Dynamic programming of the Navier–Stokes equations

S.S. Srinathan
Department of Aerospace Engineering, University of Southern California, Los Angeles, CA 90089-1191, USA

Received 3 August 1990
Revised 8 November 1990

Abstract A class of optimal control problems in viscous flow is studied. Main results are the Pontryagin maximum principle and the verification theorem for the Hamilton–Jacobi–Bellman equation.

Keywords Optimal control, dynamic programming, Hamilton–Jacobi theory, viscosity solutions, Pontryagin maximum principle.

1. Introduction

Optimal control theory of viscous fluid motion has important applications in engineering science. We will report on certain fundamental results concerning the mathematical resolution of the viscous flow control problem. A unified mathematical formulation covering several control problems is presented here. In [12], a mathematical theory for the optimal acceleration of an obstacle in incompressible viscous fluid is presented. The main result there was the existence theorem for optimal control. The present work can be considered as a sequel to the above paper. However, the general formulation we provide here also covers a number of other problems such as the control of flow in bounded containers and minimization of dissipation in channel flows. We will begin with a general nonlinear evolution problem with unbounded nonlinearity. Theorems concerning the mathematical structure of various operators involved in this equation can be found in [12, 11]. A solvability theorem is provided for this general system. A general cost functional which covers specific practical problems is defined. The existence theorem provided here establishes the existence of an optimal control which corresponds to the minimum value of the cost. The value function is defined as the minimum value of the cost. It is shown that the value function is locally Lipschitz. An important result proved in this context is that the value function is a viscosity solution (in the sense of Crandall and Lions [6, 5]) of the Hamilton–Jacobi–Bellman equation associated with our control problem. We then use the method in [4, 2] to derive the Pontryagin maximum principle for the Navier–Stokes equations. Finally, we establish the verification theorem for the Hamilton–Jacobi–Bellman equation. This theorem provides the mathematical resolution of the feedback control problem for the Navier–Stokes equations. The present work can also be considered as a generalization of the works of Barbu and his colleagues [2, 1, 3] which deal with evolution systems with bounded nonlinearities.

2. Governing equations and optimal control formulation

In this section we will show that several viscous flow control problems can be unified to a problem of the following type in a Hilbert space $H$:

\begin{align}
\partial_t y + \mathcal{A} y + \mathcal{N}(y) &= \mathcal{B} \Theta, \quad t \in (\tau, T], \\
y(\tau) &= \xi \in D(\mathcal{A}).
\end{align}

(1a)

(1b)
where \( D(\mathcal{A}) \) is the domain of \( \mathcal{A} \). Here \( \mathcal{A} \) is a positive and self adjoint operator. Moreover, \(-\mathcal{A}\) generates a holomorphic semigroup. The nonlinear map \( \mathcal{N}(\cdot): D(\mathcal{A}^{\alpha}) \to D(\mathcal{A}^{-\beta}) \), \( \frac{1}{2} < \alpha \leq 1, 0 \leq \beta \) such that \( \alpha + \beta < 1 \), is a locally bounded Fréchet differentiable operator. The linear operator \( \mathcal{B} \in L^2(\tau, T); L^2(\tau, T; H) \) is continuous. The task is to find the optimal control \( \Theta \in L^2(0, T) \) such that the cost functional

\[
\mathcal{C}(\tau, \xi, \Theta) = \phi_0(y(T, \tau, \xi, \Theta)) + \int_{\tau}^{T} (f(y(t, \tau, \xi, \Theta) + h(\Theta))) \, dt \rightarrow \inf
\]

where \( y \) is the solution of (1). Here, \( \phi_0(\cdot): H \to R^+ \) and \( h(\cdot): R \to R^+ \) are quadratic functionals. The functional \( f(\cdot): D(\mathcal{A}^{1/2}) \to R^+ \) can be written as

\[
f(\xi) = g(\mathcal{A}^{1/2}\xi), \quad \forall \xi \in D(\mathcal{A}^{1/2}),
\]

where \( g(\cdot) \) is again a quadratic functional. The value function \( \mathcal{V}(\cdot): [0, T] \times H \to R^+ \) is defined as

\[
\mathcal{V}(\tau, \xi) = \min_{\Theta \in L^2(0, T)} \mathcal{C}(\tau, \xi, \Theta)
\]

We will associate to the above control problem the Hamilton-Jacobi-Bellman equation:

\[
\begin{align*}
\partial_t \mathcal{V} - h^* (-\mathcal{B} \cdot \partial_x \mathcal{V}) - (\mathcal{A} \mathcal{V} + \mathcal{N}(\xi), \partial_x \mathcal{V})_H + f(\xi) &= 0, \quad t \in [\tau, T], \quad (2a) \\
\mathcal{V}(T, \xi) &= \phi_0(\xi), \quad \forall \xi \in H. \quad (2b)
\end{align*}
\]

Here, \( \mathcal{B}^* \in L^2(L^1(\tau, T); L^1(\tau, T)) \) is the adjoint of \( \mathcal{B} \) and \( h^*(\cdot): R \to R \) is the conjugate of \( h \).

\[
h^*(p) = \sup_{\Theta \in R} [p \cdot \Theta - h(\Theta)].
\]

Before stating the central theorems of this paper, we will derive equation (1) in detail for a specific example. Other more involved examples can be found in [12,7].

