$ over $A \cap Q$. By sub-polynomial growth there exists a constant K such that

$$|W(s,y)| \leqslant K(1 + |y|^{2m}), \quad \forall \, (s,y) \in \overline{Q}$$

for some $m > 0$. Then for some ξ given by Urysohn's lemma we can define

$$\overline{w}(s,y) = \xi(s,y)w(s,y) - (1 - \xi(s,y))K(1 + |y|^{2m}),$$

and get a minimum for $W - \overline{w}$ in (t,x). Then we are in the supersolution case.
Let $w \in C^\infty(Q)$ and (t,x) a local minimum of $W - w$ where it is null. There exists an open subset $\mathcal{N}$ of $\mathbb{R}^n$ such that $W \geqslant w$ over $\mathcal{N} \cap \overline{Q}$, and by Urysohn's lemma there exists a function $\xi \in C^\infty(\overline{Q})$ taking values in $[0,1]$, constantly 1 in a neighborhood $\tilde{\mathcal{N}}$ of (t,x) and zero outside $\mathcal{N}$. Since W is bounded below we define

$$\tilde{w}(s,y) = \xi(s,y)w(s,y) - (1 - \xi(s,y))K, \tag{24}$$

where $-K$ is the lower bound of W. Then, by hypotheses on W, w, and ξ we have

$$W = \xi W + (1 - \xi)W \geqslant \xi w + (1 - \xi)W \geqslant \xi w - (1 - \xi)K = \overline{w}.$$

Furthermore, $\overline{w} \in \mathcal{D}$ since 19 holds, (t,x) is a minimum of $W - \overline{w}$, and $W(t,x) = \overline{w}(t,x)$ then

$$-\frac{\partial}{\partial t}\overline{w}(t,x) + F(t,x,D_y\overline{w}(t,x), D_y^2\overline{w}(t,x), \overline{w}(t,x)) \geqslant 0,$$

but $\overline{w}$ and w have same derivative in (t,x), since ξ is constantly unitary in a neighborhood of it. Then we get

$$-\frac{\partial}{\partial t}w(t,x) + F(t,x,D_y w(t,x), D_y^2 w(t,x), w(t,x)) \geqslant 0.$$

Let us suppose that W is a subsolution for 0.1.3. Let $w \in \mathcal{D}$ and (t,x) a strict maximum of $W - w$ such that $(W - w)(t,x) = 0$. Since $\mathcal{D} \subset C^{1,2}(Q)$ we can find an open set Q^* over which w is $C^{1,2}$ and then extend it to be null over $\mathbb{R}^{n+1}\backslash Q^*$. Let us define a sequence of mollifications of w defined as w^n such that $w^n, w_t^n, w_{x_i}^n, w_{x_i x_j}^n$ converge to their respective of w. Since the convergence of w^n to w is uniform there must be a subsequence (t_n, x_n) converging to (t,x), such that $W - w^n$ has a maximum. If we set $\overline{w}^n = w^n - w^n(t_n, x_n) + W(t_n, x_n)$ we still have a maximum in (t_n, x_n) for $W - \overline{w}^n$ and $W(t_n, x_n) = \overline{w}^n(t_n, x_n)$. Then

$$-\frac{\partial}{\partial t}\overline{w}^n(t_n,x_n) + F(t_n,x_n,D_x\overline{w}^n(t_n,x_n), D_x^2\overline{w}^n(t_n,x_n), \overline{w}^n(t_n,x_n)) \leqslant 0.$$

Because of continuity of F and smoothness of $\overline{w}^n$ we get

$$-\frac{\partial}{\partial t}w(t,x) + F(t,x,D_y w(t,x), D_y^2 w(t,x), w(t,x)) \leqslant 0.$$

$\qquad \square$

0.1.5 Value Function as Viscosity Solution

In the first section we proved that the value function of a dynamic programming problem is the classical solution of the dynamic programming equation, under differentiability assumptions. We now prove that it is a viscosity solution of the dynamic programming equation given two sets of assumptions.

Theorem 0.1.6. *Let U be a bounded space of control and $f \in C(\overline{Q} \times U)$ such that $|f(t, x, v)| \leqslant K(1 + |x|)$. The for every $w \in C^1(Q) \cap \mathcal{M}(\overline{Q})$*

$$\lim_{h \to 0} \frac{1}{h} \left[(\mathcal{T}_{tt+h} w(t + h, \cdot))(x) - w(t, x) \right] = \frac{\partial}{\partial t} w(t, x) - H(t, x, D_x w(t, x)), \ \forall \ (t, x) \in \overline{Q}. \quad (25)$$

Note that equation 0.1.6 implies that V is a viscosity solution of **??** via theorem 0.1.2.

