Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

Preprint Nr. ???

Viscosity solutions of HJB-PDE's arising in control problems

Mohamed Aziz Lakhal

Saarbrücken 2020

Fachrichtung 6.1 – Mathematik
Universität des Saarlandes

Preprint No. ???
submitted: October 25, 2020

Viscosity solutions of HJB-PDE's arising in control problems

Mohamed Aziz Lakhal

Saarland University
Department of Mathematics
P.O. Box 15 11 50
66041 Saarbrücken
Germany
s8molakh@stud.uni-saarland.de

Edited by
FR 6.1 – Mathematik
Universität des Saarlandes
Postfach 15 11 50
66041 Saarbrücken
Germany

Fax:	+ 49 681 302 4443
e-Mail:	preprint@math.uni-sb.de
WWW:	http://www.math.uni-sb.de/

Chapter 1

A very brief introduction to control problems

Control problems arise every time a system has to be manipulated in order to satisfy a particular objective. Naturally, they find numerous applications across a broad range of domains, going from finance and economics to actual rocket science. Aside from being a powerful toolbox for the pragmatic engineer, the diverse intersections with the different areas of mathematics also offer an aesthetic landscape for the dreamiest mathematician in his quest for intellectual distraction.

1.1 Finite-time control problems

We describe, rather informally, the specific type of control problems we will deal with. Consider a system starting in state $x_0 \in \mathbb{R}^N$ at time $t_0 \geq 0$. At each point of time $t \geq t_0$, we can manipulate the current state of the system $x(t) \in \mathbb{R}^N$, by executing an external control $\alpha(t) \in \mathcal{A}$, where $\mathcal{A}$ denotes the topological space of all possible controls. The resulting trajectory of the system $x_{x_0}^{\alpha}$ is described by the dynamics $b : [0, T] \times \mathbb{R}^N \times \mathcal{A} \to \mathbb{R}$, meaning $x_{x_0}^{\alpha}$ is the unique, absolutely continuous solution to the IVP

$$\begin{aligned} x(t_0) &= x_0 \\ \dot{x}(t) &= b(t, x(t), \alpha(t)) \ . \end{aligned}$$

A manipulation of the system at time t in state $x(t)$ with control $\alpha(t)$, is penalized with costs $f(t, x(t), \alpha(t)) \in \mathbb{R}$. The system is observed during a finite, fixed time span $[t_0, T]$. The final state of the system, $x_{x_0}^{\alpha}(T)$ is then penalized with another costs $h(x_{x_0}^{\alpha}(T)) \in \mathbb{R}$, which results in totals costs of

$$J(t_0, x_0, \alpha) := \int_{t_0}^{T} f(t, x_{x_0}^{\alpha}(t), \alpha(t))dt + h(x_{x_0}^{\alpha}(T)) \ .$$

Our objective is now to minimize the total costs over all measurable controls $\alpha : [t_0, T] \to \mathcal{A}$, denoting latter set by $\mathcal{V}\,[t_0, T]$.
We remark at this point, that our problem model is not the only, and certainly not, the most general model. Suppose for example, we wanted to regulate a rocket's fuel consumption on its way to mars. As the amount of fuel injected affects the rocket's velocity, it affects the time needed to reach our destination. The time span, over which we consider our control problem, can therefore depend on the chosen control. Still, we forsake generality for a concise presentation of the principle ideas.

1.2 The Dynamic Programming Principle (DPP)

We make in this section a quite intuitive but, nonetheless, fundamental observation which relates the control problems with different starting points and initial states to each other.
We consider a system with initial state $x_0 \in \mathbb{R}^N$ at time $t_0 \geq 0$. Following the standard procedure in optimization, we derive necessary conditions for a control $\alpha \in \mathcal{V}\,[t_0, T]$ to be optimal.
Assume $\alpha^* \in \mathcal{V}\,[t_0, T]$ is an optimal control, and let $x_{x_0}^{\alpha^*} : [t_0, T] \to \mathbb{R}^N$ be its resulting trajectory. That is, if we define

$$v(t_0, x_0) := \inf_{\alpha \in \mathcal{V}[t_0,T]} J(t_0, x_0, \alpha) \ ,$$

we have $v(t_0, x_0) = J(t_0, x_0, \alpha^*)$. As α^* transfers the system to state $x_{x_0}^{\alpha^*}(t_0+h)$ at a latter point of time $t_0 + h \geq t_0$, we differentiate between the costs arisen over the time span $[t_0, t_0 + h)$, and those arisen over $[t_0 + h, T]$. Doing so, yields that

$$\begin{aligned} J(t_0, x_0, \alpha^*) = & \int_{t_0}^{t_0+h} f(t, x_{x_0}^{\alpha^*}(t), \alpha^*(t))dt \\ & + \int_{t_0+h}^{T} f(t, x_{x_0}^{\alpha^*}(t), \alpha^*(t))dt + h(x_{x_0}^{\alpha^*}(T)) \end{aligned}$$

which leads to the interesting equation:

$$J(t_0, x_0, \alpha^*) = \int_{t_0}^{t_0+h} f(t, x_{x_0}^{\alpha^*}(t), \alpha^*(t))dt + J(t_0+h, x_{x_0}^{\alpha^*}(t_0+h), \alpha^*_{|[t_0+h,T]}) \ . \quad (1.1)$$

It is natural to ask whether $\alpha^*_{|[t_0+h,T]}$, can be used to solve the control problem starting $t_0 + h$ in state $x_{x_0}^{\alpha^*}(t_0 + h)$, i.e whether

$$J(t_0 + h, x_{x_0}^{\alpha^*}(t_0 + h), \alpha^*_{|[t_0+h,T]}) = v(t_0 + h, x_{x_0}^{\alpha^*}(t_0 + h)) \ .$$

A simple exchange argument gives a positive answer to that question. Suppose the contrary, meaning

$$J(t_0 + h, x_{x_0}^{\alpha^*}(t_0 + h), \alpha^*_{|[t_0+h,T]}) > v(t_0 + h, x_{x_0}^{\alpha^*}(t_0 + h)) \ .$$

In this case there exists a control $\beta \in \mathcal{V}\,[t_0 + h, T]$, s.t.

$$J(t_0 + h, x_{x_0}^{\alpha^*}(t_0 + h), \alpha^*_{|[t_0+h,T]}) > J(t_0 + h, x_{x_0}^{\alpha^*}(t_0 + h), \beta) \ .$$

But then we can concatenate the controls $\alpha^*_{[t_0,t_0+h]}$ and β to obtain a control in $\mathcal{V}\,[t_0, T]$ that beats α^* over the whole time span $[t_0, T]$, which is of course a contradiction.
The function $v : [0, T] \times \mathbb{R}^N \to \mathbb{R}$, the so-called *value function* associated with our family of problems, consequently satisfies the functional equation

$$v(t_0, x_0) = \int_{t_0}^{t_0+h} f(t, x_{x_0}^{\alpha^*}(t), \alpha^*(t))dt + v(t_0 + h, x_{x_0}^{\alpha^*}(t_0 + h)) \ , \tag{1.2}$$

which exactly describes the so-called *Dynamic Programming Principle*, commonly abbreviated by **DPP**. The DPP is an idea that was first introduced by Richard Bellmann in a discrete-time setting [3], he then later extended to problems in continuous time [4]. Zhou and Jiongming captured the DPP quite well in their concise statement:

$$\text{“globally optimal} \implies \text{locally optimal”} \ .$$

As our functional equation 1.2 satisfied by v relies on the existence of an optimal control, we try to rewrite 1.2 as

$$v(t_0, x_0) = \inf_{\alpha \in \mathcal{V}[0,T]} \left\{ \int_{t_0}^{t_0+h} f(t, x_{x_0}^{\alpha}(t), \alpha(t))dt + v(t_0 + h, x_{x_0}^{\alpha}(t_0 + h)) \right\} \ . \tag{1.3}$$

An exchange argument, similar to the one used to justify 1.2, shows 1.3 holds, even if an optimal control does not exist.
What seems promising at the first glance, is quite unsatisfactory, as “[the dynamic programming equation] is very difficult to handle, since the operation involved on the right-hand side [...] is too complicated.”, Zhou and Jiongming confirm in [10, p. 160]. Instead of directly solving 1.3, we consider its implications for v.

