Verification theorem and construction of epsilon-optimal controls for control of abstract evolution equations

Fabbri, Giorgio; Gozzi, Fausto and Swiech, Andrzej

May 2007

Verification theorem and construction of ϵ-optimal controls for control of abstract evolution equations

G. Fabbri* F. Gozzi† and A. Święch‡

May 16, 2007

Abstract

We study several aspects of the dynamic programming approach to optimal control of abstract evolution equations, including a class of semilinear partial differential equations. We introduce and prove a verification theorem which provides a sufficient condition for optimality. Moreover we prove sub- and superoptimality principles of dynamic programming and give an explicit construction of ϵ-optimal controls.

Key words: optimal control of PDE, verification theorem, dynamic programming, ϵ-optimal controls, Hamilton-Jacobi-Bellman equations.

MSC 2000: 35R15, 49L20, 49L25, 49K20.

1 Introduction

In this paper we investigate several aspects of the dynamic programming approach to optimal control of abstract evolution equations. The optimal control problem we have in mind has the following form. The state equation is

$$\begin{cases} \dot{x}(t) = Ax(t) + b(t, x(t), u(t)), \\ x(0) = x, \end{cases} \tag{1}$$

*DPTEA, Università *LUISS - Guido Carli* Roma and School of Mathematics and Statistics, UNSW, Sydney e-mail: gfabbri@luiss.it, G.Fabbri was supported by the ARC Discovery project DP0558539.

†Dipartimento di Scienze Economiche ed Aziendali, Università *LUISS - Guido Carli* Roma, e-mail: fgozzi@luiss.it

‡School of Mathematics, Georgia Institute of Technology Atlanta, GA 30332, U.S.A., e-mail: swiech@math.gatech.edu. A. Święch was supported by NSF grant DMS 0500270.

A is a linear, densely defined maximal dissipative operator in a real separable Hilbert space $\mathcal{H}$, and we want to minimize a cost functional

$$J(x;u(\cdot)) = \int_0^T L(t,x(t),u(t))\mathrm{d}t + h(x(T)) \tag{2}$$

over all controls

$$u(\cdot) \in \mathcal{U}[0,T] = \{u\colon [0,T] \to U :\ u \text{ is measurable}\},$$

where U is a metric space.

The dynamic programming approach studies the properties of the so called value function for the problem, identifies it as a solution of the associated Hamilton-Jacobi-Bellman (HJB) equation through the dynamic programming principle, and then tries to use this PDE to construct optimal feedback controls, obtain conditions for optimality, do numerical computations, etc.. There exists an extensive literature on the subject for optimal control of ordinary differential equations, i.e. when the HJB equations are finite dimensional (see for instance the books [12, 26, 36, 37, 47, 55, 56] and the references therein). The situation is much more complicated for optimal control of partial differential equations (PDE) or abstract evolution equations, i.e. when the HJB equations are infinite dimensional, nevertheless there is by now a large body of results on such HJB equations and the dynamic programming approach ([2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 28, 29, 30, 31, 32, 38, 41, 45, 46, 50, 51, 53, 54] and the references therein). Numerous notions of solutions are introduced in these works, the value functions are proved to be solutions of the dynamic programming equations, and various verification theorems and results on existence and explicit forms of optimal feedback controls in particular cases are established. However, despite of these results, so far the use of the dynamic programming approach in the resolution of the general optimal control problems in infinite dimensions has been rather limited. Infinite dimensionality of the state space, unboundedness in the equations, lack of regularity of solutions, and often complicated notions of solutions requiring the use of sophisticated test functions are only some of the difficulties.

We will discuss two aspects of the dynamic programming approach for a fairly general control problem: a verification theorem which gives a sufficient condition for optimality, and the problem of construction of ϵ-optimal feedback controls.

The verification theorem we prove in this paper is an infinite dimensional version of such a result for finite dimensional problems obtained in [57]. It is based on the notion of viscosity solution (see Definitions 2.4-2.6). Regarding previous result in this direction we mention [21, 22] and the material in Chapter 6 §5 of [46], in particular Theorem 5.5 there which is based on [21]. We briefly discuss this result in Remark 3.6.

The construction of ϵ-optimal controls we present here is a fairly explicit procedure which relies on the proof of superoptimality inequality of dynamic programming for viscosity supersolutions of the corresponding Hamilton-Jacobi-Bellman equation. It is a delicate generalization of such a method for the finite dimensional case from [52]. Similar method has been used in [25] to construct stabilizing feedbacks for nonlinear systems and later in [42] for state constraint problems. The idea here is to approximate the value function by its appropriate inf-convolution which is more regular and satisfies a slightly perturbed HJB inequality pointwise. One can then use this inequality to construct ϵ-optimal piecewise constant controls. This procedure in fact gives the superoptimality inequality of dynamic programming and the suboptimality inequality can be proved similarly. There are other possible approaches to construction of ϵ-optimal controls. For instance under compactness assumption on the operator B (see Section 4) one can approximate the value function by solutions of finite dimensional HJB equations with the operator A replaced by some finite dimensional operators A_n (see [28]) and then use results of [52] directly to construct near optimal controls. Other approximation procedures are also possible. The method we present in this paper seems to have some advantages: it uses only one layer of approximations, it is very explicit and the errors in many cases can be made precise, and it does not require any compactness of the operator B. It does however require some weak continuity of the Hamiltonian and uniform continuity of the trajectories, uniformly in $u(\cdot)$. Finally we mention that the sub- and superoptimality inequalities of dynamic programming are interesting on their own.

The paper is organized as follows. Definitions and the preliminary material is presented in Section 2. Section 3 is devoted to the verification theorem and an example where it applies in a nonsmooth case. In Section 4 we prove sub- and superoptimality principles of dynamic programming and show how to construct ϵ-optimal controls.

2 Notation, definitions and background

Throughout this paper $\mathcal{H}$ is a real separable Hilbert space equipped with the inner product $\langle \cdot, \cdot \rangle$ and the norm $\| \cdot \|$. We recall that A is a linear, densely defined operator such that $-A$ is maximal monotone, i.e. A generates a C_0 semigroup of contractions e^{sA}, i.e.

$$\|e^{sA}\| \le 1 \quad \text{for all } s \ge 0 \tag{3}$$

We make the following assumptions on b and L.

Hypothesis 2.1.

$$b\colon [0,T] \times \mathcal{H} \times U \to \mathcal{H} \textit{ is continuous}$$

and there exist a constant $M > 0$ and a local modulus of continuity $\omega(\cdot,\cdot)$ such that

$$\begin{aligned} &\|b(t,x,u) - b(s,y,u)\| \le M\|x-y\| + \omega(|t-s|, \|x\| \vee \|y\|) \\ &\qquad\qquad\qquad\qquad \textit{for all } t,s \in [0,T],\ u \in U\ x,y \in \mathcal{H} \\ &\|b(t,0,u)\| \le M \ \textit{ for all } (t,u) \in [0,T] \times U \end{aligned}$$

Hypothesis 2.2.

$$L\colon [0,T]\times\mathcal{H}\times U\to\mathbb{R} \quad \textit{and} \quad h\colon\mathcal{H}\to\mathbb{R} \quad \textit{are continuous}$$

and there exist $M>0$ and a local modulus of continuity $\omega(\cdot,\cdot)$ such that

$$\begin{aligned} &|L(t,x,u)-L(s,y,u)|,\ |h(x)-h(y)|\le\omega(\|x-y\|+|t-s|,\|x\|\vee\|y\|)\\ &\qquad\qquad\qquad\qquad\qquad\qquad \textit{for all } t,s\in[0,T],\ u\in U\ x,y\in\mathcal{H}\\ &|L(t,0,u)|,|h(0)|\le M \ \textit{ for all } (t,u)\in[0,T]\times U \end{aligned}$$

Remark 2.3. *Notice that if we replace A and b by $\tilde{A}=A-\omega I$ and $b(t,x,u)$ with $\tilde{b}(t,x,u)=b(t,x,u)+\omega x$ the above assumptions would cover a more general case*

$$\|e^{sA}\|\le e^{\omega s} \quad \textit{for all } s\ge 0 \tag{4}$$

for some $\omega\ge 0$. However such $\tilde{b}$ does not satisfy the assumptions of Section 4 and may not satisfy the assumptions needed for comparison for equation (8). Alternatively, by making a change of variables $\tilde{v}(t,x)=v(t,e^{\omega t}x)$ in equation (8) (see [28], page 275) we can always reduce the case (4) to the case when A satisfies (3).

Following the dynamic programming approach we consider a family of problems for every $t\in[0,T], y\in\mathcal{H}$

$$\begin{cases} \dot{x}_{t,x}(s)=Ax_{t,x}(s)+b(s,x_{t,x}(s),u(s))\\ x_{t,x}(t)=x \end{cases} \tag{5}$$

We will write $x(\cdot)$ for $x_{t,x}(\cdot)$ when there is no possibility of confusion. We consider the function

$$J(t,x;u(\cdot))=\int_t^T L(s,x(s),u(s))\mathrm{d}t+h(x(T)), \tag{6}$$

where $u(\cdot)$ is in the set of admissible controls

$$\mathcal{U}[t,T]=\{u\colon[t,T]\to U:\ u \text{ is measurable}\}.$$

The associated value function $V\colon[0,T]\times\mathcal{H}\to\mathbb{R}$ is defined by

$$V(t,x)=\inf_{u(\cdot)\in\mathcal{U}[t,T]} J(t,x;u(\cdot)). \tag{7}$$

The Hamilton-Jacobi-Bellman (HJB) equation related to such optimal control problems is

$$\begin{cases} v_t(t,x)+\langle Dv(t,x),Ax\rangle+H(t,x,Dv(t,x))=0\\ v(T,x)=h(x), \end{cases} \tag{8}$$

where

$$\begin{cases} H\colon[0,T]\times\mathcal{H}\times\mathcal{H}\to\mathbb{R},\\ H(t,x,p)=\inf_{u\in U}\left(\langle p,b(t,x,u)\rangle+L(t,x,u)\right) \end{cases}$$

The solution of the above HJB equation is understood in the viscosity sense of Crandall and Lions [28, 29] which is slightly modified here. We consider two sets of tests functions:

$$test1 = \{\varphi \in C^1((0,T)\times\mathcal{H}) \;:\; \varphi \text{ is weakly sequentially lower semicontinuous and } A^*D\varphi \in C((0,T)\times\mathcal{H})\}$$

and

$$\begin{aligned} test2 = \{g \in C^1((0,T)\times\mathcal{H}) \;:\; & \exists g_0, : [0,+\infty) \to [0,+\infty), \\ & and\ \eta \in C^1((0,T)) \text{ positive } s.t. \\ & g_0 \in C^1([0,+\infty)),\ g_0'(r) \geq 0\ \forall r \geq 0, \\ & g_0'(0) = 0\ and\ g(t,x) = \eta(t) g_0(\|x\|) \\ & \forall (t,x) \in (0,T)\times\mathcal{H}\} \end{aligned}$$

We use test2 functions that are a little different from the ones used in [28]. The extra term $\eta(\cdot)$ in test2 functions is added to deal with unbounded solutions. We recall that $D\varphi$ and Dg stand for the Frechet derivatives of these functions.

