A GLOBAL MAXIMUM PRINCIPLE FOR PERIODIC CONTROL OF FUNCTIONAL DIFFERENTIAL EQUATIONS

F. Colonius

Present address: Institut für Dynamische Systeme, Universität Bremen,
Postfach 330 440, D-2800 Bremen 33, West Germany

Summary

The paper presents a global maximum principle for optimal periodic control of retarded functional differential systems. The proof is based on Ekeland's Variational Principle, and a structural theory of linear functional differential equations.

1. Introduction

Consider the following optimal periodic control problem:

Minimize (1.1) $1/(t_2-t_1)\int_{t_1}^{t_2} g(x(t),u(t),t)dt$ s.t.

(1.2) $\dot{x}(t) = f(x_t,u(t),t)$ a.a. $t \in T := [t_1,t_2]$

(1.3) $x_{t_1} = x_{t_2}$

(1.4) $u \in \mathcal{U}_{ad} := \{u: T \to \Omega, \text{ measurable}\}$

where $x_t(s) := x(t+s) \in \mathbb{R}^n$, $s \in [-r,0]$, $r > 0$, $\mathcal{O}_x \subset \mathbb{R}^n$, $\mathcal{O}_u \subset \mathbb{R}^m$, $\mathcal{O}_\varphi \subset C(-r,0;\mathbb{R}^n)$ are open sets, $g: \mathcal{O}_x \times \mathcal{O}_u \to \mathbb{R}$, $f: \mathcal{O} = \mathcal{O}_\varphi \times \mathcal{O}_u \times T \to \mathbb{R}^n$ and $\Omega \subset \mathcal{O}_u$, closed.

Optimal control problems for retarded equations have been studied since a long time, initially with targets in $\mathbb{R}^n$. Problems with fixed targets $x_{t_1} = \psi \in C(-r,0;\mathbb{R}^n)$ were studied when it became clear, that an appropriate notion of state is the function segment x_t. However, for this latter problem, a maximum principle is valid only under very restrictive assumptions. In contrast, the periodic boundary condition (1.3) allows to develop a theory of necessary optimality conditions under similarly weak conditions as in the finite dimensional case. The present paper gives a global maximum principle for this problem.

Our hypotheses are:

(1.5) The functions $g(x,u,t)$ and $f(\varphi,u,t)$ are measurable in t and jointly continuous in (x,u) and (φ,u), resp., and continuously Fréchet differentiable in x and φ, resp.

(1.6) There exists a monotonically increasing function $q: \mathbb{R}_+ \to \mathbb{R}_+$ such that for all $x \in \mathcal{O}_x$, $\varphi \in \mathcal{O}_\varphi$, $\omega \in \Omega$ and a.a. $t \in T$

$$|g(x,\omega,t)| + |g_x(x,\omega,t)| \le q(|x|)$$
$$|f(\varphi,\omega,t)| + |D_1 f(\varphi,\omega,t)| \le q(|\varphi|)$$

(1.7) For every $(u,\varphi) \in \mathcal{U}_{ad} \times \mathcal{O}_\varphi$ the initial value problem

$$x_{t_1} = \varphi, \quad \dot{x}(t) = f(x_t,u(t),t), \text{ a.a. } t \in T$$

has a solution $x(u,\varphi)$, and $x(u,\varphi)_t \in \mathcal{O}_\varphi$ is uniformly bounded for $u \in \mathcal{U}_{ad} := \{u: T \to \Omega, \text{measurable}\}$.

Under these assumptions we can reformulate problem (1.1) - (1.4) as

$$\underset{u,\varphi}{\text{Minimize}} \quad J(u,\varphi) = 1/(t_2-t_1)\int_{t_1}^{t_2} g(x(t,u,\varphi),u(t))dt$$
$$\text{s.t.} \quad \varphi = x(u,\varphi)_{t_2}$$
$$u \in \mathcal{U}_{ad}.$$

One can show that for fixed u the map

$$\varphi \to x(u,\varphi): W_2^1(-r,0;\mathbb{R}^n) \to W_2^1(T;\mathbb{R}^n)$$

is continuously Fréchet differentiable.