Chapter 2

Smooth value functions

2.1 From functional to differential equation

As mentioned in the title, we aim to derive a differential equation satisfied by the value function, starting from the DPP 1.3, which is in fact a functional equation. Before doing so, we want to make some assumptions ensuring the well-definedness of the objects appearing in our control problem. We impose:

- $\mathcal{A}$ to be a metric space.
- The functions f, b and h to be uniformly continuous. Let ψ be a placeholder for latter functions. We additionally require:

$$|\psi(t,x,a) - \psi(t,x_0,a)| \leq L|x - x_0| \ , \tag{2.1}$$

for every fixed $t \in [0,T]$ and $a \in \mathcal{A}$, and for every x, $x_0 \in \mathbb{R}^N$, as well as

$$|\psi(t,0,a)| \leq L \ . \tag{2.2}$$

Under these conditions, the prerequisites for Caratheodory's existence and uniqueness theorem (cf. [9, Theorem 1.45 p. 25]) are satisfied, and therefore any control indeed induces a unique, absolutely continuous, state trajectory. Furthermore, it is ensured, that the running costs along a given trajectory are integrable over a finite time span. Finally, the value function is known to be finite valued and at least locally Lipschitz-continuous (cf. [10, Inequality (2.68) from Theorem 2.5 p. 165]). Latter assertion relies on Gronwall-type-arguments, applied to a given controlled trajectory.
As the DPP 1.3 relates the function values of v along controlled trajectories, we want to study the variations of v along those. Provided the problem parameters do not exhibit variations too volatile, and more importantly, provided that v is smooth, our analysis reveals the value function to be a solution to some boundary value problem.

Theorem 2.1.1. *Let $\mathcal{A}$ be a metric space. Suppose the running cost function f, the dynamics b and the terminal cost function h are uniformly continuous, and admit estimates as in 2.2 and 2.1 for some constant $L \geq 0$. If the value function v is continuously differentiable over $(0,T) \times \mathbb{R}^N$, then v is a solution to the boundary problem with PDE*

$$- v_t + H(t, x, -D_x v) = 0 \ , \tag{2.3}$$

and boundary condition

$$v(T, \cdot) = h \ , \tag{2.4}$$

where H denotes the Hamiltonian of the control problem, given by the mapping $(t, x, p) \mapsto \sup\limits_{a \in \mathcal{A}} \left\{ p^T b(t, x, a) - f(t, x, a) \right\}$.

Proof. Proceed in two steps, respectively showing that v is a sub-and supersolution of 2.3, which means that v satisfies the inequalities:

$$- v_t + H(t, x, -D_x v) \leq 0 \ , \tag{2.5}$$

and

$$- v_t + H(t, x, -D_x v) \geq 0 \ . \tag{2.6}$$

For the sake of suggestive notation, define for any control $\alpha \in \mathcal{V}[0,T]$ and $x_0 \in \mathbb{R}^N$, the map

$$v_{x_0}^{\alpha} : (0,T) \to \mathbb{R}, \ t \mapsto v(t, x_{x_0}^{\alpha}(t)) \ ,$$

and introduce for $a \in \mathcal{A}$, the map

$$H^a : [0,T] \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}, \ (t, x, p) \mapsto p^T b(t, x, a) - f(t, x, a) \ ,$$

such that $H = \sup\limits_{a \in \mathcal{A}} H^a$.

In order to show 2.5, fix an arbitrary point (t_0, x_0) and consider an arbitrary constant control $\overline{a}$. According to the DPP 1.3, we have that

$$v(t_0, x_0) \leq \int_{t_0}^{t_0+h} f(t, x_{x_0}^{\overline{a}}(t), \overline{a}(t)) dt + v(t_0 + h, x_{x_0}^{\overline{a}}(t_0 + h)) \ , \text{ for any } h \geq 0 \ .$$

Rearranging the inequality, and dividing both sides by $1/h > 0$ gives:

$$\frac{v(t_0 + h, x_{x_0}^{\overline{a}}(t_0 + h)) - v(t_0, x_0)}{h} \geq -\frac{1}{h} \int_{t_0}^{t_0+h} f(t, x_{x_0}^{\overline{a}}(t), \overline{a}(t)) dt \ , \tag{2.7}$$

which, using the definition of $v_{x_0}^{\overline{a}}$, can be rewritten as:

$$\frac{v^{\overline{a}}_{x_0}(t_0+h)-v^{\overline{a}}_{x_0}(t_0)}{h} \geq \frac{1}{h}\int_{t_0}^{t_0+h} f(t,x^{\overline{a}}_{x_0}(t),\overline{a}(t))dt \ .$$

Passing to the limit $h \searrow 0$, yields

$$\dot{v}^{\overline{a}}_{x_0}(t_0) \geq -f(t_0,x_0,a) \ ,$$

with $\dot{v}^{\overline{a}}_{x_0}(t_0) = v_t(t_0,x_0) + D_x v(t_0,x_0,a)b(t_0,x_0,a)$, as the variations of $x^{\overline{a}}_{t_0}$ are described by the dynamics b. It follows that

$$v_t(t_0,x_0) \geq H^a(t_0,x_0,-D_x v(t_0,x_0)) \ ,$$

which implies

$$v_t(t_0,x_0) \geq H(t_0,x_0,D_x v(t_0,x_0)) \ ,$$

by taking the supremum over $a \in \mathcal{A}$. Latter inequality is obviously equivalent to 2.5.
Now for the more delicate proof of inequality 2.6. Fix again an arbitrary $(t_0,x_0) \in (0,T) \times \mathbb{R}^N$. For any $\epsilon, h > 0$, there exists, by the DPP 1.3, some control $\alpha = \alpha(\epsilon,h)$, such that

$$v(t_0,x_0) \geq \left\{ \int_{t_0}^{t_0+h} f(t,x^{\alpha}_{x_0}(t),\alpha(t))dt + v(t_0+h,x^{\alpha}_{x_0}(t_0+h)) \right\} - \epsilon h \ .$$

Isolate the terms corresponding to v on the left-hand-side, and divide both sides by $-1/h < 0$, to obtain that

$$\frac{v(t_0+h,x^{\alpha}_{x_0}(t_0+h))-v(t_0,x_0)}{h} \leq -\frac{1}{h}\int_{t_0}^{t_0+h} f(t,x^{\alpha}_{x_0}(t),\alpha(t))dt + \epsilon \ . \tag{2.8}$$

Use again the definition of $v^{\alpha}_{x_0}$ to rewrite 2.8 as:

$$\frac{v^{\alpha}_{x_0}(t_0+h)-v^{\alpha}_{x_0}(t_0)}{h} \leq -\frac{1}{h}\int_{t_0}^{t_0+h} f(t,x^{\alpha}_{x_0}(t),\alpha(t))dt + \epsilon \ . \tag{2.9}$$

Note that the considered control α depends on h, while the relationship between the different controls is unknown. In view of the circumstances, showing the uniform convergence of the appearing difference quotients w.r.t α as $h \searrow 0$, seems rather intimidating. Instead, we want to resolve the dependency from α. To this purpose, apply the fundamental theorem of calculus to the left-hand-side of 2.9 and subtract the resulting integral from both sides to obtain that

$$-\epsilon \leq \frac{1}{h} \int_{t_0}^{t_0+h} g(t, x^{\alpha}_{x_0}(t), \alpha(t))dt \ , \tag{2.10}$$

where g, defined by $g(t,x,a) = -v_t(t,x) + H^a(t,x,-D_x v(t,x))$, is continuous, as b and f are uniformly continuous. It is easy to deduce the equicontinuity of $\{H^a\}_{a\in\mathcal{A}}$ from the uniform continuity of f and b and condition 2.1 imposed upon them. Consequently H and the function $\hat{g}$, defined by

$$\hat{g}(t,x) \coloneqq \sup_{a\in\mathcal{A}} g(t,x,a) = -v_t(t,x) + H(t,x,-D_x v(t,x)) \ ,$$

are continuous.
Taking the supremum over $a \in \mathcal{A}$ in 2.10 yields that

$$-\epsilon \leq \frac{1}{h} \int_{t_0}^{t_0+h} \hat{g}(t, x^{\alpha}_{x_0}(t))dt \ .$$

