Journal of Convex Analysis
Volume 17 (2010), No. 2, 611–642

Verification Theorem and Construction of ϵ-Optimal Controls for Control of Abstract Evolution Equations

G. Fabbri*
Dipartimento di Studi Economici,
Università di Napoli Parthenope, Napoli, Italy;
and: School of Mathematics and Statistics, UNSW, Sydney, Australia
giorgio.fabbri@uniparthenope.it

F. Gozzi
Dipartimento di Scienze Economiche ed Aziendali,
Università LUISS - Guido Carli, Roma, Italy;
and: Centro De Giorgi, Scuola Normale Superiore, Pisa, Italy
fgozzi@luiss.it

A. Święch†
School of Mathematics, Georgia Institute of Technology,
Atlanta, GA 30332, USA
swiech@math.gatech.edu

Received: August 3, 2008

We study several aspects of the dynamic programming approach to optimal control of abstract evolution equations, including a class of semi-linear 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 present an explicit construction of ϵ-optimal controls.

Keywords: Optimal control of PDE's, verification theorem, dynamic programming, ϵ-optimal controls, Hamilton-Jacobi-Bellman equations

2000 Mathematics Subject Classification: 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

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

*Supported by the ARC Discovery project DP0558539.
†Supported by NSF grant DMS 0500270.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

where 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

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

where U is a metric space.

Our problem (1)–(2) is formulated in the standard abstract form (see e.g. [7, 15, 32, 51] and papers included in the references regarding infinite dimensional systems). It is very general and it includes optimal control problems of various semilinear partial differential equations (PDE) treated as abstract evolution equations. The reader may consult for instance [51] and papers cited in the next paragraph for many concrete examples of optimal control problems which fall into the framework of (1)–(2) (see also the example in Section 3.1).

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, 29, 41, 42, 52, 60, 61] and the references therein). The situation is much more complicated for optimal control of 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, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 31, 32, 33, 34, 35, 43, 46, 50, 51, 55, 56, 58, 59] and the references therein). Above works deal with optimal control problems for parabolic, semilinear and nonlinear evolutionary PDE, Navier-Stokes equations to name a few. There is also a growing literature on optimal control of age structured equations arising in population dynamics and economics [1, 13, 14, 37, 40]. 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 (Theorem 3.4) is an infinite dimensional version of such a result for finite dimensional problems obtained in [62]. It is based on the notion of viscosity solution (see Definitions 2.4–2.6). The proof in finite dimensions uses the Lipschitz continuity of the value function and of the state trajectories. However these properties, in particular the Lipschitz continuity of trajectories, are rather rare in the infinite dimensional case. To overcome this difficulty previous results in this direction assumed that the candidate optimal trajectory belonged to the domain of the differential operator A; see [23, 24] and the material in Chapter 6 §5 of [51], in particular Theorem 5.5 there which is based on [23]. Our Theorem 3.4 does not require this. We briefly discuss these issues in Remarks 3.6 and 3.7.

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 [57]. Similar method has been used in [28] to construct stabilizing feedbacks for nonlinear systems and later in [47] 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 2) 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 [31]) and then use results of [57] 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.

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

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

$$\langle (A^*B + c_0B)x, x\rangle \leq 0 \quad \text{for all } x \in \mathcal{H}. \tag{4}$$

Such an operator always exists, for instance

$$B = ((-A + I)(-A^* + I))^{-1/2} \tag{5}$$

(see [53], we refer to [31] 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

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

Let $\Omega \subset [0,T] \times \mathcal{H}$. We say that $u : \Omega \to \mathbb{R}$ is B-upper-semicontinuous (respectively, B-lower-semicontinuous) on Ω if whenever $t_n \to t$, $x_n \rightharpoonup x$, $Bx_n \to Bx$, $(t,x) \in \Omega$, then $\limsup_{n\to+\infty} u(t_n, x_n) \le u(t,x)$ (respectively, $\liminf_{n\to+\infty} u(t_n,x_n) \ge u(t,x)$). The function u is B-continuous on Ω if it is B-upper-semicontinuous and B-lower-semicontinuous on Ω.

We will say that a function v is B^k-semiconvex (respectively, B^k-semiconcave) for $k > 0$ if there exists a constant $C \ge 0$ such that $v(t,x) + C(\|x\|^2_{-k} + t^2)$ is convex (respectively, $v(t,x) - C(\|x\|^2_{-k} + t^2)$ is concave). If $k = 1$ we will call such functions simply B-semiconvex and B-semiconcave.

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

We make the following assumptions on b and L.

Hypothesis 2.1.

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

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

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

Hypothesis 2.2.

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

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

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

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

$$\|e^{sA}\| \le e^{\omega s} \ \text{ for all } s \ge 0 \tag{6}$$

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 (10). Alternatively, by making a change of variables $\tilde{v}(t,x) = v(t, e^{\omega t}x)$ in equation (10) (see [31], page 275) we can always reduce the case (6) to the case when A satisfies (3).

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

$$\begin{cases} \dot{x}_{t,x}(s) = Ax_{t,x}(s) + b(s, x_{t,x}(s), u(s)) \\ x_{t,x}(t) = x. \end{cases} \tag{7}$$

We recall that if Hypothesis 2.1 is satisfied then (7) has a unique continuous mild solution. We will write $x(\cdot)$ for $x_{t,x}(\cdot)$ when there is no possibility of confusion. We consider the function

$$J(t,x;u(\cdot)) = \int_t^T L(s, x(s), u(s))\mathrm{d}t + h(x(T)), \tag{8}$$

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

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

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

$$V(t,x) = \inf_{u(\cdot)\in\mathcal{U}[t,T]} J(t,x;u(\cdot)). \tag{9}$$

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

$$\begin{cases} v_t(t,x) + \langle Dv(t,x), Ax\rangle + H(t,x,Dv(t,x)) = 0 \\ v(T,x) = h(x), \end{cases} \tag{10}$$

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 [31, 32] which is slightly modified here. We consider two sets of tests functions:

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

and

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

We use test2 functions that are a little different from the ones used in [31]. The extra term $\eta(\cdot)$ in test2 functions is added to deal with unbounded solutions. Sometimes it is more convenient to take the finite linear combinations of functions $\eta(t)g_0(\|x\|)$ above as test2 functions, however this will not be needed here. We recall that $D\varphi$ and Dg stand for the Fréchet derivatives in space of these functions.

Definition 2.4. A function $v \in C((0,T] \times \mathcal{H})$ is a (viscosity) *subsolution* of the HJB equation (10) if

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

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

$$\varphi_t(\bar{t},\bar{x}) + g_t(\bar{t},\bar{x}) + \langle A^* D\varphi(\bar{t},\bar{x}), \bar{x}\rangle + H(\bar{t},\bar{x}, D\varphi(\bar{t},\bar{x}) + Dg(\bar{t},\bar{x})) \geq 0. \tag{11}$$

Definition 2.5. A function $v \in C((0,T] \times \mathcal{H})$ is a (viscosity) *supersolution* of the HJB equation (10) if

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

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

$$-\varphi_t(\bar{t},\bar{x}) - g_t(\bar{t},\bar{x}) - \langle A^* D\varphi(\bar{t},\bar{x}), \bar{x}\rangle + H(\bar{t},\bar{x}, -D\varphi(\bar{t},\bar{x}) - Dg(\bar{t},\bar{x})) \leq 0. \tag{12}$$

Definition 2.6. A function $v \in C((0,T] \times \mathcal{H})$ is a (viscosity) *solution* of the HJB equation (10) 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 (10). We will then call a viscosity subsolution (respectively, supersolution) simply a function that satisfies (11) (respectively, (12)).

Lemma 2.7. *Let Hypotheses* 2.1 *and* 2.2 *hold. Let* $\varphi \in$ test1 *and* $(t,x) \in (0,T) \times \mathcal{H}$. *Then for* $t < s \leq T$

$$\begin{aligned}\varphi(s, x_{t,x}(s)) - \varphi(t,x) = \int_t^s &(\varphi_t(r, x_{t,x}(r)) + \langle A^* D\varphi(r, x_{t,x}(r)), x_{t,x}(r)\rangle \\ &+ \langle D\varphi(r, x_{t,x}(r)), b(r, x_{t,x}(r), u(r))\rangle)\, dr\end{aligned} \tag{13}$$

and

$$\begin{aligned}&\frac{1}{s-t}\left(\varphi(s, x_{t,x}(s)) - \varphi(t,x)\right) \\ &= \varphi_t(t,x) + \langle A^* D\varphi(t,x), x\rangle + \frac{1}{s-t}\int_t^s \langle D\varphi(t,x), b(t,x,u(r))\rangle\, dr + \sigma_x(s-t),\end{aligned} \tag{14}$$

uniformly in $u(\cdot) \in \mathcal{U}[t,T]$ *for some modulus* σ_x.

