Journal of Convex Analysis
Volume 17 (2010), No. 3&4, 1007–1017

On Pontryagin's Principle for the Optimal Control of some State Equations with Memory

G. Carlier
Université Paris Dauphine, CEREMADE,
Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, France
carlier@ceremade.dauphine.fr

A. Houmia
Université Paris Dauphine, CEREMADE,
Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, France
houmia@ceremade.dauphine.fr

R. Tahraoui
Université Paris Dauphine, CEREMADE,
Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, France
tahraoui@ceremade.dauphine.fr

Dedicated to Hedy Attouch on the occasion of his 60th birthday.

Received: August 8, 2008
Revised manuscript received: March 21, 2009

We prove a form of Pontryagin's principle for a class of optimal control problems governed by a state equation with memory. Several examples and applications are then considered.

1. Introduction

The aim of this paper is to study deterministic optimal control problems of the form

$$\inf_{u(.)\in K} J(u) \quad \text{with } J(u) := \int_0^T L(s, y_u(s), u(s))ds + g(y_u(T)). \tag{1}$$

Here, u is the control variable, taking values in the admissible set K, and the state equation governing the dynamics of the state variable $y = y_u \in \mathbb{R}^d$ is an integrodifferential equation modelling quite general memory or delay effects. More precisely, y_u is the solution of the Cauchy problem:

$$\dot{x}(t) = f(t, \langle x, \nu_t \rangle, u(t)), \qquad x(0) = x_0 \tag{2}$$

where for each t, ν_t is a nonnegative measure supported by $[0, t]$ and $\langle x, \nu_t \rangle$ denotes:

$$\langle x, \nu_t \rangle := \int_0^T x(s)d\nu_t(s) = \int_0^t x(s)d\nu_t(s).$$

ISSN 0944-6532 / $ 2.50

Problems of the form above arise in various applied fields in engineering, economics, biology... It is typically the case when studying the optimal performances of a system in which the response to a given input occurs not instantaneously but only after a certain elapse of time. Such problems have in general been modelled by delayed or deviating arguments differential equations which both are particular cases of (2). We refer for instance to the classical book of Bellman and Cooke [2] for a general overview of such functional equations. In the state equation (2), introducing general parametrized measures ν_t (that can be varying sums of Dirac masses, given by some kernel, combinations of both...) allows us to treat more general memory effects. Let us also mention that, in [3], the first and third authors, studied the optimal control of a state equation with memory of the following form which is linear in the control:

$$\dot{x}(t) = \int_0^t f(s, x(s)) d\nu_t(s) + u(t).$$

The paper is organized as follows. In Section 2, we briefly discuss the Cauchy problem for (2). Our main result is then formulated in Section 3 where a generalization of Pontryagin's principle for the optimal control problem (1) is established. Finally, in Section 4, we apply our results to several examples where the Pontryagin's principle enables us to derive qualitative properties of optimal controls. In particular, we consider an economic model of consumption where the payment of interests may involve some delays and we discuss an extension to more general state equations.

2. On the Cauchy problem

In the sequel, we shall simply write C^0, L^p, $W^{1,p}$ instead respectively of $C^0([0,T],\mathbb{R}^d)$, $L^p((0,T),\mathbb{R}^d)$, $W^{1,p}((0,T),\mathbb{R}^d)$. For x and y in $\mathbb{R}^d$, the usual inner product of x and y will be denoted $x \cdot y$ and the euclidean norm of x, $|x|$ and for any matrix A, A^T denotes the transpose of A.

Throughout the paper, we will make the following assumptions

(H1) the control space K is a compact metric space, $f \in C^0([0,1] \times \mathbb{R}^d \times K, \mathbb{R}^d)$, and there exists a constant $C > 0$ such that:

$$|f(t,\eta,u) - f(t,\eta',u)| \leq C|\eta - \eta'|, \quad \forall (t,\eta,\eta',u) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \times K; \tag{3}$$

(H2) for each $t \in [0,T]$, ν_t is a nonnegative finite measure such that $\nu_t((t,T]) = 0$ and $t \mapsto \nu_t$ is measurable in the sense that $t \mapsto \langle g, \nu_t \rangle$ is measurable for every $g \in C^0([0,T],\mathbb{R})$;

(H3) defining $\alpha(t) := \nu_t([0,T]) = \nu_t([0,t])$, we assume $\alpha \in L^1(0,T)$.

Compactness of the control space K is not really necessary for the rest of the paper, this assumption is only made for the sake of simplicity. When K is not compact, it is easy to check that all our results remain true provided f satisfies some linear growth condition with respect to η uniformly in t and u.