Proof. Let $(t, x) \in Q$, $v \in U$ and $w \in C^1(Q) \cap \mathcal{M}(\overline{Q})$. We prove that

$$\limsup_{h \to 0} \frac{1}{h} \left[(\mathcal{T}_{tt+h} w(t + h, \cdot))(x) - w(t, x) \right] \leqslant \frac{\partial}{\partial t} w(t, x) - H(t, x, D_x w(t, x)), \quad (26)$$

and

$$\liminf_{h \to 0} \frac{1}{h} \left[(\mathcal{T}_{tt+h} w(t + h, \cdot))(x) - w(t, x) \right] \geqslant \frac{\partial}{\partial t} w(t, x) - H(t, x, D_x w(t, x)). \quad (27)$$

Equation 26 can be proven without assuming boundedness of the control space U and sub-linearity of the flow f. Indeed, there exists an admissible control $u(\cdot) \in \mathcal{U}(t, x)$ such that $\lim_{s \to t} u(s) = v$. Let $x(\cdot)$ be the state solution of the Cauchy problem associated to the control $u(\cdot)$ such that $x(t) = x$. Let also τ the exit time of $(s, x(s))$ from Q. If we take a sufficiently small $h > 0$ we can compute

$$\frac{1}{h} \left[(\mathcal{T}_{tt+h} w(t + h, \cdot))(x) - w(t, x) \right] \leqslant \frac{1}{h} \int_t^{t+h} L(s, x(s), u(s)) \, ds + \frac{1}{h} \left[w(t + h, x(t + h)) - w(t, x) \right]$$

$$= \frac{1}{h} \int_t^{t+h} \left[L(s, x(s), u(s)) \right.$$

$$\left. + \frac{\partial}{\partial t} w(s, x(s)) + f(s, x(s), u(s)) \cdot D_x w(s, x(s)) \right] ds.$$

Then 26 holds true.

Let us prove 27. Using $|f(t, x, v)| \leqslant K(1 + |x|)$ we get

$$|x(r) - x| \leqslant \int_r^t |f(s, x(s), u(s))| \, ds$$

$$\leqslant K(r - t) + K \int_t^r |x|(s) \, ds$$

$$\leqslant K(r - t)(1 + |x|) + K \int_t^r |x(s) - x| \, ds,$$

which implies by Gronwall's inequality

$$|x(r) - x| \leqslant (1 + |x|) e^{K(t-r) - 1}, \ \forall \ r \geqslant t.$$

Then for suitably small r that $\tau \geqslant r \wedge t_1$ for all $u(\cdot) \in \mathcal{U}(t, x)$. Then for n sufficiently large by definition of $\mathcal{T}_{tr}$ we can find a control $u^n(\cdot)$ such that

$$\left(\mathcal{T}_{tt+\frac{1}{n}}w\left(t+\frac{1}{n}\right)\right)(x) \geqslant \int_t^{t+1/n} L(s,x(s),u(s))\,ds + w\left(t+\frac{1}{n},x^n\left(t+\frac{1}{n}\right)\right) - \frac{1}{n^2},$$

which implies that

$$n\left[\left(\mathcal{T}_{tt+\frac{1}{n}}w\left(t+\frac{1}{n}\right)\right)(x) - w(t,x)\right] \geqslant \frac{\partial}{\partial t}w(t,x) + n\int_t^{t+1/n} L(s,x(s),u^n(s))\,ds$$

$$+ n\int_t^{t+1/n} f(t,x,u^n(s))\,ds \cdot D_x w(t,x) + e(n), \tag{28}$$

where

$$e(n) = -\frac{1}{n} + n\int_t^{t+1/n} \frac{\partial}{\partial t}\big(w(s,x^n(s)) - w(t,x)\big)\,ds$$

$$+ n\int_t^{t+1/n} L((s,x^n(s),u^n(s))) - L(t,x,u^n(s))\,ds \tag{29}$$

$$+ n\int_t^{t+1/n} f(s,x^n(s),u^n(s)) \cdot D_x w(s,x^n(s)) - f(t,x,u^n(s)) \cdot D_x w(t,x)\,ds.$$

Because of the continuity of L, of f and smoothness of w we have that $\lim_{n\to+\infty} e(n) = 0$. If we prove that

$$n\int_t^{t+1/n} L(s,x(s),u^n(s))\,ds + n\int_t^{t+1/n} f(t,x,u^n(s))\,ds \cdot D_x w(t,x) \leqslant \sup_{v\in U}\{f(t,x,v)\cdot p + L(t,x,v)\}, \tag{30}$$

then from 28 we get

$$\liminf_{h\to 0} \frac{1}{h}\left[(\mathcal{T}_{tt+h}w(t+h))(x) - w(t,x)\right] \geqslant \frac{\partial}{\partial t}w(t,x) - H(t,x,D_x w(t,x)).$$

PROVE THAT 30 VIA CONVEXHULL STUFF. $\qquad\square$

It is important to notice that the previous result holds under quite stringent hypotheses. We both require the space control to be bounded and the flow f to have sublinear growth in its state variable. We can relax these assumptions and get a similar result, but we need to require the existence of an optimal control.