Proof. See [51] Proposition 5.5, page 67 and Lemma 3.3, page 240. □

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* $t < s \leq T$

$$\begin{aligned}&g(s, x_{t,x}(s)) - g(t,x) \\ &\leq \int_t^s (g_t(r, x_{t,x}(r)) + \langle Dg(r, x_{t,x}(r)), b(r, x_{t,x}(r), u(r))\rangle)\, dr\end{aligned} \tag{15}$$

and

$$\frac{1}{s-t}\left(g(s,x_{t,x}(s))-g(t,x)\right)$$
$$\leq g_t(t,x)+\frac{1}{s-t}\int_t^s \langle Dg(t,x), b(t,x,u(r))\rangle \, dr + \sigma_x(s-t) \tag{16}$$

uniformly in $u(\cdot)\in\mathcal{U}[t,T]$ *for some modulus* σ_x.

Proof. The proof is also standard. (15) is obtained by considering solutions $x^n_{t,x}(\cdot)$ of (7) with A replaced by its Yosida approximations A_n, writing the formula (15) for $x^n_{t,x}(\cdot)$, using dissipativity of the A_n and then passing to the limit as $n\to+\infty$. Formula (16) follows from (15) and the (uniform in $u(\cdot)\in\mathcal{U}[t,T]$) continuity of $x_{t,x}(s)$ at $s=t$. □

Theorem 2.9. *Let Hypotheses* 2.1 *and* 2.2 *hold. Then the value function* V *(defined in* (9)*) is a viscosity solution of the HJB equation* (10).

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 [51] and the proof of Theorem 3.2, page 240 of [51] (or from [32]). □

We will need a comparison result in the proof of the verification theorem. There are various versions of such results for equation (10) available in the literature, several sufficient sets of hypotheses can be found in [31, 32]. 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 exist sets of functions $\mathcal{G}_1,\mathcal{G}_2$ on $(0,T]\times\mathcal{H}$ such that:

(i) the value function V is in $\mathcal{G}_1\cap\mathcal{G}_2$;

(ii) if $v_1\in\mathcal{G}_1, v_2\in\mathcal{G}_2$, v_1 is a viscosity subsolution of the HJB equation (10) and v_2 is a viscosity supersolution of the HJB equation (10) then $v_1\leq v_2$.

Note that from (i) and (ii) we know that V is the only solution of the HJB equation (10) in $\mathcal{G}_1\cap\mathcal{G}_2$.

Remark 2.11. Conditions that guarantee comparison are well known. For instance, Hypothesis 2.10 is satisfied (see [32, 48]) if in addition to Hypothesis 2.1 we assume Hypothesis 4.1 for some B satisfying (4)[1],

$$|h(x)-h(y)|\leq\omega(\|x-y\|_{-1})$$

for some modulus ω and all $x,y\in\mathcal{H}$, and

$$|h(x)|,|L(t,x,u)|\leq C(1+\|x\|^m) \ \text{ for all } (t,x,u)\in[0,T]\times\mathcal{H}\times U$$

for some $C,m\geq 0$. We can then take $\mathcal{G}_1$ to be the set of B-upper-semicontinuous functions v on $(0,T]\times\mathcal{H}$ such that

$$v(t,x)\leq c(1+\|x\|^k) \ \text{ for all } (t,x)\in(0,T]\times\mathcal{H} \tag{17}$$

[1]Recall that $\|x\|_{-1}=\|B^{1/2}x\|$.

for some $c, k \geq 0$, and

$$\lim_{t \to T} (v(t,x) - v(T,x))_+ = 0$$

uniformly on bounded subsets of $\mathcal{H}$, and $\mathcal{G}_2$ to be the set of B-lower-semicontinuous functions v on $(0,T] \times \mathcal{H}$ such that $-v$ satisfy (17) and

$$\lim_{t \to T} (v(t,x) - v(T,x))_- = 0$$

uniformly on bounded subsets of $\mathcal{H}$. These conditions are satisfied for instance for the vintage capital problem in Section 3.1, where A is the differentiation operator with a zero order term.

Conditions for Hypothesis 2.10 without Hypothesis 4.1 can be found in [32], Theorem 7.2 under a stronger assumption

$$\langle (A^*B + c_0 B)x, x \rangle \leq -\|x\|^2 \quad \text{for all } x \in \mathcal{H}$$

and some $c_0 \leq 0$, which is typically satisfied by second order elliptic operators (see [31, 51, 53]). For bounded domains and Dirichlet boundary conditions B can then be taken to be $-\lambda\Delta^{-1}$ for some $\lambda > 0$. Other sets of conditions that guarantee Hypothesis 2.10 can be found in [31, 32, 48, 51].

For $f \in C([t,T];\mathbb{R})$ and $s \in (t,T)$ we denote

$$\overline{D}f(s) = \limsup_{h \to 0} \frac{f(s+h) - f(s)}{h}$$

and

$$D_+ f(s) = \liminf_{h \to 0^+} \frac{f(s+h) - f(s)}{h}.$$

We state below a hypothesis for an abstract function $f \in C([t,T];\mathbb{R})$. We will require it to be satisfied for a certain function in the verification theorem of the next section.

Hypothesis 2.12. If

$$\overline{D}f(s) \leq g(s) \quad a.e.\ s \in (t,T) \tag{18}$$

for some $g \in L^1(t,T;\mathbb{R})$ then for every $t \leq \alpha < \beta \leq T$

$$f(\beta) - f(\alpha) \leq \int_\alpha^\beta g(s)ds. \tag{19}$$

Remark 2.13. Hypothesis 2.12 guarantees that we can control increments of f by its upper Dini derivative $\overline{D}f(s)$. It is obviously satisfied when f is absolutely continuous. In general (18) is not enough to obtain (19). However, by Theorem 7.3, Chapter VI, page 204 of [54], if (18) holds and $D_+f(s) < +\infty$ at every point s except at most at those of a countable set then (19) is satisfied. Estimate (19) also holds if there exists $\rho \in L^1(t,T;\mathbb{R})$ such that, for some $h_0 > 0$, we have $\frac{f(s+h)-f(s)}{h} \leq \rho(s)$, for $0 < h \leq h_0$ and $s + h \leq T$.

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 $v(s,y)-v(t,x)-\langle p,y-x\rangle-q(s-t)\le o(\|x-y\|+|t-s|)$, 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 $v(s,y)-v(t,x)-\langle p,y-x\rangle-q(s-t)\ge o(\|x-y\|+|t-s|)$.

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

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

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 (7).

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. Its main strength is that it holds under relatively weak assumptions on A and it does not assume that the (optimal) trajectory lies in the domain of A.

Theorem 3.4. *Let Hypotheses* 2.1, 2.2 *and* 2.10 *hold. Let $v\in\mathcal{G}_1$ be a subsolution of the HJB equation* (10) *such that*

$$v(T,x)=h(x)\quad \text{for all } x \text{ in } \mathcal{H}. \tag{20}$$

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

(*b*) *Let $(t,x)\in(0,T)\times H$ and let $(x_{t,x}(\cdot),u(\cdot))$ be an admissible couple for (t,x). Assume that the function $f(s):=v(s,x_{t,x}(s))$ satisfies Hypothesis* 2.12 *and 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*

$$(q(s),p_1(s),p_2(s))\in E^{1,+}v(s,x_{t,x}(s))\quad \text{for almost all } s\in(t,T) \tag{21}$$

and that

$$\begin{aligned}&\int_t^T (\langle p_1(s)+p_2(s), b(s,x_{t,x}(s),u(s))\rangle + q(s) + \langle A^*p_1(s), x_{t,x}(s)\rangle)\mathrm{d}s\\ &\leq \int_t^T -L(s,x_{t,x}(s),u(s))\mathrm{d}s. \end{aligned}\tag{22}$$

Then $(x_{t,x}(\cdot),u(\cdot))$ *is an optimal couple for* (t,x) *and* $v(t,x)=V(t,x)$. *Moreover we have equality in* (22).

Remark 3.5. It is tempting to try to prove, along the lines of Theorem 3.9, p. 243 of [61], that a condition like (22) 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 subdifferential object[2] $E^{1,-}$ are fundamentally different and it is not clear if a broad natural generalization of a result like Theorem 3.9, p. 243 of [61] is possible in this context. The reader may notice that the p_2 components of $E^{1,+}$ consist of derivatives of radially increasing functions while the p_2 components of $E^{1,-}$ consist of derivatives of radially decreasing functions.

Remark 3.6. Our verification theorem has some drawbacks.

- Hypothesis 2.12 for $f(s) := v(s,x_{t,x}(s))$ is not easy to check. Therefore it is important to know when it is verified. We mention three cases in which this can be done:
  - (i) When u,q,p_1,p_2,A^*p_1 are continuous and (21) is satisifed for every $s\in(t,T)$ except at most for those of a countable set. If the candidate optimal couple is found solving a closed loop equation as in [39], Section 4.3 and the feedback map is continuous, this guarantees the continuity of the candidate optimal control u. The continuity of the other functions q,p_1,p_2,A^*p_1 can follow (as in [39], Section 4.3) from mild regularity assumptions on the superdifferentials of the value function. This is the case of our example developed in Subsection 3.1.
  - (ii) When the value function and the candidate optimal state trajectories are Lipschitz continuous. The Lipschitz continuity of the candidate optimal state trajectories is guaranteed e.g. when they belong to the domain of A in particular when the semigroup generated by A is analytic.
  - (iii) When v in Theorem 3.4 is Lipschitz in $|\cdot|\times\|\cdot\|_{-2}$ norm and B-semiconcave. In this case $E^{1,+}v(t,x)$ is always nonempty and it contains a closed and convex subset of $\mathbb{R}\times D(A^*)\times\{0\}$. It is then easy to see using the definition of B-semiconcavity that $D_+f(s)<+\infty$ for every s. Since $E^{1,+}v(t,x)$ is nonempty for every (t,x) we can thus always apply the verification theorem.