We first have the following easy Lemma (see [3] for details):

Lemma 2.1. *Let $\lambda > 0$ and define for every $t \in [0,T]$,*

$$\varphi_\lambda(t) := \int_0^t e^{\lambda(s-t)}\alpha(s)ds.$$

Then φ_λ converges uniformly to 0 on $[0,T]$ as $\lambda \to +\infty$.

We then have existence and uniqueness for the Cauchy problem:

Proposition 2.2. *Under the assumptions above, for any measurable function $u : [0,T] \to K$, the Cauchy problem*

$$\dot{x}(t) = f(t, \langle x, \nu_t \rangle, u(t)), \qquad x(0) = x_0 \tag{4}$$

possesses a unique $W^{1,1}$ solution which will be denoted y_u from now on.

Proof. Let us rewrite (4) as

$$x = Tx, \quad \text{with } Tx(t) := x_0 + \int_0^t f(s, \langle x, \nu_s \rangle, u(s))ds$$

and equip the space $C^0([0,T], \mathbb{R}^d)$ with the norm

$$\|x\|_\lambda := \sup_{t\in[0,T]} e^{-\lambda t}|x(t)| \tag{5}$$

where $\lambda > 0$ will be chosen later on. Of course, $(C^0, \|.\|_\lambda)$ is a Banach space and $T(C^0) \subset C^0$. Defining φ_λ as in Lemma 2.1, our assumptions ensure that for x_1 and x_2 in C^0, one has:

$$\|Tx_1 - Tx_2\|_\lambda \leq C \max_{t\in[0,T]} \varphi_\lambda(t)\|x_1 - x_2\|_\lambda.$$

By Lemma 2.1, choosing λ large enough, we deduce that T is a strict contraction of $(C^0, \|.\|_\lambda)$ and therefore admits a unique fixed point that we denote y_u. The fact that y_u is in $W^{1,1}$ follows from the inequality

$$|Tx(t_2) - Tx(t_1)| \leq \int_{t_1}^{t_2} (a + C\|x\|_\infty \alpha(s))ds \tag{6}$$

for $a := \sup\{|f(t, 0, u)|,\ t \in [0,T],\ u \in K\}$, any x in C^0 and any t_1 and t_2 such that $0 \leq t_1 \leq t_2 \leq T$. □

3. Pontryagin principle

Now, we consider the optimal control problem

$$\inf_{u(.)\in K} J(u) \ \text{ with } J(u) := \int_0^T L(s, y_u(s), u(s))ds + g(y_u(T)), \tag{7}$$

where as before, y_u denotes the solution of (4). In addition to assumptions **(H1)**, **(H2)** and **(H3)** we further assume:

(H4) for every $t \in [0,T]$ and $u \in K$, $\eta \in \mathbb{R}^d \to f(t,\eta,u)$ is of class C^1 and $D_\eta f$ is continuous in all its arguments;

(H5) g is of class C^1 on $\mathbb{R}^d$, L is continuous in all its arguments and for every $t \in [0,T]$ and $u \in K$, $x \in \mathbb{R}^d \mapsto L(t,x,u)$ is of class C^1 and $\nabla_x L$ is continuous in all its arguments.

To define the adjoint state, we need a few definitions and notations. Let us denote by $\mathcal{L}^1$ the Lebesgue measure on $[0,T]$ and let us introduce the nonnegative measure $\gamma := \nu_t \otimes \mathcal{L}^1$ on $[0,T]^2$, and define ν as the second marginal of γ. Using test-functions, γ and ν are then defined by:

$$\int_{[0,T]^2} \phi(t,s)d\gamma(t,s) = \int_0^T \left(\int_0^T \phi(t,s)d\nu_t(s) \right) dt, \quad \forall \phi \in C^0([0,T]^2, \mathbb{R}),$$
$$\int_0^T \psi(s)d\nu(s) = \int_0^T \left(\int_0^T \psi(s)d\nu_t(s) \right) dt, \quad \forall \psi \in C^0([0,T], \mathbb{R}).$$

Using the disintegration theorem (see for instance the book of Dellacherie and Meyer [4] or the appendix in the lecture notes of Ambrosio [1]) we may also write $\gamma = \nu_s^* \otimes \nu$ where ν_s^* is a measurable family of probability measures on $[0,T]$. We recall that $\phi \in L^1(\gamma)$ if and only if:

- for $\mathcal{L}^1$-a.e. $t \in [0,T]$, $\phi(t,.) \in L^1(\nu_t)$, and
- $t \mapsto \langle \phi(t,.), \nu_t \rangle \in L^1(\mathcal{L}^1)$

which is also equivalent to

- for ν-a.e. $s \in [0,T]$, $\phi(.,s) \in L^1(\nu_s^*)$, and
- $s \mapsto \langle \phi(.,s), \nu_s^* \rangle \in L^1(\nu)$.