Definition 2.4. *A function $v \in C((0,T]\times\mathcal{H})$ is a (viscosity)* subsolution *of the HJB equation (8) if*

$$v(T,x) \leq h(x) \quad for\ all\ x \in \mathcal{H}$$

and whenever $v - \varphi - g$ has a local maximum at $(\bar{t},\bar{x}) \in [0,T)\times\mathcal{H}$ for $\varphi \in test1$ and $g \in test2$, we have

$$\varphi_t(\bar{t},\bar{x}) + g_t(\bar{t},\bar{x}) + \langle A^*D\varphi(\bar{t},\bar{x}),\bar{x}\rangle + H(\bar{t},\bar{x},D\varphi(\bar{t},\bar{x}) + Dg(\bar{t},\bar{x})) \geq 0. \tag{9}$$

Definition 2.5. *A function $v \in C((0,T]\times\mathcal{H})$ is a (viscosity)* supersolution *of the HJB equation (8) if*

$$v(T,x) \geq h(x) \quad for\ all\ x \in \mathcal{H}$$

and whenever $v + \varphi + g$ has a local minimum at $(\bar{t},\bar{x}) \in [0,T)\times\mathcal{H}$ for $\varphi \in test1$ and $g \in test2$, we have

$$-\varphi_t(\bar{t},\bar{x}) - g_t(\bar{t},\bar{x}) - \langle A^*D\varphi(\bar{t},\bar{x}),\bar{x}\rangle + H(\bar{t},\bar{x},-D\varphi(\bar{t},\bar{x}) - Dg(\bar{t},\bar{x})) \leq 0. \tag{10}$$

Definition 2.6. *A function $v \in C((0,T]\times\mathcal{H})$ is a (viscosity)* solution *of the HJB equation (8) if it is at the same time a subsolution and a supersolution.*

We will be also using viscosity sub- and supersolutions in situations where no terminal values are given in (8). We will then call a viscosity subsolution (respectively, supersolution) simply a function that satisfies (9) (respectively, (10)).

Lemma 2.7. *Let Hypotheses 2.1 and 2.2 hold. Let $\phi \in test1$ and $(t,x) \in (0,T) \times \mathcal{H}$. Then the following convergence holds uniformly in $u(\cdot) \in \mathcal{U}[t,T]$:*

$$\lim_{s \downarrow t} \left(\frac{1}{s-t} \left(\varphi(s, x_{t,x}(s)) - \varphi(t,x) \right) - \varphi_t(t,x) - \langle A^* D\varphi(t,x), x \rangle \right.$$
$$\left. - \frac{1}{s-t} \int_t^s \langle D\varphi(t,x), b(t,x,u(r)) \rangle \, \mathrm{d}r \right) = 0 \quad (11)$$

Moreover we have for $s-t$ sufficiently small

$$\varphi(s, x_{t,x}(s)) - \varphi(t,x) = \int_t^s \varphi_t(r, x_{t,x}(r)) + \langle A^* D\varphi(r, x_{t,x}(r)), x_{t,x}(r) \rangle$$
$$+ \langle D\varphi(r, x_{t,x}(r)), b(r, x_{t,x}(r), u(r)) \rangle \, \mathrm{d}r \quad (12)$$

Proof. See [46] Lemma 3.3 page 240 and Proposition 5.5 page 67. □

Lemma 2.8. *Let Hypotheses 2.1 and 2.2 hold. Let $g \in test2$ and $(t,x) \in (0,T) \times \mathcal{H}$. Then for $s - t \to 0^+$*

$$\frac{1}{s-t} \left(g(s, x_{t,x}(s)) - g(t,x) \right) \le g_t(t,x)$$
$$+ \frac{1}{s-t} \int_t^s \langle Dg(t,x), b(t,x,u(r)) \rangle \, \mathrm{d}r + o(1) \quad (13)$$

where $o(1)$ is uniform in $u(\cdot) \in \mathcal{U}[t,T]$

Proof. To prove the statement when $x \neq 0$ we use the fact that, in this case (see [46] page 241, equation (3.11)),

$$\|x_{t,x}(s)\| \le \|x\| + \int_t^s \left\langle \frac{x}{\|x\|}, b(t,x,u(r)) \right\rangle \mathrm{d}r + o(s-t)$$

So we have

$$g(s, x_{t,x}(s)) - g(t,x) = \eta(s) g_0(\|x_{t,x}(s)\|) - \eta(t) g_0(\|x\|)$$
$$\le \eta(s) g_0 \left(\|x\| + \int_t^s \left\langle \frac{x}{\|x\|}, b(t,x,u(r)) \right\rangle \mathrm{d}r + o(s-t) \right) - \eta(t) g_0(\|x\|)$$
$$\le \eta'(t) g_0(\|x\|)(s-t) + \eta(t) g_0'(\|x\|) \left(\int_t^s \left\langle \frac{x}{\|x\|}, b(t,x,u(r)) \right\rangle \mathrm{d}r \right) + o(s-t)$$
$$= g_t(t,x)(s-t) + \int_t^s \langle Dg(t,x), b(t,x,u(r)) \rangle \, \mathrm{d}r + o(s-t) \quad (14)$$

where $o(s-t)$ is uniform in $u(\cdot)$. When $x = 0$, using the fact that $g_0'(0) = 0$, we get

$$g(s, x_{t,x}(s)) - g(t,x) = g_t(t,x)(s-t) + o(s - t + \|x_{t,x}(s)\|)$$

and (13) follows upon noticing that $\|x_{t,x}(s)\| \le C(s-t)$ for some C independent of $u(\cdot) \in \mathcal{U}[t,T]$. □

Theorem 2.9. *Let Hypotheses 2.1 and 2.2 hold. Then the value function V (defined in (7)) is a viscosity solution of the HJB equation (8).*

Proof. The proof is quite standard and can be obtained with small changes (due to the small differences in the definition of test2 functions) from Theorem 2.2, page 229 of [46] and the proof of Theorem 3.2, page 240 of [46] (or from [29]). □

We will need a comparison result in the proof of the verification theorem. There are various versions of such results for equation (8) available in the literature, several sufficient sets of hypotheses can be found in [28, 29]. Since we are not interested in the comparison result itself we choose to assume a form of comparison theorem as a hypothesis.

Hypothesis 2.10. *There exists a set $\mathcal{G} \subseteq C([0,T] \times \mathcal{H})$ such that:*

(i) the value function V is in $\mathcal{G}$;

(ii) if $v_1, v_2 \in \mathcal{G}$, v_1 is a subsolution of the HJB equation (8) and v_2 is a supersolution of the HJB equation (8) then $v_1 \leq v_2$.

Note that from (i) and (ii) we know that V is the only solution of the HJB equation (8) in $\mathcal{G}$.

We will use the following lemma whose proof can be found in [56], page 270.

Lemma 2.11. *Let $g \in C([0,T];\mathbb{R})$. We extend g to a function (still denoted by g) on $(-\infty, +\infty)$ by setting $g(t) = g(T)$ for $t > T$ and $g(t) = g(0)$ for $t < 0$. Suppose there is a function $\rho \in L^1(0,T;\mathbb{R})$ such that*

$$\limsup_{h \to 0^+} \frac{g(t+h) - g(t)}{h} \leq \rho(t) \quad a.e.\ t \in [0,T].$$

Then

$$g(\beta) - g(\alpha) \leq \int_\alpha^\beta \limsup_{h \to 0^+} \frac{g(t+h) - g(t)}{h} \mathrm{d}t \quad \forall\, 0 \leq \alpha \leq \beta \leq T.$$

We will denote by B_R the open ball of radius R centered at 0 in $\mathcal{H}$.

3 The verification theorem

We first introduce a set related to a subset of the superdifferential of a function in $C((0,T) \times \mathcal{H})$. Its definition is suggested by the definition of a sub/super solution. We recall that the superdifferential $D^{1,+}v(t,x)$ of $v \in C((0,T) \times \mathcal{H})$ at (t,x) is given by the pairs $(q,p) \in \mathbb{R} \times \mathcal{H}$ such that $v(s,y) - v(t,x) - \langle p, y-x \rangle - q(s-t) \leq o(\|x-y\| + |t-s|)$, and the subdifferential $D^{1,-}v(t,x)$ at (t,x) is the set of all $(q,p) \in \mathbb{R} \times \mathcal{H}$ such that $v(s,y) - v(t,x) - \langle p, y-x \rangle - q(s-t) \geq o(\|x-y\| + |t-s|)$.

Definition 3.1. *Given* $v \in C((0,T)\times\mathcal{H})$ *and* $(t,x)\in(0,T)\times\mathcal{H}$ *we define* $E^{1,+}v(t,x)$ *as*

$$E^{1,+}v(t,x) = \{(q,p_1,p_2)\in\mathbb{R}\times D(A^*)\times\mathcal{H} : \begin{array}{l}\exists\varphi\in test1,\ g\in test2 \ s.t.\\ v-\varphi-g \text{ attains a local}\\ \text{maximum at } (t,x),\\ \partial_t(\varphi+g)(t,x)=q,\\ D\varphi(t,x)=p_1,\ \ Dg(t,x)=p_2\\ \text{and } v(t,x)=\varphi(t,x)+g(t,x)\}\end{array}$$

Remark 3.2. *If we define*

$$E_1^{1,+}v(t,x)=\{(q,p)\in\mathbb{R}\times\mathcal{H} \ : \ p=p_1+p_2 \text{ with } (q,p_1,p_2)\in E^{1,+}v(t,x)\}$$

then $E_1^{1,+}v(t,x)\subseteq D^{1,+}v(t,x)$ *and in the finite dimensional case we have* $E_1^{1,+}v(t,x)=D^{1,+}v(t,x)$*. Here we have to use* $E^{1,+}v(t,x)$ *instead of* $E_1^{1,+}v(t,x)$ *because of the different roles of* g *and* φ*. It is not clear if the sets* $E^{1,+}v(t,x)$ *and* $E_1^{1,+}v(t,x)$ *are convex. However if we took finite sums of functions* $\eta(t)g_0(\|x\|)$ *as* $test2$ *functions then they would be convex. All the results obtained are unchanged if we use the definition of viscosity solution with this enlarged class of* $test2$ *functions.*

Definition 3.3. *A trajectory-strategy pair* $(x(\cdot),u(\cdot))$ *will be called an* admissible couple *for* (t,x) *if* $u\in\mathcal{U}[t,T]$ *and* $x(\cdot)$ *is the corresponding solution of the state equation (5).*

A trajectory-strategy pair $(x^*(\cdot),u^*(\cdot))$ *will be called an* optimal couple *for* (t,x) *if it is admissible for* (t,x) *and if we have*

$$-\infty < J(t,x;u^*(\cdot)) \le J(t,x;u(\cdot))$$

for every admissible control $u(\cdot)\in\mathcal{U}[t,T]$*.*

We can now state and prove the verification theorem.