[2]Following the definition of viscosity supersolution we define

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

We remark that for a minimization problem it is often natural to expect that the value function may be B-semiconcave (see Example 3.11).

- Condition (22) implicitly implies that $\langle p_2(r), Ax_{t,x}(r)\rangle = 0$ a.e. if the trajectory is in the domain of A. This follows from the fact that in such a case the limit in the left hand side of (24) in the proof of Theorem 3.4 can be explicitly computed and the result is the right hand side of (24) plus the negative term $\eta(r)g_0'(|x_{t,x}(r)|)\,\langle Ax_{t,x}(r), x_{t,x}(r)\rangle$. Using this one would get from (26) that

$$V(t,x) \geq J(t,x,u) - \int_t^T \eta(s)g_0'(|x_{t,x}(s)|)\,\langle Ax_{t,x}(s), x_{t,x}(s)\rangle\,\mathrm{d}s$$

  which implies the claim. Therefore the applicability of the theorem is somehow limited as in practice (22) may be satisfied only if the function is “nice” (i.e. its superdifferential should really only consist of points p_1 belonging to the domain of A^*).
- From the statement of the theorem it follows that the superdifferential of v must be nonempty at almost every point of the admissible trajectory considered. So if an optimal trajectory on a set of positive measure goes through a set where the superdifferential of v is empty it cannot be discovered with such a condition.

However, despite all of the drawbacks our verification theorem still can be applied in some cases where other results fail (see Remarks 3.7 and 3.9). Perhaps when the operator A is more coercive a better result is possible. We plan to investigate such cases in the future. However our assumptions on A here are very weak.

Proof. The first statement ($v \leq V$) is just a restatement of Hypothesis 2.10. It remains to prove the second one. The function

$$s \mapsto b(s, x_{t,x}(s), u(s))$$

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 points which are both left- and right-Lebesgue points of this function that in addition satisfy (21) is of full measure. We choose $t < r < T$ 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) $v \leq \varphi + g$ 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, 2.8 and the continuity of

$x_{t,x}(s)$

$$
\begin{aligned}
&\frac{v(\tau, x_{t,x}(\tau)) - v(r,y)}{\tau - r} \leq \frac{g(\tau, x_{t,x}(\tau)) - g(r,y)}{\tau - r} + \frac{\varphi(\tau, x_{t,x}(\tau)) - \varphi(r,y)}{\tau - r} \\
&\leq \frac{1}{\tau - r}\int_r^\tau \big(g_t(s, x_{t,x}(s)) + \langle Dg(s, x_{t,x}(s)), b(s, x_{t,x}(s), u(s))\rangle \\
&\qquad + \varphi_t(s, x_{t,x}(s)) + \langle D\varphi(s, x_{t,x}(s)), b(s, x_{t,x}(s), u(s))\rangle \\
&\qquad + \langle A^* D\varphi(s, x_{t,x}(s)), x_{t,x}(s)\rangle\big)\, ds \\
&\leq \frac{1}{\tau - r}\bigg(\int_r^\tau \big(g_t(r,y) + \langle Dg(r,y), b(s, x_{t,x}(s), u(s))\rangle + \varphi_t(r,y) \\
&\qquad + \langle D\varphi(r,y), b(s, x_{t,x}(s), u(s))\rangle + \langle A^* D\varphi(r,y), y\rangle\big)\, ds + o(\tau - s)\bigg)
\end{aligned}
\tag{23}
$$

In view of the choice of r we know that

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

and

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

and thus we can pass to the limsup in (23) to obtain that

$$
\begin{aligned}
&\limsup_{\tau \downarrow r} \frac{v(\tau, x_{t,x}(\tau)) - v(r, x_{t,x}(r)))}{\tau - r} \\
&\leq \langle Dg(r, x_{t,x}(r)) + D\varphi(r, x_{t,x}(r)), b(r, x_{t,x}(r), u(r))\rangle \\
&\qquad + g_t(r, x_{t,x}(r)) + \varphi_t(r, x_{t,x}(r)) + \langle A^* D\varphi(r, x_{t,x}(r)), x_{t,x}(r)\rangle \\
&= \langle p_1(r) + p_2(r), b(r, x_{t,x}(r), u(r))\rangle + q(r) + \langle A^* p_1(r), x_{t,x}(r)\rangle\,.
\end{aligned}
\tag{24}
$$

Repeating the above procedure for $\tau < r$ we can also obtain (24) with $\limsup_{\tau\downarrow r}$ replaced by $\limsup_{\tau\uparrow r}$. Then, Hypothesis 2.12 and (22) yield

$$
\begin{aligned}
&v(T, x_{t,x}(T)) - v(t,x) \\
&\leq \int_t^T \big(\langle p(r), b(r, x_{t,x}(r), u(r))\rangle + q(r) + \langle A^* p_1(r), x_{t,x}(r)\rangle\big)\mathrm{d}r \\
&\leq \int_t^T -L(r, x_{t,x}(r), u(r))\mathrm{d}r.
\end{aligned}
\tag{25}
$$

Thus, using (a), we finally arrive at

$$
\begin{aligned}
V(T, x_{t,x}(T)) - V(t,x) &= h(x_{t,x}(T)) - V(t,x) \leq h(x_{t,x}(T)) - v(t,x) \\
&= v(T, x_{t,x}(T)) - v(t,x) \leq \int_t^T -L(r, x_{t,x}(r), u(r))\mathrm{d}r
\end{aligned}
\tag{26}
$$

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

Remark 3.7. In the book [51] (page 263, Theorem 5.5) the authors present a verification theorem (based on a previous result of [24], see also [23] 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 [51] (under assumptions similar to and in some aspects slightly stronger than Hypotheses 2.1 and 2.2) that the couple $(x(\cdot), u(\cdot))$ is optimal if and only if

$$u(s) \in \left\{ u \in U \,:\, \lim_{\delta \to 0} \frac{V((s+\delta), x(s) + \delta(Ax(s) + b(s, x(s), u))) - V(s, x(s))}{\delta} \right.$$
$$\left. = -L(s, x(s), u) \right\} \tag{27}$$

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 [1, 44, 45]), in population economics [40], optimal technology adoption in a vintage capital context [13, 14].

Consider the state equation

$$\begin{cases} \dot{x}(s) = Ax(s) + Ru(s) \\ x(t) = x \end{cases} \tag{28}$$

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 2 satisfying (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

$$J(x, u(\cdot)) = \int_t^T \left(- |\langle \alpha, x(s) \rangle| + \frac{1}{2} u(s)^2 \right) \mathrm{d}s. \tag{29}$$

We define

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

and we take $M \overset{def}{=} \sup_{t \in [0,T]} |\langle \bar{\alpha}(t), \beta \rangle|$. 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 $b(t, x, u) = Ru$, $L(t, x, u) = -|\langle \alpha, x(t) \rangle| + (1/2)u(t)^2$, $h = 0$, $U = [-M-1, M+1]$.

The HJB equation (10) becomes

$$\begin{cases} v_t + \langle Dv, Ax\rangle - |\langle \alpha, x\rangle| + \inf_{u\in U}\left(\langle u, R^* Dv\rangle_{\mathbb{R}} + \frac{1}{2}u^2\right) = 0 \\ v(T,x) = 0. \end{cases} \tag{30}$$

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 $\int_t^T \left(\langle \alpha, x(s)\rangle + \frac{1}{2}u(s)^2\right) \mathrm{d}s$ (respectively $\int_t^T \left(-\langle \alpha, x(s)\rangle + \frac{1}{2}u(s)^2\right) \mathrm{d}s$) and it is known in the literature (see [38] 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 $\langle \alpha, x\rangle = 0$,

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

Therefore we can glue v_1 and v_2 writing

$$W(t,x) = \begin{cases} v_1(t,x) & \text{if } \langle \alpha, x\rangle \le 0 \\ v_2(t,x) & \text{if } \langle \alpha, x\rangle > 0. \end{cases} \tag{31}$$

It is easy to see that W is continuous and concave in x. We claim that W is a viscosity solution of (30). 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

$$\partial_t(\varphi+g)(t,x) = \partial_t W(t,x) = \frac{1}{2}\left(R^*\bar\alpha(t)\right)^2, \tag{32}$$

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

$$\inf_{u\in U}\left(\langle Ru, p\rangle + \frac{1}{2}u^2\right) = -\frac{1}{2}\gamma^2\left(R^*\bar\alpha(t)\right)^2. \tag{33}$$

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

$$\begin{aligned}\langle A^*p_1, x\rangle = &\ \langle A^*(p-p_2), x\rangle = \langle A^*\gamma G(t)\alpha, x\rangle - \frac{g_0'(|x|)}{|x|}\langle A^*x, x\rangle \\ &\geq \gamma G(t)\langle A^*\alpha, x\rangle = 0. \end{aligned}\tag{34}$$

