2

Optimal control of ordinary differential systems. Optimality conditions

This chapter and the next one are devoted to some basic ideas and techniques in optimal control theory of ordinary differential systems. We do not treat the optimal control problem or Pontryagin's principle in their most general form; instead we prefer a direct approach for some significant optimal control problems in life sciences and economics governed by ordinary differential systems. We point out the main steps in the study of an optimal control problem for each investigated example. These steps are similar for all examples. There are, however, specific technical difficulties for each investigated problem.

The main goal of this chapter is to prove the existence of an optimal control and to obtain first-order necessary conditions of optimality (Pontryagin's principle) for some significant optimal control problems. The necessary optimality conditions give valuable information about the structure of the optimal control. Numerical algorithms to approximate the optimal control and corresponding MATLAB® programs are indicated.

A general formulation of Pontryagin's principle for optimal control problems related to ordinary differential systems can be found in [Bar93] and [Bar94].

2.1 Basic problem. Pontryagin's principle

A quite general optimal control problem governed by an ordinary differential system can be formulated in the following form,

$$\text{Maximize } \mathcal{L}(u, x^u) = \int_0^T G(t, u(t), x^u(t))dt + \varphi(x^u(T)), \qquad \textbf{(P1)}$$

subject to $u \in K \subset L^2(0,T;\mathbb{R}^m)$ $(T > 0)$, where x^u is the Carathéodory solution to

$$\begin{cases} x'(t) = f(t, u(t), x(t)), & t \in (0,T) \\ x(0) = x_0. \end{cases} \qquad (2.1)$$

S. Aniţa et al. *An Introduction to Optimal Control Problems in Life Sciences and Economics*, Modeling and Simulation in Science, Engineering and Technology, DOI 10.1007/978-0-8176-8098-5_2, © Springer Science+Business Media, LLC 2011

Here

$$G : [0,T] \times I\!R^m \times I\!R^N \to I\!R,$$
$$\varphi : I\!R^N \to I\!R,$$
$$f : [0,T] \times I\!R^m \times I\!R^N \to I\!R^N,$$

$x_0 \in I\!R^N$, $m, N \in I\!N^*$, and $K \subset L^2(0,T; I\!R^m)$ is a closed convex subset. From now all elements of an $I\!R^n$, $n \in I\!N^*$, are considered as column vectors.
Recall that a Carathéodory solution (we call it simply a solution) to (2.1) is a function x^u that belongs to $AC([0,T]; I\!R^N)$ (see Appendix A.3), and satisfies

$$\begin{cases} (x^u)'(t) = f(t, u(t), x^u(t)) \ \text{ a.e. } t \in (0,T) \\ x^u(0) = x_0. \end{cases}$$

$L^2(0,T; I\!R^m)$ is the set of the controllers.

This optimal control problem can be reformulated as the following minimization problem,

$$\text{Minimize } \{-\mathcal{L}(u, x^u)\},$$

subject to $u \in K \subset L^2(0,T; I\!R^m)$.

We assume here that for any $u \in L^2(0,T; I\!R^m)$, Problem (2.1) admits a unique solution, denoted by x^u. Equation (2.1) is called the state problem (equation).

- u ($\in K$) is called the control (or controller). This is a constrained control because $u \in K$, and K is a subset of $L^2(0,T; I\!R^m)$.
- x^u is the state corresponding to the control u, and the mapping.
- $u \mapsto \mathcal{L}(u, x^u) = \Phi(u)$ is the cost functional.

We say that $u^* \in K$ is an optimal control for Problem (P1) if

$$\mathcal{L}(u^*, x^{u^*}) \geq \mathcal{L}(u, x^u),$$

for any $u \in K$. The pair (u^*, x^{u^*}) is called an optimal pair and $\mathcal{L}(u^*, x^{u^*})$ is the optimal value of the cost functional. We also say that (u^*, x^*) is an optimal pair if u^* is an optimal control and $x^* = x^{u^*}$.

Let $u^* \in K$ be an optimal control for (P1); that is,

$$\int_0^T G(t, u^*(t), x^{u^*}(t))dt + \varphi(x^{u^*}(T)) \geq \int_0^T G(t, u(t), x^u(t))dt + \varphi(x^u(T)),$$

for any $u \in K$.