2. Structure Theory of Linear Retarded Equations

Linearizing the system equation (1.2) one obtains a retarded equation of the form

(2.1) $\dot{x}(t) = L(t)x_t$,

where $L(t): C := C(-r,0;\mathbb{R}^n) \to \mathbb{R}^n$ is linear and bounded and for all $\varphi \in C$, $\varphi \to L(t)\varphi$ is measurable, with $|L(t)|$ essentially bounded. For the formulation of the maximum principle an adjoint equation is employed. As is well-known, the functional-analytic adjoint and the so-called transposed equation do not coincide. This section sketches the required structural theory developed in [2]cp. also [6].

First observe that $L(\cdot)$ can be represented by a measurable nxn-matrix valued function $\eta(t,\tau)$ in the form

(2.2) $L(t)\varphi = \int_{-r}^{0} [d_\tau \eta(t,\tau)]\varphi(\tau)$, $\varphi \in C$,

where $\eta(t,\cdot) \in NBV(-r,0;\mathbb{R}^{n\times n})$, i.e. $\eta(t,\cdot)$ is of bounded variation and left continuous for $-r < \tau < 0$ and $\eta(t,\tau) = 0$ for $\tau \ge 0$. The evolution of the state $x_t \in C$ is given by a family of strongly continuous evolution operators $\Phi(t,s)$ with $\Phi(t,s)$ compact for $t \ge s+r$.

Associated with (2.1) are the structural operators $F(t): C(-r,0;\mathbb{R}^n) \to C(0,r;\mathbb{R}^n)$

(2.3) $[F(t)\varphi](s) = \varphi(0) + \int_0^s \int_{[-r,-\sigma]} [d_\tau \eta(t+\sigma,\tau)]\varphi(\sigma+\tau)\,d\sigma$

and $G(t) : C(0,r;\mathbb{R}^n) \to C(-r,0;\mathbb{R}^n)$ given by

(2.4) $[G(t)^{-1}\varphi](s) = \varphi(s-r) - \int_0^s \int_{[-\sigma,0]} [d_\tau \eta(t+\sigma,\tau)]\varphi(\sigma+\tau-r)\,d\sigma.$

The operator $F(t_1)$ maps the initial function $x_{t_1} = \varphi$ into the corresponding "forcing term" f^{t_1} of the integrated version of (2.1); and the operator $G(t_1)$ maps this forcing term into the corresponding solution segment x_{t_1+r}. The system (2.1) can also be described by the evolution of the forcing terms $f^t = F(t)x_t$. The corresponding operators are denoted by $\Psi(t,s)$. Then for $t \geq s$

(2.5) $\Phi(t+r,t)=G(t)F(t),\ \Psi(t+r,t)=F(t+r)G(t)$

(2.6) $F(t)\Phi(t,s)=\Psi(t,s)F(s), \Phi(t+r,s+r)G(s)=G(t)\Psi(s,t).$

The adjoints of $F(t)$, $G(t)$ are given by

$[F^*(t)\psi](\tau) = \psi(0) - \int_0^r [\eta^T(t+s,\tau-s)-\eta^T(t+s,-s)]\psi(s)\,ds,$

$[G^*(t)^{-1}\psi](\tau) = \psi(\tau+r) + \int_t^0 \eta^T(t+r+\sigma,\tau-\sigma)\psi(\sigma+r)\,d\sigma,$

$-r \leq \tau < 0,\ \psi \in NBV(0,r;\mathbb{R}^n).$

They are related to the transposed equation

(2.7) $z(t) - z(t_2) = -\int_t^{t_2+r} [\eta^T(\alpha,t-\alpha)-\eta^T(\alpha,t_2-\alpha)]z(\alpha)\,d\alpha \qquad t \leq t_2.$

The evolution of the states $z^t \in NBV(0,r;\mathbb{R}^n)$

(2.8) $z^t(s) = \begin{cases} z(t+s) & , \quad 0 \leq s < r \\ 0 & , \quad s = r \end{cases}$

is described by $\Psi^*(t,s)$:

(2.9) $z^t = \Psi^*(t_2,t)\psi$

for the solution of (2.8) with final condition $z^{t_2} = \psi$.