So, by (32), (33) and (34),

$$\begin{aligned}&\partial_t(\varphi+g)(t,x) + \langle A^*p_1, x\rangle - |\langle \alpha, x\rangle| + \inf_{u\in U}\left(\langle Ru, D(\varphi+g)(t,x)\rangle + \frac{1}{2}u^2\right) \\ &\geq \frac{1}{2}(1-\gamma^2)\left(R^*\bar\alpha(s)\right)^2 \geq 0\end{aligned}\tag{35}$$

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

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

for some $C \geq 0$. In particular W and V are B-continuous and 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 (30) in $(0,T]\times\mathcal{H}$. Moreover, since $\alpha = By$ for some $y \in \mathcal{H}$, comparison holds for equation (30) which yields $W = V$ on $[0,T]\times\mathcal{H}$. (Comparison theorem can be easily obtained by a modification of techniques of [32] 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 [48], see Remark 2.11 here. The reader can also consult the proof of Theorem 4.4 of [49]. We point out that our assumptions are different from the assumptions of the uniqueness Theorem 4.6 of [51], 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

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

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

Proposition 3.8. *The feedback map given by*

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

is optimal. Similarly, also the feedback map

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

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 is 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 an 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 $s \ge t$.

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

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

and $p_2(s) = 0$. It is easy to see that $(q(s), p_1(s), p_2(s)) \in E^{1,+}V(s, x^{op}(s))$ and that the right hand side of (24) is finite at every point in this case. Therefore Hypothesis 2.12 is satisfied. 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 in particular with the model used in [13] (see also [14, 40]): $x(t)[s]$ represents the amount of capital of vintage s at time t, $\bar{s}$ is the maximal age of the capital (we take $\bar{s} = 1$ for simplicity), $u(t)$ is the amount of investment at time t, its contribution to the capital is $u(t)\beta(s)$, and $\mu > 0$ is the rate of depreciation. Let $\mathcal{H} = W^{1,2}(0,1)$ and let $\{e^{tA};\ t \geq 0\}$ be the semigroup on $W^{1,2}(0,1)$ defined as follows: for every $f \in W^{1,2}(0,1)$ and every $s \in [0,1]$

$$(e^{tA}f)[s] := \begin{cases} e^{-\mu t} f(s-t) & \text{if } s-t \geq 0 \\ e^{-\mu t} f(0) & \text{if } s-t < 0. \end{cases}$$

The domain of A will be

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

and, for all f in $D(A)$, $A(f)[s] = -\frac{\mathrm{d}}{\mathrm{d}s} f(s) - \mu f(s)$. The expression for A^* is

$$\begin{cases} D(A^*) = \{ f \in W^{2,2}(0,1) \ : \ f'(1) = 0 \} \\ A^*(f)[s] = f'(s) - f(1)h_1(s) + f(0)h_0(s) - \mu f(s) \end{cases}$$

(as proved in [13] Appendix B), where

$$h_0(s) := \frac{\cosh(1-s)}{\sinh(1)}, \quad h_1(s) := \frac{\cosh(s)}{\sinh(1)} \ \text{ for } s \in [0,1].$$

Using the fact that $\langle h_0, f \rangle_{W^{1,2}} = f(0)$ for all $f \in W^{1,2}(0,1)$ it is easy to check that A is maximal dissipative for $\mu \geq \mu_0 = \frac{1}{2}\|h_0\|^2_{W^{1,2}} = \frac{1}{2}\coth(1)$. By [53], if $\mu > \mu_0$, we can simply take $B = (AA^*)^{-1/2}$. We choose α to be the constant function equal to 1 at every point of the interval $[0,1]$. Again it is easy to verify that α is in the image of B and the hypotheses of Remark 2.11 are satisfied. Consequently (thanks to the explicit form of V given in (31)) we can conclude that Hypothesis 2.10 is satisfied with the sets $\mathcal{G}_1$ and $\mathcal{G}_2$ described in Remark 2.11.

We now choose $\beta(s) = \chi_{[0,\frac{1}{2}]}(s) - \chi_{[\frac{1}{2},1]}(s)$ (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 3.8 an optimal strategy u^{op} is

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

The related optimal trajectory is

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

Remark 3.9. We observe that, using such strategy, if the initial datum x is such that $\langle x, \alpha \rangle = 0$ then $\langle \alpha, x^{op}(t) \rangle = 0$ for all $s \geq t$. So the trajectory remains for the 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 [51] (page 263, Theorem 5.5) are not verified in this case.

Remark 3.10. Note that, if we consider the functional

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

with $\sigma \in (0,1)$, when $\langle \alpha, \beta \rangle = 0$, we can easily see that the optimal control is always $u \equiv 0$ and then the value function (that is the only viscosity solution of the HJB equation) is

$$V(t,x) = \int_t^T -|\left\langle \alpha, e^{sA}x \right\rangle|^\sigma \mathrm{d}s.$$

V is not semiconcave but Theorem 3.4 can be applied as well. In particular if for the initial datum x we have $\langle \alpha, x \rangle = 0$ the same pathology described in Remark 3.9 holds.

Example 3.11. In this example we present a class of problems for which the value functions are B-semiconcave. We do not consider the most general case as we just want to show that this property is not unnatural.

Consider the dynamics

$$\begin{cases} \dot{x}(s) = Ax(s) + b(u(s)), \\ x(t) = x, \end{cases} \tag{36}$$

where $A = A^*$ on $\mathcal{H}$, and $b : U \to \mathcal{H}$ is bounded and continuous. We will denote the solution of (36) by $x(s; t, x, u(\cdot))$. Suppose there is no running cost and we want to minimize a cost functional $J(t, x; u(\cdot)) = h(x(T; t, x, u(\cdot)))$. We assume that A commutes with B, h is B^2-semiconcave and

$$|h(x) - h(y)| \le C\|x - y\|_{-4}. \tag{37}$$

Using (4) it is easy to see that

$$\|x(s; t, x, u(\cdot)) - x(s; t, y, u(\cdot))\|_{-k} \le C\|x - y\|_{-k} \tag{38}$$

for $s \in [t, T]$ for every $k \ge 1$. Therefore it follows from (37) and (38) that

$$|V(t,x) - V(t,y)| \le C\|x - y\|_{-4} \tag{39}$$

for all $t \in [0, T], x, y \in \mathcal{H}$, where V is the value function.

Let $t \in [0, T], x, h \in \mathcal{H}$ and let $u(\cdot)$ be an ϵ-optimal control for (t, x). Then using the B^2-semiconcavity of h we have

$$\begin{aligned} &V(t, x+h) + V(t, x-h) - 2V(t,x) \\ &\le h(x(T; t, x+h, u(\cdot))) + h(x(T; t, x-h, u(\cdot))) - 2h(x(T; t, x, u(\cdot))) + 2\epsilon \\ &\le C\|x(T; t, x+h, u(\cdot)) - x(T; t, x-h, u(\cdot)\|_{-2}^2 + 2\epsilon \le C\|h\|_{-2}^2 + 2\epsilon. \end{aligned} \tag{40}$$

To show the full B-semiconcavity of V we adapt the strategy of [25], p. 198–199.

Let $x, h \in \mathcal{H}, 0 \le t - \tau \le t + \tau \le T$, and let $u(\cdot)$ be a control from the dynamic programming principle such that

$$V(t,x) \ge V(t+\tau, x(t+\tau; t, x, u(\cdot))) - \epsilon. \tag{41}$$

Define

$$\bar{u}(s) = u\left(\frac{t+\tau+s}{2}\right), \quad s \in [t-\tau, t+\tau].$$

After some computations we obtain

$$\|x+h-x(t+\tau;t-\tau,x-h,\bar{u}(\cdot))\|_{-2} \le 2\|h\|_{-2}+C\tau, \tag{42}$$

and

$$\begin{aligned}
&\|x+h+x(t+\tau;t-\tau,x-h,\bar{u}(\cdot))-2x(t+\tau;t,x,u(\cdot))\|_{-4} \\
&= 2\|\int_t^{t+\tau} AB^2(x(2s-t-\tau;t-\tau,x-h,\bar{u}(\cdot))-x(s;t,x,u(\cdot)))ds\| \\
&\le C\tau(\|h\|_{-2}+\tau).
\end{aligned} \tag{43}$$

It therefore follows from (39)–(43) that

$$\begin{aligned}
&V(t+\tau,x+h)+V(t-\tau,x-h)-2V(t,x) \\
&\le V(t+\tau,x+h)+V(t+\tau,x(t+\tau;t-\tau,x-h,\bar{u}(\cdot))) \\
&\qquad -2V(t+\tau,x(t+\tau;t,x,u(\cdot)))+2\epsilon \\
&\le V(t+\tau,x+h)+V(t+\tau,x(t+\tau;t-\tau,x-h,\bar{u}(\cdot))) \\
&\qquad -2V\left(t+\tau,\frac{x+h+x(t+\tau;t-\tau,x-h,\bar{u}(\cdot))}{2}\right)+2\epsilon \\
&\qquad +2V\left(t+\tau,\frac{x+h+x(t+\tau;t-\tau,x-h,\bar{u}(\cdot))}{2}\right) \\
&\qquad -2V(t+\tau,x(t+\tau;t,x,u(\cdot))) \\
&\le C\|x+h+x(t+\tau;t-\tau,x-h,\bar{u}(\cdot))\|_{-2}^2+4\epsilon \\
&\qquad +C\|x+h+x(t+\tau;t-\tau,x-h,\bar{u}(\cdot))-2x(t+\tau;t,x,u(\cdot))\|_{-4} \\
&\le C(\|h\|_{-2}+\tau)^2+C\tau(\|h\|_{-2}+\tau)+4\epsilon
\end{aligned} \tag{44}$$

and we conclude sending $\epsilon \to 0$. This inequality shows that V is B^2-semiconcave. Similar result would also hold for a more general case under appropriate assumptions on b and L.