**Example.** Let us consider the problem of driving the velocity field of a viscous incompressible fluid inside a bounded container, to a desired field by suction and blowing on the boundary. Let \( \Omega \subset R^n, n = 2 \) or 3, be an arbitrary simply connected bounded open set with class \( C^2 \) boundary \( \partial \Omega \). Let \( (u, p): \Omega \times [0, T] \to R^n \times R \) be velocity and pressure fields respectively. The governing equations are

\[
\begin{align*}
\partial_t u + (u \cdot \nabla) u &= -\nabla p + \nu\Delta u \quad \text{in } \Omega \times (0, T), \\
\nabla u &= 0 \quad \text{in } \Omega \times (0, T), \\
u \cdot u(x, t) &= U(t)g(x) \quad \text{for } (x, t) \in \partial \Omega \times [0, T],
\end{align*}
\]

with \( \int_{\partial \Omega} g \cdot n \, dS = 0 \) and \( u(x, 0) = u_0(x) \), for \( x \in \Omega \). Here \( \nu \) is the coefficient of kinematic viscosity \( U(\cdot): [0, T] \to R \) and \( g(\cdot): \partial \Omega \to R^n \) represent respectively the temporal dependence and spatial distribution of the fluid velocity at the boundary. The task is to find \( \partial U \) such that the cost functional

\[
\int_0^T \int_{\Omega} |
\nabla(u - u^d)|^2 \, dx \, dt + \lambda \int_0^T (\partial U - \partial U^d)^2 \, dt \rightarrow \inf.
\]

Here \( u^d(\cdot, \cdot): \Omega \times [0, T] \to R^n \) is a smooth desired field, \( U^d(\cdot): [0, T] \to R \) is a smooth nominal forcing speed and \( \lambda > 0 \) is some given number. In order to handle the nonzero boundary condition, we will use the well known Leray–Hopf cut-off function method (see for example [13,11]).

**Proposition 1.** Let the flux distribution satisfy \( g \in H^{1/2}(\partial \Omega) \) and \( \int_{\partial \Omega} g \cdot n \, dS = 0 \). Then \( \forall \delta > 0, \exists w^\delta \in H^2(\Omega) \) such that \( \nabla \cdot w^\delta = 0; \ w^\delta \mid_{\partial \Omega} = g \) and

\[
\left| \sum_{i,j=1}^{n} \int_{\Omega} \frac{\partial w^\delta_j}{\partial x_i} z_j \, dx \right| \leq \delta \int_{\Omega} |\nabla z|^2 \, dx, \quad \forall z \in H^1_0(\Omega) \text{ such that } \nabla \cdot z = 0.
\]
We then set \( u(x, t) = v(x, t) + U(t)w_S(x) \) to get

\[
\begin{align*}
\partial_t v + (v \cdot \nabla)v + U(t)(v \cdot \nabla)w_S + U(t)(w_S \cdot \nabla)v &= -\nabla p + \nu \Delta v + vU(t)\Delta w_S - U(t)^2 (w_S \cdot \nabla)v - \partial_t U w_S & \text{in } \Omega \times (0, T), \\
\nabla \cdot v &= 0 & \text{in } \Omega \times (0, T), \\
v(x, t) &= 0, \quad (x, t) \in \partial \Omega \times [0, T] & \text{and} \quad v(x, 0) = u_0(x) - U(0)w_S(x), \quad x \in \Omega.
\end{align*}
\]

(3)

Let us define \( f(\Omega) = \{ \phi : \Omega \to \mathbb{R}; \phi \in C^\infty(\Omega) \text{ and } \nabla \cdot \phi = 0 \} \) and let \( H(\Omega), V_1(\Omega) \) be respectively the completions of \( f(\Omega) \) in the norms of \( L^2(\Omega) \) and \( H'(\Omega) \). Applying the Hodge orthogonal projection \( P_\mu : L^2(\Omega) \to H(\Omega) \) to the system (3) we get

\[
\begin{align*}
\partial_t v + Av + B(v, v) + UBl(v) &= f_U + f_U^2 + f_\mu, & v(t) &= v(0) \\
\end{align*}
\]