Theorem 3.4. *Let Hypotheses 2.1, 2.2 and 2.10 hold. Let* $v\in\mathcal{G}$ *be a subsolution of the HJB equation (8) such that*

$$v(T,x)=h(x) \quad \text{for all } x \text{ in } \mathcal{H}. \tag{15}$$

(a) We have $v(t,x)\le V(t,x)\le J(t,x,u(\cdot)) \quad \forall (t,x)\in(0,T]\times\mathcal{H},\ u(\cdot)\in\mathcal{U}[t,T]$*.*

(b) Let $(t,x)\in(0,T)\times H$ *and let* $(x_{t,x}(\cdot),u(\cdot))$ *be an admissible couple at* (t,x)*. Assume that there exist* $q\in L^1(t,T;\mathbb{R})$*,* $p_1\in L^1(t,T;D(A^*))$ *and* $p_2\in L^1(t,T;\mathcal{H})$ *such that*

$$(q(s),p_1(s),p_2(s))\in E^{1,+}v(s,x_{t,x}(s)) \quad \text{for almost all } s\in(t,T) \tag{16}$$

and that

$$\int_t^T (\langle p_1(s)+p_2(s), b(s,x_{t,x}(s),u(s))\rangle + q(s) + \langle A^*p_1(s), x_{t,x}(s)\rangle)\mathrm{d}t \le \int_t^T -L(s,x_{t,x}(s),u(s))\mathrm{d}s. \tag{17}$$

Then $(x_{t,x}(\cdot), u(\cdot))$ is an optimal couple at (t, x) and $v(t, x) = V(t, x)$. Moreover we have equality in (17).

Remark 3.5. *It is tempting to try to prove, along the lines of Theorem 3.9, p.243 of [56], that a condition like (17) can also be necessary if v is a viscosity solution (or maybe simply a supersolution). However this is not an easy task: the main problem is that $E^{1,+}$ and the analogous object $E^{1,-}$ are fundamentally different so a natural generalization of a result like Theorem 3.9, p.243 of [56] does not seem possible. Moreover our verification theorem has some drawbacks. Condition (17) implicitly implies that $< p_2(r), Ax_{t,x}(r) >= 0$ a.e. if the trajectory is in the domain of A. This follows from the fact that we would then have an additional term $< p_2(r), Ax_{t,x}(r) >$ in the integrand of the middle line of (20) so (17) would also have to be an equality with this additional term. Therefore the applicability of the theorem is somehow limited as in practice (17) may be satisfied only if the function is "nice" (i.e. its superdifferential should really only consist of p_1). Still it applies in some cases where other results fail (see Remarks 3.6 and 3.8). Many issues are not fully resolved yet and we plan to work on them in the future.*

Proof. The first statement ($v \le V$) follows from Hypothesis 2.10, it remains to prove second one. The function

$$\begin{cases} [t, T] \to \mathcal{H} \times \mathbb{R} \\ s \mapsto (b(s, x_{t,x}(s), u(s)), L(s, x_{t,x}(s), u(s)) \end{cases}$$

in view of Hypotheses 2.1 and 2.2 is in $L^1(t, T; \mathcal{H} \times \mathbb{R})$ (in fact it is bounded). So the set of the right-Lebesgue points of this function that in addition satisfy (16) is of full measure. We choose r to be a point in this set. We will denote $y = x_{t,x}(r)$.

Consider now two functions $\varphi^{r,y} \in test1$ and $g^{r,y} \in test2$ such that (we will avoid the index ^{r,y} in the sequel) $v \le \varphi + g$ in a neighborhood of (r, y), $v(r, y) - \varphi(r, y) - g(r, y) = 0, (\partial_t)(\varphi + g)(r, y)) = q(r)$, $D\phi(r, y) = p_1(r)$ and $Dg(r, y) = p_2(r)$. Then for $\tau \in (r, T]$ such that $(\tau - r)$ is small enough we have by Lemmas 2.7 and 2.8

$$\frac{v(\tau, x_{t,x}(\tau)) - v(r, y)}{\tau - r} \le \frac{g(\tau, x_{t,x}(\tau)) - g(r, y)}{\tau - r} + \frac{\varphi(\tau, x_{t,x}(\tau)) - \varphi(r, y)}{\tau - r}$$

$$\begin{aligned} &\le g_t(r, y) + \frac{\int_r^\tau \langle Dg(r, y), b(r, y, u(s)) \rangle \, \mathrm{d}s}{\tau - r} \\ &\quad + \varphi_t(r, y) + \frac{\int_r^\tau \langle D\varphi(r, y), b(r, y, u(s)) \rangle \, \mathrm{d}s}{\tau - r} + \langle A^* D\varphi(r, y), y \rangle + o(1). \quad (18) \end{aligned}$$

In view of the choice of r we know that

$$\frac{\int_r^\tau \langle Dg(r, y), b(r, y, u(s)) \rangle \, \mathrm{d}s}{\tau - r} \xrightarrow{\tau \to r} \langle Dg(r, y), b(r, y, u(r)) \rangle$$

and

$$\frac{\int_r^\tau \langle D\varphi(r,y), b(r,y,u(s))\rangle \, \mathrm{d}s}{\tau - r} \xrightarrow{\tau\to r} \langle D\varphi(r,y), b(r,y,u(r))\rangle \, .$$

Therefore for almost every r in $[t,T]$ we have

$$\begin{aligned}
\limsup_{\tau\downarrow r} & \frac{v(\tau, x_{t,x}(\tau)) - v(r, x_{t,x}(r)))}{\tau - r} \\
& \le \langle Dg(r, x_{t,x}(r)) + D\varphi(r, x_{t,x}(r)), b(r, x_{t,x}(r), u(r))\rangle \\
& + g_t(r, x_{t,x}(r)) + \varphi_t(r, x_{t,x}(r)) + \langle A^* D\varphi(r, x_{t,x}(r)), x_{t,x}(r)\rangle \\
& = \langle p_1(r) + p_2(r), b(r, x_{t,x}(r), u(r))\rangle + q(r) + \langle A^* p_1(r), x_{t,x}(r)\rangle \, . \quad (19)
\end{aligned}$$

We can then use Lemma 2.11 and (17) to obtain

$$\begin{aligned}
v(T, & x_{t,x}(T)) - v(t,x) \\
& \le \int_t^T (\langle p(r), b(r, x_{t,x}(r), u(r))\rangle + q(r) + \langle A^* p_1(r), x_{t,x}(r)\rangle) \mathrm{d}r \\
& \qquad \le \int_t^T -L(r, x_{t,x}(r), u(r)) \mathrm{d}r. \quad (20)
\end{aligned}$$

Thus, using (a), we finally arrive at

$$\begin{aligned}
V(T, x_{t,x}(T)) - V(t,x) &= h(x_{t,x}(T)) - V(t,x) \le h(x_{t,x}(T)) - v(t,x) \\
&= v(T, x_{t,x}(T)) - v(t,x) \le \int_t^T -L(r, x_{t,x}(r), u(r)) \mathrm{d}r \quad (21)
\end{aligned}$$

which implies that $(x_{t,x}(\cdot), u(\cdot))$ is an optimal pair and that $v(t,x) = V(t,x)$. □

Remark 3.6. *In the book [46] (page 263, Theorem 5.5) the authors present a verification theorem (based on a previous result of [22], see also [21] for similar results) in which it is required that the trajectory of the system remains in the domain of A a.e. for the admissible control $u(\cdot)$ in question. This is not required here and in fact this is not satisfied in the example of the next section.*

It is shown in [46] (under assumptions similar to Hypotheses 2.1 and 2.2) that the couple $x(\cdot), u(\cdot))$ is optimal if and only if

$$u(s) \in \left\{ u \in U \, : \, \lim_{\delta\to 0} \frac{V((s+\delta), x(s) + \delta(Ax(s) + b(s, x(s), u))) - V(s, x(s))}{\delta} = -L(s, x(s), u) \right\} \quad (22)$$

for almost every $s \in [t,T]$, where V is the value function.

3.1 An example

We present an example of a control problem for which the value function is a nonsmooth viscosity solution of the corresponding HJB equation, however we can apply our verification theorem. The problem can model a number of phenomena, for example in age-structured population models (see [39, 1, 40]), in population economics [35], optimal technology adoption in a vintage capital context [13, 14].

Consider the state equation

$$\begin{cases} \dot{x}(s) = Ax(s) + Ru(s) \\ x(t) = x \end{cases} \tag{23}$$

where$A$ is a linear, densely defined maximal dissipative operator in $\mathcal{H}$, R is a continuous linear operator $R\colon \mathbb{R} \to \mathcal{H}$, so it is of the form $R\colon u \mapsto u\beta$ for some $\beta \in \mathcal{H}$. Let B be an operator as in Section 4 satisfying (30). We will be using the notation of Section 4.

We will assume that A^* has an eigenvalue λ with an eigenvector α belonging to the range of B.

We consider the functional to be minimized

$$J(x, u(\cdot)) = \int_t^T - |\langle \alpha, x(s) \rangle| + \frac{1}{2} u(s)^2 \mathrm{d}s. \tag{24}$$

We define

$$\bar{\alpha}(t) \stackrel{def}{=} \int_t^T e^{(s-t)A^*} \alpha \mathrm{d}s$$

and we take $M \stackrel{def}{=} \sup_{t\in[0,T]} | \langle \bar{\alpha}(t), \beta \rangle |$. We consider as control set U the compact subset of $\mathbb{R}$ given by $U = [-M-1, M+1]$. So we specify the general problem characterized by (1) and (2) taking $b(t, x, u) = Ru$, $L(t, x, u) = - |\langle \alpha, x(s) \rangle| + 1/2u(t)^2$, $h = 0$, $U = [-M-1, M+1]$.

The HJB equation (8) becomes

$$\begin{cases} v_t + \langle Dv, Ax \rangle - |\langle \alpha, x \rangle| + \inf_{u\in U} \left(\langle u, R^* Dv \rangle_{\mathbb{R}} + \frac{1}{2}u^2 \right) = 0 \\ v(T, x) = 0 \end{cases} \tag{25}$$

Note that the operator $R^* \colon \mathcal{H} \to \mathbb{R}$ can be explicitly expressed using β which was used to define the operator R: $R^*x = \langle \beta, x \rangle$.

Now we observe that for $\langle \alpha, x \rangle < 0$ (respectively > 0) the HJB equation is the same as the one for the optimal control problem with the objective functional $\int_t^T \langle \alpha, x(s) \rangle + \frac{1}{2}u(s)^2 \mathrm{d}s$ (respectively $\int_t^T - \langle \alpha, x(s) \rangle + \frac{1}{2}u(s)^2 \mathrm{d}s$) and it is known in the literature (see [34] Theorem 5.5) that its solution is

$$v_1(t, x) = \langle \bar{\alpha}(t), x \rangle - \int_t^T \frac{1}{2} \left(R^* \bar{\alpha}(s) \right)^2 \mathrm{d}s$$

(respectively

$$v_2(t, x) = - \langle \bar{\alpha}(t), x \rangle - \int_t^T \frac{1}{2} \left(R^* \bar{\alpha}(s) \right)^2 \mathrm{d}s).$$

Note that on the separating hyperplane $\langle \alpha, x\rangle = 0$ the two functions assume the same values. Indeed, since α an eigenvector for A^*,

$$\bar{\alpha}(t) = G(t)\alpha$$

where

$$G(t) = \int_t^T e^{\lambda(s-t)} \mathrm{d}s$$

So, if $\langle \alpha, x\rangle = 0$,

$$\langle \bar{\alpha}(t), x\rangle = 0 \qquad \text{for all } t \in [0, T].$$

Therefore we can glue v_1 and v_2 writing

$$W(t,x) = \begin{cases} v_1(t,x) & \text{if } \langle \alpha, x\rangle \le 0 \\ v_2(t,x) & \text{if } \langle \alpha, x\rangle > 0 \end{cases}$$

It is easy to see that W is continuous and concave in x. We claim that W is a viscosity solution of (25). For $\langle \alpha, x\rangle < 0$ and $\langle \alpha, x\rangle > 0$ it follows from the fact that v_1 and v_2 are explicit regular solutions of the corresponding HJB equations.