Moreover, if $\phi \in L^1(\gamma)$, then:

$$\int_{[0,T]^2} \phi(t,s)d\gamma(t,s) := \int_0^T \langle \phi(t,.), \nu_t \rangle \, dt = \int_0^T \langle \phi(.,s), \nu_s^* \rangle \, d\nu(s)$$

Let us also remark that since ν_t is supported by $[0,t]$, ν_s^* is supported by $[s,T]$. Then our last assumption is:

(H6) ν is absolutely continuous with respect to $\mathcal{L}^1$.

Next, let us remark that if $\phi \in L^\infty([0,T],\mathbb{R}^d)$, then $(t,s) \mapsto \phi(t) \in L^1(\gamma)$ hence for ν-a.e. $s \in [0,T]$, $\phi \in L^1(\nu_s^*)$ and $s \mapsto \langle \phi, \nu_s^* \rangle \in L^1(\nu)$. By **(H6)** (and slightly abusing notations denoting by ν the density of ν with respect to $\mathcal{L}^1$), we also have that $s \mapsto \langle \phi, \nu_s^* \rangle \, \nu(s)$ is in L^1 and the inequality

$$\| \langle \phi, \nu_.^* \rangle \, \nu \|_{L^1} = \int_0^T | \langle \phi, \nu_s^* \rangle | \nu(s) ds \leq \|\phi\|_\infty \|\nu\|_{L^1}.$$

Let us assume that $\overline{u}$ is an optimal control, i.e., solves (7) and let us denote by $\overline{x} := y_{\overline{u}}$ the corresponding optimal trajectory. Let $t \in (0,T)$ satisfy

$$t \text{ is a Lebesgue point of } s \mapsto (\langle \overline{x}, \nu_s \rangle, L(s, \overline{x}(s), \overline{u}(s)), f(s, \langle \overline{x}, \nu_s \rangle, \overline{u}(s))). \tag{8}$$

Now let $v \in K$ be an arbitrary admissible control and $\varepsilon \in (0, t)$. Let us remark that with (8) and **(H1)**, t is a Lebesgue point of $s \mapsto f(s, \langle \overline{x}, \nu_s \rangle, v))$. Let us then define

$$u_\varepsilon(s) = \begin{cases} v & \text{if } s \in (t-\varepsilon, t] \\ \overline{u}(s) & \text{otherwise.} \end{cases}$$

and denote by $x_\varepsilon := y_{u_\varepsilon}$ the associated state.

To shorten notations, let us also set:

$$A(s) := D_\eta f(s, \langle \overline{x}, \nu_s \rangle, \overline{u}(s)), \quad \text{and} \quad \theta(s) := \nabla_x L(s, \overline{x}(s), \overline{u}(s)) \tag{9}$$

and let us remark that both A and θ are in L^∞.

Lemma 3.1. *Let us define $z_\varepsilon := \varepsilon^{-1}(x_\varepsilon - \overline{x})$. Then z_ε is bounded in L^∞, z_ε converges pointwise to $z = 0$ on $[0, t)$ and z_ε converges uniformly on $[t, T]$ to the function z that solves:*

$$\dot{z}(s) = A(s) \langle z, \nu_s \rangle \quad \text{on } (t, T], \; z(t) = f(t, \langle \overline{x}, \nu_t \rangle, v) - f(t, \langle \overline{x}, \nu_t \rangle, \overline{u}(t)). \tag{10}$$

Proof. By uniqueness for the Cauchy problem, $z_\varepsilon = 0$ on $[0, t-\varepsilon]$ so that z_ε converges pointwise to 0 on $[0, t)$.