The restriction of $\Phi(t,s)$ to $W_2^1(-r,0;\mathbb{R}^n)$ is bounded and denoted by $\Phi_W(t,s)$; note that $\Phi_W(t+r,t)$ is compact. The dual space of W_2^1 is in this context appropriately identified with $\mathbb{R}^n \times L_2$ and for $\varphi \in W_2^1$ and $\psi = (\psi^0,\psi^1) \in (W_2^1)^* = \mathbb{R}^n \times L_2$ the duality is given by

$$\langle\varphi,\psi\rangle = \varphi(0)^T\psi^0 + \int \dot{\varphi}\,\psi\,d\theta\ .$$

cp. [2].

Note that $\Psi^*(t+r,t)$ can be continuously extended to a map from $\mathbb{R}^n \times L^2(0,r;\mathbb{R}^n)$ $(= W_2^1(0,r;\mathbb{R}^n)^*)$ into $C(0,r;\mathbb{R}^n)^* = NBV(0,r;\mathbb{R}^n)$, replacing the final condition $z^{t_2} = \psi$ by $(z(t_2),z^{t_2}) = (\psi^0,\psi^1) \in \mathbb{R}^n \times L^2$.

Finally let for $\alpha \in \mathbb{R}^n$

$$(X_0\alpha)(s) := \begin{cases} 0 & -r \leq s < 0 \\ \alpha & s = 0\ . \end{cases}$$

3. The Global Maximum Principle

The proof of the following theorem is based on Ekeland's Variational Principle, and uses arguments similar to those in [4,5].

Theorem 3.1 Suppose that (x^0,u^0) is a strong local minimum with (1.5) - (1.7) holding. Then there exist $\lambda_0 \geq 0$ and $y^* \in W_2^1(-r,0;\mathbb{R}^n)^*$ such that

(i) $y^*(s) := \Phi_W^*(t_2,s)y^* + \lambda_0 \int_s^{t_2} \Phi_W^*(t,s)X_0 g_x(x^0(t),u^0(t),t)\,dt \quad s \in T$

satisfies

(ii) $y^*(t_1) = y^*(t_2) = y^*$

and

(iii) $\lambda_0 g(x^0(s),u^0(s),s) + \langle y^*(s),X_0 f(x_s^0,u^0(s),s)\rangle$

$\leq \lambda_0 g(x^0(s),\omega,s) \qquad + \langle y^*(s),X_0 f(x_s^0,\omega,s)\rangle$

for a.a. $s \in [t_1,t_2-r]$ and all $\omega \in \Omega$.

Here $\Phi(s,t)$ is the evolution operator associated with

$$\dot{x}(t) = \mathcal{D}_1 f(x_t^0,u^0(t),t)x_t,\ t \in T.$$

Proof: The proof uses strong variations of the optimal control u^0 of the following form:

(3.1) $u_{s,\omega}^{\rho}(t) := \begin{cases} u^0(t) & \text{for } t_1 \leq t < s-\rho \text{ and } s < t \leq t_2 \\ \omega & s-\rho \leq t < s; \end{cases}$

here $\rho > 0$, $\omega \in \Omega$ and $s \in (t_1,t_2-r]$.

Where no confusion should possibly arise the shorthand u^ρ for $u_{s,\omega}^\rho$ is used. Note that for ρ small enough, $u^\rho \in \mathcal{U}_{ad}$ and let x^ρ be the corresponding trajectory with initial state $x_{t_1}^\rho := x_{t_1}^0 := \varphi^0$.

First two lemmata are cited from [1, Lemma III.2.2,2.3] (slightly strengthened here). Let $\omega \in \Omega$, $s \in [t_1,t_2]$.

Lemma 3.2 There exists a sequence (ρ_i) tending to zero such that for all $t \in T$

$$\bar{x}(t) := \lim 1/\rho_i\ [x^{\rho_i}(t) - x^0(t)]$$

exists, vanishes on $[t_1,s)$ and coincides on $[s,t_2]$ with the unique solution of

(3.2) $x(t) = 0,\ t < s,\ x(s) = f(x_s^0,\omega,s) - f(x_s^0,u^0(s),s)$

(3.3) $\dot{x}(t) = \mathcal{D}_1 f(x^o_t, u^o(t), t)x_t$, a.a. $t \in [s, t_2]$.