Theorem 0.1.7. *If for each $(t,x)\in Q$ there exists a $u^*\mathcal{U}(t,x)$ be an optimal control, then a continuous value function is a viscosity solution of its dynamic programming equation.*

Proof. Let us drop the $*$. As we did before in theorem 0.1.6, for every $w\in C^\infty(Q)$ we have

$$\limsup_{h\to 0} \frac{1}{h}\left[(\mathcal{T}_{tt+h}w(t+h,\cdot))(x) - w(t,x)\right] \leqslant \frac{\partial}{\partial t}w(t,x) - H(t,x,D_x w(t,x)).$$

If $V - w$ has a maximum at (t,x) and $V(t,x) = w(t,x)$ then

$$\limsup_{h\to 0} \frac{1}{h}\left[\underbrace{\int_t^{t_1} L(s,x(s),u(s))\,ds}_{=V(t,x)-w(t,x)=0} + w(t+h,x) - V(t+h,x)\right] = 0,$$

which proves subsolution. Let the subsolution hypotheses hold, then from the dynamic programming principle we get

$$0 \geqslant \int_t^{t+h} L(s, x(s), u(s)) \, ds + w(t+h, x(t+h)) - w(t, x)$$

$$\geqslant \int_t^{t+h} L(s, x(s), u(s)) + \frac{\partial}{\partial t} w(s, x(s)) + f(s, x(s), u(s)) \cdot D_x w(s, x(s)) \, ds$$

$$\geqslant \int_t^{t+h} L(s, x(s), u(s)) - H(s, x(s), D_x w(s, x(s))) \, ds.$$

Finally, if we let $h \to 0$ we get

$$-\frac{\partial}{\partial t} w(t, x) + H(t, x, D_x w(t, x)) \geqslant 0.$$

$\square$

We now turn our attention to uniqueness property of viscosity solution to the dynamic programming equation. Let us consider the dynamic programming equation in a more general form

$$-\frac{\partial}{\partial t} V(t, x) + H(t, x, D_x V(t, x)) = 0, \ (t, x) \in Q. \tag{31}$$

We will show that under technical assumptions on H, which are satisfied by the Hamiltonian, the supremum norm of the difference between two viscosity solution can be computed only on the frontier. Therefore, if we impose boundary conditions we get uniqueness of the viscosity solution. The assumptions on H are

- There exists $K \in \mathbb{R}$ and $h \in C([0, \infty))$ with $h(0) = 0$ such that $\forall \ (t, x), (s, y) \in \overline{Q}$ and $\forall \ p, p' \in \mathbb{R}^n$ we have

$$H(t, x, p) - H(s, y, p') \leqslant h(|t - s| + |x - y|) + h(|t - s|)|p| + K|x - y||p| + K|p - p'|.$$

- There exists $K' \in \mathbb{R}$ such that

$$|H_t| + |H_x| \leqslant K'(1 + |p|),$$

and

$$|H_p| \leqslant K'.$$

Under this hypotheses we can state the following theorem.

Theorem 0.1.8. *Let W and V viscosity subsolution and supersolution of 31 in Q, respectively. If Q is unbounded we assume W, V to be bounded and uniformly continuous on its closure. Then*

$$\sup_{\overline{Q}}[W - V] = \sup_{\partial * Q}[W - V].$$

This Theorem implies that if V and W are two viscosity solution of 31 satisfying the boundary conditions

$$V(t, x) = g(t, x), \ \forall \ (t, x) \in [T_0, t_1) \times O,$$
$$V(t_1, x) = \psi(x), x \in \overline{O}, \tag{32}$$

then, if the assumption of theorem 0.1.8 are satisfied, we get

$$\|W - V\|_\infty = \sup_{\overline{Q}}[W - V] = \sup_{\partial *Q}[W - V] = 0.$$

The proof of Theorem 0.1.8 is rather long; we first prove the $\overline{Q}$ bounded case and then the unbounded one.

Proof. Let Q be bounded. For $\epsilon, \delta, \beta > 0$ define

$$\Phi(t, x, s, y) = W(t, x) - V(s, y) - \frac{1}{2\epsilon}|x - y|^2 - \frac{1}{2\delta}|t - s|^2 + \beta(s - t_1). \tag{33}$$

Since $\overline{Q}$ is bounded the function Φ achieves its supremum at $(t', x'), (s', y') \in \overline{Q} \times \overline{Q}$. We will essentially prove that Φ divides $W - V$ from its supremum over the boundary in an inequality chain.

1. Let $\rho >$ and define

$$D_\rho = \left\{ (t, x), (s, y) \in \overline{Q} \times \overline{Q} \,|\, |t - s|^2 + |x - y|^2 \leqslant \rho \right\}, \tag{34}$$

   and

$$m_W(\rho) = 2 \left\{ |W(t, x) - W(s, y)| : ((t, x), (s, y)) \in D_\rho \right\},$$

$$m_V(\rho) = 2 \left\{ |V(t, x) - V(s, y)| : ((t, x), (s, y)) \in D_\rho \right\},$$

   and

$$K_1 = \sup\{m_w(\rho) : \rho \geqslant 0\}.$$

Since $\overline{Q}$ is compact and W, V are continuous they are uniformly continuous, therefore $m_W, m_V \in C([0, \infty))$. Clearly $m_W(0) = m_V(0) = 0$. We now claim two estimates on $|x - y|$ and $|t - s|$:

$$|t' - s'| \leqslant \sqrt{K_1 \delta}$$
$$|x' - y'| \leqslant \sqrt{\epsilon m_W(K_1[\epsilon + \delta])}. \tag{35}$$