(4)

Here \( A \) is the Stokes operator which is positive definite, selfadjoint and \(-A \) generates a holomorphic semigroup. The operators \( B(\cdot, \cdot) \) and \( B_l(\cdot) \) are standard in Navier–Stokes theory \([13,11]\). Note that

\[
L^2(\Omega) \ni f \to L^2(\Omega);
\]

\[
\text{and } V_1 = D(A^{1/2}).
\]

Let us now set

\[
y = \left( \begin{array}{c} v \\ u \end{array} \right) \quad \text{and} \quad \Theta = \partial_t U.
\]

Then (4) reduces to our general evolution equation (1) with

\[
\begin{align*}
\mathcal{L} = & \begin{pmatrix} v A & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathcal{N}(y) = \begin{pmatrix} B(v, v) + UB_l(v) - f_U - f_U^2 \\ -U \end{pmatrix} \quad \text{and} \quad \mathcal{A} = \begin{pmatrix} f_U \\ 1 \end{pmatrix}.
\end{align*}
\]

Moreover comparing with the general cost functional we get \( \phi_0 = 0, \quad f(y) = \| \nabla(v + Uw_S - u^d) \|_{L^2(\Omega)}^2 \) and \( h(\Theta) = \lambda(\Theta - \partial_t U^d)^2 \)

The working function spaces in this case are \( H = H(\Omega) \times \mathbb{R}, \quad V_1 = V_1(\Omega) \times \mathbb{R} \) and \( D(\mathcal{A}) = D(A) \times \mathbb{R}. \)

Note that \( \mathcal{A} \) is positive, selfadjoint and \(-\mathcal{A} \) generates a holomorphic semigroup. Moreover \( V_1 = D(A^{1/2}). \)

From the properties of the maps \( B(\cdot, \cdot) \) and \( B_l(\cdot) \) it is easy to show that \( \mathcal{N}(\cdot) : D(\mathcal{A}^{1/2}) \to H \) for \( \alpha > \frac{1}{2} \) in three dimensions and for \( \alpha > \frac{1}{4} \) for two dimensions. The unique solvability theorem stated below can be deduced from the results in \([9,10,12,8,14]\) for various geometries

**Theorem 1.** Let \( \Theta \in L^2(\tau, T) \) and \( \xi \in H \). Then for the system (1) there exists a unique solution \( y \in L^2(\tau, T; D(\mathcal{A}^{1/2})) \cap C([\tau, T]; H) \). Moreover, if \( \xi \in D(\mathcal{A}^{1/2}) \), then \( y \in L^2(\tau, T; D(\mathcal{A}^{1/2})) \) and \( \partial_t y \in L^2(\tau, T; H) \).

**Theorem 2.** Existence theorem: Let \( \tau \in [0, T] \) and \( \xi \in H \) be given. Then, there exists an optimal control \( \hat{\theta} \in L^2(\tau, T) \) such that \( (\hat{y}(t, \tau; \xi; \hat{\theta}), \hat{\theta}) \) solves (1) and

\[
\mathcal{L}(\tau, \xi, \hat{\theta}) = \inf_{\Theta \in L^2(0, T)} \mathcal{L}(\tau, \xi, \Theta)
\]

Proof of this theorem for the case of exterior hydrodynamics is given in [12]. Other cases can be proven by similar arguments. Let us now state the Pontryagin maximum principle for hydrodynamics:

**Theorem 3.** Let \( \hat{\Theta} \in L^2(\tau, T) \) be the optimal control and \( \hat{y}(t, \tau, \xi; \hat{\Theta}) \) be the optimal trajectory corresponding to the initial data \( \xi \in D(\mathcal{A}) \) at \( t = \tau \). Then, there exists \( p \in C([\tau, T]; H) \) such that

\[
-\partial_t p + \mathcal{A} p + \left( D\mathcal{N} \right)^*(\hat{y}) = \left( Df \right)(\hat{y}) \quad \text{and} \quad p(T) = [D\phi_0](\hat{y}(T)).
\]

(5)

Moreover,

\[
h^*(-\mathcal{B}^* p(t)) = -\left( \mathcal{B}^* p(t) \cdot \hat{\Theta}(t) \right) - h(\hat{\Theta}(t)) \quad \text{for } t \text{ a.e. in } [\tau, T]
\]
\[ \hat{\Theta}(t) = \nabla h^*(-A^*p(t)), \quad t \in [\tau, T]. \]

Here \([D',\mathcal{F}]^*(\hat{y})\) is the adjoint of the Fréchet derivative of \(\mathcal{F}(\cdot)\) at \(\hat{y}\). The proof of this theorem will be given in Section 4. We will then prove in Section 5 the verification theorem.