Furthermore

$$\lim |\bar{x} - 1/\rho_i [x^{\rho i} - x^o]|_{W^1_2[s,t_2]; \mathbb{R}^n)} = 0.$$

Lemma 3.3 For a subsequence (ρ^i)

$$\lim 1/\rho_i [J(u^{\rho i}, \varphi^o) - J(u^o, \varphi^o)]$$

$$= g(x^o(s), \omega, s) - g(x^o(s), u^o(s), s) + \int_s^{t_2} g_x(x^o_\sigma, u^o(\sigma), \sigma)$$

$$\{\Phi(\sigma, s) X_o [f_o(x^o_s, \omega, s) - f(x^o_s, u^o(s), s)]\}\ (0) d\sigma .$$

Next we set the stage for an application of Ekeland's variational principle. Consider $\mathcal{U}_{ad}$ in the Ekeland metric

$$d(u,v) := \text{meas}\ \{t: u(t) \neq v(t)\}\ .$$

Then $V := \mathcal{U}_{ad} \times \mathcal{O}'_\varphi$ is a complete metric space where $\mathcal{O}'_\varphi \subset \mathcal{O}_\varphi$ is closed, endowed with the metric induced by the W^1_2-norm. Take a sequence $\delta_n \to 0$, $\delta_n > 0$, and define functionals $F_n: V \to \mathbb{R}$ by

$$F_n(u,\varphi) := [|x(u,\varphi)_{t_2} - \varphi|_2^2 + |J(u,\varphi) - (m - \delta_n)|^2]^{1/2}$$

where $m := J(u^o, \varphi^o)$, and the norm in the first summand is taken in $W^1_2(-r,0; \mathbb{R}^n)$. Then F_n is continuous on V and $F_n(u,\varphi) > 0$ for all $(u,\varphi) \in V$. The sequence (ε_n) with $\varepsilon_n := F_n(u^o, \varphi^o)$ converges to zero and

$$F_n(u^o, \varphi^o) \le \inf \{F_n(u,\varphi) : (u,\varphi) \in V\} + \varepsilon_n .$$

Thus Ekeland's Variational Principle [4, p.456] implies that there exist $(u^n, \varphi^n) \in V$ with the following properties:

(3.4) $0 \le F_n(u^n, \varphi^n) \le \varepsilon_n$

(3.5) $d(u^o, u^n) \le \varepsilon_n^{1/2}$, $|\varphi^o - \varphi^n| \le \varepsilon_n^{1/2}$

(3.6) $F_n(u,\varphi) \ge F_n(u^n, \varphi^n) - \varepsilon_n^{1/2}[d(u,u^n) + |\varphi - \varphi^n|]$ for all $(u,\varphi) \in V$.

Let $s \in T$, $\omega \in \Omega$ and $\rho > 0$, small enough. We shall use the relations above for $(u,\varphi) = (u^{n,\rho}_{s,\omega}, \varphi^n)$ and for $(u,\varphi) = (u^n, \varphi^o + \rho\psi)$.

From (3.6) one obtains for $(u,\varphi) \neq (u^n, \varphi^n)$ the following important inequality:

$$(3.7)\quad -\varepsilon_n^{1/2} \le \{[|x(u,\varphi)_{t_2} - \varphi|_2^2 + |J(u,\varphi) - (m-\delta_n)|^2]^{1/2}$$

$$- [|x(u^n,\varphi^n)_{t_2} - \varphi^n|_2^2 + |J(u^n,\varphi^n) - (m-\delta_n)|^2]^{1/2}\}$$

$$/ | [d(u,u^n) + |\varphi - \varphi^n|] .$$

Let $x^n := x(u^n, \varphi^n)$ and define $y^n \in W^1_2(-r,0; \mathbb{R}^n)^* = \mathbb{R}^n \times L^2(-r,0; \mathbb{R}^n)$ and $\lambda^n \in \mathbb{R}_+$ by