From the maximum in $((t', x'), (s', y'))$ we have

$$\Phi(t', x', s', y') \leqslant \Phi(s', y', s', y'),$$

implies

$$\frac{1}{\epsilon}|x' - y'|^2 + \frac{1}{\delta}|t' - s'|^2 \leqslant 2(W(t', x') - W(s', y')) \leqslant m_W(|t' - s'| + |x' - y'|).$$

Then we have

$$\frac{1}{\delta}|t' - s'|^2 \leqslant K_1, \ \frac{1}{\epsilon}|x' - y'|^2 \leqslant K_1,$$

which also imply

$$\frac{1}{\epsilon}|x' - y'|^2 \leqslant m_w(K_1[\epsilon + \delta]).$$

2. We now show that if $(t', x') \in \partial^* Q$ or $(s', y') \in \partial^* Q$ we have

$$\Phi(t', x', s', y') \leqslant \frac{1}{2} m_V(K_1(\epsilon + \delta)) + \sup_{\partial^* Q}[W - V], \qquad (36)$$

or

$$\Phi(t', x', s', y') \leqslant \frac{1}{2} m_W(K_1(\epsilon + \delta)) + \sup_{\partial^* Q}[W - V], \qquad (37)$$

respectively. Indeed, in the first case, we have

$$
\begin{aligned}
\Phi(t', x', s', y') &\leqslant W(t', x') - W(s', y') \\
&\leqslant V(t', x') - V(s', y') + \sup_{\partial^* Q}[W - V] \\
&\leqslant \frac{1}{2} m_W(|t' - s'|^2 + |x' - y'|^2) + \sup_{\partial^* Q}[W - V].
\end{aligned}
$$

Same reasoning for the latter.

3. We now construct an estimate over the parameter β if $(t', x'), (s', y') \in Q$. Let us consider the test function

$$w(t, x) = \frac{1}{2\delta}|t - s'|^2 + \frac{1}{2\epsilon}|x - y'|^2,$$

then (t', x') is where $W - w$ is maximized and $W(t', x') = w(t', x')$, then

$$\underbrace{-\frac{1}{\delta}(t' - s')}_{\stackrel{def}{=} q_\delta} + H(t', x', \underbrace{\frac{1}{\epsilon}(x' - y')}_{\stackrel{def}{=} p_\epsilon}) \leqslant 0. \qquad (38)$$

Analogously, defining

$$w(s, y) = -\frac{1}{2\delta}|t' - s|^2 - \frac{1}{2\epsilon}|x' - y|^2 + \beta(s - t_1),$$

we get:

$$-\beta - q_\delta + H(s', y', p_\epsilon) \geqslant 0. \qquad (39)$$

Combining 38 and 39 we get

$$
\begin{aligned}
\beta &\leqslant H(s', y', p_\epsilon) - H(t', x', p_\epsilon) \\
&\leqslant h(|t' - s'| + |x' - y'|) + h(|t' - s'|)|p_\epsilon| + K|x' - y'||p_\epsilon| \\
&\leqslant h(\sqrt{K_1 \epsilon} + \sqrt{K_1 \delta}) + h(\sqrt{K_1 \delta})\frac{1}{\epsilon}\sqrt{K_1 \epsilon} + K\frac{|x' - y'|^2}{\epsilon} \\
&\leqslant h(\sqrt{K_1 \epsilon} + \sqrt{K_1 \delta}) + h(\sqrt{K_1 \delta})\sqrt{\frac{K_1 \epsilon}{\epsilon}} + K m_W(K_1[\epsilon + \delta]).
\end{aligned}
$$

4. We denote

$$k(\epsilon, \delta) = h(\sqrt{K_1\epsilon} + \sqrt{K_1\delta}) + h(\sqrt{K_1\delta})\sqrt{\frac{K_1\epsilon}{\epsilon}} + Km_W(K_1[\epsilon + \delta]), \tag{40}$$

then

$$\lim_{\epsilon \to 0} \lim_{\delta \to 0} k(\epsilon, \delta) = 0.$$

Then for suitably small ϵ and δ we have

$$k(\epsilon, \delta) < \beta,$$

which implies that either (t', x') or (s', y') is in $\partial^* Q$. Then by 36 or 37 we have

$$\limsup_{\epsilon \to 0} \limsup_{\delta \to 0} \Phi(t', x', s', y') \leqslant \sup_{\partial^* Q}[W - V], \tag{41}$$

for every $\beta > 0$. Finally, we have

$$W(t, x) - V(s, y) + \beta(t - t_1) = \Phi(t, x, t, x) \leqslant \Phi(t', x', s', y'),$$

then by 41 we get

$$W(t, x) - V(s, y) = \lim_{\beta \to 0} W(t, x) - V(s, y) + \beta(t - t_1) \leqslant \lim_{\beta \to 0} \Phi(t', x', s', y') \leqslant \sup_{\partial^* Q}[W - V].$$