**Theorem 4.** The value function \(\mathcal{V} \in C([0, T] \times \mathcal{H})\). For each \(t \in [\tau, T]\), \(\mathcal{V}(t, \cdot)\) is Lipschitz in \(\xi \in \mathcal{H}\) and for each \(\xi \in D(\mathcal{A})\), \(\mathcal{V}(\cdot, \xi)\) is absolutely continuous in \(t \in [\tau, T]\). The super differential \(\partial_{\xi}^+ \mathcal{V}(t, \xi)\) is nonempty \(\forall (\tau, T)\), \(\forall \xi \in D(\mathcal{A})\),

\[
\partial_{\xi}^+ \mathcal{V}(t, \xi) = \phi_0(\xi), \quad \forall \xi \in \mathcal{H}.
\]

Moreover the optimal control \(\hat{\Theta}\) is given by the feedback relation

\[
\hat{\Theta}(t) = \nabla h^*(-A^*p(t)) \quad \text{for some} \quad p(t) \in \partial_{\xi}^+ \mathcal{V}(t, \xi), \quad \forall t \in [\tau, T].
\]

The main result of this section is that the value function \(\mathcal{V}(\cdot, \cdot) \in [0, T] \times \mathcal{H} \rightarrow \mathbb{R}\) is a viscosity solution to the Hamilton–Jacobi–Bellman equation (2). We will begin with a series of lemmas concerning the continuity properties of \(\varphi\) and \(\mathcal{V}\). Detailed proofs of these results can be found in [7].

**Lemma 1.** The solution \(\varphi\) of (1) satisfies the a priori estimate

\[
\| \varphi(t) \|_{H}^2 + \int_{\tau}^{t} \| \mathcal{A}^{1/2}_{e}(\varphi(s)) \|_{H}^2 \, ds \leq C_1 \left( \| \varphi \|_{H}, \| \Theta \|_{L^2(\tau, T)} \right), \quad \forall \varphi \in \mathcal{H} \quad \text{and} \quad \forall \Theta \in L^2(\tau, T).
\]

**Lemma 2.** Let (1) correspond to viscous flow in the two-dimensional (interior, exterior or channel type) domain. Then we have

\[
\| \varphi(t, \tau; \xi, \Theta) - \varphi(t, \tau; \xi, \Theta) \|_{H}^2 + \int_{\tau}^{t} \| \mathcal{A}^{1/2}_{e}(\varphi(s, \tau; \xi, \Theta) - \varphi(s, \tau; \xi, \Theta)) \|_{H}^2 \, ds \leq C_2 \left( \| \xi \|_{H}, \| \xi \|_{H}, \| \Theta \|_{L^2(\tau, T)} \right) \| \xi - \xi \|_{H}, \quad \forall \xi, \xi \in \mathcal{H} \quad \text{and} \quad \forall \Theta \in L^2(\tau, T).
\]

The solution \(\varphi\) is strongly continuous in time: \(\forall \xi \in \mathcal{H} \quad \text{and} \quad \forall \Theta \in L^2(\tau, T),\)

\[
\| \varphi(t, \tau; \xi, \Theta) - \xi \|_{H} \rightarrow 0 \quad \text{as} \quad t \rightarrow \tau.
\]

Moreover, for some \(\Theta_0 \in \mathbb{R},\)

\[
\| \varphi(t, \tau; \xi, \Theta_0) - \xi \|_{H} \leq C_3 \left( \| \xi \|_{D(\mathcal{A}^{1/2})}, \Theta_0 \right) |t - \tau|, \quad \forall \xi \in D(\mathcal{A}^{1/2}) \quad \text{and} \quad \forall t \in [\tau, T].
\]
and
\[ \| \mathcal{A}^{1/2} y(t, \tau; \xi; \theta_0) \|_H \leq C_d ( \| \xi \|_{D(\mathcal{A}^{1/2})}, \Theta_0), \quad \forall \xi \in D(\mathcal{A}^{1/2}) \text{ and } \forall t \in [\tau, T] \]

Using these results we can prove:

**Lemma 3.** The value function satisfies
\[ | \mathcal{V}'(t, \xi) - \mathcal{V}'(t, \xi') | \leq C_5 ( \| \xi \|_H, \| \xi' \|_H) \| \xi - \xi' \|_H, \quad \forall \xi, \xi \in H \text{ and } \forall t \in [\tau, T], \]
and
\[ | \mathcal{V}'(t, \xi) - \mathcal{V}'(t, \xi) | \leq C_6 | t - \tau | + C_7 ( \| \xi \|_H) \| y(t, \tau, \xi; \Theta_0) - \xi \|_H \]
\[ + C_8 \int_{\tau}^{T} \| \mathcal{A}^{1/2} y(s, \tau; \xi; \Theta_0) \|_H^2 ds, \quad \forall \xi \in H, \forall \Theta_0 \in R \text{ and } \forall t \in [\tau, T]. \]