Note how the integrand does not depend anymore on the function values of α, but only on its induced trajectory.
Apply the mean value theorem to rewrite latter inequality as

$$-\epsilon \leq \hat{g}(\xi, x^{\alpha}_{x_0}(\xi)) \ ,$$

with $\xi = \xi(\epsilon,h) \in (t_0, t_0+h)$. If we show that $\hat{g}(\xi, x^{\alpha}_{x_0}(\xi))$ converges towards $\hat{g}(t_0,x_0)$, as h tends to zero, we also get that

$$-\epsilon \leq \hat{g}(t_0,x_0) = -v_t(t_0,x_0) + H(t_0,x_0,-D_x v(t_0,x_0)) \ ,$$

and since ϵ was arbitrary, our proof for 2.6 would be complete. This is actually the case, as the trajectories $\{x^{\alpha}_{x_0}\}$ are equicontinuous in t_0. To see this, use assumptions 2.2 and 2.1 satisfied by the dynamics b, to derive that

$$|x^{\alpha}_{x_0}(\xi) - x^{\alpha}_{x_0}(t_0)| \leq \int_{t_0}^{\xi} L|x^{\alpha}_{x_0}(t) - x^{\alpha}_{x_0}(t_0)|dt + L(1+|x_0|)(\xi - t_0) \ . \tag{2.11}$$

By Gronwall's inequality, 2.11 implies

$$|x^{\alpha}_{x_0}(\xi) - x^{\alpha}_{x_0}(t_0)| \leq L(1+|x_0|)(\xi - t_0)\exp(L(\xi - t_0)) \ ,$$

and the trajectories $\{x^{\alpha}_{x_0}\}$ are indeed equicontinuous in t_0, which, according to our previous remark, brings our proof of 2.6 to an end.
The boundary condition $v(T,\cdot) = h$, follows immediately from the definition of v. □

2.2 Uniqueness among classical solutions

On the importance of uniqueness results While the question of whether or not the boundary value problem, consisting of PDE 2.3 and boundary condition 2.4, admits a unique continuously differentiable solution, might seem natural for a seasoned specialist in PDEs, it is actually also of practical interest. Although we have not explicitly computed the value function yet, we know at the very least, that it satisfies some boundary value problem, whenever smooth. Now suppose we somehow computed a solution of the boundary value problem. Without any uniqueness result, we are left wondering whether or not the computed solution is also the sought value function. We immediately want to make up for this inconvenience lest the more practically inclined reader quit reading with a deprecating smile.

2.2.1 A local comparison result

Comparison results are often the key relation when showing the uniqueness of a given boundary value problem. This case is no different. Given a subsolution u and a supersolution v, satisfying $u \leq v$ on the boundary $\{T\} \times \mathbb{R}^N$, we extend latter relationship on some closed region. In order to do so, we examine the difference $w \coloneqq u - v$, to find out that w is a subsolution to some other PDE described by F, that is

$$F(t, x, Dw) \leq 0$$

for all $(t, x) \in (0, T) \times \mathbb{R}^N$. To conclude $u \leq v$, i.e $w \leq 0$ on mentioned region, we locally compare w to a strict supersolution $\phi : [t_\phi, T] \times \overline{\Omega'} \to \mathbb{R}$ of F, meaning

$$F(t, x, D\phi) > 0$$

for all $(t, x) \in (t_\phi, T) \times \Omega'$. More precisely, we exclude the possibility of $w - \phi$ attaining a global maximum in $[t_\phi, T) \times \Omega'$. Excluding local extrema in $(t_\phi, T) \times \Omega'$ is an immediate consequence of the conflicting sub-and supersolution inequalities of w and ϕ together with Fermat's rule, which yields $Dw = D\phi$ for *interior* extrema of $w - \phi$. Unfortunately Fermat's rule does not necessarily apply to extrema of the form (t_ϕ, x^*) located in $\{t_\phi\} \times \Omega'$ and another strategy is needed to derive the contradiction $F(t_\phi, x^*, D\phi(t_\phi, x^*)) \leq 0$. This is achieved by the means of the following, more general lemma.

Lemma 2.2.1. *Let $\Omega' \subset \mathbb{R}^{N-1}$ be an open domain. Set $\Omega \coloneqq (a, b) \times \Omega'$. Suppose $w \in C^1(\Omega) \cap C(\overline{\Omega})$ satisfies*

$$F(t, x, w, Dw) \leq 0$$

for all $(t, x) \in \Omega$, where $F : [a, b) \times \Omega' \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ is a continuous function, and the mapping

$$s \mapsto F(t, x, r, s, p_1, \ldots, p_{N-1})$$

is non-increasing for every $(t, x, r, p_1, \dots, p_{N-1})$*. Let* $\phi \in C^1(\Omega)$*, s.t. its total differential continuously extends to* $[a, b) \times \Omega'$*.*
If $w - \phi$ *attains a local maximum in some point* $(\bar{t}, \overline{x}) \in [a, b) \times \Omega'$*, then*

$$F(\bar{t}, \overline{x}, w(\bar{t}, \overline{x}), D\phi(\bar{t}, \overline{x}))) \leq 0.$$

Proof. The proof is analogous to the one given in [1, p. 41]. As discussed above, the claim is an immediate consequence of Fermat's rule, if $(\bar{t}, \overline{x})$ is an interior extrema.
Now for the critical case, in which $(\bar{t}, \overline{x}) = (a, \overline{x})$ is a local maximum of $w - \phi$ with respect to some neighbourhood $(a, a+r] \times B_r(\overline{x}) \subset\subset [a, b) \times \Omega'$.
Add the penalty term $-\frac{1}{n(t-a)}$ to $w - \phi$, i.e consider the functions

$$\psi_n(t, x) = (w(t, x) - \phi(t, x)) - \frac{1}{n(t-a)} \ .$$

Since $w - \phi$ is bounded from above by $w(a, \overline{x}) - \phi(a, \overline{x})$ when restricted to $(a, a+r] \times B_r(\overline{x})$,

$$\lim_{t \searrow a} \psi_n(t, x) = -\infty$$

uniformly with respect to $x \in B_r(\overline{x})$, and ψ_n is known to admit a local maximizer $(t_n, x_n) \in (a, a+r] \times B_r(\overline{x})$.
After taking a subsequence if necessary, $(t_n, x_n)_{n \in \mathbb{N}}$ converges to some point $(t^*, x^*) \in [a, a+r] \times \overline{B_r(\overline{x})}$. The goal is to show that $(t_n, x_n)_{n \in \mathbb{N}}$ actually converges to $(a, \overline{x})$. For this purpose assume $(a, \overline{x})$ was a *strict* local maximum. Otherwise replace ϕ by

$$(t, x) \mapsto \phi(t, x) - \frac{\|(t, x) - (a, \overline{x})\|^2}{2},$$

which leaves the total derivative of ϕ in $(a, \overline{x})$ unchanged.
Observe that $(\psi_n(t_n, x_n))_{n \in \mathbb{N}}$ is a non-decreasing sequence bounded from above; consequently converging to some limit l. Firstly, note

$$\psi_n(t_n, x_n) \leq (w - \phi)(t_n, x_n) \to (w - \phi)(t^*, x^*)$$

,
and therefore $l \leq (w - \phi)(t^*, x^*)$. Secondly, consider the sequence (s_n, z_n), defined by

$$(s_n, z_n) \coloneqq \left(a + \frac{1}{\sqrt{n}}, \overline{x}_1, \dots, \overline{x}_{N-1}\right).$$