For the points x where $\langle \alpha, x\rangle = 0$ it is not difficult to see that

$$\begin{cases} D^{1,+}W(t,x) = \left\{ \left(\frac{1}{2} \left(R^*\bar{\alpha}(t)\right)^2, \gamma G(t)\alpha \right) \ : \ \gamma \in [-1,1] \right\} \subseteq D(A^*) \\ D^{1,-}W(t,x) = \emptyset \end{cases}$$

So we have to verify that W is a subsolution on $\langle \alpha, x\rangle = 0$. If $W - \varphi - g$ attains a maximum at (t,x) with $\langle \alpha, x\rangle = 0$ we have that $p \overset{def}{=} (p_1 + p_2) \overset{def}{=} D(\varphi + g)(t,x) \in \{\gamma G(t)\alpha \ : \ \gamma \in [-1,1]\} \subseteq D(A^*)$. From the definition of test1 function $p_1 = D\varphi(t,x) \in D(A^*)$ so $\eta(t) g_0'(|x|) \frac{x}{|x|} = p_2 = Dg(t,x) \in D(A^*)$. $W(\cdot, x)$ is a C^1 function and then, recalling that $\langle \bar{\alpha}(t), x\rangle_t = \langle G'(t)\alpha, x\rangle = 0$, we have

$$\partial_t(\varphi + g)(t,x) = \partial_t W(t,x) = \frac{1}{2} \left(R^*\bar{\alpha}(t)\right)^2, \tag{26}$$

and for $p = \gamma\bar{\alpha}(t)$ we have

$$\inf_{u\in U} \left(\langle Ru, p\rangle + \frac{1}{2}u^2 \right) = -\frac{1}{2}\gamma^2 \left(R^*\bar{\alpha}(t)\right)^2 \tag{27}$$

Moreover, recalling that $g_0'(|x|) \ge 0$ and $-A^*$ is monotone, we have

$$\langle A^* p_1, x\rangle = \langle A^*(p - p_2), x\rangle = \langle A^*\gamma G(t)\alpha, x\rangle - \frac{g_0'(|x|)}{|x|} \langle A^* x, x\rangle \ge \\ \ge \gamma G(t) \langle A^*\alpha, x\rangle = 0 \tag{28}$$

So, by (26), (27) and (28),

$$\partial_t(\varphi + g)(t,x) + \langle A^* p_1, x\rangle - |\langle \alpha, x\rangle| + \\ + \inf_{u\in U} \left(\langle Ru, D(\varphi + g)(t,x)\rangle + \frac{1}{2}u^2 \right) \ge \frac{1}{2}(1 - \gamma^2) \left(R^*\bar{\alpha}(s)\right)^2 \ge 0 \tag{29}$$

and so the claim in proved.

It is easy to see that both W and the value function V for the problem are continuous on $[0,T]\times\mathcal{H}$ and moreover $\psi = W$ and $\psi = V$ satisfy

$$|\psi(t,x)-\psi(t,y)| \le C\|x-y\|_{-1} \quad \text{for all } t\in[0,T], x,y\in\mathcal{H}$$

for some $C \ge 0$. In particular W and V have at most linear growth as $\|x\| \to \infty$. By Theorem 2.9, the value function V is a a viscosity solution of the HJB equation (25) in $(0,T]\times\mathcal{H}$. Moreover, since $\alpha = By$ for some $y \in \mathcal{H}$, comparison holds for equation (25) which yields $W = V$ on $[0,T]\times\mathcal{H}$. (Comparison theorem can be easily obtained by a modification of techniques of [29] but we cannot refer to any result there since both V and W are unbounded. However the result follows directly from Theorem 3.1 together with Remark 3.3 of [43]. The reader can also consult the proof of Theorem 4.4 of [44]. We point out that our assumptions are different from the assumptions of the uniqueness Theorem 4.6 of [46], page 250).

Therefore we have an explicit formula for the value function V given by $V(t,x) = W(t,x)$. We see that V is differentiable at points (t,x) if $\langle\alpha,x\rangle \neq 0$ and

$$DV(t,x) = \begin{cases} \bar{\alpha}(t) & if\ \langle\alpha,x\rangle < 0 \\ -\bar{\alpha}(t) & if\ \langle\alpha,x\rangle > 0 \end{cases}$$

and is not differentiable whenever $\langle\alpha,x\rangle = 0$. However we can apply Theorem 3.4 and prove the following result.

Proposition 3.7. *The feedback map given by*

$$u^{op}(t,x) = \begin{cases} -\langle\beta,\bar{\alpha}(t)\rangle & if\ \langle\alpha,x\rangle \le 0 \\ \langle\beta,\bar{\alpha}(t)\rangle & if\ \langle\alpha,x\rangle > 0 \end{cases}$$

is optimal. Similarly, also the feedback map

$$\bar{u}^{op}(t,x) = \begin{cases} -\langle\beta,\bar{\alpha}(t)\rangle & if\ \langle\alpha,x\rangle < 0 \\ \langle\beta,\bar{\alpha}(t)\rangle & if\ \langle\alpha,x\rangle \ge 0 \end{cases}$$

is optimal.

Proof. Let $(t,x) \in (0,T]\times\mathcal{H}$ be the initial datum. If $\langle\alpha,x\rangle \le 0$, taking the control $-\langle\beta,\bar{\alpha}(t)\rangle$ the associated state trajectory is

$$x^{op}(s) = e^{(s-t)A}x - \int_t^s e^{(s-r)A}R(\langle\beta,\bar{\alpha}(r)\rangle)\mathrm{d}r$$

and it easy to check that it satisfies $\langle\alpha,x^{op}(s)\rangle \le 0$ for every $s \ge t$. Indeed, using the form of R and the fact that α is eigenvector of A^* we get

$$\begin{aligned}\langle\alpha,x^{op}(s)\rangle &= e^{\lambda(s-t)}\langle\alpha,x\rangle - \langle\alpha,\beta\rangle\int_t^s e^{\lambda(s-r)}\langle\beta,\bar{\alpha}(r)\rangle\,\mathrm{d}r \\ &= e^{\lambda(s-t)}\langle\alpha,x\rangle - \langle\alpha,\beta\rangle^2\int_t^s e^{\lambda(s-r)}G(r)\mathrm{d}r.\end{aligned}$$

Similarly if $\langle \alpha, x \rangle > 0$, taking the control $\langle \beta, \bar{\alpha}(t) \rangle$ the associated state trajectory is

$$x^{op}(s) = e^{(s-t)A}x + \int_t^s e^{(s-r)A} R(\langle \beta, \bar{\alpha}(r) \rangle) \mathrm{d}r$$

and it easy to check that it satisfies $\langle \alpha, x^{op}(s) \rangle > 0$ for every $s \geq t$.

We now apply Theorem 3.4 taking $q(s) = \partial_t V(s, x^{op}(s))$,

$$p_1(s) = \begin{cases} \bar{\alpha}(s) & if \ \langle \alpha, x^{op}(s) \rangle \leq 0 \\ -\bar{\alpha}(s) & if \ \langle \alpha, x^{op}(s) \rangle > 0 \end{cases}$$

and $p_2(s) = 0$. It is easy to see that $(q(s), p_1(s), p_2(s)) \in E^{1,+}V(s, x^{op}(s))$. The argument for $\bar{u}^{op}$ is completely analogous. □

We continue by giving a specific example of the Hilbert space $\mathcal{H}$, the operator A, and the data α and β. This example is related to the vintage capital problem in economics, see e.g. [14, 13]. Let $\mathcal{H} = L^2(0,1)$. Let $\{e^{tA};\ t \geq 0\}$ be the semigroup that, if we identify the points 0 and 1 of the interval $[0,1]$, "rotates" the function:

$$e^{tA} f(s) = f(t + s - [t+s])$$

where $[\cdot]$ is the greatest natural number n such that $n \leq t + s$. The domain of A will be

$$D(A) = \left\{ f \in W^{1,2}(0,1) \ : \ f(0) = f(1) \right\}$$

and for all f in $D(A)$ $A(f)(s) = \frac{\mathrm{d}}{\mathrm{d}s} f(s)$. We choose α to be the constant function equal to 1 at every point of the interval $[0,1]$. (We can take for instance $B = (I - \Delta)^{-\frac{1}{2}}$.) Moreover we choose $\beta(s) = \chi_{[0,\frac{1}{2}]}(s) - \chi_{[0,\frac{1}{2}]}(s)$ (χ_Ω is the characteristic function of a set Ω). Consider an initial datum (t, x) such that $\langle \alpha, x \rangle = 0$. In view of Proposition 3.7 an optimal strategy u^{op} is

$$u^{op}(s) = - \langle \beta, \bar{\alpha}(s) \rangle = 0$$

The related optimal trajectory is

$$x^{op}(s) = e^{(s-t)A} y.$$

Remark 3.8. *We observe that, using such strategy, $\langle \alpha, x^{op}(t) \rangle = 0$ for all $s \geq t$. So the trajectory remains for a whole interval in a set in which the value function is not differentiable. Anyway, applying Theorem 3.4, the optimality is proved. Moreover x can be chosen out of the domain of A and so the assumptions of the verification theorem given in [46] (page 263, Theorem 5.5) are not verified in this case.*

4 Sub- and superoptimality principles and construction of ϵ-optimal controls

Let B be a bounded linear positive self-adjoint operator on $\mathcal{H}$ such that A^*B bounded on $\mathcal{H}$ and let $c_0 \leq 0$ be a constant such that

$$\langle (A^*B + c_0 B)x, x \rangle \leq 0 \qquad for\ all\ x \in \mathcal{H}. \tag{30}$$

Such an operator always exists [49] and we refer to [28] for various examples. Using the operator B we define for $\gamma > 0$ the space $\mathcal{H}_{-\gamma}$ to be the completion of $\mathcal{H}$ under the norm

$$\|x\|_{-\gamma} = \|B^{\frac{\gamma}{2}} x\|.$$

We need to impose another set of assumptions on b and L.