From this result and the estimates for \( y \), we can conclude that the value function \( \mathcal{V} \in C([0, T] \times H) \).
For each \( t \in [\tau, T] \), \( \mathcal{V}(t, \cdot) \) is Lipschitz in \( \xi \in H \) and for each \( \xi \in D(\mathcal{A}^{1/2}) \), \( \mathcal{V}(\cdot, \xi) \) is Lipschitz in \( t \in [\tau, T] \). We can also prove the Bellman principle of optimality.

**Theorem 5.** For \( 0 \leq \tau \leq t \leq T \),
\[ \mathcal{V}'(t, \xi) = \inf \left\{ \int_{\tau}^{t} \left[ f(y(s, \tau; \xi; \Theta)) + h(\Theta(s)) \right] ds + \mathcal{V}'(t, y(t, \tau, \xi; \Theta)); \Theta \in L^2(\tau, t) \right\}. \]

The main theorem stated below can be proved using the above results and the methods in [1,2].

**Theorem 6.** The value function \( \mathcal{V}', [\tau, T] \times H \rightarrow R \) is a viscosity solution to the Hamilton–Jacobi–Bellman equation (2). That is, \( \mathcal{V}' \in C^1([\tau, T] \times H) \), if \( \mathcal{V}' - \phi \) attains a local maximum at \( (t_0, \xi_0) \in [\tau, T] \times D(\mathcal{A}) \), then
\[ -\partial_t \phi(t_0, \xi_0) + h^* \left( -\mathcal{B}^* \left[ D_t \phi \right](t_0, \xi_0) \right) + \left( \mathcal{A} \xi_0 + \mathcal{N}^{\prime}(\xi_0), \left[ D_t \phi \right](t_0, \xi_0) \right)_H - f(\xi_0) \leq 0 \]
and if \( \mathcal{V}' - \phi \) attains a local minimum at \( (t_0, \xi_0) \in [\tau, T] \times D(\mathcal{A}) \), then
\[ -\partial_t \phi(t_0, \xi_0) + h^* \left( -\mathcal{B}^* \left[ D_t \phi \right](t_0, \xi_0) \right) + \left( \mathcal{A} \xi_0 + \mathcal{N}^{\prime}(\xi_0), \left[ D_t \phi \right](t_0, \xi_0) \right)_H - f(\xi_0) \geq 0. \]

4. **Proof of the Pontryagin maximum principle**

We will derive the maximum principle using a method introduced by Barron and Jensen [4] (see also [2] which deals with infinite dimensional nonlinear evolution systems with bounded nonlinearity). The main advantage of this method is that it closely follows the formal derivation of the maximum principle from the Hamilton–Jacobi–Bellman equation. However, unlike the classical procedure which requires \( C^2 \) regularity on the value function, this method only requires the value function to be continuous. Let us begin with the evolution system,
\[ \partial_t z + \mathcal{A} z + \mathcal{N}(z) = \mathcal{B} \hat{\Theta}, \quad t \in [s, T], \quad \text{and} \quad z(s) = \xi \in D(\mathcal{A}), \]
where \( \hat{\Theta} \) is the optimal control in \([\tau, T]\). Note that in general \( z(t, s; \xi, \hat{\Theta}) \) is the optimal trajectory only if \( s = \tau \). Let us define \( \mathcal{W} \), \([\tau, T] \times H \rightarrow R \) as
\[ \mathcal{W}(s, \xi) := \phi_0 (z(T, s; \xi; \hat{\Theta})) + \int_{T}^{T} \left( f(z(r, s; \xi; \hat{\Theta})) + h(\hat{\Theta}) \right) dr \]
with \( \mathcal{W}(T, \xi) = \phi_0 (\xi), \quad \xi \in H. \)
Theorem 7. Let $\hat{\Theta} \in L^2(\tau, T)$ be an optimal control and $\hat{y}(t, \tau; \xi; \hat{\Theta})$ be the corresponding optimal trajectory. Then, for $t$ almost everywhere in $[\tau, T]$,

$$h^*\left(-\mathcal{B}^*\left[D, \mathcal{W}\right](t, \hat{y}(t, \tau; \xi; \hat{\Theta}))\right) = -\left(\mathcal{B}^*\left[D, \mathcal{W}\right](t, \hat{y}(t, \tau; \xi; \hat{\Theta}))\right) \cdot \hat{\Theta} - h(\hat{\Theta}). \quad (7)$$

Moreover, if $p(\cdot) \in C([\tau, T]; H)$ solves the adjoint system (5) then

$$p(t) = \left[D, \mathcal{W}\right](t, \hat{y}(t, \tau; \xi; \hat{\Theta})), \quad \forall t \in [\tau, T].$$

Proof. The proof of this theorem uses the fact that the value function is a viscosity solution. We will first show that the function $\mathcal{W}(\cdot, \cdot)$ has enough regularity properties to serve as the test function for the viscosity solution technique.