Using that

$$\psi_n(t_n, x_n) \geq \psi_n(s_n, z_n) = (w - \phi)(s_n, z_n) - \frac{\sqrt{n}}{n}$$

conclude that $l \geq (w-\phi)(a,\overline{x})$ by taking the limit $n \to \infty$. Altogether we obtain that $(w-\phi)(a,\overline{x}) \leq (w-\phi)(t^*,x^*)$. Since $(a,\overline{x})$ was supposed to be a *strict* maximum, we know that $(t^*,x^*) = (a,\overline{x})$.
To complete the proof, note that (t_n,x_n) is an *interior* local maximum of ψ_n, and therefore the total derivative of w in (t_n,x_n) equals the total derivative of the mapping

$$(t,x) \mapsto \phi(t,x) + \frac{1}{n(t-a)}$$

in (t_n,x_n), i.e

$$Dw(t_n,x_n) = \left(\phi_t(t_n,x_n) - \frac{1}{n(t-a)^2}, D_x\phi(t_n,x_n)\right) .$$

Since w is a subsolution, we have

$$F\left(t_n,x_n,w(t_n,x_n),\phi_t(t_n,x_n) - \frac{1}{n(t-a)^2}, D_x\phi(t_n,x_n)\right) \leq 0 \ ,$$

and the monotonicity of

$$s \mapsto F(t_n,x_n,w(t_n,x_n),s,D_x\phi(t_n,x_n))$$

yields

$$F(t_n,x_n,w(t_n,x_n),\phi_t(t_n,x_n),D_x\phi(t_n,x_n)) \leq 0 \ .$$

Taking the limit $n \to \infty$ proves the claim. □

Having laid the technical groundwork to compare sub-and supersolutions, the proof of the following comparison result boils down to derive an appropriate PDE for w and construct suitable strict supersolutions. The closed regions to which will apply the theorem, are the closed cones

$$\mathcal{C}_{z_0,R} \coloneqq \left\{(t,x) \in \left[T - \frac{R}{C}, T\right] : |x - z_0| + C(T-t) \leq R\right\} .$$

Theorem 2.2.2 (Local comparison of classical solutions)**.** *Consider the open domain $\Omega \coloneqq (0,T) \times B_R(z_0)$ and let u, $v \in C^1(\Omega) \cap C(\overline{\Omega})$ be respectively sub-and supersolutions of some PDE of the form*

$$-u_t + H(t,x,-D_x u) = 0 \ .$$

If $H : \Omega \times \mathbb{R}^N \to \mathbb{R}$ is C-Lipschitz continuous w.r.t the gradient variable, that is if H satisfies

$$|H(t,x,p) - H(t,x,q)| \leq C|p-q| \tag{H1}$$

then $w \coloneqq u - v$ satisfies

$$- w_t - C|D_x w| \leq 0 \ . \tag{2.12}$$

If additionally the subsolution u is lesser or equal to the supersolution v on $\{T\} \times \overline{B_R(z_0)}$, then the same applies to the whole cone $\mathcal{C}_{z_0,R}$.

Proof. We begin the proof by showing that the difference w satisfies 2.12. Since u and v are respectively sub-and supersolutions, we have by definition

$$-u_t + H(t, x, -D_x u) \leq 0 \leq -v_t + H(t, x, -D_x v).$$

Reordering the Hamiltonians on the right-hand-side, and the derivatives on the left one, yields

$$-w_t \leq |H(t, x, -D_x u) - H(t, x, -D_x v)| \ .$$

The Lipschitz-continuity w.r.t the gradient variable (H1) immediately gives

$$-w_t \leq C|D_x w| \ ,$$

which is obviously equivalent to inequality 2.12.
To conclude $u \leq v$, i.e $w \leq 0$ on $\mathcal{C}_{x_0,R}$, compare w with strict supersolutions

$$\begin{aligned} \phi_\delta : [t_\delta, T] \times \overline{B_R(z_0)} &\to \mathbb{R} \\ (t, x) &\mapsto \chi_\delta\left(|x - x_0| + C(T - t)\right) + \delta(T - t) \end{aligned}$$

of 2.12, where $t_\delta := T - \frac{R-\delta}{C}$ and χ_δ is some smooth function, with

$$\chi_\delta(y) = \begin{cases} 0 & \text{, if } y \leq R - \delta \\ \max\limits_{[0,T]\times\overline{B_z(x_0)}} w & \text{, if } y \geq R \end{cases}$$

as described in [2, p. 73]. An explicit computation of the difference quotients of ϕ_δ w.r.t the state variables, shows that ϕ_δ is indeed a smooth, strict supersolution of 2.12 in $(t_\delta, T) \times B_R(z_0)$, and that $D\phi_\delta$ continuously extends to $[t_\delta, T) \times B_R(z_0)$. In view of the conflicting sub-and supersolution conditions of w and ϕ_δ in combination with lemma 2.2.1, the continuous function $w - \phi_\delta$ attains its global maximum over the compact set $[t_\delta, T] \times \overline{B_R(z_0)}$, in $[t_\delta, T] \times \partial B_R(z_0)$. By construction of ϕ_δ, we have $w \leq \phi_\delta$, i.e $w - \phi_\delta \leq 0$ in $[t_\delta, T] \times \partial B_R(z_0)$, and therefore $w - \phi_\delta \leq 0$ on the whole domain $[t_\delta, T] \times \overline{B_R(z_0)}$. Now observe that $\phi_\delta \leq 0$ in $\mathcal{C}_{z_0,R-\delta}$ by construction, and since $w \leq \phi_\delta$, the same applies to w. Taking the limit $\delta \searrow 0$, completes the proof. □

2.2.2 From local to global

In this section we want to extend our local comparison result, namely theorem 2.2.2, to a global one. Using the Lipschitz-conditions imposed onto our problem-parameters, we can easily verify

$$|H(t,x,p) - H(t,x,q)| \leq L\,(1+|x|)\,|p-q| \ , \tag{2.13}$$

for all $x \in \mathbb{R}^N$, as stated in [10, p. 167]. The most obvious Lipschitz-constant from inequality (H1) is now given by $C = L(1+|z_0|+R)$, when considering the domain $(0,T) \times B_R(z_0)$. The time interval spanned by the cone $\mathcal{C}_{z_0,R}$ is consequently tending towards zero with increasing $|z_0|$, if the choice of the radius R is left to be constant. The proof of the following corollary essentially consists in choosing a suitable radius R for each z_0, in order to cover the whole domain.

Corollary 2.2.2.1. *Let u and $v \in C^1((0,T) \times \mathbb{R}^N) \cap C([0,T] \times \mathbb{R}^N)$ be respectively sub-and supersolutions of some PDE of the form*

$$-u_t + H(t,x,-D_x u) = 0$$

where H satisfies 2.13.
If the subsolution u is lesser or equal to the supersolution v in $\{T\} \times \mathbb{R}^N$, then the same applies to the whole domain $[0,T] \times \mathbb{R}^N$.