We assume that the following succession of operations and arguments is allowed (under certain hypotheses – including Gâteaux differentiability (see Appendix A.1.5) – on G, φ, and f). We use the notations:

$$\begin{cases} f_u = \dfrac{\partial f}{\partial u}, \quad f_x = \dfrac{\partial f}{\partial x} \\[2ex] G_u = \dfrac{\partial G}{\partial u}, \quad G_x = \dfrac{\partial G}{\partial x} \\[2ex] \varphi_x = \dfrac{\partial \varphi}{\partial x} \end{cases}$$

(see also Appendix A.1.5). Here G_u, G_x, and φ_x are considered as column vectors.

Assume that the function defined on $L^2(0, T; \mathbb{R}^m)$, $u \mapsto x^u$ is everywhere Gâteaux differentiable. We denote this differential by dx^u.

Consider

$$V = \{v \in L^2(0, T; \mathbb{R}^m);\ u^* + \varepsilon v \in K \text{ for any } \varepsilon > 0 \text{ sufficiently small}\}.$$

For any $v \in V$, we define $z = dx^{u^*}(v)$; z is the solution to

$$\begin{cases} z'(t) = f_u(t, u^*(t), x^{u^*}(t))v(t) + f_x(t, u^*(t), x^{u^*}(t))z(t), \quad t \in (0, T) \\ z(0) = 0. \end{cases} \tag{2.2}$$

For an arbitrary but fixed $v \in V$ we have that

$$\int_0^T G(t, u^*(t), x^{u^*}(t))dt + \varphi(x^{u^*}(T)) \geq \int_0^T G(t, u^*(t) + \varepsilon v(t), x^{u^*+\varepsilon v}(t))dt + \varphi(x^{u^*+\varepsilon v}(T)),$$

and consequently

$$\int_0^T \frac{1}{\varepsilon}\left[G(t, u^*(t) + \varepsilon v(t), x^{u^*+\varepsilon v}(t)) - G(t, u^*(t), x^{u^*}(t))\right] dt + \frac{1}{\varepsilon}\left[\varphi(x^{u^*+\varepsilon v}(T)) - \varphi(x^{u^*}(T))\right] \leq 0,$$

for any $v \in V$, and for any $\varepsilon > 0$ sufficiently small.
We pass to the limit in the last inequality ($\varepsilon \to 0+$) and we get that

$$\int_0^T [v(t) \cdot G_u(t, u^*(t), x^{u^*}(t)) + z(t) \cdot G_x(t, u^*(t), x^{u^*}(t))]dt + z(T) \cdot \varphi_x(x^{u^*}(T)) \leq 0 \tag{2.3}$$

(here $\cdot$ denotes the usual scalar product on $\mathbb{R}^m$ as well as on $\mathbb{R}^N$), for any $v \in V$.

Let p be the Carathéodory solution (we assume that this solution exists and is unique), that we simply call the solution, to the adjoint problem (equation):

$$\begin{cases} p'(t) = -f_x^*(t, u^*(t), x^{u^*}(t))p(t) - G_x(t, u^*(t), x^{u^*}(t)), \quad t \in (0, T) \\ p(T) = \varphi_x(x^{u^*}(T)) \end{cases} \tag{2.4}$$

(p is called the adjoint state; the equation in (2.4) is linear).
Recall that if $A : I\!R^k \longrightarrow I\!R^s$ is a linear (and bounded) operator (A may be identified with a matrix, also denoted by A), then its adjoint operator $A^* : I\!R^s \longrightarrow I\!R^k$ (also linear and bounded) may be identified with the transpose of matrix A, and denoted also by A^* (or A^T).

By multiplying (2.2) by p and integrating by parts on $[0, T]$ we get that

$$\begin{aligned}
& z(T) \cdot p(T) - \int_0^T z(t) \cdot p'(t) dt \\
& = \int_0^T [f_u(t, u^*(t), x^{u^*}(t))v(t) + f_x(t, u^*(t), x^{u^*}(t))z(t)] \cdot p(t)dt,
\end{aligned}$$

for any $v \in V$. By (2.4) we obtain that

$$\begin{aligned}
& z(T) \cdot \varphi_x(x^{u^*}(T)) \\
& + \int_0^T z(t) \cdot [f_x^*(t, u^*(t), x^{u^*}(t))p(t) + G_x(t, u^*(t), x^{u^*}(t))]dt \\
& = \int_0^T [v(t) \cdot f_u^*(t, u^*(t), x^{u^*}(t))p(t) + z(t) \cdot f_x^*(t, u^*(t), x^{u^*}(t))p(t)]dt,
\end{aligned}$$