Step 1: z_ε is bounded in L^∞. For $s > t$, we have:

$$z_\varepsilon(s) = I(\varepsilon) + \frac{1}{\varepsilon} \int_t^s [f(\theta, \langle x_\varepsilon, \nu_\theta \rangle, \overline{u}(\theta)) - f(\theta, \langle \overline{x}, \nu_\theta \rangle, \overline{u}(\theta))] d\theta \tag{11}$$

where

$$I(\varepsilon) = \frac{1}{\varepsilon} \int_{t-\varepsilon}^t [f(\theta, \langle x_\varepsilon, \nu_\theta \rangle, v) - f(\theta, \langle \overline{x}, \nu_\theta \rangle, \overline{u}(\theta))] d\theta.$$

It is easy to check that x_ε is bounded and then so is the integrand in $I(\varepsilon)$ and then $|I(\varepsilon)| \leq M$ for some constant M. Defining the norm, $\|.\|_\lambda$ as in the proof of Proposition 2.2, φ_λ as in Lemma 2.1 and using **(H1)** we then have

$$|z_\varepsilon(s)| e^{-\lambda s} \leq M e^{-\lambda s} + C \|z_\varepsilon\|_\lambda \varphi_\lambda(s). \tag{12}$$

Thanks to Lemma 2.1, we easily deduce that z_ε is bounded in L^∞. Let then C_0 be such that $\|x_\varepsilon - \overline{x}\|_\infty \leq C_0 \varepsilon$ for every ε.

Step 2: Convergence of $z_\varepsilon(t)$. Let us write

$$\begin{aligned} z_\varepsilon(t) = & \frac{1}{\varepsilon} \int_{t-\varepsilon}^t [f(s, \langle \overline{x}, \nu_s \rangle, v) - f(s, \langle \overline{x}, \nu_s \rangle, \overline{u}(s))] ds \\ & + \frac{1}{\varepsilon} \int_{t-\varepsilon}^t [f(s, \langle x_\varepsilon, \nu_s \rangle, v) - f(s, \langle \overline{x}, \nu_s \rangle, v)] ds. \end{aligned}$$

The second term is bounded by $C_0 C \int_{t-\varepsilon}^t \alpha$ hence converges to 0. Since t is a Lebesgue point of $s \mapsto f(s, \langle \overline{x}, \nu_s \rangle, v)$ and $s \mapsto f(s, \langle \overline{x}, \nu_s \rangle, \overline{u}(s))$, we deduce that

$$\lim_{\varepsilon \to 0^+} z_\varepsilon(t) = f(t, \langle \overline{x}, \nu_t \rangle, v) - f(t, \langle \overline{x}, \nu_t \rangle, \overline{u}(t)). \tag{13}$$

Step 3: z_ε is uniformly equicontinuous on $[t,T]$. Let t_1 and t_2 be such that $T \geq t_2 \geq t_1 \geq t$, we then have

$$|z_\varepsilon(t_2) - z_\varepsilon(t_1)| \leq \frac{C}{\varepsilon} \int_{t_1}^{t_2} \langle |x_\varepsilon - \overline{x}|, \nu_s \rangle \, ds \leq C_0 C \int_{t_1}^{t_2} \alpha(s) ds. \tag{14}$$

Ascoli's theorem and (14) then prove that the family $(z_\varepsilon)_\varepsilon$ is relatively compact in $C^0([t,T], \mathbb{R}^d)$.

Step 4: z_ε converges on $[t,T]$ to the solution of the linearized equation. Thanks to *Step 3*, there is a sequence $\varepsilon_n \to 0$ such that $z_n := z_{\varepsilon_n}$ converges uniformly to some $z \in C^0([t,T])$. Let t_1 and t_2 be such that $T \geq t_2 \geq t_1 > t$, we have

$$z_n(t_2) - z_n(t_1) = \frac{1}{\varepsilon_n} \int_{t_1}^{t_2} [f(s, \langle \overline{x} + \varepsilon_n z_n, \nu_s \rangle, \overline{u}(s)) - f(s, \langle \overline{x}, \nu_s \rangle, \overline{u}(s)] ds. \tag{15}$$

Thanks to **(H4)** and Lebesgue's dominated convergence theorem, passing to the limit in (15), yields

$$\begin{aligned} z(t_2) - z(t_1) &= \int_{t_1}^{t_2} [D_\eta f(s, \langle \overline{x}, \nu_s \rangle, \overline{u}(s)) \langle z, \nu_s \rangle] ds \\ &= \int_{t_1}^{t_2} A(s) \langle z, \nu_s \rangle \, ds. \end{aligned}$$

This proves that z solves

$$\dot{z}(s) = A(s) \langle z, \nu_s \rangle \quad \text{on } (t,T]. \tag{16}$$

And with (13), z also satisfies the initial condition

$$z(t) = f(t, \langle \overline{x}, \nu_t \rangle, v) - f(t, \langle \overline{x}, \nu_t \rangle, \overline{u}(t)). \tag{17}$$

Finally, the system (16)–(17) admits z as unique solution, together with the relative compactness of the family z_ε, this implies that the whole family z_ε converges uniformly to z on $[t,T]$ as $\varepsilon \to 0^+$ and the proof is complete. □