Proof. We want to choose for every $x_0 \in \mathbb{R}^N$, a radius $R(z_0) > 0$, such that the time interval spanned by the cone $\mathcal{C}_{z_0,R(z_0)}$ covers the whole interval $[0,T]$. By the definition of $\mathcal{C}_{z_0,R}$, with $C = L(1+|z_0|+R)$, this is nothing else than solving the equation

$$\frac{R}{L(1+|z_0|+R)} = T \ . \tag{2.14}$$

Note that we may assume w.l.o.g that $T < 1/L$, since we could otherwise partition the interval $[0,T]$ accordingly, and iterate the following argument. If T is strictly less than $1/L$, equation 2.14 admits exactly one positive solution, $R(z_0)$, given by

$$R(z_0) = \frac{LT(1+|z_0|)}{(1-LT)} \ .$$

Choosing $R(z_0)$ this way, yields $[0,T] \times \mathbb{R}^N = \bigcup_{z_0 \in \mathbb{R}^N} \mathcal{C}_{z_0,R(z_0)}$. Applying theorem 2.2.2 to the domains $[0,T] \times \overline{B}_{R(z_0)}(z_0)$ completes the proof. □

2.3 A quick reality check

We want to put the practicability of our results to the test. More precisely, we want to check whether or not it is sensible to assume the value function being smooth. We therefore consider the very simple control problem, starting in $(t_0, x_0) \in [0,T] \times \mathbb{R}^N$, with elementary dynamics

$$\dot{x} = \alpha$$

for measurable controls $\alpha : [t_0, T] \to [-1, 1]$, having zero running costs, and terminal costs $h = |\cdot|$. This means our objective is to minimize $h(x_{x_0}^{\alpha}(T))$ over all possible controls α.
It is quite natural to perceive the considered control problem, as to move the system's state as close as possible to the zero state with only limited velocity. Intuitively, we would transform the system with maximal velocity, and abruptly stop as soon as we reach the desired state. We formalize this procedure by defining the control

$$\alpha^* : t \mapsto \begin{cases} \operatorname{sign}(x_0) & \text{if } |x_0| \geq t - t_0 \\ 0 & \text{otherwise} \end{cases} .$$

Indeed our intuition is not misleading and we can explicitly compute the value function in (t_0, x_0), which gives:

$$\begin{aligned} v(t_0, x_0) &= \begin{cases} 0, & \text{if } |x_0| \leq T - t_0 \\ |x_0| - (T - t_0), & \text{otherwise} \end{cases} \\ &= \max(0, |x_0| - (T - t_0)) \ . \end{aligned}$$

The value function is obviously not smooth. As our assumption already fails on such a simple example, we must weaken our notion of solution.
Recall that under the assumptions imposed onto our problem parameters, the value function is at least known to be locally Lipschitz-continuous, and consequently differentiable almost everywhere by Rademacher's theorem, found in [7]. We are therefore tempted to interpret 2.3 in the so-called *generalized sense*, which means we require solutions to be locally Lipschitz-continuous and to satisfy 2.3 almost everywhere. Our simple example still reveals itself problematic in this case, as the associated PDE 2.3, now given by

$$-u_t + |u_x| = 0 \ ,$$

admits the functions v and $(t_0, x_0) \mapsto |x_0| - (T - t_0)$ as generalized solutions, both satisfying the boundary condition $u(T, \cdot) = h$, as noted by Barles in [2]. Another sense of solution is therefore needed, if we insist on characterizing the value function as the unique solution of the considered boundary problem.

Chapter 3

Merely continuous value functions

3.1 Generalizing the total derivative

The example from section 2.3 shows that it is impractical to assume the smoothness of the value function. We therefore extract the essential properties of its classical total derivative, used to derive the HJB-equation 2.3 and the subsequent comparison results in the smooth case. Lemma 2.2.1, the key to our previous comparison and uniqueness results, proves itself particularly helpful in our quest.
Given a smooth subsolution w of

$$F(t, x, w, Dw) \leq 0 \;, \tag{3.1}$$

the proof of lemma 2.2.1 was based on the fact, that for every smooth function ϕ, the inequality

$$F(t_0, x_0, w(x_0, x_0), D\phi(t_0, x_0)) \leq 0$$

is satisfied, whenever $w - \phi$ attains an interior local maximum in (t_0, x_0). That is, we can replace the classical total derivative of w in 3.1, by the derivative of ϕ, in local maxima of $w-\phi$. In view of this observation, the following generalization seems sensible.

Definition 3.1.1. Let $\Omega \subset \mathbb{R}^N$ be an open domain. A real-valued continuous function $u : \Omega \to \mathbb{R}$ is called a *viscosity subsolution* of $F : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$, if for any smooth function $\phi \in C^1(\Omega)$, we have that

$$F(x_0, u(x_0), Du(x_0)) \leq 0 \;, \tag{3.2}$$

whenever $u - \phi$ attains a local *maximum* in $x_0 \in \Omega$.
We respectively call u a *viscosity supersolution* of F, if

$$F(x_0, u(x_0), Du(x_0)) \geq 0 \tag{3.3}$$

whenever $u - \phi$ attains local *minimum* in $x_0 \in \Omega$.
We finally call u a *viscosity solution* of F, if it is a viscosity sub-and supersolution of F.

Of course the notion of viscosity solution is only of interest, if the value function is a viscosity solution of 2.3.

Theorem 3.1.1. *Suppose the same assumptions as in theorem 2.1.1, except for the smoothness of the value function v, hold. Then v is a viscosity solution of 2.3.*

Proof. The proof of theorem 2.1.1 only requires minor modifications.
We have already stated that v is known to be locally Lipschitz-continuous, and consequently continuous, under the given assumptions.
In order to show that v is a viscosity subsolution of 2.3, note that

$$\phi(t,x) - \phi(t_0, x_0) \geq v(t,x) - v(t_0, x_0) \ ,$$

in a sufficiently small neighbourhood of (t_0, x_0), whenever $v - \phi$ attains a local maximum in (t_0, x_0). The left-hand-side of 2.7, can therefore be replaced by

$$\frac{\phi(t_0 + h, x_{x_0}^{\overline{a}}(t_0 + h)) - \phi(t_0, x_0)}{h} \ .$$

Now proceed as in theorem 2.1.1 to show that v is viscosity subsolution of 2.3. Showing that v is a supersolution of 2.3 can be done in an analogous manner. □

Up to now, the notion of viscosity solutions, first introduced in [5] by M.G. Crandall and P.L. Lions, seems to fit our needs. Of course we still have to check whether the comparison and uniqueness results proven in the smooth case still remain valid, if replaced by viscosity solutions.

3.2 From classical solutions to viscosity solutions

Our goal is to transfer the results proven for classical solutions, namely lemma 2.2.1, theorem 2.2.2 and corollary 2.2.2.1 to viscosity solutions. We start with lemma 2.2.1.

Lemma 3.2.1. *Lemma 2.2.1 still holds for merely continuous viscosity subsolutions.*

Proof. The claim holds immediately for interior extrema by the definition of viscosity subsolutions, which replaces the use of Fermat's rule in the smooth case. For boundary points, proceed exactly as for smooth subsolutions. □

Now for theorem 2.2.2. The proof given consisted of exactly two steps: firstly showing that the difference w is a subsolution to equation 2.12, namely

$$-w_t - C|D_x w| \leq 0 \ ,$$

and secondly comparing w to some supersolutions of the same equation. Latter step simply relies on lemma 2.2.1 which has already been further generalized to continuous viscosity subsolutions in lemma 3.2.1. In contrast, the generalization of the first step proves itself trickier.

3.2.1 A naive approach

We want to deduce that $w := u - v$ is a viscosity subsolution of 2.12, provided that u and v are respectively viscosity sub-and supersolutions of

$$-u_t + H(t, x, -D_x u) = 0 \ .$$

For this purpose, suppose we are given an arbitrary subdifferential in some point (t_0, x_0). We could naively try to decompose this subdifferential of w as the difference of a subdifferential of u and a superdifferential of v in the *same* point (t_0, x_0), and proceed exactly as in the smooth case. This approach reveals itself being too optimistic, as we consider the continuous function

$$\begin{aligned} u : \mathbb{R} &\to \mathbb{R} \\ r &\mapsto \begin{cases} |r|^{1/2} \sin(\frac{1}{r^2}), & \text{if } r \neq 0 \\ 0 & \text{if } r = 0 \ , \end{cases} \end{aligned}$$

taken from [1, p. 32, exercise 1.5]. Suppose for the sake of contradiction, that $Du^+(0)$ was not empty. Then there exists a function $\phi \in C^1(\mathbb{R})$, such that $f - \phi$ attains a local maximum in zero, i.e

$$f(0) - f(0+h) \geq \phi(0) - \phi(0+h)$$

for small h. For strictly positive h, we additionally get

$$\frac{f(0+h) - f(0)}{h} \leq \frac{\phi(0+h) - \phi(0)}{h} \ ,$$

and consequently

$$\limsup_{h \searrow 0} \frac{f(0+h) - f(0)}{h} \leq \limsup_{h \searrow 0} \frac{\phi(0+h) - \phi(0)}{h} \ .$$

But

$$\limsup_{h \searrow 0} \frac{f(0+h) - f(0)}{h} = \infty \ ,$$

contradicting the differentiability of ϕ in zero. An analogous argument shows that also $D^{-}u(0)$ must be empty.
On the other hand, the function $u - u = \hat{0}$ admits a superdifferential in zero, namely its classical derivative.
The next subsection shows a way to circumvent this problem.