Let us suppose that Q is unbounded, and V, W are bounded and uniformly continuous. We keep the same notation as before. For every $\gamma > 0$ there exist $(t_\gamma, x_\gamma), (s_\gamma, y_\gamma) \in \overline{Q}$ such that

$$\Phi(t_\gamma, x_\gamma, s_\gamma, y_\gamma) \geqslant \sup_{\overline{Q} \times \overline{Q}} \Phi - \gamma, \tag{42}$$

since W and V are bounded. We therefore define

$$\Phi_\gamma(t, x, s, y) = \Phi(t, x, s, y) - \frac{\gamma}{2}\left[|t - t_\gamma|^2 + |s - s_\gamma|^2 + |x - x_\gamma|^2 + |y - y_\gamma|^2\right], \tag{43}$$

for $(t, x), (s, y) \in \overline{Q}$.
We now follow a similar procedure to the one used in the bounded case.

1') Let the moduli of continuity m_W, m_V and the bound K_1 as before. Since we assume the value functions to be bounded and absolutely continuous we have, as before, that

$$m_W, m_V \in C([0, \infty)), \ m_W(0) = m_V(0) = 0, \ K_1 = \sup\{m_W(\rho) : \rho \geqslant 0\} < +\infty.$$

Let us note that for $(t, x), (s, y) \in \overline{Q} \times \overline{Q}$ such that

$$|t - t_\gamma|^2 + |s - s_\gamma|^2 + |x - x_\gamma|^2 + |y - y_\gamma|^2 > 2,$$

then

$$\Phi_\gamma(t, x, s, y) \leqslant \Phi(t, x, s, y) - \gamma \leqslant \Phi(t_\gamma, x_\gamma, s_\gamma, y_\gamma) = \Phi_\gamma\Phi(t_\gamma, x_\gamma, s_\gamma, y_\gamma).$$

The supremum of Φ_γ is achieved since Φ_γ is bounded by $\Phi(t_\gamma, x_\gamma, s_\gamma, y_\gamma)$. It is achieved at $(t', x'), (s', y')$. Then

$$|t - t_\gamma|^2 + |s - s_\gamma|^2 + |x - x_\gamma|^2 + |y - y_\gamma|^2 \leqslant 2. \tag{44}$$

We now claim two estimates on the distance between the time variables and space variables. For $\gamma, \delta, \epsilon \leqslant \frac{1}{2}$ we have

$$|t' - s'| \leqslant \sqrt{2(K_1 + 1)\delta}, \ |x' - y'| \leqslant \sqrt{2\epsilon(2\gamma + m_W(2\epsilon(K_1 + 1)))}. \tag{45}$$

Indeed, $\Phi_\gamma(t', x', s', y') \geqslant \Phi_\gamma(s', x', s', y')$ implies

$$\begin{aligned}
\frac{1}{2\delta}|t' - s'|^2 &\leqslant W(t', x') - W(s', x') + \frac{\gamma}{2}[|s' - t_\gamma|^2 - |t' - t_\gamma|^2] \\
&\leqslant \frac{1}{2}K_1 + \gamma|t' - s'|^2 + \frac{\gamma}{2}|t' - t_\gamma|^2 \\
&\leqslant \frac{1}{2}K_1 + \gamma|t' - s'|^2 + \gamma,
\end{aligned}$$

which implies, recalling the bound on δ and γ, that

$$\left(\frac{1}{4\delta}\right)|t' - s'|^2 \leqslant \left(\frac{1 - \delta}{2\delta}\right)|t' - s'|^2 \leqslant \left(\frac{1}{2\delta} - \gamma\right)|t' - s'|^2 \leqslant \frac{1}{2}K_1 + \frac{1}{2}.$$

From $\Phi_\gamma(t', x', s', y') \geqslant \Phi_\gamma(t', y', s', y')$ we get

$$\begin{aligned}
\frac{1}{2\epsilon}|x' - y'|^2 &\leqslant W(t', x') - W(t', x') + \frac{\gamma}{2}[|y' - x_\gamma|^2 - |x' - x_\gamma|^2] \\
&\leqslant \frac{1}{2}m_W(|x' - y'|^2) + \gamma|x' - y'|^2 + \frac{\gamma}{2}|x' - x_\gamma|^2 \\
&\leqslant \frac{1}{2}m_W(|x' - y'|^2) + \gamma|x' - y'|^2 + \gamma,
\end{aligned}$$

which implies

$$\frac{1}{4\epsilon}|x' - y'|^2 \leqslant \frac{1}{2}m_W(|x' - y'|^2) + \gamma \Rightarrow |x' - y'|^2 \leqslant 2(K_1 + 1)\epsilon.$$

Therefore, we get

$$\frac{1}{4\epsilon}|x' - y'|^2 \leqslant \left(\frac{1}{2\epsilon} - \gamma\right)|x' - y'|^2 \leqslant \frac{1}{2}m_W(2(K_1 + 1)\epsilon) + \gamma,$$

hence the second equation in 45.