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

The construction of ϵ-optimal feedback controls is a consequence of the proof of the superoptimality principle of dynamic programming. It is similar to a method used in the finite dimensional case in [57]. This method relied on the regularization of the value function by its inf-convolution which satisfies a slightly perturbed HJB inequality, and integration along trajectories obtained by choosing piecewise constant controls that approximately minimized the Hamiltonian. A generalization of this procedure to the infinite dimensional case is very delicate because of the presence of the unbounded operator A. First we have to employ an appropriate modification of the $\|\cdot\|_{-1}$ norm inf-convolution introduced in [32]. We then show that this inf-convolution is a supersolution of a perturbed HJB inequality in a sense that a pointwise inequality holds at every point at properly chosen elements of its superdifferential. This is part of Lemmas 4.3 and 4.5 which generalize a result from [32]. The last step is the proper selection of piecewise constant feedback controls and a careful integration along trajectories.

Overall the procedure is very technical and is not a straightforward generalization of the finite dimensional technique.

Throughout this section B is a bounded, linear, positive, self-adjoint operator introduced in Section 2 satisfying (4).

To study the construction of ϵ-optimal controls we need to impose another set of assumptions on b and L:

Hypothesis 4.1. The functions $b, L\colon [0,T] \times \mathcal{H} \times U \to \mathbb{R}$ are continuous and 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

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

for all $t, s \in [0,T], x, y \in \mathcal{H}, u \in U$.

Let $m \ge 2$. Modifying slightly the functions introduced in [32] 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*

$$w(t,x) \le C(1+\|x\|^k) \quad (\textit{respectively, } w(t,x) \ge -C(1+\|x\|^k)) \tag{45}$$

on $(0,T)\times\mathcal{H}$ for some $k \ge 0$. Let $m > k, m \ge 2$. 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

$$|v(t,x) - v(s,y)| \le M_{R,\epsilon,\beta}(|t-s| + \|x-y\|_{-2}) \quad \textit{on } (0,T)\times B_R. \tag{46}$$

(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). In particular $w^{\lambda,\epsilon,\beta}$ (respectively, $w_{\lambda,\epsilon,\beta}$) is B-semiconvex (respectively, B-semiconcave).

(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

$$w^{\lambda,\epsilon,\beta}(t,x) = \sup_{(s,y)\in(0,T)\times\mathcal{H},\ \|y\|\le N} \left\{ w(s,y) - \frac{\|x-y\|_{-1}^2}{2\epsilon} - \frac{(t-s)^2}{2\beta} - \lambda e^{2mK(T-s)}\|y\|^m \right\}, \tag{47}$$

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

$$\begin{aligned} &|w^{\lambda,\epsilon,\beta}(t,x) - w^{\lambda,\epsilon,\beta}(s,y)| \\ &\le \sigma - \frac{\|x-\tilde{x}\|_{-1}^2}{2\epsilon} - \frac{(t-\tilde{t})^2}{2\beta} + \frac{\|\tilde{x}-y\|_{-1}^2}{2\epsilon} + \frac{(\tilde{t}-s)^2}{2\beta} \\ &\le \sigma - \frac{\langle B(x-y), x+y\rangle}{2\epsilon} + \frac{\langle B(x-y),\tilde{x}\rangle}{\epsilon} + \frac{(2\tilde{t}-t-s)(t-s)}{2\beta} \\ &\le \frac{(2R+N)}{2\epsilon}\|B(x-y)\| + \frac{2T}{2\beta}|t-s| + \sigma \end{aligned} \tag{48}$$

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 [30].

(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 B-upper-semicontinuous viscosity subsolution (respectively, a B-lower-semicontinuous viscosity supersolution) of* (10) *satisfying* (45). *Let $m > k, m \ge 2$. Then for every $R, \delta > 0$ there exists a non-negative function $\gamma_{R,\delta}(\lambda,\epsilon,\beta)$, where*

$$\lim_{\lambda\to 0}\limsup_{\epsilon\to 0}\limsup_{\beta\to 0}\gamma_{R,\delta}(\lambda,\epsilon,\beta) = 0, \tag{49}$$

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

$$v_t(t,x) + \langle Dv(t,x), Ax\rangle + H(t,x,Dv(t,x)) = -\gamma_{R,\delta}(\lambda,\epsilon,\beta) \quad \text{in } (\delta, T-\delta)\times B_R \tag{50}$$

(respectively,

$$v_t(t,x) + \langle Dv(t,x), Ax\rangle + H(t,x,Dv(t,x)) = \gamma_{R,\delta}(\lambda,\epsilon,\beta) \quad \text{in } (\delta, T-\delta)\times B_R) \quad (51)$$

for β sufficiently small (depending on δ), in the sense that if $v-\psi$ has a local maximum (respectively, $v+\psi$ has a local minimum) at (t,x) for a test function $\psi = \varphi + g$ then

$$\psi_t(t,x) + \langle A^* D\psi(t,x), x\rangle + H(t,x,D\psi(t,x)) \ge -\gamma_{R,\delta}(\lambda,\epsilon,\beta) \quad (52)$$

(respectively,

$$-\psi_t(t,x) - \langle A^* D\psi(t,x), x\rangle + H(t,x,-D\psi(t,x)) \le \gamma_{R,\delta}(\lambda,\epsilon,\beta)). \quad (53)$$

Proof. The proof is similar to the proof of Proposition 5.3 of [32]. 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 $w^{\lambda,\epsilon,\beta}-\varphi-g$. We can assume that the maximum is global and strict (see Proposition 2.4 of [32]) and that $w^{\lambda,\epsilon,\beta}-\varphi-g \to -\infty$ as $\|x\|\to\infty$ uniformly in t. In view of these facts and (47) we can choose $S > 2\|x_0\|$, depending on λ such that, for all $\|x\|+\|y\| > S-1$ and $s,t\in(0,T)$,

$$\begin{aligned} &w(s,y) - \frac{1}{2\epsilon}\|(x-y)\|_{-1}^2 - \frac{(t-s)^2}{2\beta} - \lambda e^{2mK(T-s)}\|y\|^m - \varphi(t,x) - g(t,x) \\ &\le w(t_0,x_0) - \lambda e^{2mK(T-t_0)}\|x_0\|^m - \varphi(t_0,x_0) - g(t_0,x_0) - 1. \end{aligned} \quad (54)$$

We can then use a perturbed optimization technique of [32] (see page 424 there), which is a version of the Ekeland-Lebourg Lemma [36], 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

$$\begin{aligned} \phi(t,x,s,y) \overset{def}{=} \; &w(s,y) - \frac{1}{2\epsilon}\|(x-y)\|_{-1}^2 - \frac{(t-s)^2}{2\beta} - \lambda e^{2mK(T-s)}\|y\|^m \\ &- g(t,x) - \varphi(t,x) - \langle Bp, y\rangle - \langle Bq, x\rangle - at - bs \end{aligned} \quad (55)$$

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 (54) 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

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

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

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

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 [46]) give us

$$\lim_{\beta\to 0}\limsup_{\alpha\to 0}\frac{|\bar{s}-\bar{t}|^2}{2\beta}=0, \tag{56}$$

$$\lim_{\epsilon\to 0}\limsup_{\beta\to 0}\limsup_{\alpha\to 0}\frac{|\bar{x}-\bar{y}|_{-1}^2}{2\epsilon}=0. \tag{57}$$

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

$$\begin{aligned}-\frac{(\bar{t}-\bar{s})}{\beta}-2\lambda mKe^{2mK(T-\bar{s})}\|\bar{y}\|^m+b-\frac{\langle A^*B(\bar{x}-\bar{y}),\bar{y}\rangle}{\epsilon}+\langle A^*Bp,\bar{y}\rangle\\ +H\left(\bar{s},\bar{y},\frac{1}{\epsilon}B(\bar{y}-\bar{x})+\lambda me^{2mK(T-\bar{s})}\|y\|^{m-1}\frac{y}{\|y\|}+Bp\right)\geq 0.\end{aligned} \tag{58}$$