3.2.2 A fundamental technique

The example from subsection 3.2.1 deters us from decomposing a superdifferential of $w = u - v$ in (t_0, x_0), using super-and subdifferentials of u and v in the same point (t_0, x_0), since $D^{+}u(t_0, x_0)$ or $D^{-}v(t_0, x_0)$ might be empty. Suppose we only required to find some (t, x) in an arbitrarily small vicinity $B_R(t_0, x_0)$ of (t_0, x_0), s.t $D^{+}u(t, x)$ is not empty. For any smooth function ϕ, we are sure that $u-\phi$ achieves a global maximum over $\overline{B}_R(t_0, x_0)$ in some point (t, x). But since (t, x) might lie on the boundary of $\overline{B}_R(t, x)$, it is not clear whether (t, x) is also a local maximizer with respect to the whole domain $(0, T) \times \mathbb{R}^N$. To avoid this phenomena, construct ϕ in a way which penalizes the distance from (t_0, x_0). To make this idea more concrete, consider the proof of the following theorem.

Theorem 3.2.2. *Let $\Omega \subset \mathbb{R}^N$ be an open domain, and $u : \Omega \to \mathbb{R}$ a continuous function. The set*

$$\left\{x \in \Omega : D^{+}u(x) \neq \emptyset\right\}$$

is dense in Ω. The same applies when $D^{+}u(x)$ is replaced by $D^{-}u(x)$.

Proof. We restate the proof of Lemma 1.8 in [1, p. 30]. Let $\overline{x} \in \Omega$ and consider the smooth function $\phi_\epsilon(x) = |x - \overline{x}|^2/\epsilon$. For any $\epsilon > 0$, the function $u - \phi_\epsilon$ attains its maximum over $\overline{B} = \overline{B}_R(\overline{x})$, in some point x_ϵ. From the inequality

$$\phi_\epsilon(x_\epsilon) \geq \phi_\epsilon(\overline{x}) = u(\overline{x})$$

we get, for all $\epsilon > 0$, that

$$|x_\epsilon - \overline{x}|^2 \leq 2\epsilon \sup_{x \in \overline{B}} |u(x)|.$$

Thus x_ϵ does not lie on the boundary of $\overline{B}$ for ϵ small enough, and $D\phi(x_\epsilon) = 2(x_\epsilon - \overline{x})/\epsilon$ belongs to $D^{+}u(x_\epsilon)$. Similar arguments prove the claim for subdifferentials. □

Remark. The proof of theorem 3.2.2 actually shows that in any vicinity of $\overline{x}$, there is a point x_ϵ, which admits a sub-(super)differential of the form $|x_\epsilon - \overline{x}|/\epsilon$.

We now elaborate how to modify our first, simplistic attempt. Note that owing to theorem 3.2.2, we can choose points $(\overline{t}, \overline{x})$ and $(\overline{s}, \overline{y})$ in some arbitrarily small vicinity of (t_0, x_0), s.t. $D^{+}u(\overline{t}, \overline{x})$ and $D^{-}v(\overline{s}, \overline{y})$ are non-empty. Select a superdifferential $p = p(\overline{t}, \overline{x})$ from $D^{+}u(\overline{t}, \overline{x})$ and a subdifferential $q = q(\overline{s}, \overline{y})$

from $D^-v(\overline{s},\overline{y})$. Use the sub-and supersolution inequalities of u and v as in the smooth case to derive that

$$-(p_t - q_s) \leq |H(\overline{t},\overline{x},-p_x) - H(\overline{s},\overline{y},-q_y)| \ .$$

In order to benefit from the Lipschitz-condition (H1) satisfied by H, vary the arguments of H on the right-hand-side in an iterative manner and apply the triangle inequality to obtain the estimate:

$$\begin{aligned}|H(\overline{t},\overline{x},-p_x) - H(\overline{s},\overline{y},-q_y)| \leq\ & |H(\overline{t},\overline{x},-p_x) - H(\overline{s},\overline{x},-p_x)| \\ & +|H(\overline{s},\overline{x},-p_x) - H(\overline{s},\overline{y},-p_x)| \\ & +|H(\overline{s},\overline{y},-p_x) - H(\overline{s},\overline{y},-q_y)| \ .\end{aligned}$$

In view of assumption (H1), the last summand can be replaced by $C|p_x - q_y|$, and we conclude that

$$\begin{aligned}|H(\overline{t},\overline{x},-p_x) - H(\overline{s},\overline{y},-q_y)| \leq\ & |H(\overline{t},\overline{x},-p_x) - H(\overline{s},\overline{x},-p_x)| \\ & +|H(\overline{s},\overline{x},-p_x) - H(\overline{s},\overline{y},-p_x)| \\ & +C|p_x - q_s| \ .\end{aligned} \tag{3.4}$$

Equation 3.4 suggests that we should select p and q in way which ensures that p and q both tend towards the given superdifferential of w, as $(\overline{t},\overline{x})$ and $(\overline{s},\overline{y})$ approach (t_0,x_0). According to exercise 2.4 c) in [1, p. 49] we have

$$D^+(u-\phi)(\overline{t},\overline{x}) = D^+u(\overline{t},\overline{x}) - D\phi(\overline{t},\overline{x}) \ ,$$

for any $(\overline{t},\overline{x})$. Applying remark 3.2.2 to $u-\phi$ and v, we could find for sufficiently small $\epsilon > 0$, points $(\overline{t},\overline{x}) = (\overline{t},\overline{x})_\epsilon$ and $(\overline{s},\overline{y}) = (\overline{s},\overline{y})_\epsilon$ respectively admitting super-and subdifferentials

$$\begin{aligned}p &= D\phi(\overline{t},\overline{x}) + \frac{(\overline{t},\overline{x}) - (t_0,x_0)}{\epsilon} \\ q &= \frac{(\overline{s},\overline{y}) - (t_0,x_0)}{\epsilon} \ ,\end{aligned}$$

with respect to u and v. Unfortunately, it is not clear whether or not

$$\frac{(\overline{t},\overline{x}) - (\overline{s},\overline{y})}{\epsilon}$$

and consequently $|p-q|$ respectively converge to zero and the given subdifferential of w, as ϵ tends towards zero. Seeking to proceed with $u-\phi$ and v as described in the proof of theorem 3.2.2 while penalizing the distance between $(\overline{t},\overline{x})$ and $(\overline{s},\overline{x})$, we assume that $w-\phi$ attains a local maximum in (t_0,x_0) w.r.t the neighbourhood $B_R(t_0,x_0) \subset\subset (0,T) \times \mathbb{R}^N$ and consider the continuous function

$$\Phi = \Phi_{\epsilon,\alpha} : \overline{B}_{R/2}(t_0, x_0)^2 \to \mathbb{R}$$
$$(t, x, s, y) \mapsto (u(t, x) - \phi(t, x)) - v(s, y) - \frac{|x - y|^2}{2\epsilon} - \frac{|t - s|^2}{2\alpha} .$$

The purpose of the additional parameter $\alpha > 0$ will reveal itself, once we study the convergence of the right-hand-side of inequality 3.4 with respect to ϵ and α. If we additionally assume w.l.o.g that (t_0, x_0) is even a *strict* local maximizer, the functions $\Phi_{\epsilon,\alpha}$ implicitly penalize the distance to (t_0, x_0), as the penalty for (t, x) and (s, y) gets harsher and harsher with ϵ and α tending towards zero. As the next lemma confirms, we can indeed generate points $(\overline{t}, \overline{x})$ and $(\overline{s}, \overline{y})$ with the desired features, by maximizing the functions $\Phi_{\epsilon,\alpha}$.