2') Replicating the same argument of the bounded case we get if $(t', x') \in \partial^* Q$ then

$$\Phi_\gamma(t', x', s', y') \leqslant \frac{1}{2}m_V(2(K_1 + 1)(\epsilon + \delta)) + \sup_{\partial^* Q}[W - V], \tag{46}$$

and if $(s', y') \in \partial^* Q$ then

$$\Phi_\gamma(t', x', s', y') \leqslant \frac{1}{2}m_W(2(K_1 + 1)(\epsilon + \delta)) + \sup_{\partial^* Q}[W - V]. \tag{47}$$

3') Finally, if we find a bound over β for $(t', x'), (s', y') \in Q$ we then prove that this bound is not satisfied and then we can invoke point $2'$ and get

$$\limsup_{\epsilon \to 0} \limsup_{\delta \to 0} \limsup_{\gamma \to 0} \Phi_\gamma(t', x', s', y') \leqslant \sup_{\partial^* Q}[W - V], \qquad (48)$$

which implies, as before, that

$$W(t, x) - V(s, y) \leqslant \sup_{\partial^* Q}[W - V].$$

Let $(t', x'), (s', y') \in Q$ and the test function

$$\overline{w}(t, x) = \frac{1}{2\delta}|t - s'|^2 + \frac{1}{2\epsilon}|x - y'|^2 + \frac{\gamma}{2}\left(|t - t_\gamma|^2 + |x - x_\gamma|^2\right),$$

defined over Q. Then (t', x') is where $W - \overline{w}$ is null and maximized. Therefore

$$-\underbrace{\frac{1}{\delta}(t' - s')}_{\overset{def}{=} q_\delta} - \underbrace{\gamma(t' - t_\gamma)}_{\overset{def}{=} q_\gamma} + H(t', x', p_\epsilon + p_\gamma) \leqslant 0, \qquad (49)$$

where

$$p_\epsilon = \frac{1}{\epsilon}(x' - y'), \ p_\gamma = \gamma(x' - x_\gamma).$$

With the same reasoning we define the test function

$$w^*(s, y) = -\frac{1}{2\delta}|t' - s|^2 - \frac{1}{2\epsilon}|x' - y|^2 + \frac{\gamma}{2}\left(|s - s_\gamma|^2 + |y - y_\gamma|^2\right) + \beta(s - t_1),$$

from which we get

$$-\beta - q_\delta - \underbrace{\gamma(s' - s_\gamma)}_{\overset{def}{=} \overline{q}_\gamma} + H(s', y', p_\epsilon + \overline{p}_\gamma) \geqslant 0, \qquad (50)$$

where $\overline{p}_\gamma = \gamma(y_\gamma - y')$.

We can find the upper bound to β using the estimates in 45 and equations 50 and 49:

$$\begin{aligned}
\beta &\leqslant H(s', y', p_\epsilon + \overline{p}_\gamma) - H(t', x', p_\epsilon + p_\gamma) + q_\gamma - \overline{q}_\gamma \\
&\leqslant h(|t' - s'| + |x' - y'|) + h(|t' - s'|)\left[|p_\epsilon + |\overline{p}_\gamma||\right] \\
&\quad + K|x' - y'|\left[|p_\epsilon + |\overline{p}_\gamma||\right] + K|\overline{p}_\gamma - p_\gamma| + q_\gamma - \overline{q}_\gamma \\
&\leqslant h(\sqrt{2(K_1 + 1)\delta} + \sqrt{2(K_1 + 1)\epsilon}) + h(\sqrt{2(K_1 + 1)\delta})\left[\sqrt{\frac{|2(K_1 + 1)|}{\epsilon}} + 2\gamma\right] \\
&\quad + K2\left[2\gamma + m_W(2(K_1 + 1))\epsilon + \sqrt{2(K_1 + 1)\epsilon\delta}\right] + 2(K + 1)\gamma.
\end{aligned}$$

$\square$

0.1.6 Continuity Properties of the Value Function

We now investigate continuity properties of the value function. In particular, we consider two different sets of assumptions which guarantee continuity, actually something stronger: Lipschitz continuity and Uniform continuity. Let us consider f, L, ψ as usual, recalling that there exists $K_\rho > 0$ such that

$$|f(t, x, v) - f(t, y, v)| \leqslant K_\rho |x - y|, \tag{51}$$

for all $|v| \leqslant \rho$. Note that these are minimal assumption to start working with the optimal control problem.

Theorem 0.1.9. *Let a bounded control space U, $Q = [t_0, t_1) \times \mathbb{R}^n$. Assume that f, L, ψ are bounded, f satisfies 51 and L, ψ uniformly continuous. Then the value function V is bounded and uniformly continuous.*