Hypothesis 4.1. *There exist a constant $K > 0$ and a local modulus of continuity $\omega(\cdot, \cdot)$ such that:*

$$\|b(t, x, u) - b(s, y, u)\| \le K\|x - y\|_{-1} + \omega(|t - s|, \|x\| \vee \|y\|)$$

and

$$|L(t, x, u) - L(s, y, u)| \le \omega(\|x - y\|_{-1} + |t - s|, \|x\| \vee \|y\|)$$

Let $m \ge 2$. Modifying slightly the functions introduced in [29] we define for a function $w : (0, T) \times \mathcal{H} \to \mathbb{R}$ and $\epsilon, \beta, \lambda > 0$ its sup- and inf-convolutions by

$$w^{\lambda,\epsilon,\beta}(t, x) = \sup_{(s,y)\in(0,T)\times\mathcal{H}} \left\{ w(s, y) - \frac{\|x - y\|_{-1}^2}{2\epsilon} - \frac{(t - s)^2}{2\beta} - \lambda e^{2mK(T-s)}\|y\|^m \right\},$$

$$w_{\lambda,\epsilon,\beta}(t, x) = \inf_{(s,y)\in(0,T)\times\mathcal{H}} \left\{ w(s, y) + \frac{\|x - y\|_{-1}^2}{2\epsilon} + \frac{(t - s)^2}{2\beta} + \lambda e^{2mK(T-s)}\|y\|^m \right\}.$$

Lemma 4.2. *Let w be such that*

$$w(t, x) \le C(1 + \|x\|^k) \quad (\textit{respectively, } w(t, x) \ge -C(1 + \|x\|^k)) \tag{31}$$

on $(0, T) \times \mathcal{H}$ for some $k \ge 0$. Let $m > k$. Then:

(i) For every $R > 0$ there exists $M_{R,\epsilon,\beta}$ such that if $v = w^{\lambda,\epsilon,\beta}$ (respectively, $v = w_{\lambda,\epsilon,\beta}$) then

$$|v(t, x) - v(s, y)| \le M_{R,\epsilon,\beta}(|t - s| + \|x - y\|_{-2}) \quad \textit{on } (0, T) \times B_R \tag{32}$$

(ii) The function

$$w^{\lambda,\epsilon,\beta}(t, x) + \frac{\|x\|_{-1}^2}{2\epsilon} + \frac{t^2}{2\beta}$$

is convex (respectively,

$$w_{\lambda,\epsilon,\beta}(t, x) - \frac{\|x\|_{-1}^2}{2\epsilon} - \frac{t^2}{2\beta}$$

is concave).

(iii) If $v = w^{\lambda,\epsilon,\beta}$ (respectively, $v = w_{\lambda,\epsilon,\beta}$) and v is differentiable at $(t, x) \in (0, T) \times B_R$ then $|v_t(t, x)| \le M_{R,\epsilon,\beta}$, and $Dv(t, x) = Bq$, where $\|q\| \le M_{R,\epsilon,\beta}$

Proof. **(i)** Consider the case $v = w^{\lambda,\epsilon,\beta}$. Observe first that if $\|x\| \le R$ then

$$w^{\lambda,\epsilon,\beta}(t,x) = $$
$$= \sup_{(s,y)\in(0,T)\times\mathcal{H},\ \|y\|\le N} \left\{ w(s,y) - \frac{\|x-y\|_{-1}^2}{2\epsilon} - \frac{(t-s)^2}{2\beta} - \lambda e^{2mK(T-s)}\|y\|^m \right\}, \tag{33}$$

where N depends only on R and λ.

Now suppose $w^{\lambda,\epsilon,\beta}(t,x) \ge w^{\lambda,\epsilon,\beta}(s,y)$. We choose a small $\sigma > 0$ and $(\tilde{t},\tilde{x})$ such that

$$w^{\lambda,\epsilon,\beta}(t,x) \le \sigma + w(\tilde{t},\tilde{x}) - \frac{\|x-\tilde{x}\|_{-1}^2}{2\epsilon} - \frac{(t-\tilde{t})^2}{2\beta} - \lambda e^{2mK(T-\tilde{t})}\|\tilde{x}\|^m.$$

Then

$$\begin{aligned}|w^{\lambda,\epsilon,\beta}(t,x) - w^{\lambda,\epsilon,\beta}(s,y)| &\le \sigma - \frac{\|x-\tilde{x}\|_{-1}^2}{2\epsilon} - \frac{(t-\tilde{t})^2}{2\beta} + \frac{\|\tilde{x}-y\|_{-1}^2}{2\epsilon} + \frac{(\tilde{t}-s)^2}{2\beta} \\ &\le \sigma - \frac{\langle B(x-y), x+y\rangle}{2\epsilon} + \frac{\langle B(x-y),\tilde{x}\rangle}{\epsilon} + \frac{(2\tilde{t}-t-s)(t-s)}{2\beta} \\ &\le \frac{(2R+N)}{2\epsilon}\|B(x-y)\| + \frac{2T}{2\beta}|t-s| + \sigma \end{aligned} \tag{34}$$

and we conclude because of the arbitrariness of σ. The case of $w_{\lambda,\epsilon,\beta}$ is similar.

(ii) It is a standard fact, see for example the Appendix of [27].

(iii) The fact that $|v_t(t,x)| \le M_{R,\epsilon,\beta}$ is obvious. Moreover if $\alpha > 0$ is small and $\|y\| = 1$ then

$$\alpha M_{R,\epsilon,\beta}\|y\|_{-2} \ge |v(t,x+\alpha y) - v(x)| = \alpha|\langle Dv(t,x), y\rangle| + o(\alpha)$$

which upon dividing by α and letting $\alpha \to 0$ gives

$$|\langle Dv(t,x), y\rangle| \le M_{R,\epsilon,\beta}\|y\|_{-2}$$

which then holds for every $y \in \mathcal{H}$. This implies that $\langle Dv(t,x), y\rangle$ is a bounded linear functional in $\mathcal{H}_{-2}$ and so $Dv(t,x) = Bq$ for some $q \in \mathcal{H}$. Since $|\langle q, By\rangle| \le M_{R,\epsilon,\beta}\|By\|$ we obtain $\|q\| \le M_{R,\epsilon,\beta}$. □

Lemma 4.3. *Let Hypotheses 2.1, 2.2 and 4.1 be satisfied. Let w be a locally bounded viscosity subsolution (respectively, supersolution) of (8) satisfying (31). Let $m > k$. Then for every $R, \delta > 0$ there exists a non-negative function $\gamma_{R,\delta}(\lambda,\epsilon,\beta)$, where*

$$\lim_{\lambda\to 0}\limsup_{\epsilon\to 0}\limsup_{\beta\to 0}\gamma_{R,\delta}(\lambda,\epsilon,\beta) = 0, \tag{35}$$

such that $w^{\lambda,\epsilon,\beta}$ (respectively, $w_{\lambda,\epsilon,\beta}$) is a viscosity subsolution (respectively, supersolution) of

$$v_t(t,x) + \langle Dv(t,x), Ax\rangle + H(t,x,Dv(t,x)) = -\gamma_{R,\delta}(\lambda,\epsilon,\beta) \quad \textit{in } (\delta, T-\delta)\times B_R \tag{36}$$

(respectively,

$$v_t(t,x) + \langle Dv(t,x), Ax\rangle + H(t,x,Dv(t,x)) = \gamma_{R,\delta}(\lambda,\epsilon,\beta) \quad \text{in } (\delta, T-\delta)\times B_R) \tag{37}$$

for β sufficiently small (depending on δ).

Proof. The proof is similar to the proof of Proposition 5.3 of [29]. We notice that $w^{\lambda,\epsilon,\beta}$ is bounded from above.

Let $(t_0, x_0) \in (\delta, T-\delta)\times \mathcal{H}$ be a local maximum of $w^{\lambda,\epsilon,\beta} - \phi - g$. We can assume that the maximum is global and strict (see Proposition 2.4 of [29]) and that $w^{\lambda,\epsilon,\beta} - \phi - g \to -\infty$ as $\|x\| \to \infty$ uniformly in t. In view of these facts and (33) we can choose $S > 2\|x_0\|$, depending on λ such that, for all $\|x\| + \|y\| > S - 1$ and $s, t \in (0,T)$,

$$\begin{aligned} w(s,y) - \frac{1}{2\epsilon}\|(x-y)\|_{-1}^2 - \frac{(t-s)^2}{2\beta} - \lambda e^{2mK(T-s)}\|y\|^m - \phi(t,x) - g(t,x) \\ \le w(t_0,x_0) - \lambda e^{2mK(T-t_0)}\|x_0\|^m - \phi(t_0,x_0) - g(t_0,x_0) - 1. \end{aligned} \tag{38}$$

We can then use a perturbed optimization technique of [29] (see page 424 there) which is a version of the Ekeland-Lebourg Lemma [33] to obtain for every $\alpha > 0$ elements $p, q \in \mathcal{H}$ and $a, b \in \mathbb{R}$ with $\|p\|, \|q\| \le \alpha$ and $|a|, |b| \le \alpha$ such that the function

$$\begin{aligned} \psi(t,x,s,y) \stackrel{def}{=} w(s,y) - \frac{1}{2\epsilon}\|(x-y)\|_{-1}^2 - \frac{(t-s)^2}{2\beta} - \lambda e^{2mK(T-s)}\|y\|^m \\ - g(t,x) - \phi(t,x) - \langle Bp, y\rangle - \langle Bq, x\rangle - at - bs \end{aligned} \tag{39}$$

attains a local maximum $(\bar t, \bar x, \bar s, \bar y)$ over $[\delta/2, T-\delta/2]\times B_S \times [\delta/2, T-\delta/2]\times B_S$. It follows from (38) that if α is sufficiently small then $\|\bar x\|, \|\bar y\| \le S - 1$.

By possibly making S bigger we can assume that $(0,T)\times B_S$ contains a maximizing sequence for

$$\sup_{(s,y)\in(0,T),\ \|y\|\le N} \left\{ w(s,y) - \frac{\|x_0 - y\|_{-1}^2}{2\epsilon} - \frac{(t_0 - s)^2}{2\beta} - \lambda e^{2mK(T-s)}\|y\|^m \right\}.$$

Then

$$\psi(\bar t, \bar x, \bar s, \bar y) \ge w^{\lambda,\epsilon,\beta}(t_0, x_0) - \phi(t_0,x_0) - g(t_0,x_0) - C\alpha$$

where the constant C does not depend on $\alpha > 0$, and

$$\psi(\bar t, \bar x, \bar s, \bar y) \le w^{\lambda,\epsilon,\beta}(\bar t, \bar x) - \phi(\bar t, \bar x) - g(\bar t, \bar x) + C\alpha.$$

Therefore, since (t_0, x_0) is a strict maximum, we have that $(\bar t, \bar x) \xrightarrow{\alpha\downarrow 0} (t_0, x_0)$ and so for small α $\bar t \in (\delta, T-\delta)$. It then easily follows that if β is big enough (depending on λ and δ) then $\bar s \in (\delta/2, T-\delta/2)$.