Lemma 3.2.3. *Let $(\overline{t}, \overline{x}, \overline{s}, \overline{y}) = (\overline{t}, \overline{x}, \overline{s}, \overline{y})_{\epsilon,\alpha}$ be a global maximizer of the continuous function $\Phi_{\epsilon,\alpha}$, defined on the compact set $\overline{B}_R(t_0, x_0)^2$. Then:*

(i)

$$|\overline{x} - \overline{y}| \to 0 \ \text{as } \epsilon \searrow 0, \ \text{uniformly w.r.t} \ \alpha$$
$$|\overline{t} - \overline{s}| \to 0 \ \text{as } \alpha \searrow 0, \ \text{uniformly w.r.t} \ \epsilon$$

(ii) $(u-\phi)(\overline{t}, \overline{x}) - v(\overline{s}, \overline{y}) \to (w-\phi)(t_0, x_0)$ *and the components* $(\overline{t}, \overline{x})$ *and* $(\overline{s}, \overline{y})$ *of the maximizers converge towards* (t_0, x_0) *as* $\epsilon, \alpha \searrow 0$ *for some subnet of* $(\overline{t}, \overline{x}, \overline{s}, \overline{y})_{\epsilon,\alpha}$*. For the same subnet, the vectors* p *and* q*, with*

$$p = \left(\frac{\overline{t} - \overline{s}}{\alpha}, \frac{\overline{x} - \overline{y}}{\epsilon} \right) + D\phi(\overline{t}, \overline{x})$$
$$q = \left(\frac{\overline{t} - \overline{s}}{\alpha}, \frac{\overline{x} - \overline{y}}{\epsilon} \right)$$

respectively belong to $D^+u(\overline{t}, \overline{x})$ *and* $D^-v(\overline{s}, \overline{y})$*, for sufficiently small* ϵ *and* α*.*

(iii) For the subnet mentioned in claim (ii), we even have that

$$\frac{|\overline{x} - \overline{y}|^2}{\epsilon} \to 0$$
$$\frac{|\overline{t} - \overline{s}|^2}{\alpha} \to 0$$

as $\epsilon, \alpha \searrow 0$.

Proof. We adapt the proof of Lemma 5.2 in [2, p. 69].
Note that modifying ϕ by an additive constant neither changes the location of its extrema nor its total derivative. We might therefore assume in the following that $M := (w - \phi)(t_0, x_0)$ is strictly positive.

To show claim (i), use that

$$\Phi(\bar{t},\bar{x},\bar{s},\bar{y}) \geq \Phi(t_0,x_0,t_0,x_0) = M > 0 \tag{3.5}$$

implies that

$$\frac{|x-y|^2}{2\epsilon} + \frac{|t-s|^2}{2\alpha} \leq u(\bar{t},\bar{x}) - \phi(\bar{t},\bar{x}) - v(\bar{s},\bar{y})$$

and since the right-hand side is bounded by the suprema of u, v and ϕ over the compact set $\overline{B}_{R/2}(t_0,x_0)$, we get that

$$|\bar{t}-\bar{s}| \leq \sqrt{2K\alpha}$$
$$|\bar{x}-\bar{y}| \leq \sqrt{2K\epsilon}$$

for some constant $K \geq 0$.
We continue with claim (ii). Note that

$$\Phi(\bar{t},\bar{x},\bar{s},\bar{y}) \leq (u-\phi)(\bar{t},\bar{x}) - v(\bar{s},\bar{y}) \tag{3.6}$$

by stripping the non-negative penalty terms off $\Phi(\bar{t},\bar{x},\bar{s},\bar{y})$. By combining inequalities 3.5 and 3.6, deduce

$$(u-\phi)(\bar{t},\bar{x}) - v(\bar{s},\bar{y}) \geq M > 0\ . \tag{3.7}$$

Since $(\bar{t},\bar{x},\bar{s},\bar{y})$ is contained in the compact set $\overline{B}_{R/2}(t_0,x_0)$, it admits a convergent subnet. In view of claim (i), its limit point is of the form (t^*,x^*,t^*,x^*). As ϵ and α tend towards zero, the left-hand-side of inequality 3.7 converges to $(w-\phi)(t^*,x^*)$, and therefore

$$(w-\phi)(t^*,x^*) \geq M\ .$$

Since (t_0,x_0) is a strict local maximum, we have that $(t^*,x^*) = (t_0,x_0)$ which proves

$$(u-\phi)(\bar{t},\bar{x}) - v(\bar{s},\bar{y}) \to (w-\phi)(t_0,x_0)$$
$$(\bar{t},\bar{x}),\ (\bar{s},\bar{y}) \to (t_0,x_0)\ .$$

To see that the vectors p and q are indeed sub-and superdifferentials, note that $(\bar{t},\bar{x})$ and $(\bar{s},\bar{y})$ are interior points of $\overline{B}_{R/2}(t_0,x_0)$ for sufficiently small ϵ and α, since $(\bar{t},\bar{x})$ and $(\bar{s},\bar{y})$ converge to (t_0,x_0). The mappings $\Phi(\cdot,\cdot,\bar{s},\bar{y}) = u - \phi_1$ and $\Phi(\bar{t},\bar{x},\cdot,\cdot) = v - \phi_2 : (0,T)\times\mathbb{R}^N \to \mathbb{R}$, where

$$\phi_1(t,x) := \phi(t,x) + \frac{|x-\bar{y}|^2}{2\epsilon} + \frac{|t-\bar{s}|^2}{2\alpha} + v(\bar{s},\bar{y})$$
$$\phi_2(s,y) := -\frac{|\bar{x}-y|^2}{2\epsilon} - \frac{|\bar{x}-s|^2}{2\alpha} + u(\bar{t},\bar{x}) + \phi(\bar{t},\bar{x})\ ,$$

therefore respectively attain a local maximum in $(\bar{t}, \bar{x})$ and a local minimum in $(\bar{s}, \bar{y})$. Observing that $p = D\phi_1(\bar{t}, \bar{x})$ and $q = D\phi_2(\bar{s}, \bar{y})$ completes the proof of claim (ii).
Prove claim (iii) by using inequality (3.5), which implies that

$$\frac{|\bar{x}-\bar{y}|^2}{2\epsilon} + \frac{|\bar{t}-\bar{s}|^2}{2\alpha} \leq \left[(u-\phi)(\bar{t}, \bar{x}) - v(\bar{s}, \bar{y})\right] - M \ .$$

For the subnet mentioned in claim (ii), the right-hand-side converges to zero, as ϵ, α tend towards zero. Therefore claim (iii) holds, and lemma 3.2.3 is proven. $\square$

This method of *doubling the variables* goes back to Kružkov who employed a similar function in [8]. It is this technique which also lies at the heart of Lions' and Crandall's uniqueness results for viscosity solutions, first introduced in [6]. We come back to our task of showing that w is a viscosity subsolution of 2.12. We considered inequality 3.4 to determine suitable candidates for superdifferentials p of u, and subdifferentials q of v, which we have now generated using lemma 3.2.3 (ii). After plugging the corresponding values for p and q into inequality 3.4, the first summand $|H(\bar{t}, \bar{x}, -p_x) - H(\bar{s}, \bar{x}, -p_x)|$, on its right-hand-side, becomes

$$\left| H\left(\bar{t}, \bar{x}, -\left(\frac{\bar{x}-\bar{y}}{\epsilon} + D_x\phi(\bar{t}, \bar{x})\right)\right) - H\left(\bar{s}, \bar{x}, -\left(\frac{\bar{x}-\bar{y}}{\epsilon} + D_x\phi(\bar{t}, \bar{x})\right)\right)\right|$$

and the second summand $|H(\bar{s}, \bar{x}, -p_x) - H(\bar{s}, \bar{y}, -p_x)|$, becomes

$$\left| H\left(\bar{s}, \bar{x}, -\left(\frac{\bar{x}-\bar{y}}{\epsilon} + D_x\phi(\bar{t}, \bar{x})\right)\right) - H\left(\bar{s}, \bar{y}, -\left(\frac{\bar{x}-\bar{y}}{\epsilon} + D_x\phi(\bar{t}, \bar{x})\right)\right)\right| \ .$$