Lemma 4. If $\xi \in D(\mathcal{A})$, the function $\mathcal{W}(-, \cdot)$ is absolutely continuous in $s \in [\tau, T]$ and $\forall s \in [\tau, T]$, $\mathcal{W}(s, \cdot)$ is continuously Fréchet differentiable for $\xi \in D(\mathcal{A})$.

Proof. Let us define for $t \in [s, T]$, $\Psi(t) = [D_x\mathcal{W}](t, s; \xi; \hat{\Theta})\gamma$, for some $\gamma \in H$ and $\Phi(t) = \partial_z\xi(t, s; \xi; \hat{\Theta})$. Then $\Psi(\cdot)$ solves

$$\partial_t\Psi + \mathcal{A}\Psi + \left[D, \mathcal{N}(\cdot)\right]\Psi = 0, \quad t \in [s, T]. \quad (8a)$$

$$\Psi(s) = \gamma \in H. \quad (8b)$$

Similarly, $\Phi(\cdot)$ solves

$$\partial_t\Phi + \mathcal{A}\Phi + \left[D, \mathcal{N}(\cdot)\right]\Phi = 0, \quad t \in [s, T]. \quad (9a)$$

$$\Phi(s) = \xi'(\cdot) + \mathcal{N}(\xi) - \mathcal{B}\hat{\Theta}(s). \quad (9b)$$

Using methods similar to that used in the solvability theorem for (1), it is possible to show that if $\xi \in D(\mathcal{A})$ then, (8) and (9) have unique solutions $\Psi, \Phi \in C([s, T]; H) \cap L^2(s, T; D(\mathcal{A}^{1/2}))$. Now, for $\gamma \in H$,

$$\langle [D_x\mathcal{W}](s, \xi), \gamma \rangle = \langle [D\Phi_0](z(T)), \Psi(T) \rangle_H + \int_s^T \langle [Df](z(r)), \Psi(r) \rangle dr. \quad (10)$$

Hence, using the regularity of $\Psi$ and $z$,

$$\|\langle [D_x\mathcal{W}](s, \xi), \gamma \rangle\| \leq C_{10}(\|\xi\|_{D(\mathcal{A})}) \|\gamma\|_H, \quad \forall \gamma \in H.$$ 

Thus $[D_x\mathcal{W}](s, \xi) \in H, \forall \xi \in D(\mathcal{A})$ and $\forall s \in [\tau, T]$. Now, if $\xi \in D(\mathcal{A})$, then for $s$ almost everywhere in $[\tau, T]$,

$$\partial_s\mathcal{W}(s, \xi) = \langle [D\Phi_0](z(T)), \Phi(T) \rangle_H + \int_s^T \langle [Df](z(r)), \Phi(r) \rangle dr - f(\xi) - h(\hat{\Theta}(s)). \quad (11)$$

Now, using the properties of $\Phi$ we conclude that $\partial_s\mathcal{W}(\cdot, \cdot) \in L^1(0, T)$.

Let us now proceed with the proof of the theorem. Consider

$$\mathcal{W}(t, z(t, s; \xi; \hat{\Theta})) = \phi_0(z(T, s; \xi; \hat{\Theta})) + \int_t^{T}(f(z(r)) + h(\hat{\Theta}(r))) dr \quad + \int_t^{t+\delta t}(f(z(r)) + h(\hat{\Theta}(r))) dr$$

$$= \mathcal{W}(t + \delta t, z(t + \delta t, s; \xi; \hat{\Theta})) + \int_t^{t+\delta t}(f(z(r)) + h(\hat{\Theta}(r))) dr$$
Hence, for $t$ almost everywhere in $[s, T]$,

$$\frac{d}{dt} W(t, z(t, s; \xi, \Theta)) = -f(z(t, s; \xi, \Theta)) - h(\Theta(t)).$$