Moreover, standard arguments (see for instance [41]) give us

$$\lim_{\beta\to 0}\limsup_{\alpha\to 0}\frac{|\bar s-\bar t|^2}{2\beta}=0, \tag{40}$$

$$\lim_{\epsilon\to 0}\limsup_{\beta\to 0}\limsup_{\alpha\to 0}\frac{|\bar x-\bar y|_{-1}^2}{2\epsilon}=0. \tag{41}$$

We can now use the fact that w is a subsolution to obtain

$$-\frac{(\bar t-\bar s)}{\beta}-2\lambda mKe^{2mK(T-\bar s)}\|\bar y\|^m+b-\frac{\langle A^*B(\bar x-\bar y),\bar y\rangle}{\epsilon}+\langle A^*Bp,\bar y\rangle$$
$$+H\left(\bar s,\bar y,\frac{1}{\epsilon}B(\bar y-\bar x)+\lambda me^{2mK(T-\bar s)}\|y\|^{m-1}\frac{y}{\|y\|}+Bp\right)\ge 0. \tag{42}$$

We notice that

$$-\frac{(\bar t-\bar s)}{\beta}=\phi_t(\bar t,\bar x)+g_t(\bar t,\bar x)+a$$

and

$$\frac{1}{\epsilon}B(\bar y-\bar x)=D\phi(\bar t,\bar x)+Dg(\bar t,\bar x)+Bq$$

which in particular implies that $Dg(\bar t,\bar x)\in D(A^*)$, i.e. $\bar x\in D(A^*)$, and so it follows that $\langle A^*\bar x,Dg(\bar t,\bar x)\rangle\le 0$. Therefore using this, the assumptions on b and L, and (40) and (41) we have

$$\begin{aligned}
&\phi_t(\bar t,\bar x)+g_t(\bar t,\bar x)+\langle\bar x,A^*D\phi(\bar t,\bar x)\rangle+H\left(\bar t,\bar x,D\phi(\bar t,\bar x)+Dg(\bar t,\bar x)\right)\\
&\qquad\ge 2\lambda mKe^{2mK(T-\bar s)}\|\bar y\|^m-\langle A^*Bp,\bar y\rangle-a-b\\
&\qquad-\left\langle(\bar y-\bar x),A^*\frac{1}{\epsilon}B(\bar y-\bar x)\right\rangle-\langle\bar x,A^*Dg(\bar t,\bar x)+A^*Bq)\rangle\\
&+H\left(\bar t,\bar x,\frac{1}{\epsilon}B(\bar y-\bar x)-Bq\right)-H\left(\bar s,\bar y,\frac{1}{\epsilon}B(\bar y-\bar x)+\lambda me^{2mK(T-\bar s)}\|y\|^{m-1}\frac{y}{\|y\|}\right)\\
&\qquad\ge 2\lambda mKe^{2mK(T-\bar s)}\|\bar y\|^m-C_{\lambda,\epsilon}\alpha+\frac{c_0}{\epsilon}\|\bar x-\bar y\|_{-1}^2\\
&-K\|\bar x-\bar y\|_{-1}\frac{\|B(\bar x-\bar y)\|}{\epsilon}-\gamma_{\lambda,\epsilon}(|\bar t-\bar s|)-\lambda m(M+K\|\bar y\|)e^{2mK(T-\bar s)}\|\bar y\|^{m-1}\\
&\qquad\ge -C_{\lambda,\epsilon}\alpha-\gamma(\lambda,\epsilon,\beta,\alpha)
\end{aligned} \tag{43}$$

for some $\gamma(\lambda,\epsilon,\beta,\alpha)$ such that

$$\lim_{\lambda\to 0}\limsup_{\epsilon\to 0}\limsup_{\beta\to 0}\limsup_{\alpha\to 0}\gamma(\lambda,\epsilon,\beta,\alpha)=0.$$

We obtain the claim by letting $\alpha\to 0$. The proof for $w_{\lambda,\beta,\epsilon}$ is similar. □

Remark 4.4. *Similar argument would also work for problems with discounting if w was uniformly continuous in $|\cdot|\times\|\cdot\|_{-1}$ norm uniformly on bounded sets* *of* $(0,T)\times\mathcal{H}$. *Moreover in some cases the function* $\gamma_{R,\delta}$ *could be explicitly computed. For instance if* w *is bounded and*

$$|w(t,x)-w(s,y)|\leq\sigma(\|x-y\|_{-1})+\sigma_1(|t-s|;\|x\|\vee\|y\|) \tag{44}$$

for $t,s\in(0,T),\|x\|,\|y\|\in\mathcal{H}$, *we can replace* $\lambda e^{2mK(T-\bar{s})}\|\bar{y}\|^m$ *by* $\lambda\mu(y)$ *for some radial nondecreasing function* μ *such that* $D\mu$ *is bounded and* $\mu(y)\to+\infty$ *as* $\|y\|\to\infty$ *(see [29], page 446). If we then replace the order in which we pass to the limits we can get an explicit (but complicated) form for* $\gamma_{R,\delta}$ *satisfying*

$$\lim_{\epsilon\to 0}\limsup_{\lambda\to 0}\limsup_{\beta\to 0}\gamma_{R,\delta}(\epsilon,\lambda,\beta)=0.$$

The proofs of Theorem 3.7 and Proposition 5.3 in [29] can give hints how to do this.

Lemma 4.5. *Let the assumptions of Lemma 4.3 be satisfied. Then:*

(a) If $(a,p)\in D^{1,-}w^{\lambda,\epsilon,\beta}(t,x)$ *for* $(t,x)\in(\delta,T-\delta)\times B_R$ *then*

$$a+\langle A^*p,x\rangle+H(t,x,p)\geq-\gamma_{R,\delta}(\lambda,\epsilon,\beta) \tag{45}$$

for β *sufficiently small.*

(b) If in addition $H(s,y,q)$ *is weakly lower-semicontinuous with respect to the* q*-variable and* $(a,p)\in D^{1,+}w_{\lambda,\epsilon,\beta}(t,x)$ *for* $(t,x)\in(\delta,T-\delta)\times B_R$ *is such that* $Dw_{\lambda,\epsilon,\beta}(t_n,x_n)\rightharpoonup p$ *for some* $(t_n,x_n)\to(t,x)$, $(t_n,x_n)\in(\delta,T-\delta)\times B_R$, *then*

$$a+\langle A^*p,x\rangle+H(t,x,p)\leq\gamma_{R,\delta}(\lambda,\epsilon,\beta)$$

for β *sufficiently small.*

Remark 4.6. *The Hamiltonian* H *is weakly lower-semicontinuous with respect to the* q*-variable for instance if* U *is compact. To see this we observe that thanks to the compactness of* U *the infimum in the definition of the Hamiltonian is a minimum. Let now* $q_n\rightharpoonup q$ *and let*

$$H(s,y,q_n)=\langle q_n,b(s,y,u_n)\rangle+L(s,y,u_n)$$

for some $u_n\in U$. *Passing to a subsequence if necessary we can assume that* $u_n\longrightarrow\bar{u}$, *and then passing to the limit in the above expression we obtain*

$$\liminf_{n\to\infty}H(s,y,q_n)=\langle q,b(s,y,\bar{u})\rangle+L(s,y,\bar{u})\geq H(s,y,q).$$

We also remark that since H *is concave in* q *it is weakly upper-semicontinuous in* q. *Therefore in (b) the Hamiltonian* H *is assumed to be weakly continuous in* q.

Proof. *(of Lemma 4.5)* Recall first that for a convex/concave function v its sub/super-differential at a point (s, z) is equal to

$$\overline{\text{conv}}\{((a,p) : v_t(s_n, z_n) \to a, Dv(s_n, z_n) \rightharpoonup p, s_n \to s, z_n \to z\}$$

(see [48], page 319).

(a) Step 1: Denote $v = w^{\lambda,\epsilon,\beta}$. At points of differentiability, it follows from Lemma 4.2(iii) and the "semiconvexity" (see Lemma 4.2(ii)) of $w^{\lambda,\epsilon,\beta}$ that there exists a test1 function φ such that $v - \varphi$ has a local maximum and the result then follows from Lemma 4.3.

Step 2: Consider first the case $Dv(t_n, x_n) \rightharpoonup p$ with $(t_n, x_n) \to (t, x)$. From Lemma 4.2 (iii) $Dv(t_n, x_n) = Bq_n$ with $\|q_n\| \le M_{R,\epsilon,\beta}$, so, it is always possible to extract a subsequence $q_{n_k} \rightharpoonup q$ for some $q \in \mathcal{H}$. Then $Dv(t_{n_k}, x_{n_k}) = Bq_{n_k} \rightharpoonup Bq$ and $Bq = p$. Therefore

$$\langle A^* Bq_{n_k}, x_{n_k}\rangle = \langle q_{n_k}, (A^*B)^* x_{n_k}\rangle \longrightarrow \langle q, (A^*B)^* x\rangle = \langle A^*Bq, x\rangle = \langle A^*p, x\rangle$$

Moreover, since H is concave in p it is weakly upper-semicontinuous so we have

$$H(t,x,p) \ge \limsup_{k\to+\infty} H(t_{n_k}, x_{n_k}, Dv(t_{n_k}, x_{n_k}))$$

and we conclude from Step 1.

Step 3: If p is a generic point of $\overline{\text{conv}}\{p : Dv(t_n, x_n) \rightharpoonup p, (t_n, x_n) \to (t, x)\}$, i.e. $p = \lim_{n\to\infty} \sum_{i=1}^n \lambda_i^n Bq_i^n$, where $\sum_{i=1}^n \lambda_i^n = 1$, $\|q_i^n\| \le M_{R,\epsilon,\beta}$, and the Bq_i^n are weak limits of gradients. By passing to a subsequence if necessary we can assume that $\sum_{i=1}^n \lambda_i^n q_i^n \rightharpoonup q$ and $p = Bq$. But then

$$\left\langle A^* \left(\sum_{i=1}^n \lambda_i^n B q_i^n\right), x_n \right\rangle = \left\langle A^* B \left(\sum_{i=1}^n \lambda_i^n q_i^n\right), x_n \right\rangle \to \langle A^*Bq, x\rangle = \langle A^*p, x\rangle$$

as $n \to \infty$. The result now follows from Step 2 and the concavity of

$$p \mapsto \langle A^*p, x\rangle + H(t,x,p).$$

(b) As in (a) at the points of differentiability the claim follows from Lemmas 4.2 and 4.3. Denote $v = w_{\lambda,\epsilon,\beta}$. If $Dv(t_n, x_n) \rightharpoonup p$ for some $(t_n, x_n) \to (t, x)$, $(t_n, x_n) \in (\delta, T - \delta) \times B_R$ we have that

$$v_t(t_n, x_n) + \langle A^* Dv(t_n, x_n), x_n\rangle + H(t_n, x_n, Dv(t_n, x_n)) \le \gamma_{R,\delta}(\lambda, \epsilon, \beta). \quad (46)$$

Observing as in Step 2 of (a) that

$$\langle A^* Dv(t_n, x_n), x_n\rangle \to \langle A^*p, x\rangle$$

we can pass to the limit in (46), using the weak lower semicontinuity of H with respect to the third variable, to get

$$a + \langle A^*p, x\rangle + H(t,x,p) \le \gamma_{R,\delta}(\lambda, \epsilon, \beta).$$

□

Theorem 4.7. *Let the assumptions of Lemma 4.3 be satisfied and let w be a function such that for every $R > 0$ there exists a modulus σ_R such that*

$$|w(t,x)-w(s,y)| \le \sigma_R(|t-s|+\|x-y\|_{-1}) \quad \text{for } t,s \in (0,T), \|x\|, \|y\| \le R. \tag{47}$$