Lemma 3.2. *There exists a unique $W^{1,1}$ function $\overline{p}$ that solves the adjoint system:*

$$\dot{p}(s) = -\left\langle A^T p, \nu_s^* \right\rangle \nu(s) - \theta(s) \tag{18}$$

together with the transversality condition:

$$p(T) = \nabla g(\overline{x}(T)). \tag{19}$$

Proof. The existence and uniqueness of a continuous solution to (18) can be proven by similar arguments as for Proposition 2.2. The fact that $\overline{p}$ is $W^{1,1}$ follows from the fact that, since $\nu \in L^1$, $s \mapsto \left\langle A^T \overline{p}, \nu_s^* \right\rangle \nu(s)$ is L^1 and:

$$\| \left\langle A^T \overline{p}, \nu_s^* \right\rangle \nu \|_{L^1} \leq \|A^T \overline{p}\|_\infty \int_0^T \nu = \|A^T \overline{p}\|_\infty \|\nu\|_{L^1}.$$

□

Lemma 3.3. *Let z be as in Lemma 3.1 and $\overline{p}$ be the adjoint variable defined in Lemma 3.2. Then we have:*

$$\overline{p}(T)\cdot z(T)=\overline{p}(t)\cdot z(t)-\int_t^T \theta\cdot z. \tag{20}$$

Proof. Since both z and p are $W^{1,1}$ on (t,T) we have:

$$\overline{p}(T)\cdot z(T)-\overline{p}(t)\cdot z(t)=\int_t^T \left(\overline{p}\cdot\dot{z}+\dot{\overline{p}}\cdot z\right)$$
$$=\int_t^T \overline{p}(s)\cdot A(s)\left\langle z,\nu_s\right\rangle ds-\int_t^T \left\langle A^T\overline{p},\nu_s^*\right\rangle\cdot z(s)\nu(s)ds-\int_t^T \theta\cdot z.$$

Since $\overline{p}$ and A are in L^∞ and $s\mapsto\langle z,\nu_s\rangle$ is L^1, the map

$$(s,\tau)\mapsto A(\tau)^T\overline{p}(\tau)\cdot z(s)$$

belongs to $L^1(\gamma)$. Since $z=0$ on $[0,t)$ and then $\langle z,\nu_\tau\rangle=0$ for $\tau<t$, we thus get

$$\begin{aligned}\int_t^T \overline{p}(\tau)\cdot A(\tau)\left\langle z,\nu_\tau\right\rangle d\tau &= \int_0^T \overline{p}(\tau)\cdot A(\tau)\left\langle z,\nu_\tau\right\rangle d\tau\\ &= \int_{[0,T]^2} A^T(\tau)\overline{p}(\tau)\cdot z(s)d\gamma(\tau,s)\\ &= \int_0^T \left\langle A^T\overline{p},\nu_s^*\right\rangle\cdot z(s)\nu(s)ds\\ &= \int_t^T \left\langle A^T\overline{p},\nu_s^*\right\rangle\cdot z(s)\nu(s)ds\end{aligned}$$

which proves (20). □

Lemma 3.4. *Let z be as in Lemma 3.1 and $\overline{p}$ be the adjoint variable defined in Lemma 3.2. If t satisfies (8), then we have:*

$$\lim_{\varepsilon\to 0^+}\frac{1}{\varepsilon}[J(u_\varepsilon)-J(\overline{u})]=\overline{p}(T)\cdot z(T)+\int_t^T\theta\cdot z+L(t,\overline{x}(t),v)-L(t,\overline{x}(t),\overline{u}(t)). \tag{21}$$

Proof. By construction, we have

$$\frac{1}{\varepsilon}[J(u_\varepsilon)-J(\overline{u})]=\frac{1}{\varepsilon}[g(x_\varepsilon(T))-g(\overline{x}(T))]+\frac{1}{\varepsilon}\int_{t-\varepsilon}^t[L(s,x_\varepsilon(s),v)-L(s,\overline{x}(s),\overline{u}(s))]ds$$
$$+\frac{1}{\varepsilon}\int_t^T[L(s,x_\varepsilon(s),\overline{u}(s))-L(s,\overline{x}(s),\overline{u}(s))]ds.$$