That is for $s \leq t \leq T$, and for $t$ almost everywhere in $[s, T]$,

$$-\partial W(t, z(t)) + \left(\left[D_{\mathcal{W}}(t, z(t))\right] A z(t) + N(z)\right)_{H} - f(z(t)) + \left(-B^* \left[D_{\mathcal{W}}(t, z(t))\right] \Theta\right) = 0$$

with $W(T, z(T, s; \xi, \Theta)) = \phi_0(z(T, s; \xi, \Theta))$. Thus

$$-\partial W(t, z(t)) + \left(\left[D_{\mathcal{W}}(t, z(t))\right] A z(t) + N(z)\right)_{H} - f(z(t))$$

$$+ h^*\left(-B^* \left[D_{\mathcal{W}}(t, z(t))\right] \Theta\right) \geq 0$$

(12)

Now, note that $s \in [\tau, T]$ and $\xi \in D(\mathcal{A})$, we have $W(s, \xi) \geq W(s, \xi)$ and $W(\tau, \xi) = \nu^*(s, \xi)$. In fact along the optimal trajectory, $W(t, \hat{y}(t, \tau; \xi, \Theta)) = \nu^*(t, \hat{y}(t, \tau; \xi, \Theta))$. Hence $\nu^* - \nu$ attains a maximum of zero along each point $(t, \hat{y}(t, \tau; \xi, \Theta))$. Now, using the fact that $\nu^*$ is a viscosity subsolution of (2),

$$-\partial W(t, \hat{y}(t)) + \left(\left[D_{\mathcal{W}}\right] A \hat{y}(t) + N(\hat{y})\right)_{H} - f(\hat{y}(t))$$

$$+ h^*\left(-B^* \left[D_{\mathcal{W}}\right] \Theta\right) \leq 0.$$ (14)

Comparing (14) and (13) for $s = \tau$, we deduce that

$$-\partial W(t, \hat{y}(t)) + \left(\left[D_{\mathcal{W}}\right] A \hat{y}(t) + N(\hat{y})\right)_{H} - f(\hat{y}(t))$$

$$+ h^*\left(-B^* \left[D_{\mathcal{W}}\right] \Theta\right) = 0.$$ (15)

Again comparing (15) with (12) for $s = \tau$, we deduce (7) Let us now consider the solution $\Psi \in C([\tau, T]; D(\mathcal{A}))$ of the linear problem (8) with $\gamma \in D(\mathcal{A})$. Taking duality pairing with (5),

$$\langle -\partial p, \Psi(t) \rangle + \langle A p, \Psi(t) \rangle + \langle [D \mathcal{N}]^*(\hat{y}) p, \Psi(t) \rangle = \langle [Df](\hat{y}), \Psi \rangle.$$

That is

$$-\partial p, \Psi(t) + \langle A p, \Psi(t) + [D \mathcal{N}](\hat{y}) \Psi, p \rangle = \langle [Df](\hat{y}), \Psi \rangle.$$

Noting that the second term on the left is zero, we integrate from $t$ to $T$,

$$(p(t), \Psi(t))_H - ([D\phi_0](y(T)), \Psi(T))_H = \int_t^T \langle [Df](\hat{y}(r)), \Psi(r) \rangle dr.$$

Here we used the fact that $p(T) = [D\phi_0](y(T))$. We now compare this result with (10):

$$\left([D_{\mathcal{W}}](t, \hat{y}(t)), \gamma\right)_H - ([D\phi_0](\hat{y}(T)), \Psi(T))_H = \int_t^T \langle [Df](\hat{y}(r)), \Psi(r) \rangle dr$$

and conclude that $(p(t) - [D_{\mathcal{W}}](t, \hat{y}(t)), \gamma)_H = 0, \forall \gamma \in D(\mathcal{A})$. Since $D(\mathcal{A}) \subset H$ is dense, we deduce that $p(\gamma) = [D_{\mathcal{W}}](\gamma, \hat{y}(\gamma, \tau, \xi, \Theta)) \in C([0, T]; H)$. This proves the theorem. Note also, that from (7) we get

$$\Theta(t) = \nabla h^* \left(-B^* \left[D_{\mathcal{W}}\right] (t, \hat{y}(t, \tau, \xi, \Theta))\right) = \nabla h^* (-B^* p(t)), \quad t \in [\tau, T].$$ (16)
5. Proof of the verification theorem