Then:

(a) If w is a viscosity subsolution of (8) satisfying (31) for subsolutions then for every $0 < t < t+h < T$, $x \in \mathcal{H}$

$$w(t,x) \le \inf_{u(\cdot)\in\mathcal{U}[t,T]} \left\{ \int_t^{t+h} L(s,x(s),u(s))\mathrm{d}s + w(t+h,x(t+h)) \right\}. \tag{48}$$

(b) Assume in addition that $H(s,y,q)$ is weakly lower-semicontinuous in q and that for every (t,x) there exists a modulus $\omega_{t,x}$ such that

$$\|x_{t,x}(s_2) - x_{t,x}(s_1)\| \le \omega_{t,x}(s_2 - s_1) \tag{49}$$

for all $t \le s_1 \le s_2 \le T$ and all $u(\cdot) \in \mathcal{U}[t,T]$, where $x_{t,x}(\cdot)$ is the solution of (5). If w is a viscosity supersolution of (8) satisfying (31) for supersolutions then for every $0 < t < t+h < T, x \in H$, and $\nu > 0$ there exists a piecewise constant control $u_\nu \in \mathcal{U}[t,T]$ such that

$$w(t,x) \ge \int_t^{t+h} L(s,x(s),u_\nu(s))\mathrm{d}s + w(t+h,x(t+h)) - \nu. \tag{50}$$

In particular we obtain the superoptimality principle

$$w(t,x) \ge \inf_{u(\cdot)\in\mathcal{U}[t,T]} \left\{ \int_t^{t+h} L(s,x(s),u(s))\mathrm{d}s + w(t+h,x(t+h)) \right\} \tag{51}$$

and if w is the value function V we have existence (together with the explicit construction) of piecewise constant ν-optimal controls.

Proof. We will only prove (b) as the proof of (a) follows the same strategy after we fix any control $u(\cdot)$ and is in fact much easier. We follow the ideas of [52] (that treats the finite dimensional case).

Step 1. Let $n \ge 1$. We approximate w by $w_{\lambda,\epsilon,\beta}$ with $m > k$. We notice that for any $u(\cdot)$ if $x_{t,x}(\cdot)$ is the solution of (5) then

$$\sup_{t\le s\le T} \|x_{t,x}(s)\| \le R = R(T,\|x\|).$$

Step 2. Take any $(a,p) \in D^{1,+}w_{\lambda,\epsilon,\beta}(t,x)$ as in Lemma 4.5(b) (i.e. p is the weak limit of derivatives nearby). Such elements always exist because $w_{\lambda,\epsilon,\beta}$ is "semiconcave". Then we choose $u_1 \in U$ such that

$$a + \langle A^*p, x\rangle + \langle p, b(t,x,u_1)\rangle + L(t,x,u_1) \le \gamma_{R,\delta}(\lambda,\epsilon,\beta) + \frac{1}{n^2}. \tag{52}$$

By the "semiconcavity" of $w_{\lambda,\epsilon,\beta}$

$$w_{\lambda,\epsilon,\beta}(s,y) \le w_{\lambda,\epsilon,\beta}(t,x) + a(s-t) + \langle p, y-x\rangle + \frac{\|x-y\|_{-1}^2}{2\epsilon} + \frac{(t-s)^2}{2\beta}. \quad (53)$$

But the right hand side of the above inequality is a test1 function so if $s \ge t$ and $x(s) = x_{t,x}(s)$ with constant control $u(s) = u_1$, we can use (12) and write

$$\begin{aligned} &\Bigg| \frac{a(s-t) + \langle p, x(s)-x\rangle + \frac{\|x(s)-x\|_{-1}^2}{2\epsilon} + \frac{(s-t)^2}{2\beta}}{s-t} \\ &\qquad - (a + \langle p, b(t,x,u_1)\rangle + \langle A^*p, x\rangle) \Bigg| \\ &\le \frac{|t-s|}{2\beta} + \left| \frac{\int_t^s \langle A^*p, x(r)-x\rangle \,\mathrm{d}r}{s-t} \right| \\ &+ \left| \frac{\int_t^s \langle p, b(r,x(r),u_1) - b(t,x,u_1)\rangle \,\mathrm{d}r}{s-t} \right| + \left| \frac{\int_t^s \langle A^*B(x(r)-x), x(r)\rangle \,\mathrm{d}r}{\epsilon(s-t)} \right| \\ &+ \left| \frac{\int_t^s \langle B(x(r)-x), b(r,x(r),u_1)\rangle \,\mathrm{d}r}{\epsilon(s-t)} \right| \\ &\le \omega'_{t,x}(|s-t| + \sup_{t\le r\le s} \|x(r)-x\|) \le \tilde{\omega}_{t,x}(s-t) \quad (54) \end{aligned}$$

for some moduli $\omega'_{t,x}$ and $\tilde{\omega}_{t,x}$ that depend on $(t,x), \epsilon, \beta$ but not on u_1. We can now use (52), (53) and (54) to estimate

$$\begin{aligned} &\frac{w_{\lambda,\epsilon,\beta}(t+\frac{h}{n}, x(t+\frac{h}{n})) - w_{\lambda,\epsilon,\beta}(t,x)}{h/n} \\ &\qquad \le \tilde{\omega}_{t,x}\left(\frac{h}{n}\right) + \gamma_{R,\delta}(\lambda,\epsilon,\beta) + \frac{1}{n^2} - L(t,x,u_1) \quad (55) \end{aligned}$$

Step 3. Denote $t_i = t + \frac{(t-1)h}{n}$ for $i = 1, ..., n$. We now repeat the above procedure starting at $x(t_2)$ to abtain u_2 satisfying (55) with $(t_2, x(t_2))$ replaced by $(t_3, x(t_3))$, $(t,x) = (t_1, x(t_1))$ replaced by $(t_2, x(t_2))$, and u_1 replaced by u_2. After n iterations of this process we obtain a piecewise constant control $u^{(n)}$, where $u^{(n)}(s) = u_i$ if $s \in [t_i, t_{i+1})$. Then if $x(r)$ solves (5) with the control $u^{(n)}$ we have

$$\begin{aligned} &\frac{w_{\lambda,\epsilon,\beta}(t+h, x(t+h)) - w_{\lambda,\epsilon,\beta}(t,x)}{h/n} \\ &\qquad \le \tilde{\omega}_{t,x}\left(\frac{h}{n}\right) n + \gamma_{R,\delta}(\lambda,\epsilon,\beta) n + \frac{n}{n^2} - \sum_{i=1}^{n} L(t_{i-1}, x(t_{i-1}), u_i). \end{aligned}$$

We remind that (49) is needed here to guarantee that $\sup_{t_{i-1}\le r\le t_i}\|x(r)-x(t_{i-1})\|$ is independent of u_i and $x(t_{i-1})$ and depends only on x and t. We then easily obtain

$$w_{\lambda,\epsilon,\beta}(t+h,x(t+h))-w_{\lambda,\epsilon,\beta}(t,x)$$
$$\le \tilde{\omega}_{t,x}\left(\frac{h}{n}\right)h+\gamma_{R,\delta}(\lambda,\epsilon,\beta)h+\frac{h}{n^2}-\int_t^{t+h}L(r,x(r),u^{(n)})\mathrm{d}r+\tilde{\omega}'_{t,x}\left(\frac{h}{n}\right)h \tag{56}$$

for some modulus $\tilde{\omega}'_{t,x}$, where we have used Hypothesis 4.1 and (49) to estimate how the sum converges to the integral. We now finally notice that it follows from (47) that

$$|w_{\lambda,\epsilon,\beta}(s,y)-w(s,y)|\le\tilde{\sigma}_R(\lambda+\epsilon+\beta;R)\quad\text{for }\ s\in(\delta,T-\delta),\|y\|\le R,$$

where the modulus $\tilde{\sigma}_R$ can be explicitly calculated from σ_R. Therefore, choosing β,λ,ϵ small and then n big enough, and using (35), we arrive at (50). □

We show below one example when condition (49) is satisfied.

Example 4.8. *Condition (49) holds for example if $A=A^*$, it generates a differentiable semigroup, and $\|Ae^{tA}\|\le C/t^{\delta}$ for some $\delta<2$. Indeed under these assumptions, if $u(\cdot)\in\mathcal{U}[t,T]$ and writing $x(s)=x_{t,x}(s)$, we have*

$$\|(A+I)^{\frac12}x(s)\|\le\|(A+I)^{\frac12}e^{(s-t)A}x\|+\int_t^s\|(A+I)^{\frac12}e^{(s-\tau)A}b(\tau,x(\tau),u(\tau))\|d\tau$$

However for every $y\in H$ and $0\le\tau\le T$

$$\|(A+I)^{\frac12}e^{\tau A}y\|^2\le\|(A+I)e^{\tau A}y\|\,\|y\|\le\frac{C_1}{\tau^{\delta}}\|y\|^2.$$

This yields

$$\|(A+I)^{\frac12}e^{\tau A}\|\le\frac{\sqrt{C_1}}{\tau^{\frac{\delta}{2}}}$$

and therefore

$$\|(A+I)^{\frac12}x(s)\|\le C_2\left(\frac{1}{(s-t)^{\frac{\delta}{2}}}+(s-t)^{1-\frac{\delta}{2}}\right)\le\frac{C_3}{(s-t)^{\frac{\delta}{2}}}.$$

We will first show that for every $\epsilon>0$ there exists a modulus σ_ϵ (also depending on x but independent of $u(\cdot)$) such that $\|e^{(s_2-s_1)A}x(s_1)-x(s_1)\|\le\sigma_\epsilon(s_2-s_1)$ for all $t+\epsilon\le s_1<s_2\le T$. This is now rather obvious since

$$e^{(s_2-s_1)A}x(s_1)-x(s_1)=\int_0^{s_2-s_1}Ae^{sA}x(s_1)ds$$

$$=\int_0^{s_2-s_1}(A+I)^{\frac12}e^{sA}(A+I)^{\frac12}x(s_1)ds-\int_0^{s_2-s_1}e^{sA}x(s_1)ds$$

and thus

$$\|e^{(s_2-s_1)A}x(s_1)-x(s_1)\| \le \|(A+I)^{\frac{1}{2}}x(s_1)\| \int_0^{s_2-s_1} \frac{\sqrt{C_1}}{s^{\frac{\delta}{2}}} ds + (s_2-s_1)\|x(s_1)\|$$
$$\le \frac{C_4}{\epsilon^{\frac{\delta}{2}}}(s_2-s_1)^{1-\frac{\delta}{2}} + C_5(s_2-s_1).$$

We also notice that there exists a modulus σ, depending on x and independent of $u(\cdot)$, such that

$$\|x(s)-x\| \le \sigma(s-t).$$

Let now $t \le s_1 < s_2 \le T$. Denote $\bar{s} = \max(s_1, t+\epsilon)$. If $s_2 \le t+\epsilon$ then

$$\|x(s_2)-x(s_1)\| \le 2\sigma(\epsilon).$$

Otherwise

$$\begin{aligned}\|x(s_2)-x(s_1)\| &\le 2\sigma(\epsilon) + \|x(s_2)-x(\bar{s})\| \\ &\le 2\sigma(\epsilon) + \|e^{(s_2-\bar{s})A}x(s_1)-x(\bar{s})\| + \int_{\bar{s}}^{s_2} \|e^{(s_2-\tau)A}b(\tau, x(\tau), u(\tau))\| d\tau \\ &\le 2\sigma(\epsilon) + \sigma_\epsilon(s_2-s_1) + C_4(s_2-s_1) \qquad (57)\end{aligned}$$

for some constant C_4 independent of $u(\cdot)$. Therefore (49) is satisfied with the modulus

$$\omega_{t,x}(\tau) = \inf_{0<\epsilon<T-t} \{2\sigma(\epsilon) + \sigma_\epsilon(\tau) + C(\tau)\}.$$

References

[1] S. Aniţa, *Analysis and control of age-dependent population dynamics*, Kluwer Academic Publishers, Dordrecht, (2001).