Proof. V is bounded. Indeed

$$V(t, x) = \inf_{u \in \mathcal{U}} \int_t^{t_1} L(s, x(s), u(s)) \, ds + \psi(t_1, x(t_1)) \leqslant (t_1 - t) K_L + K_\psi,$$

where K_g is the bound over the function g.
V is uniformly continuous. Standard arguments on solutions of ODE (integral form of the solution, bound on the flow and Gronwall) yield

$$\Lambda(s) \stackrel{def}{=} |x(s) - y(s)| \leqslant |x - y| e^{K(s-t)}, \tag{52}$$

where $x(s), y(s)$ are solution to the Cauchy problem

$$\frac{d}{ds} w(s) = f(s, w(s), u(s)), \ w(t) = w.$$

Then

$$|J(t, x, u) - J(t, y, u)| \leqslant \int_t^{t_1} |L(s, x(s), u(s)) - L(s, y(s), u(s))| \, ds + |\psi(x(t_1)) - \psi(y(t_1))|$$

$$\leqslant \int_t^{t_1} m_L(\Lambda(s)) \, ds + m_\psi(\Lambda(t_1))$$

$$+ (t_1 - t_0) m_L(|x - y| e^{K(t_1 - t_0)}) + m_\psi(|x - y| e^{K(t_1 - t_0)}),$$

where m_g is the modulus of continuity of the function g. Since both L and ψ are uniformly continuous $m_L, m_\psi \in C([0, \infty))$ and they are null at the origin. Then for a fixed t we have the bound

$$|V(t, x) - V(t, y)| \leqslant (t_1 - t_0) m_L(|x - y| e^{K(t_1 - t_0)}) + m_\psi(|x - y| e^{K(t_1 - t_0)}). \tag{53}$$

Similarly, if we fix x we get

$$|V(t, x) - V(s, x)| \leqslant K_1 |t - s| + (t_1 - t_0) m_L(|t - s| K_2 e^{K(t_1 - t_0)}) + m_\psi(|t - s| K_2 e^{K(t_1 - t_0)}), \tag{54}$$

where $K_1 = \sup |L|$ and K_2 is such that $|x(s) - x| \leqslant K_2 |t - s|$, which exists because f is bounded. Then by splitting $|V(t, x) - V(s, y)|$ we find a modulus of continuity for it, by summing 54 and 53. $\qquad \square$

From the proof it is clear that by taking L and ψ Lipschitz continuous we get V Lipschitz continuous, since we can use

$$m_L(\rho) = K_L \rho, \ m_\psi(\rho) = K_\psi \rho,$$

for some constants $K_L, K_\psi > 0$.
This theorem allows us to conclude the following result.

Corollary 0.1.10. *Under the previous assumptions, the value function is the unique viscosity solution of the dynamic programming equation with fixed terminal conditions* 32.

Proof. Since U is bounded the Hamiltonian H satisfies 0.1.5. Indeed, the bound on the derivative is obvious. The upper bound function h can be found as follows:

$$
\begin{aligned}
H(t,x,p) - H(s,y,p') &\leqslant \sup_{v \in U} \left\{ |L(t,x,v) - L(s,y,v) + f(t,x,v) \cdot p - f(s,y,v) \cdot p'| \right\} \\
&\leqslant \sup_{v \in U} \left\{ |L(t,x,v) - L(s,y,v)| \right\} + \sup_{v \in U} |f(s,y,v)||p - p'| \\
&\quad + \sup_{v \in U} |f(t,x,v) - f(s,y,v)||p|,
\end{aligned}
$$

but the uniform continuity of L and Lipschitz continuity of f yield 0.1.5. Then by Theorem 0.1.8 we get the thesis. $\qquad\square$

0.1.7 Pontryagin's Principle for Viscosity Solutions

In section **??** we stated the Pontryagin's principle for continuously differentiable value functions. As we come to understand in this chapter, value functions may easily fail to be C^1 which requires a weaker notion of solution. We have introduced the concept of viscosity solution, which turns out to be also a handy notion to extend Pontryagin's principle in non-differentiable settings. Recall that it states the relation between a so-called adjoint variable P and the state variable via a system of differential equations. Since may fail to be differentiable somewhere we need to carefully extend the possible values of the adjoint variable for those instances. Let us define two new key concepts that incorporate the possible value that the incremental limit may have in non-smooth points.

Definition 0.1.4. *Let $W \in C(\overline{Q})$ and $(t,x) \in Q$. We define:*

1. *The set of superdifferentials of W at (t,x) as the collection of all $(q,p) \in \mathbb{R} \times \mathbb{R}^n$ such that there exists some $w \in C^1(Q)$ for which:*

$$
(q,p) = \left(\frac{\partial}{\partial t} w(t,x), D_x w(t,x) \right), \tag{55}
$$

and

$$
(t,x) \in \arg\max \left\{ (W - w)(s,y) \,|\, (s,y) \in \overline{Q} \right\}. \tag{56}
$$

We denote it with $D^+ W(t,x)$.

2. *The set of subdifferentials of W at (t,x) as the collection of all $(q,p) \in \mathbb{R} \times \mathbb{R}^n$ such that there exists some $w \in C^1(Q)$ for which:*

$$
(q,p) = \left(\frac{\partial}{\partial t} w(t,x), D_x w(t,x) \right), \tag{57}
$$

and

$$
(t,x) \in \arg\min \left\{ (W - w)(s,y) \,|\, (s,y) \in \overline{Q} \right\}. \tag{58}
$$

We denote it with $D^- W(t,x)$.

Note that this definition is consistent with the notion of viscosity solution since if W is a viscosity subsolution in Q if and only if

$$-q + H(t, x, p) \leqslant 0, \quad \forall\, (q, p) \in D^+ W(t, x), \quad \forall\, (t, x) \in Q,$$

and similarly for a supersolution.