As opposed to the first difference, reducing the second difference to an arbitrarily small quantity, by choosing ϵ and α sufficiently small, proves itself trickier, since claim (iii) only provides that

$$\frac{|\bar{x}-\bar{y}|^2}{\epsilon} \to 0$$

as ϵ and α tend towards zero, but does not say anything about the convergence of $\bar{x} - \bar{y}/\epsilon$. The slower convergence of $|\bar{x} - \bar{y}|$, provided by claim (iii), becomes sufficient, if the Hamiltonian H additionally satisfies

$$|H(t,x,p) - H(t,y,p)| \leq \omega(R, |x-y|(1+|p|)) \text{ for all } \ x, y \in B_R(0) \ , \qquad \text{(H2)}$$

where $\omega : [0,\infty)^2 \to [0,\infty)$ describes some modulus of continuity. As assumption (H2) might seem far-fetched, note that the Hamiltonian, corresponding to our considered control problem, satisfies (H2) on account of the Lipschitz-conditions imposed onto the problem parameters (cf. [10, p. 167]). We are now ready to adapt theorem 2.2.2 to the case of viscosity solutions.

Theorem 3.2.4. *Theorem 2.2.2 still holds in the viscosity sense, if the given Hamiltonian H is continuous and additionally satisfies* (H2).

Proof. We show that w is a viscosity subsolution of 2.12, namely:

$$-w_t - C|D_x w| \le 0 \ .$$

For this purpose, consider an arbitrary point $(t_0, x_0) \in (0, T) \times B_R(z_0)$, and let $\phi : (0, T) \times B_R(z_0) \to \mathbb{R}$ be a smooth function, s.t. $w - \phi$ admits a local maximum in (t_0, x_0). Let $(\bar{t}, \overline{x})$ and $(\overline{s}, \overline{y})$ denote the points constructed in lemma 3.2.3. According to claim (ii) from latter lemma, the same points $(\bar{t}, \overline{x})$ and $(\overline{s}, \overline{y})$ respectively admit the super-and subdifferentials

$$p = \left(\frac{\bar{t} - \overline{s}}{\alpha}, \frac{\overline{x} - \overline{y}}{\epsilon} \right) + D\phi(\bar{t}, \overline{x})$$
$$q = \left(\frac{\bar{t} - \overline{s}}{\alpha}, \frac{\overline{x} - \overline{y}}{\epsilon} \right)$$

with respect to the functions u and v. Add the sub-and supersolutions inequalities satisfied by u and v to obtain the inequality:

$$-\phi_t(\bar{t}, \overline{x}) \le \left| H\left(\bar{t}, \overline{x}, -\left(\frac{\overline{x} - \overline{y}}{\epsilon} + D_x \phi(\bar{t}, \overline{x}) \right) \right) - H\left(\overline{s}, \overline{y}, -\left(\frac{\overline{x} - \overline{y}}{\epsilon} \right) \right) \right| \ .$$

Apply inequality 3.4 with the concrete values of p and q, to deduce that

$$\begin{aligned} -\phi_t(\bar{t}, \overline{x}) \le & |H(\bar{t}, \overline{x}, -p_x) - H(\overline{s}, \overline{x}, -p_x)| \\ & + |H(\overline{s}, \overline{x}, -p_x) - H(\overline{s}, \overline{y}, -p_x)| \\ & + C|D_x \phi(\bar{t}, \overline{x})| \ . \end{aligned} \tag{3.8}$$

Since $(\bar{t}, \overline{x})$ and $(\overline{s}, \overline{y})$ converge to (t_0, x_0), as ϵ and α tend to zero, the points $(\bar{t}, \overline{x})$ and $(\overline{s}, \overline{y})$ are bounded for sufficiently small ϵ and α, and then so is $D_x \phi(\bar{t}, \overline{x})$. For fixed ϵ, the arguments of H, appearing in the first difference $|H(\bar{t}, \overline{x}, -p_x) - H(\overline{s}, \overline{x}, -p_x)|$ of inequality 3.8, are bounded. Use the uniform continuity of H on bounded sets, and the uniform convergence of $|\bar{t} - \overline{s}|$ stated in lemma 3.2.3 (i) to reduce the first difference to less than some arbitrarily small quantity $\delta/2$, by further decreasing α as necessary.
Since $D_x \phi(\bar{t}, \overline{x})$ is bounded, use assumption (H2) and claim (iii) from lemma 3.2.3 to reduce the second difference $|H(\overline{s}, \overline{x}, -p_x) - H(\overline{s}, \overline{y}, -p_x)|$ to less then $\delta/2$.
We therefore have that $-\phi_t(\bar{t}, \overline{x}) \le \delta + C|D_x \phi(\bar{t}, \overline{x})|$ for sufficiently small ϵ and α. Since the quantity δ can be chosen arbitrarily small, conclude

$$-\phi(t_0, x_0) \le C|D_x \phi(t_0, x_0)|$$

by letting ϵ and α tend towards zero. Since (t_0, x_0) and ϕ were arbitrary, we conclude that w is indeed a subsolution of 2.12.
Proceed as in the proof of theorem 2.2.2 to locally compare the subsolution u with the supersolution v. $\square$

We can now simply transfer our global comparison result for classical solutions from corollary 2.2.2.1, by additionally assuming the Hamiltonian satisfies (H2), and is continuous.

Corollary 3.2.4.1. *Corollary 2.2.2.1 also applies to viscosity sub-and supersolutions, if the Hamiltonian* $H : [0,T] \times \mathbb{R}^N \times \mathbb{R}^{N+1}$ *is continuous and satisfies* (H2).

Proof. The proof of corollary 2.2.2.1 only relied on the local comparison of sub- and supersolutions provided by theorem 2.2.2.
According to theorem 3.2.4, theorem 2.2.2 still applies in the viscosity sense under the additional assumptions. Consequently, the proof of corollary 2.2.2.1 can be repeated verbatim. $\square$

Bibliography

[1] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer Science & Business Media, 2008.

[2] Guy Barles. An introduction to the theory of viscosity solutions for first-order hamilton-jacobi equations and applications. `http://www.mathematik.uni-muenchen.de/~sorensen/Lehre/WiSe2014-15/ViscSol/Literatur/9783642364327-c1.pdf`, last accessed on 09/29/2020.

[3] Richard Bellman. On the theory of dynamic programming. *Proceedings of the National Academy of Sciences of the United States of America*, 38(8):716, 1952.

[4] Richard Bellman. Dynamic programming and a new formalism in the calculus of variations. *Proceedings of the National Academy of Sciences of the United States of America*, 40(4):231, 1954.

[5] Michael G Crandall and Pierre-Louis Lions. Viscosity solutions of hamilton-jacobi equations. *Transactions of the American mathematical society*, 277(1):1–42, 1983.

[6] Michael G Crandall and Pierre-Louis Lions. Viscosity solutions of hamilton-jacobi equations. *Transactions of the American mathematical society*, 277(1):1–42, 1983.

[7] Lawrence Craig Evans and Ronald F Gariepy. *Measure theory and fine properties of functions*. CRC press, 2015.

[8] SN Kružkov. First order quasilinear equations in several independent variables. *Mathematics of the USSR-Sbornik*, 10(2):217–243, 1970.

[9] Tomáš Roubíček. *Nonlinear partial differential equations with applications*, volume 153. Springer Science & Business Media, 2013.

[10] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.