We notice that

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

and

$$\frac{1}{\epsilon}B(\bar{y}-\bar{x})=D\varphi(\bar{t},\bar{x})+Dg(\bar{t},\bar{x})+Bq \tag{59}$$

which in particular implies that $Dg(\bar{t},\bar{x})\in D(A^*)$, i.e. $\bar{x}\in D(A^*)$. Therefore using this, the assumptions on b and L, and (56) and (57) we have, denoting $\psi=\varphi+g$,

$$\begin{aligned}&\psi_t(\bar{t},\bar{x})+\langle\bar{x},A^*D\psi(\bar{t},\bar{x})\rangle+H\left(\bar{t},\bar{x},D\psi(\bar{t},\bar{x})\right)\\ &\geq 2\lambda mKe^{2mK(T-\bar{s})}\|\bar{y}\|^m-\langle A^*Bp,\bar{y}\rangle-a-b\\ &\qquad-\frac{1}{\epsilon}\langle(\bar{y}-\bar{x}),A^*B(\bar{y}-\bar{x})\rangle-\langle\bar{x},A^*Bq\rangle+H\left(\bar{t},\bar{x},\frac{1}{\epsilon}B(\bar{y}-\bar{x})-Bq\right)\\ &\qquad-H\left(\bar{s},\bar{y},\frac{1}{\epsilon}B(\bar{y}-\bar{x})+\lambda me^{2mK(T-\bar{s})}\|y\|^{m-1}\frac{y}{\|y\|}\right)\\ &\geq 2\lambda mKe^{2mK(T-\bar{s})}\|\bar{y}\|^m-C_{\lambda,\epsilon}\alpha+\frac{c_0}{\epsilon}\|\bar{x}-\bar{y}\|_{-1}^2-K\|\bar{x}-\bar{y}\|_{-1}\frac{\|B(\bar{x}-\bar{y})\|}{\epsilon}\\ &\qquad-\gamma_{\lambda,\epsilon}(|\bar{t}-\bar{s}|)-\lambda m(M+K\|\bar{y}\|)e^{2mK(T-\bar{s})}\|\bar{y}\|^{m-1}\\ &\geq\ -C_{\lambda,\epsilon}\alpha-\gamma(\lambda,\epsilon,\beta,\alpha)\end{aligned} \tag{60}$$

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

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

Since $(\bar{t},\bar{x})\to(t_0,x_0)$ as $\alpha\to 0$ we have that $D\psi(\bar{t},\bar{x})\to D\psi(t_0,x_0)$, and from (59) that, possibly along a subsequence, $A^*Dg(\bar{t},\bar{x})\rightharpoonup\bar{z}$ for some $\bar{z}$. But this obviously implies that $\bar{z}=A^*Dg(t_0,x_0)$. We therefore 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

$$|w(t,x) - w(s,y)| \leq \sigma(\|x-y\|_{-1}) + \sigma_1(|t-s|; \|x\| \vee \|y\|) \tag{61}$$

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 [32], 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 [32] 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*

$$a + \langle A^*p, x\rangle + H(t,x,p) \geq -\gamma_{R,\delta}(\lambda,\epsilon,\beta) \tag{62}$$

for β *sufficiently small.*

(b) *If in addition* $H(s,y,q)$ *is weakly sequentially 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* $(w_{\lambda,\epsilon,\beta})_t(t_n,x_n) \to a$, $Dw_{\lambda,\epsilon,\beta}(t_n,x_n) \rightharpoonup p$ *for some* $(t_n,x_n) \to (t,x)$, *where* (t_n,x_n) *are points of Fréchet differentiability of* $w_{\lambda,\epsilon,\beta}$, *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 sequentially 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 sequentially upper-semicontinuous in q. Therefore in (b) the Hamiltonian H is assumed to be weakly sequentially 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

$$\overline{\mathrm{conv}}\{(a,p) : v_t(s_n,z_n) \to a, Dv(s_n,z_n) \rightharpoonup p, s_n \to s, z_n \to z\}, \tag{63}$$

where (s_n, z_n) above are points of Fréchet differentiability of v (see [16], page 522).

(a) *Step 1.* Denote $v = w^{\lambda,\epsilon,\beta}$. Let $(t,x) \in (\delta, T-\delta) \times B_R$ be a point of differentiability of v. It is easy to see, taking test functions of the form $\psi(s,y) = \alpha(s-t_0)^2 + \beta\|y-x_0\|^2$, where $\alpha, \beta > 0, t_0 \in (0,T), x_0 \in D(A^*)$, and applying perturbed optimization technique as in the proof of Lemma 4.3, that the set $\mathcal{E}^+ = \{(s,y) : E^{1,+}v(s,y) \neq \emptyset\}$ is dense in $(0,T) \times \mathcal{H}$. Let $(t_n, x_n) \in \mathcal{E}^+$ be such that $(t_n, x_n) \to (t,x)$. By the B-semisemiconvexity, v it is differentiable at (t_n, x_n) and, from Lemma 4.2 (iii), $|v_t(t_{n_k}, x_{n_k})| \leq M_{R,\epsilon,\beta}$ and $Dv(t_n, x_n) = Bq_n$ for some $\|q_n\| \leq M_{R,\epsilon,\beta}$. Therefore we can extract a subsequence n_k such that $v_t(t_{n_k}, x_{n_k}) \to a$ and $q_{n_k} \rightharpoonup q$ for some $(a,q) \in (\delta, T-\delta) \times B_R$. Then $Dv(t_{n_k}, x_{n_k}) = Bq_{n_k} \rightharpoonup Bq$. Moreover $a = v_t(t,x), Bq = Dv(t,x)$ by (63). Therefore

$$A^*Dv(t_{n_k}, x_{n_k}) = A^*Bq_{n_k} \rightharpoonup A^*Bq = A^*Dv(t,x).$$

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

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

It now follows from Lemma 4.3 that

$$v_t(t,x) + \langle A^*Dv(t,x), x\rangle + H(t,x,Dv(t,x)) \geq -\gamma_{R,\delta}(\lambda,\epsilon,\beta). \tag{64}$$

Step 2. If (a,p) is such that $v_t(t_n,x_n) \to a, Dv(t_n,x_n) \rightharpoonup p$ with $(t_n,x_n) \to (t,x)$, the claim follows from (64) and the same arguments as in *Step 1.*

Step 3. Let (a,p) be a generic point of $D^{1,-}v(t,x)$, i.e. $p = \lim_{n\to\infty} \sum_{i=1}^n \lambda_i^n Bq_i^n$, where $\sum_{i=1}^n \lambda_i^n = 1, \|q_i^n\| \leq M_{R,\epsilon,\beta}$, 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

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

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) Denote $v = w_{\lambda,\epsilon,\beta}$. If (t,x) is a point of differentiability of v then, similarly as in *Step 1* of (a), we can choose a sequence $(t_n,x_n) \in \mathcal{E}^- = \{(s,y) : E^{1,-}v(s,y) \neq \emptyset\}$ such that $(t_n,x_n) \to (t,x)$, $v_t(t_n,x_n) \to v_t(t,x)$, $Dv(t_n,x_n) \rightharpoonup Dv(t,x)$ and $A^*Dv(t_n,x_n) \rightharpoonup A^*Dv(t,x)$. Lemma 4.3 and the weak sequential lower-semicontinuity of the Hamiltonian now yield that

$$v_t(t,x) + \langle A^*Dv(t,x), x\rangle + H(t,x,Dv(t,x)) \leq \gamma_{R,\delta}(\lambda,\epsilon,\beta). \tag{65}$$

The general case when (a,p) is such that $v_t(t_n,x_n) \to a, Dv(t_n,x_n) \rightharpoonup p$ for some $(t_n,x_n) \to (t,x)$ follows from (65) and the same arguments. $\square$

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*

$$|w(t,x) - w(s,y)| \leq \sigma_R(|t-s| + \|x-y\|_{-1}) \quad \textit{for } t,s \in (0,T), \|x\|, \|y\| \leq R. \tag{66}$$

Then:

(a) *If w is a viscosity subsolution of* (10) *satisfying $w(t,x) \leq C(1+\|x\|^k)$ for some $k \geq 0$ then for every $0 < t < t+h < T$, $x \in \mathcal{H}$*

$$w(t,x) \leq \inf_{u(\cdot)\in\mathcal{U}[t,T]} \left\{ \int_t^{t+h} L(s,x(s),u(s))\mathrm{d}s + w(t+h, x(t+h)) \right\}. \tag{67}$$

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

$$\|x_{t,x}(s_2) - x_{t,x}(s_1)\| \leq \omega_{t,x}(s_2 - s_1) \tag{68}$$

for all $t \leq s_1 \leq s_2 \leq T$ and all $u(\cdot) \in \mathcal{U}[t,T]$, where $x_{t,x}(\cdot)$ is the solution of (7)*. If w is a viscosity supersolution of* (10) *satisfying $w(t,x) \geq -C(1+\|x\|^k)$ for some $k \geq 0$ 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*

$$w(t,x) \geq \int_t^{t+h} L(s,x(s),u_\nu(s))\mathrm{d}s + w(t+h, x(t+h)) - \nu. \tag{69}$$

In particular we obtain the superoptimality principle

$$w(t,x) \geq \inf_{u(\cdot)\in\mathcal{U}[t,T]} \left\{ \int_t^{t+h} L(s,x(s),u(s))\mathrm{d}s + w(t+h, x(t+h)) \right\} \tag{70}$$

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 [57] (that treats the finite dimensional case).