The notion of superdifferential will take the place of the derivative in Pontryagin's result. We recall the definition of the adjoint variable for a state variable x defined by the flow f, a control u, a terminal condition ψ, a Lagrangian L and a Hamiltonian H:

$$\dot{p}_j^*(s) = -\sum_{i=1}^n \frac{\partial}{\partial x_j} f_i(s, x^*(s), u^*(s)) p_i(s) - \frac{\partial}{\partial x_j} L(s, x^*, u^*), \tag{59}$$

And also:

$$p(s) \cdot f(s, x^*(s), u^*(s)) + L(s, x^*(s), u^*(s)) = -H(s, x^*(s), u^*(s), p^*(s)), \tag{60}$$

With:

$$p^*(t_1) = D\psi(x^*(t_1)). \tag{61}$$

We can now state Pontryagin's maximum principle in this broader context.

Theorem 0.1.11. *Let $u^*(\cdot)$ be an optimal control at (t, x) which is right continuous at each $[t, t_1)$, and $p^*(s)$ defined by , 60 and ??. Then for each $s \in [t, t_1)$*

$$\left(H(s, x^*(s), p^*(s)), p^*(s) \right) \in D^+ V(s, x^*(s)). \tag{62}$$

Proof. By definition of value function we have

$$V(r, y) \leqslant J(r, y; u^*), \quad \forall\, (r, y) \in Q,$$

and equality for $(r, y) = (s, x^*(s))$ for each $s \in [t, t_1]$. Hence, if we show that J is continuously differentiable and

$$\frac{\partial}{\partial r} J(s, x^*(s); u^*) = p^*(s), \tag{63}$$

and

$$D_y J(s, x^*(s); u^*) = H(s, x^*(s), p^*(s)) \tag{64}$$

for $s \in [t, t_1)$, then the thesis holds by definition. We prove 63 by direct computation

$$\frac{\partial}{\partial x_i} J(s, x^*(s); u^*) = \sum_{j=1}^n \left\{ \int_s^{t_1} L_{x_j}(r, x^*(r), u^*(r))\, dr + \psi_{x_j}(x^*(t_1)) \frac{\partial}{\partial x_j} x_i(t_1, x^*(t_1)) \right\},$$

but we have

$$\frac{d}{dr} \left\{ \sum_{j=1}^n \frac{\partial}{\partial x_j} x_i(r, x^*(r)) p_i(r) \right\} = -\sum_{i=1}^n \frac{\partial}{\partial x_j} f_i(s, x^*(r), u^*(r)) \frac{\partial}{\partial x_j} x_i(r, x^*(r)), \tag{65}$$

then by integrating and recalling that

$$\frac{\partial}{\partial x_j} x_i(s, x^*(s)) = \delta_{ij}$$

we get

$$p_i(s) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j} x_i(s, x^*(s)) p_j(s)$$

$$= \sum_{j=1}^{n} \frac{\partial}{\partial x_j} x_i(t_1, x^*(t_1)) p_j(t_1) - \int_s^{t_1} \frac{d}{dr} \left\{ \sum_{j=1}^{n} \frac{\partial}{\partial x_j} x_i(r, x^*(r)) p_i(r) \right\} dr$$

$$= \frac{\partial}{\partial x_i} J(s, x^*(s); u^*)$$

$$= \frac{\partial}{\partial x_i} V(s, x^*(s)).$$

To prove 63 we define

$$z(s; r, y) = \frac{\partial}{\partial r} x(s; r, y),$$

which allows us to compute

$$\frac{\partial}{\partial r} J(r, x^*(r); u^*) = -L(r, x^*(r), u^*(r))$$

$$+ \sum_{j=1}^{n} \left\{ \int_r^{t_1} \frac{\partial}{\partial x_j} L(s, x^*(s), u^*(s)) z_j(s; r, x^*(r)) \, ds + \frac{\partial}{\partial x_j} \psi(t_1, x^*(t_1)) z_j(t_1; r, x^*(r)) \right\},$$

but

$$\frac{d}{ds} \left[z(s) \cdot p(s) \right] = -\sum_{j=1}^{n} \frac{\partial}{\partial x_j} L(s, x^*(s), u^*(s)) z_j(s),$$

then

$$-f(r, x^*(r), u^*(r)) \cdot p(r) = z(r) \cdot p(r) = z(t_1) \cdot p(t_1) - \int_r^{t_1} \frac{d}{ds} \left[z(s) \cdot p(s) \right] ds$$

$$= \sum_{j=1}^{n} \left\{ \frac{\partial}{\partial x_j} \psi(t_1, x^*(t_1)) z_j(t_1) + \int_r^{t_1} \frac{\partial}{\partial x_j} L(s, x^*(s), u^*(s)) z_j(s) \, ds \right\},$$

which implies that

$$\frac{\partial}{\partial r} J(r, x^*(r); u^*) = -L(r, x^*(r), u^*(r)) - f(r, x^*(r), u^*(r)) \cdot p(r).$$

$\square$