Recall that $\mathcal{V}(\cdot, \xi)$ for $\xi \in D(\mathcal{A})$ is Lipschitz in time. Hence by the Rademacher theorem $\mathcal{V}(\cdot, \xi)$ is differentiable almost everywhere in $[\tau, T]$. Let $t \in [\tau, T]$ be such a point. Let $\hat{\Theta}(\cdot)$ be the optimal control in the interval $[t, T]$ corresponding to the initial data $\xi \in D(\mathcal{A})$ at time $t$. Then from the definition of the value function,

$$
\mathcal{V}(t + \epsilon, \xi) - \mathcal{V}(t, \xi) \leq \int_{t+\epsilon}^{T} \left( f(y(r, t + \epsilon; \xi; \hat{\Theta})) + h(\hat{\Theta}(r)) \right) \, dr + \phi_0(y(T, t + \epsilon; \xi, \hat{\Theta}))
$$

$$
- \int_{t}^{T} \left( f(y(r, t; \xi, \hat{\Theta})) + h(\hat{\Theta}(r)) \right) \, dr - \phi_0(y(T, t; \xi, \hat{\Theta}))
$$

$$
= \int_{t+\epsilon}^{T} \left( f(y(r, t + \epsilon; \xi; \hat{\Theta})) - f(y(r, t; \xi; \hat{\Theta})) \right) \, dr
$$

$$
- \int_{t}^{t+\epsilon} \left( f(y(r, t; \xi, \hat{\Theta})) + h(\hat{\Theta}(r)) \right) \, dr
$$

$$
+ (\phi_0(y(T, t + \epsilon; \xi, \hat{\Theta})) - \phi_0(y(T, t; \xi, \hat{\Theta})))
$$

Hence,

$$
\partial_2 \mathcal{V}(t, \xi) \leq -f(\xi) - h(\hat{\Theta}(t)) + \int_{t}^{T} \langle [Df](y(r, t; \xi, \hat{\Theta})), \Phi(r) \rangle \, dr + ([D\phi_0](y(T)), \Phi(T))_H.
$$

Now substituting from the adjoint equation (5),

$$
\partial_2 \mathcal{V}(t, \xi) \leq -f(\xi) - h(\hat{\Theta}(t)) + (p(t), \Phi(t))_H.
$$

Substituting for $\Phi(t)$, and using the maximum principle, we get

$$
\partial_2 \mathcal{V}(t, \xi) \leq -f(\xi) + (p(t), A^*x + A^*(\xi))_H + h^*(-B^*p(t)) \tag{17}
$$

Now, let $\hat{\Theta}$ be the optimal control in $[t + \epsilon, T]$ corresponding to initial data $\xi \in D(\mathcal{A})$ at $t + \epsilon$. Then setting

$$
\Theta(r) = \begin{cases} 
\hat{\Theta}(r) & \text{if } r \in (t + \epsilon, T], \\
\hat{\Theta}(t + \epsilon) & \text{if } r \in (t, t + \epsilon], 
\end{cases}
$$

we get

$$
\mathcal{V}(t + \epsilon, \xi) - \mathcal{V}(t, \xi) \geq \int_{t+\epsilon}^{T} \left( f(y(r, t + \epsilon; \xi; \hat{\Theta})) + h(\hat{\Theta}(r)) \right) \, dr + \phi_0(y(T, t + \epsilon; \xi, \hat{\Theta}))
$$

$$
- \int_{t}^{T} \left( f(y(r, t; \xi, \Theta)) + h(\Theta(r)) \right) \, dr - \phi_0(y(T, t; \xi, \Theta))
$$

$$
= \int_{t+\epsilon}^{T} \left( f(y(r, t + \epsilon; \xi; \hat{\Theta})) - f(y(r, t; \xi; \Theta)) \right) \, dr
$$

$$
- \int_{t}^{t+\epsilon} \left( f(y(r, t; \xi, \hat{\Theta})) + h(\hat{\Theta}(r)) \right) \, dr
$$

$$
+ (\phi_0(y(T, t + \epsilon; \xi, \hat{\Theta})) - \phi_0(y(T, t; \xi, \Theta))).
$$

Thus noting that $\partial_2 y(r, t, \xi; \hat{\Theta}) \partial_2 y(t, t; \xi; \hat{\Theta}) = -\Phi(r)$ we get

$$
\partial_2 \mathcal{V}(t, \xi) \geq -f(\xi) - h(\hat{\Theta}(t)) + \int_{t}^{T} \langle [Df](y(r, t; \xi, \hat{\Theta})), \Phi(r) \rangle \, dr + ([D\phi_0](y(T)), \Phi(T))_H.
$$
This gives as before
\[
\partial_t \mathcal{V}(t, \xi) \geq -f(\xi) + (p(t), A\xi + N(\xi))_\mathcal{H} + h^*(-A^*p(t)).
\]

(18)

Comparing (17) with (18) we get (6) with \( p(t) \in \partial_t^* \mathcal{V}(t, \xi) \). Note also that the right hand side of (18) is integrable in time and hence \( \partial_t \mathcal{V}(\cdot, \xi) \in L^1(0, T) \).

Acknowledgement

This research has been supported by the Office of Naval Research through the URI-Grant No. N00014-86-K-0679

References