$$(3.8)\quad y^n := [|x^n_{t_2} - \varphi^n|^2 + |J(u^n,\varphi^n) - (m-\delta_n)|^2]^{-1/2}$$

$$[(x^n(t_2), \dot{x}^n_{t_2}) - (\varphi^n(0), \dot{\varphi}^n)]$$

$$(3.9)\quad \lambda^n := [|x^n_{t_2} - \varphi^n|^2 + |J(u^n,\varphi^n) - (m-\delta_n)|^2]^{-1/2}$$

$$[J(u^n,\varphi^n) - (m-\delta_n)] .$$

Let $\Phi^n(t,s)$, $n = 1,2,\ldots$ be the family of evolution operators associated with the linearized equations

(3.10) $\dot{x}(t) = \mathcal{D}_1 f(x^n_t, u^n(t), t)x_t$, a.a. $t \in T$.

Finally, define bounded linear operators

$$T^n : W^1_2(-r,0; \mathbb{R}^n) \to W^1_2(-r,0; \mathbb{R}^n) \text{ by}$$

(3.11) $T^n\psi := \Phi^n_W(t_2, t_1)\psi - \psi$.

Now take $(u,\varphi) = (u^n, \varphi^n + \rho\psi)$ in (3.7). The chain rule implies that F_n is continuously Fréchet differentiable with respect to φ. Thus in the limit for $\rho \to 0$ one gets, noting that ψ is arbitrary in W^1_2,

$$(3.12)\quad (T^{n*}y^n)\psi = -\lambda^n \int_{t_1}^{t_2} g_x(x^n(s), u^n(s), s)[\Phi^n_W(s,t_1)\psi](0)ds$$

for all $\psi \in W^1_2$.

Next we consider the limit for $n \to \infty$. By (3.5), $|\varphi^n - \varphi^o| \to 0$. Hence by continuity, resp. continuous Fréchet differentiability we get

$|x^n - x^o| \to 0$, $|\Phi^n_W(s,t_1) - \Phi_W(s,t_1)| \to 0$ for all $s \in T$

$|T^n - T^o| \to 0$ and $|\Phi_W(\sigma,s)|$ is uniformly bounded. Recall that $|\lambda^n| \le 1$ and $|y^n|_{(W^1_2)^*} \le 1$. We have to exlude that both sequences (λ^n) and (y^n) converge to zero. Suppose first that the Fredholm operator $\Phi_W(t_2,t_1) - \text{Id}$ is surjective, hence an isomorphism on $W^1_2(-r,0; \mathbb{R}^n)$. This implies that for sufficiently large n also $T^n = \Phi^n_W(t_2,t_1) - \text{Id}$ is an isomorphism cp. [3, Lemma VII.6.1] and hence also T^{n*} is an isomorphism. Suppose that there exists a subsequence of (λ^n) again denoted by (λ^n) converging to zero. This yields

$$|y^n|_{(W^1_2)^*} \to 1.$$

This implies existence of $\psi^n \in W^1_2(-r,0; \mathbb{R}^n)$ with $|\psi^n| \le c_o$ and $(T^{n*}y^n)\psi^n \ge 1/2$; but from (3.12) one obtains

$$|(T^{n*}y^n)\psi^n| \le |\lambda^n| c_1$$

here c_o, c_1 are constants independent of n. This is a contradiction.

Let λ_o be a clusterpoint of (λ^n) and y^* be a weak clusterpoint of (y^n). Then (3.12) implies for all $\varphi \in W_2^1(-r,0;\mathbb{R}^n)$

$$(T^{o*}y^*)\varphi = [(\Phi_W(t_2,t_1)^* - Id)y^*]\varphi$$
$$= -\lambda_o \int_{t_1}^{t_2} g_x(x^o(s),u^o(s),s)\ [\Phi(s,t)\varphi](0)ds.$$

Thus (i) and (ii) hold.

The other case can be treated using the Hahn-Banach-Theorem.

Now take $(u,\varphi) = (u^{n,\rho},\varphi^n)$ in (3.7) in order to derive the maximum condition. Again take first the limit for $\rho \to 0$ and then for $n \to \infty$. Then use the adjoint equation in order to get (iii).