Thanks to Lemma 3.1, assumption **(H5)** and (19), the first term above converges to

$$\nabla g(\overline{x}(T))\cdot z(T)=\overline{p}(T)\cdot z(T).$$

Since t is a Lebesgue point of $s \mapsto L(s, \overline{x}(s), \overline{u}(s))$, one easily deduces from Lemma 3.1 that the second term converges to:

$$L(t, \overline{x}(t), v) - L(t, \overline{x}(t), \overline{u}(t)).$$

Finally, recalling (9), by Lebesgue's dominated convergence theorem, **(H5)** and Lemma 3.1, the third term converges to:

$$\int_t^T \nabla_x L(s, \overline{x}(s), \overline{u}(s)) \cdot z(s) ds = \int_t^T \theta \cdot z.$$

□

Our main result then states the following form of the Pontryagin's principle:

Theorem 3.5. *If $\overline{u}$ solves* (7) *and $\overline{x} := y_{\overline{u}}$ denotes the corresponding trajectory, then for a.e. $t \in (0, T)$ one has*

$$\overline{u}(t) \in \operatorname{argmin}_{v \in K} \{\overline{p}(t) \cdot f(t, \langle \overline{x}, \nu_t \rangle, v) + L(t, \overline{x}(t), v)\} \tag{22}$$

where the adjoint variable $\overline{p}$ is the solution of

$$\begin{cases} \dot{p}(s) = -\left\langle D_\eta f(., \langle \overline{x}, \nu_. \rangle, \overline{u}(.))^T p(.), \nu_s^* \right\rangle \nu(s) - \nabla_x L(s, \overline{x}(s), \overline{u}(s)), \\ p(T) = \nabla g(\overline{x}(T)), \end{cases}$$

where $\langle \overline{x}, \nu_. \rangle$ denotes $\tau \mapsto \langle \overline{x}, \nu_\tau \rangle$.

Proof. Almost every $t \in (0, T)$ is a Lebesgue point of

$$s \mapsto \langle \overline{x}, \nu_s \rangle, \qquad s \mapsto L(s, \overline{x}(s), \overline{u}(s)), \quad \text{and} \quad s \mapsto f(s, \langle \overline{x}, \nu_s \rangle, \overline{u}(s)).$$

For such a t, we define u_ε and x_ε as above. Since $\overline{u}$ is an optimal control, then $J(u_\varepsilon) \geq J(u)$. With Lemmas 3.3 and 3.4 we then have, defining z as in Lemma 3.1 and $\overline{p}$ as in Lemma 3.2:

$$\begin{aligned} 0 \leq \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon}[J(u_\varepsilon) - J(\overline{u})] &= p(T) \cdot z(T) + \int_t^T \theta + L(t, \overline{x}(t), v) - L(t, \overline{x}(t), \overline{u}(t)) \\ &= p(t) \cdot z(t) + L(t, \overline{x}(t), v) - L(t, \overline{x}(t), \overline{u}(t)) \\ &= p(t) \cdot (f(t, \langle \overline{x}, \nu_t \rangle, v) - f(t, \langle \overline{x}, \nu_t \rangle, \overline{u}(t))) + L(t, \overline{x}(t), v) - L(t, \overline{x}(t), \overline{u}(t)) \end{aligned}$$

since v is arbitrary in K in the previous inequality, we deduce (22) and the proof is complete. □

4. Applications and extensions

The aim of this section is to consider some examples and derive some qualitative properties of the optimal controls from the optimality conditions given in Theorem 3.5. We also mention an extension of the Pontryagin's principle of Theorem 3.5 to more general state equations.

4.1. The case of a linear dynamic

Let us consider the case of the linear dynamic:

$$\dot{x}(t) = A(t)(\langle x, \nu_t \rangle) + u(t), \qquad x(0) = x_0, \tag{23}$$

where A is a continuous function from $[0, T]$ to the space of $d \times d$ matrices. We assume that the admissible set of controls K is some convex compact subset of $\mathbb{R}^d$ and as before, for any measurable u taking values in K we denote by y_u the solution of (23). We now consider the optimal control problem (7) in the convex case, i.e., when in addition to the general assumptions of the paper, g is convex, and L is convex and differentiable with respect (x, u). In this convex setting, the optimality conditions given in Theorem 3.5 are in fact necessary and sufficient. More precisely, let us assume that the control $\overline{u}$ and the associated state $\overline{x} = y_{\overline{u}}$ satisfy the conditions of the Pontryagin's principle:

$$\overline{u}(t) \in \operatorname{argmin}_{v \in K} \{\overline{p}(t) \cdot v + L(t, \overline{x}(t), v)\}, \quad \text{for a.e. } t \tag{24}$$

where the adjoint variable $\overline{p}$ solves

$$\dot{p}(s) = -\left\langle A^T(.)\overline{p}(.), \nu_s^* \right\rangle \nu(s) - \nabla_x L(s, \overline{x}(s), \overline{u}(s)), \qquad p(T) = \nabla g(\overline{x}(T)). \tag{25}$$