Step 1. Let $n \geq 1$ and let δ be such that $0 < \delta < t < t+h < T-\delta$. We approximate w by $w_{\lambda,\epsilon,\beta}$ with $m > k$. We will always take β small enough so that the conclusion of Lemma 4.5(b) is satisfied on $(\delta, T-\delta) \times B_R$. We notice that for any $u(\cdot)$ if $x_{t,x}(\cdot)$ is the solution of (7) then

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

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 B-semiconcave. Then we choose $u_1 \in U$ such that

$$a + \langle A^*p, x\rangle + \langle p, b(t,x,u_1)\rangle + L(t,x,u_1) \leq \gamma_{R,\delta}(\lambda,\epsilon,\beta) + \frac{1}{n^2}. \tag{71}$$

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

$$w_{\lambda,\epsilon,\beta}(s,y) \leq w_{\lambda,\epsilon,\beta}(t,x) + a(s-t) + \langle p, y-x\rangle + \frac{\|x-y\|_{-1}^2}{2\epsilon} + \frac{(t-s)^2}{2\beta}. \tag{72}$$

But the right hand side of the above inequality is a test1 function so if $s \geq t$ and $x(s) = x_{t,x}(s)$ with constant control $u(s) = u_1$, we can use (13) and write

$$\begin{aligned}
&\left|\frac{a(s-t) + \langle p, x(s)-x\rangle + \frac{\|x(s)-x\|_{-1}^2}{2\epsilon} + \frac{(s-t)^2}{2\beta}}{s-t} - (a + \langle p, b(t,x,u_1)\rangle + \langle A^*p, x\rangle)\right| \\
&\leq \frac{|t-s|}{2\beta} + \left|\frac{\int_t^s \langle A^*p, x(r)-x\rangle\,\mathrm{d}r}{s-t}\right| + \left|\frac{\int_t^s \langle p, b(r,x(r),u_1) - b(t,x,u_1)\rangle\,\mathrm{d}r}{s-t}\right| \\
&\qquad + \left|\frac{\int_t^s \langle A^*B(x(r)-x), x(r)\rangle\,\mathrm{d}r}{\epsilon(s-t)}\right| + \left|\frac{\int_t^s \langle B(x(r)-x), b(r,x(r),u_1)\rangle\,\mathrm{d}r}{\epsilon(s-t)}\right| \\
&\leq \omega'_{t,x}\big(|s-t| + \sup_{t\leq r\leq s}\|x(r)-x\|\big) \leq \tilde{\omega}_{t,x}(s-t)
\end{aligned} \tag{73}$$

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 (71), (72) and (73) to estimate

$$\begin{aligned}
&\frac{w_{\lambda,\epsilon,\beta}(t+\frac{h}{n}, x(t+\frac{h}{n})) - w_{\lambda,\epsilon,\beta}(t,x)}{h/n} \\
&\leq \tilde{\omega}_{t,x}\left(\frac{h}{n}\right) + \gamma_{R,\delta}(\lambda,\epsilon,\beta) + \frac{1}{n^2} - L(t,x,u_1).
\end{aligned} \tag{74}$$

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 aobtainbtain u_2 satisfying (74) 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 (7) with the control $u^{(n)}$ we have

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

We remind that (68) is needed here to guarantee that $\sup_{t_{i-1}\leq r\leq t_i}\|x(r) - x(t_{i-1})\|$ is independent of u_i and $x(t_{i-1})$ and depends only on x and t. We then easily obtain

$$\begin{aligned}
&w_{\lambda,\epsilon,\beta}(t+h, x(t+h)) - w_{\lambda,\epsilon,\beta}(t,x) \\
&\leq \tilde{\omega}_{t,x}\left(\frac{h}{n}\right)h + \gamma_{R,\delta}(\lambda,\epsilon,\beta)h + \frac{h}{n^2} - \int_t^{t+h} L(r, x(r), u^{(n)}(r))\mathrm{d}r + \tilde{\omega}'_{t,x}\left(\frac{h}{n}\right)h
\end{aligned} \tag{75}$$

for some modulus $\tilde{\omega}'_{t,x}$, where we have used Hypothesis 4.1 and (68) to estimate how the sum converges to the integral. It now remains to observe that it follows from (66) that

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

where the modulus $\tilde{\sigma}_R$ can be explicitly calculated from σ_R. We thus obtain

$$w(t+h, x(t+h)) + \int_t^{t+h} L(r, x(r), u^{(n)}(r))\mathrm{d}r$$
$$\leq w(t,x) + 2\tilde{\sigma}_R(\lambda+\epsilon+\beta) + \tilde{\omega}_{t,x}\left(\frac{h}{n}\right)h + \gamma_{R,\delta}(\lambda,\epsilon,\beta)h + \frac{h}{n^2} + \tilde{\omega}'_{t,x}\left(\frac{h}{n}\right)h.$$

Now, given $\nu > 0$, we first have to use (49) and take λ, ϵ, β small so that

$$\gamma_{R,\delta}(\lambda,\epsilon,\beta)h + 2\tilde{\sigma}_R(\lambda+\epsilon+\beta) \leq \frac{\nu}{2},$$

and then we need to take n such that

$$\tilde{\omega}_{t,x}\left(\frac{h}{n}\right)h + \frac{h}{n^2} + \tilde{\omega}'_{t,x}\left(\frac{h}{n}\right)h \leq \frac{\nu}{2}$$

to arrive at (69) and consequently (70).

Finally we notice that if w is the value function V then, in virtue of the dynamic programming principle, the control $u^{(n)}$ constructed above is ν-optimal on the interval $[t, t+h]$. If h is such that $t+h$ is sufficiently close to T then obviously this construction, together with (66), gives us a nearly optimal piecewise constant control on the whole interval $[t,T]$ whose value on $[t+h,T]$ can be arbitrary. □

Condition (68) is restrictive, however it seems necessary to obtain uniform estimates on the error terms $\tilde{\omega}_{t,x}\left(\frac{h}{n}\right)$ and $\tilde{\omega}'_{t,x}\left(\frac{h}{n}\right)$ in (74) and (75). We present below an example when it is satisfied. In general one may expect it to hold when the semigroup e^{tA} has some regularizing properties.

Example 4.8. Condition (68) holds for example if $A = A^*$, it generates a differentiable semigroup, and $\|Ae^{tA}\| \leq C/t^\delta$ 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

$$\|(-A+I)^{\frac{1}{2}}x(s)\| \leq \|(-A+I)^{\frac{1}{2}}e^{(s-t)A}x\| + \int_t^s \|(-A+I)^{\frac{1}{2}}e^{(s-\tau)A}b(\tau,x(\tau),u(\tau))\|d\tau.$$

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

$$\|(-A+I)^{\frac{1}{2}}e^{\tau A}y\|^2 \leq \|(-A+I)e^{\tau A}y\|\,\|y\| \leq \frac{C_1}{\tau^\delta}\|y\|^2.$$

This yields

$$\|(-A+I)^{\frac{1}{2}}e^{\tau A}\| \leq \frac{\sqrt{C_1}}{\tau^{\frac{\delta}{2}}}$$

and therefore

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

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

$$\begin{aligned} e^{(s_2-s_1)A}x(s_1) - x(s_1) &= \int_0^{s_2-s_1} Ae^{sA}x(s_1)ds \\ &= \int_0^{s_2-s_1} (-A+I)^{\frac{1}{2}}e^{sA}(-A+I)^{\frac{1}{2}}x(s_1)ds - \int_0^{s_2-s_1} e^{sA}x(s_1)ds \end{aligned}$$

and thus

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

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

$$\begin{aligned} \|x(s_2) - x(s_1)\| &\le 2\sigma(\epsilon) + \|x(s_2) - x(\bar{s})\| \\ &\le 2\sigma(\epsilon) + \|e^{(s_2-\bar{s})A}x(s_1) - x(\bar{s})\| + \int_{\bar{s}}^{s_2} \|e^{(s_2-\tau)A}b(\tau, x(\tau), u(\tau))\|d\tau \\ &\le 2\sigma(\epsilon) + \sigma_\epsilon(s_2-s_1) + C_4(s_2-s_1) \end{aligned} \tag{76}$$

for some constant C_4 independent of $u(\cdot)$. Therefore (68) is satisfied with the modulus

$$\omega_{t,x}(\tau) = \inf_{0<\epsilon<T-t} \{2\sigma(\epsilon) + \sigma_\epsilon(\tau) + C_4\tau\}.$$

References

[1] S. Aniţa: Analysis and Control of Age-Dependent Population Dynamics, Kluwer, Dordrecht (2001).

[2] V. Barbu: Optimal feedback controls for a class of nonlinear distributed parameter systems, SIAM J. Control Optimization 21(6) (1983) 871–894.

[3] V. Barbu: Hamilton-Jacobi equations and nonlinear control problems, J. Math. Anal. Appl. 120(2) (1986) 494–509.

[4] V. Barbu: Analysis and Control of Nonlinear Infinite-Dimensional Systems, Mathematics in Science and Engineering 190, Academic Press, Boston (1993).