[2] V. Barbu, *Optimal feedback controls for a class of nonlinear distributed parameter systems* SIAM J. Control Optim. 21 (1983), no. 6, 871–894.

[3] V. Barbu, *Hamilton-Jacobi equations and nonlinear control problems*, J. Math. Anal. Appl 120 (1986), no. 2, 494–509.

[4] V. Barbu, *Analysis and control of nonlinear infinite-dimensional systems*, Mathematics in Science and Engineering, 190. Academic Press, Inc., Boston, MA, 1993.

[5] V. Barbu, *Approximation Hamilton-Jacobi equations and suboptimal feedback controllers*, Nonlinear analysis and applications (Warsaw, 1994), 27–47, GAKUTO Internat. Ser. Math. Sci. Appl., 7, Gakkōtosho, Tokyo, 1996.

[6] V. Barbu, E. N. Barron and R. Jensen, *The necessary conditions for optimal control in Hilbert spaces*, J. Math. Anal. Appl. 133 (1988), no. 1, 151–162.

[7] V. Barbu and G. Da Prato, *Hamilton-Jacobi equations in Hilbert spaces*, Research Notes in Mathematics 86, Pitman, Boston, MA, 1983.

[8] V. Barbu and G. Da Prato, *Hamilton-Jacobi equations and synthesis of nonlinear control processes in Hilbert spaces*, J. Differential Equations 48 (1983), no. 3, 350–372.

[9] V. Barbu and G. Da Prato, *Hamilton-Jacobi equations in Hilbert spaces: variational and semigroup approach*, Ann. Mat. Pura Appl. (4) 142 (1985), 303–349 (1986).

[10] V. Barbu and G. Da Prato, *A note on a Hamilton-Jacobi equation in Hilbert space* Nonlinear Anal. 9 (1985), no. 12, 1337–1345.

[11] V. Barbu, G. Da Prato and C. Popa, *Existence and uniqueness of the dynamic programming equation in Hilbert space* Nonlinear Anal. 7 (1983), no. 3, 283–299.

[12] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Systems & Control: Foundations & Applications, Birkhauser Boston, Inc., Boston, MA, 1997.

[13] E. Barucci and F. Gozzi, *Optimal investment in a vintage capital model*, Res. Econ. 52 (1998), 159-188.

[14] E. Barucci and F. Gozzi, *Technology adoption and accumulation in a vintage capital model*, J. Econ. 74 (2001) no. 1, 1-30.

[15] P. Cannarsa, *Regularity properties of solutions to Hamilton-Jacobi equations in infinite dimensions and nonlinear optimal control*, Differential Integral Equations 2 (1989), no. 4, 479–493.

[16] P. Cannarsa and O. Carjă, *On the Bellman equation for the minimum time problem in infinite dimensions*, SIAM J. Control Optim. 43 (2004), no. 2, 532–548

[17] P. Cannarsa and Da Prato, *G. Nonlinear optimal control with infinite horizon for distributed parameter systems and stationary Hamilton-Jacobi equations*, SIAM J. Control Optim. 27 (1989), no. 4, 861–875.

[18] P. Cannarsa and Da Prato, *Some results on nonlinear optimal control problems and Hamilton-Jacobi equations in infinite dimensions*, J. Funct. Anal. 90 (1990), no. 1, 27–47.

[19] P. Cannarsa and G. Di Blasio, *A direct approach to infinite-dimensional Hamilton-Jacobi equations and applications to convex control with state constraints* Differential Integral Equations 8 (1995), no. 2, 225–246.

[20] P. Cannarsa, F. Gozzi and H. M. Soner, *A dynamic programming approach to nonlinear boundary control problems of parabolic type* J. Funct. Anal. 117 (1993), no. 1, 25–61.

[21] P. Cannarsa and H. Frankowska, *Value function and optimality conditions for semilinear control problems*, Appl. Math. Optim. 26 (1992), no. 2, 139–169.

[22] P. Cannarsa and H. Frankowska, *Value function and optimality condition for semilinear control problems. II. Parabolic case*, Appl. Math. Optim. 33 (1996), no. 1, 1–33.

[23] P. Cannarsa and M. E. Tessitore, *Optimality conditions for boundary control problems of parabolic type*, Control and estimation of distributed parameter systems: nonlinear phenomena (Vorau, 1993), 79–96, Internat. Ser. Numer. Math., 118, Birkhauser, Basel, 1994.

[24] P. Cannarsa and M. E. Tessitore, *Infinite dimensional Hamilton–Jacobi equations and Dirichlet boundary control problems of parabolic type*, SIAM J. Control Optim. 34 (1996), no. 6, pp. 1831-1847.

[25] F. H. Clarke, Yu. S. Ledyaev, E. D. Sontag and A. I. Subbotin, *Asymptotic controllability implies feedback stabilization*, IEEE Trans. Automat. Control 42 (1997), no. 10, 1394–1407.

[26] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.

[27] M. G. Crandall, H. Ishii P. L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Am. Math. Soc., New Ser. 27 (1992), no. 1, 1-67.

[28] M. G. Crandall and P. L. Lions, *Hamilton-Jacobi equations in infinite dimensions. IV. Hamiltonians with unbounded linear terms*, J. Funct. Anal. 90, (1990), 237-283.

[29] M. G. Crandall and P. L. Lions, *Hamilton-Jacobi equations in infinite dimensions. V. Unbounded linear terms and B-continuous solutions*, J. Funct. Anal. 97, (1991), 417-465.

[30] M. G. Crandall and P. L. Lions, *Hamilton-Jacobi equations in infinite dimensions. VI: Nonlinear A and Tataru's method refined*, Evolution equations, control theory, and biomathematics (Han sur Lesse, 1991), pp. 51-89, Lecture Notes in Pure and Appl. Math., 155, Dekker, New York, 1994.

[31] M. G. Crandall and P. L. Lions, *Hamilton-Jacobi equations in infinite dimensions. VII: The HJB equation is not always satisfied*, J. Funct. Anal. 125 (1994) pp. 111-148.

[32] G. Di Blasio, *Global solutions for a class of Hamilton-Jacobi equations in Hilbert spaces*, Numer. Funct. Anal. Optim. 8 (1985/86), no. 3-4, 261–300.

[33] I. Ekeland and G. Lebourg, *Generic Frechet-differentiability and perturbed optimization problems in Banach spaces.*, Trans. Amer. Math. Soc., 224, (1977), 193-216.

[34] S. Faggian and F. Gozzi, *On the dynamic programming approach for optimal control problems of PDE's with age structure*, Math. Popul. Stud. 11 (2004), no. 3-4, 233-270.

[35] G. Feichtinger, A. Prskawetz and V. M. Veliov, *Age-structured optimal control in population economics*, Math. Popul. Stud. 65 (2004), no. 4, 373-387.

[36] W. Fleming and R. W. Rishel, *Deterministic and stochastic optimal control*, Applications of Mathematics, No. 1. Springer-Verlag, Berlin-New York, 1975.

[37] W. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, Second edition, Stochastic Modelling and Applied Probability, 25, Springer, New York, 2006.

[38] F. Gozzi, S. S. Sritharan and A. Święch, *Viscosity solutions of dynamic-programming equations for the optimal control of the two-dimensional Navier-Stokes equations*, Arch. Ration. Mech. Anal. 163 (2002), no. 4, 295–327.

[39] M. Iannelli, *Mathematical theory of age-structured population dynamics*, Giardini, Pisa, (1995)

[40] M. Iannelli, M. Martcheva and F.A. Milner, *Gender-structured population modeling*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, (2005)

[41] H. Ishii, *Viscosity solutions for a class of Hamilton–Jacobi equations in Hilbert spaces*, J. Funct. Anal., 105 (1992), pp. 301–341.

[42] H. Ishii and S. Koike, *On ϵ-optimal controls for state constraint problems*, Ann. Inst. H. Poincare Anal. Non Lineaire 17 (2000), no. 4, 473–502.

[43] D. Kelome, *Viscosity solutions of second order equations in a separable Hilbert space and applications to stochastic optimal control*, Ph.D. Thesis, 2002.

[44] D. Kelome and A. Święch, *Perron's method and the method of relaxed limits for "unbounded" PDE in Hilbert spaces*, Studia Math. 176 (2006), no. 3, 249–277.

[45] M. Kocan and P. Soravia, *A viscosity approach to infinite-dimensional Hamilton-Jacobi equations arising in optimal control with state constraints*, SIAM J. Control Optim. 36 (1998), 1348–1375.

[46] X.Y. Li and J.M. Yong, *Optimal control theory for infinite-dimensional systems*, Birkhauser Boston, Cambridge, MA, 1995.

[47] P. D. Loewen, *Optimal control via nonsmooth analysis*, CRM Proceedings & Lecture Notes, 2. American Mathematical Society, Providence, RI, 1993. x+153 pp.

[48] D. Preiss, *Differentiability of Lipschitz functions on Banach spaces*, J. Funct. Anal. 91 (1990), no. 2, 312-345.

[49] M. Renardy, *Polar decomposition of positive operators and problem od Crandall and Lions*, Appl. Anal., 57-3 (1995), 383-385

[50] K. Shimano, *A class of Hamilton-Jacobi equations with unbounded coefficients in Hilbert spaces*, Appl. Math. Optim. 45 (2002), no. 1, 75–98.

[51] S. S. Sritharan, *Dynamic programming of the Navier-Stokes equations*, Systems Control Lett. 16 (1991), no. 4, 299–307.

[52] A. Święch, *Sub- and superoptimality principles of dynamic programming revisited*, Nonlinear Anal. 26 (1996), no. 8, 1429-1436.

[53] D. Tataru, *Viscosity solutions of Hamilton–Jacobi equations with unbounded linear terms*, J. Math. Anal. Appl. 163 (1992), pp. 345–392.

[54] D. Tataru, *Viscosity solutions for the dynamic programming equation*, Appl. Math. Optim., 25 (1992), pp. 109–126.

[55] R. Vinter, *Optimal control*, Systems & Control: Foundations & Applications. Birkhauser Boston, Inc., Boston, MA, 2000.

[56] J. Yong and X.Y. Zhou, *Stochastic controls*, Springer, New York, (1999).

[57] X. Y. Zhou, *Verification theorems within the framework of viscosity solutions*, J. Math. Anal. Appl. 177 (1993), no. 1, 208–225.