Then, we claim that $\overline{u}$ is an optimal control. Indeed, let u be another admissible control and $x = y_u$ be the corresponding state, then on the one hand, the convexity of K and (24) yield

$$\overline{p}(t) \cdot (u(t) - \overline{u}(t)) + \nabla_u L(t, \overline{x}(t), \overline{u}(t)) \cdot (u(t) - \overline{u}(t)) \geq 0, \quad \text{for a.e. } t. \tag{26}$$

On the other hand, using our convexity assumptions, (23), (25) and (26) and an integration by parts, we have

$$\begin{aligned} J(u) - J(\overline{u}) \geq & \int_0^T \nabla_x L(t, \overline{x}, \overline{u}) \cdot (x - \overline{x}) + \nabla_u L(t, \overline{x}, \overline{u}) \cdot (u - \overline{u}) \\ & + \overline{p}(T) \cdot (x(T) - \overline{x}(T)) \\ \geq & \int_0^T \overline{p} \cdot (\dot{x} - \dot{\overline{x}}) - \int_0^T \left\langle A^T \overline{p}, \nu_.^* \right\rangle \cdot (x - \overline{x})\nu - \int_0^T \overline{p} \cdot (u - \overline{u}) \\ = & \int_0^T \overline{p} \cdot A(\langle x - \overline{x}, \nu_. \rangle) - \int_0^T \left\langle A^T \overline{p}, \nu_.^* \right\rangle \cdot (x - \overline{x})\nu = 0. \end{aligned}$$

Let us further consider the case where

$$L(t, x, u) := h(t, x) + \frac{1}{2}|u|^2$$

with h convex in x. Then there is a unique optimal control, $\overline{u}$, that is characterized as follows. First, condition (24) simplifies to

$$\overline{u}(t) = \pi_K(-\overline{p}(t)) \tag{27}$$

where π_K denotes the projection on K. Note that since, π_K is one-Lipschitz then the optimal control is absolutely continuous. The optimal state and the corresponding adjoint state are then obtained by solving:

$$\begin{cases} \dot{\overline{x}}(t) = A(t)(\langle \overline{x}, \nu_t \rangle) + \pi_K(-\overline{p}(t)), \ \overline{x}(0) = x_0, \\ \dot{\overline{p}}(s) = -\left\langle A^T(.)\overline{p}(.), \nu_s^* \right\rangle \nu(s) - \nabla_x h(s, \overline{x}(s)), \overline{p}(T) = \nabla g(\overline{x}(T)). \end{cases}$$

4.2. An economic example

A classical approach to study the tradeoff between consumption and saving in continuous time is to consider the problem

$$\sup_{c(.)} \int_0^T e^{-\delta t} U(c(t))dt + V(x(T)) \tag{28}$$

where c denotes the consumption (control) variable, x denotes the agent's wealth, $\delta > 0$ is a discount factor and U and V are strictly concave increasing utility functions. Usually, the state equation governing the evolution of wealth is the linear equation:

$$\dot{x} = rx - c \tag{29}$$

where r denotes the (possibly nonconstant) interest rate. In the case of a constant interest rate r, it is well-known that the optimal consumption is increasing when $\delta < r$ (i.e. when the agent is sufficiently patient), constant when $\delta = r$ and decreasing when $\delta > r$. Now, let us consider the more general case where there may be some delay in the payment of interests and where the state equation is given by

$$\dot{x}(t) = \langle rx, \mu_t \rangle - c(t) = \langle x, \nu_t \rangle - c(t), \quad \text{with } \nu_t = r\mu_t. \tag{30}$$

The (necessary and sufficient) optimality conditions for the optimal consumption $\overline{c}$ are as follows:

$$U'(\overline{c}(t)) = e^{\delta t}\overline{q}(t)$$

where the adjoint state $\overline{q}$ is the solution of

$$\dot{q}(s) = -\langle q, \nu_s^* \rangle \nu(s), \qquad q(T) = V'(\overline{x}(T)) > 0. \tag{31}$$

It is easy to check that $\overline{q}$ is positive and decreasing and since U' is decreasing, we deduce that the optimal consumption is increasing exactly when $e^{\delta t}\overline{q}$ is decreasing (note that this condition does dot depend on the value $\overline{q}(T)$ and is satisfied for instance in the case of a large constant interest rate r).