Theorem 3.1 has an abstract form. The structural theory of retarded equations exposed in section 2 allows to derive the concrete form of the optimality condition. The following lemma is crucial.

Lemma 3.4 Suppose that $t_2 \geq t_1 + r$. Then $y^* \in NBV(0,r;\mathbb{R}^n)$ and the equations (i) and (ii) imply that there exists $\psi \in NBV(0,r;\mathbb{R}^n)$ with $y^* = F^*(t_2)\psi$.

Proof: This follows by an analysis of

$$y^* = y^*(t_1) = \Phi_W^*(t_2,t_1)y^*$$
$$+ \lambda_o \int_{t_1}^{t_2} \Phi_W^*(t,t_1)X_o g_x(x^o(t),u^o(t),t)dt$$

and (2.5). Observe that $y^*(t_1) \in NBV(0,r;\mathbb{R}^n)$.

Using (2.5), (2.6) one obtains the following global maximum principle.

Corollary 3.5 Let $t_2 \geq t_1 + r$ and suppose that (x^o,u^o) is a strong local minimum with (1.5) - (1.7) holding.

Then there exist $\lambda_o \geq 0$ and a solution y of the transposed equation

(i) $$\frac{d}{ds}\{y(s)+\int_s^{t_2+r}[\eta^T(\alpha,s-\alpha)-\eta^T(\alpha,t_2-\alpha)]y(\alpha)d\alpha\}$$
$$= \lambda_o g_x(x^o(s),u^o(s),s), \quad s \in T$$

such that

(ii) $y^{t_1} = y^{t_2}$ and $(\lambda_o,y^{t_2}) \neq (0,0)$ in $\mathbb{R}\times NBV(0,r;\mathbb{R}^n)$.

(iii) $$\lambda_o g(x^o(s),u^o(s),s) + y(s)^T f(x_s^o,u^o(s),s)$$
$$\leq \lambda_o g(x^o(s),\omega,s) + y(s)^T f(x_s^o,\omega,s)$$

for a.a. $s \in [t_1,t_2-r]$ and all $\omega \in \Omega$;

here η is given by the representation (cp. (2.2))

$$\mathcal{D}_1 f(x_s^o,u^o(s),s)\varphi = \int_{-r}^{0}[d_s\eta(t,s)]\varphi(s),\ \varphi \in C(-r,0;\mathbb{R}^n).$$

Remark In the theorems above, the minimum condition (iii) is obtained only on $[t_1,t_2-r]$. Suppose, however, that $|\Phi^n(t_2,t_1)x|_{\mathbb{R}^n\times L^2(-r,0;\mathbb{R}^n)}$, $n = 1,2,\ldots$, is uniformly bounded for all $x \in C(-r,0;\mathbb{R}^n)$ with $|x|_{\mathbb{R}^n\times L^2} \leq 1$ (this is e.g. the case for systems with constant delays, cp. [2]).

Then one can show that (iii) holds on the whole interval $T = [t_1,t_2]$. One has to modify the definition of F_n by taking the $\mathbb{R}^n \times L^2$-norm instead of the W_2^1-norm and then argue with the continuous extensions of T^n to $\mathbb{R}^n \times L^2$.

References

[1] F. Colonius, Optimal Periodic Control, Habilitationsschrift, Universität Bremen, Bremen 1986.

[2] F. Colonius, A. Manitius, D. Salamon, Structure theory and duality for time varying functional differential equations (submitted).

[3] N. Dunford, J. Schwartz, Linear Operators, Part I: General Theory, Wiley 1967.

[4] I. Ekeland, Nonconvex minimization problems, Bull. Amer. Math. Soc., 1 (New Series)(1979), 443-474.

[5] H.O. Fattorini, The maximum principle for nonlinear nonconvex systems in infinite dimensional spaces, Distributed Parameter Systems, F. Kappel, K. Kunisch, W. Schappacher, eds., Springer-Verlag 1986.

[6] J. Hale, Theory of Functional Differential Equations, Springer-Verlag, 1977.

[7] Li, X.J., Maximum principle of optimal periodic control for functional differential systems, to appear in: J. Optim. Theory Appl.