[5] V. Barbu: Approximation Hamilton-Jacobi equations and suboptimal feedback controllers, in: Nonlinear Analysis and Applications (Warsaw, 1994), N. Kenmochi et al. (ed.), GAKUTO Int. Ser., Math. Sci. Appl. 7, Gakkōtosho, Tokyo (1996) 27–47.

[6] V. Barbu, E. N. Barron, R. Jensen: The necessary conditions for optimal control in Hilbert spaces, J. Math. Anal. Appl. 133(1) (1988) 151–162.

[7] V. Barbu, G. Da Prato: Hamilton-Jacobi Equations in Hilbert Spaces, Research Notes in Mathematics 86, Pitman, Boston (1983).

[8] V. Barbu, G. Da Prato: Hamilton-Jacobi equations and synthesis of nonlinear control processes in Hilbert spaces, J. Differ. Equations 48(3) (1983) 350–372.

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

[10] V. Barbu, G. Da Prato: A note on a Hamilton-Jacobi equation in Hilbert space, Nonlinear Anal. 9(12) (1985) 1337–1345.

[11] V. Barbu, G. Da Prato, C. Popa: Existence and uniqueness of the dynamic programming equation in Hilbert space, Nonlinear Anal. 7(3) (1983) 283–299.

[12] M. Bardi, I. Capuzzo-Dolcetta: Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations, Systems & Control: Foundations & Applications, Birkhäuser, Boston (1997).

[13] E. Barucci, F. Gozzi: Investment in a vintage capital model, Res. Econ. 52 (1998) 159–188.

[14] E. Barucci, F. Gozzi: Technology adoption and accumulation in a vintage capital model, J. Econ. 74(1) (2001) 1–30.

[15] A. Bensoussan, G. Da Prato, M. C. Delfour, S. K. Mitter: Representation and Control of Infinite Dimensional Systems, Birkhäuser, Boston (1992).

[16] J. M. Borwein, D. Preiss: A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, Trans. Amer. Math. Soc. 303(2) (1987) 517–527.

[17] P. Cannarsa: Regularity properties of solutions to Hamilton-Jacobi equations in infinite dimensions and nonlinear optimal control, Differ. Integral Equ. 2(4) (1989) 479–493.

[18] P. Cannarsa, O. Carjă: On the Bellman equation for the minimum time problem in infinite dimensions, SIAM J. Control Optimization 43(2) (2004) 532–548

[19] P. Cannarsa, G. Da Prato: Nonlinear optimal control with infinite horizon for distributed parameter systems and stationary Hamilton-Jacobi equations, SIAM J. Control Optimization 27(4) (1989) 861–875.

[20] P. Cannarsa, G. Da Prato: Some results on non-linear optimal control problems and Hamilton-Jacobi equations in infinite dimensions, J. Funct. Anal. 90(1) (1990) 27–47.

[21] P. Cannarsa, G. Di Blasio: A direct approach to infinite-dimensional Hamilton-Jacobi equations and applications to convex control with state constraints, Differ. Integral Equ. 8(2) (1995) 225–246.

[22] P. Cannarsa, F. Gozzi, H. M. Soner: A dynamic programming approach to nonlinear boundary control problems of parabolic type, J. Funct. Anal. 117(1) (1993) 25–61.

[23] P. Cannarsa, H. Frankowska: Value function and optimality conditions for semilinear control problems, Appl. Math. Optimization 26(2) (1992) 139–169.

[24] P. Cannarsa, H. Frankowska: Value function and optimality condition for semilinear control problems. II: parabolic case, Appl. Math. Optimization 33(1) (1996) 1–33.

[25] P. Cannarsa, C. Sinestrari: Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control, Progress in Nonlinear Differential Equations and their Applications 58, Birkhäuser, Boston (2004).

[26] P. Cannarsa, M. E. Tessitore: Optimality conditions for boundary control problems of parabolic type, in: Control and Estimation of Distributed Parameter Systems: Nonlinear Phenomena (Vorau, 1993), W. Desch et al. (ed.), Int. Ser. Numer. Math. 118, Birkhäuser, Basel (1994) 79–96.

[27] P. Cannarsa, M. E. Tessitore: Infinite dimensional Hamilton–Jacobi equations and Dirichlet boundary control problems of parabolic type, SIAM J. Control Optimization 34(6) (1996) 1831–1847.

[28] F. H. Clarke, Yu. S. Ledyaev, E. D. Sontag, A. I. Subbotin: Asymptotic controllability implies feedback stabilization, IEEE Trans. Autom. Control 42(10) (1997) 1394–1407.

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

[30] M. G. Crandall, H. Ishii, P. L. Lions: User's guide to viscosity solutions of second order partial differential equations, Bull. Amer. Math. Soc., New Ser. 27(1) (1992) 1–67.

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

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

[33] M. G. Crandall, P. L. Lions: Hamilton-Jacobi equations in infinite dimensions. VI. Nonlinear A and Tataru's method refined, in: Evolution Equations, Control Theory, and Biomathematics (Han sur Lesse, 1991), P. Clément et al. (ed.), Lect. Notes Pure Appl. Math. 155, Dekker, New York (1993) 51–89.

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

[35] G. Di Blasio: Global solutions for a class of Hamilton-Jacobi equations in Hilbert spaces, Numer. Funct. Anal. Optimization 8(3–4) (1985/86) 261–300.

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

[37] G. Fabbri: A viscosity solution approach to the infinite-dimensional HJB equation related to a boundary control problem in a transport equation, SIAM J. Control Optimization 47(2) (2008) 1022–1052.

[38] S. Faggian, F. Gozzi: On the dynamic programming approach for optimal control problems of PDE's with age structure, Math. Popul. Stud. 11(3–4) (2004) 233–270.

[39] S. Federico, B. Goldys, F. Gozzi: HJB Equations for the optimal control of differential equations with delays and state constraints II: Optimal feedbacks and approximations, arXiv:0905.1825 (2009).

[40] G. Feichtinger, A. Prskawetz, V. M. Veliov: Age-structured optimal control in population economics, Math. Popul. Stud. 65(4) (2004) 373–387.

[41] W. Fleming, R. W. Rishel: Deterministic and Stochastic Optimal Control, Applications of Mathematics 1, Springer, Berlin (1975).

[42] W. Fleming, H. M. Soner: Controlled Markov Processes and Viscosity Solutions, 2nd Ed., Stochastic Modelling and Applied Probability 25, Springer, New York (2006).

[43] F. Gozzi, S. S. Sritharan, A. Święch: Viscosity solutions of dynamic-programming equations for the optimal control of the two-dimensional Navier-Stokes equations, Arch. Ration. Mech. Anal. 163(4) (2002) 295–327.

[44] M. Iannelli: Mathematical Theory of Age-Structured Population Dynamics, Giardini, Pisa (1995).

[45] M. Iannelli, M. Martcheva, F. A. Milner: Gender-Structured Population Modeling, SIAM, Philadelphia (2005).

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

[47] H. Ishii, S. Koike: On ϵ-optimal controls for state constraint problems, Ann. Inst. H. Poincaré, Anal. Non Linéaire 17(4) (2000) 473–502.

[48] D. Kelome: Viscosity Solutions of Second Order Equations in a Separable Hilbert Space and Applications to Stochastic Optimal Control, Ph.D. Thesis, Georgia Institute of Technology, Atlanta (2002).

[49] D. Kelome, A. Święch: Perron's method and the method of relaxed limits for "unbounded" PDE in Hilbert spaces, Stud. Math. 176(3) (2006) 249–277.

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

[51] X. Y. Li, J. M. Yong: Optimal Control Theory for Infinite-Dimensional Systems, Birkhäuser, Basel (1994).

[52] P. D. Loewen: Optimal control via Nonsmooth Analysis, CRM Proceedings & Lecture Notes 2, AMS, Providence (1993).

[53] M. Renardy: Polar decomposition of positive operators and problem of Crandall and Lions, Appl. Anal. 57(3–4) (1995) 383–385.

[54] S. Saks: Theory of the Integral, 2nd rev. Ed., Dover, New York (1964).

[55] K. Shimano: A class of Hamilton-Jacobi equations with unbounded coefficients in Hilbert spaces, Appl. Math. Optimization 45(1) (2002) 75–98.

[56] S. S. Sritharan: Dynamic programming of the Navier-Stokes equations, Syst. Control Lett. 16(4) (1991) 299–307.

[57] A. Święch: Sub- and superoptimality principles of dynamic programming revisited, Nonlinear Anal. 26(8) (1996) 1429–1436.

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

[59] D. Tataru: Viscosity solutions for the dynamic programming equation, Appl. Math. Optimization 25 (1992) 109–126.

[60] R. Vinter: Optimal Control, Systems & Control: Foundations & Applications, Birkhäuser, Boston (2000).

[61] J. Yong, X. Y. Zhou: Stochastic Controls. Hamiltonian Systems and HJB Equations, Springer, New York (1999).

[62] X. Y. Zhou: Verification theorems within the framework of viscosity solutions, J. Math. Anal. Appl. 177(1) (1993) 208–225.