4.3. Bang-bang control

As a first example, let us consider the case where $L = 0$ and the state equation has the separable form:

$$\dot{x}_i(t) = f_i(t, \langle x, \nu_t \rangle, u_i(t)), \quad i = 1, ..., d$$

where each f_i is increasing with respect to u_i and the admissible control set K is of the form $K = [a_1, b_1] \times ... \times [a_d, b_d]$. In this case, the optimality condition (22) implies some bang-bang property of the optimal controls: $\overline{u}_i(t) = a_i$ for a.e. t such that $\overline{p}_i(t) > 0$ and $\overline{u}_i(t) = b_i$ for a.e. t such that $\overline{p}_i(t) < 0$.

As a second example, let us assume that g satisfies $\nabla g(x) \in \mathbb{R}^d_+$ for all x, and f is such that $D_\eta f(t, \eta, u)$ has nonnegative entries for every (t, η, u). It is easy to check that this implies that the adjoint state $\overline{p}$ necessarily has nonnegative components. If, in addition, K is convex, L is strictly concave with respect to u and each component of f is concave with respect to u, then the optimality condition (22) implies that any optimal control $\overline{u}(t)$ has to minimize over K, for a.e. t, the strictly concave Hamiltonian $v \mapsto L(t, \overline{x}(t), v) + \overline{p}(t) \cdot f(t, \langle \overline{x}, \nu_t \rangle, v)$. This implies that optimal controls necessarily take values in the set of extreme points of K. In particular if K is a convex polytope, optimal controls take a finite number of values.

4.4. Extension to more general state equations

By the same arguments as above, one can consider the optimal control of the more general state equation than (2):

$$\dot{x}(t) = f(t, \langle h_1(., x(.)), \nu_t^1 \rangle, ..., \langle h_m(., x(.)), \nu_t^m \rangle, u(t)), \qquad x(0) = x_0 \tag{32}$$

provided all the functions h_1, ..., h_m satisfy suitable Lipschitz and differentiability conditions and the measures $\nu_t^1, ..., \nu_t^m$ satisfy the same conditions as ν_t in the previous sections. For the sake of simplicity, let us assume that x, f and each function h_j are real-valued and denoting y_u the solution of (32), we consider the optimal control problem

$$\inf_{u(.)\in K} J(u) \ \text{ with } J(u) := \int_0^T L(s, y_u(s), u(s))ds + g(y_u(T)). \tag{33}$$

It is easy to check that the following generalization of Theorem 3.5 holds (under natural assumptions that we do not make precise here). If $\overline{u}$ solves (33) and $\overline{x} := y_{\overline{u}}$ then for a.e. t, one has

$$\overline{u}(t) \in \text{argmin}_{v\in K}\{\overline{p}(t)\cdot f(t, \langle h_1(., \overline{x}(.)), \nu_t^1\rangle, ..., \langle h_m(., \overline{x}(.)), \nu_t^m\rangle, v) + L(t, \overline{x}(t), v)\}$$

for the adjoint variable $\overline{p}$ that solves:

$$\dot{p}(s) = -\sum_{j=1}^m \langle \partial_{\eta_j} f(., \eta(.), \overline{u}(.))p(.), \nu_s^{j*}\rangle \, \partial_x h_j(s, \overline{x}(s))\nu^j(s) - \partial_x L(s, \overline{x}(s), \overline{u}(s)),$$

$$p(T) = g'(\overline{x}(T))$$

where we have set:

$$\eta(\tau) := (\langle h_1(., \overline{x}(.)), \nu_\tau^1\rangle, ..., \langle h_m(., \overline{x}(.)), \nu_\tau^m\rangle)$$

and the measures ν^j and ν_s^{j*} are obtained as before by desintegrating the measures $\gamma^j := \nu_t^j \otimes \mathcal{L}^1$.

References

[1] L. Ambrosio: Lecture notes on optimal transport problems, in: Mathematical Aspects of Evolving Interfaces (Madeira, 2000), P. Colli et al. (ed.), Lect. Notes Math. 1812, Springer, Berlin (2003) 1–52.

[2] R. Bellman, K. L. Cooke: Differential-Difference Equations, Mathematics in Science and Engineering 6, Academic Press, New York (1963).

[3] G. Carlier, R. Tahraoui: On some optimal control problems governed by a state equation with memory, ESAIM, Control Optim. Calc. Var. 14(4) (2008) 725–743.

[4] C. Dellacherie, P.-A. Meyer: Probabilities and Potential, Mathematical Studies 29, North-Holland, Amsterdam (1978).