and consequently

$$\begin{aligned}
& \int_0^T z(t) \cdot G_x(t, u^*(t), x^{u^*}(t))dt + z(T) \cdot \varphi_x(x^{u^*}(T)) \\
& = \int_0^T v(t) \cdot f_u^*(t, u^*(t), x^{u^*}(t))p(t)dt,
\end{aligned}$$

for any $v \in V$. By (2.3) we finally get that

$$\int_0^T v(t) \cdot [G_u(t, u^*(t), x^{u^*}(t)) + f_u^*(t, u^*(t), x^{u^*}(t))p(t)]dt \leq 0,$$

for any $v \in V$, which means

$$G_u(\cdot, u^*, x^{u^*}) + f_u^*(\cdot, u^*, x^{u^*})p \in N_K(u^*), \tag{2.5}$$

where $N_K(u^*)$ is the normal cone at K in u^* (see Appendix A.1.4).
We get the same conclusion if we multiply (2.4) by z (after a similar argumentation).

Equations (2.1), (2.4), and (2.5) represent Pontryagin's (or maximum) principle and (2.4) and (2.5) are the first-order necessary conditions of optimality (optimality conditions) for the given optimal control problem.

The main goal now is to use the maximum principle in order to calculate an optimal control u^* or to approximate it by using an appropriate numerical scheme. In order to use Condition (2.5) we need to determine the set $N_K(u^*)$.

If we take, for example, $K = L^2(0,T;\mathbb{R}^m)$, then for any $u \in K = L^2(0,T;\mathbb{R}^m)$, $N_K(u) = \{0\} \subset L^2(0,T;\mathbb{R}^m)$.

If we take $m = 1$, and

$$K = \{w \in L^2(0,T);\ L_1 \le w(t) \le L_2 \text{ a.e. } t \in (0,T)\},$$

where $L_1, L_2 \in \mathbb{R}$, $L_1 < L_2$, then for any $u \in K$ we have

$$N_K(u) = \{w \in L^2(0,T); \quad w(t) \ge 0 \quad \text{if } u(t) = L_2, w(t) \le 0 \text{ if } u(t) = L_1, \\ w(t) = 0 \quad \text{if } L_1 < u(t) < L_2 \text{ a.e. } t \in (0,T)\}$$

(see Appendix A.1.4).

A general scheme to prove the existence of an optimal control u^* is the following one.
Let

$$d = \sup_{u \in K} \mathcal{L}(u, x^u) \in \mathbb{R}.$$

For any $n \in \mathbb{N}^*$, there exists $u_n \in K$, such that

$$d - \frac{1}{n} < \mathcal{L}(u_n, x^{u_n}) \le d.$$

Step 1: Prove that there exists a subsequence $\{u_{n_k}\}$ such that

$$u_{n_k} \longrightarrow u^* \text{ weakly in } L^2(0,T;\mathbb{R}^m).$$

If for example, K is bounded, then the last conclusion follows immediately.

Inasmuch as K is a closed convex subset of $L^2(0,T;\mathbb{R}^m)$, K is also weakly closed, and consequently $u^* \in K$.

Step 2: Prove that there exists a subsequence of $\{x^{u_{n_k}}\}$, denoted by $\{x^{u_{n_r}}\}$, convergent to x^{u^*} in $C([0,T];\mathbb{R}^N)$ (sometimes the convergence in $L^2(0,T;\mathbb{R}^N)$ is enough).

Step 3: From

$$d - \frac{1}{n_r} < \mathcal{L}(u_{n_r}, x^{u_{n_r}}) \le d,$$

we get (by passing to the limit) that

$$\mathcal{L}(u^*, x^{u^*}) = d,$$

and consequently u^* is an optimal control for problem (P1).
Notice that we can derive (2.4) and (2.5) by using the Hamiltonian H, defined by

$$H(t,u,x,p) = G(t,u,x) + f(t,u,x) \cdot p.$$

If we take

$$x' = H_p$$

we get the state equation. By

$$p' = -H_x$$

we get the adjoint equation and by

$$H_u \in N_K(u^*),$$

we get (2.5).

Let us mention that some authors consider the following problem as the adjoint problem:

$$\begin{cases} p'(t) = -f_x^*(t, u^*(t), x^{u^*}(t))p(t) + G_x(t, u^*(t), x^{u^*}(t)), \;\; t \in (0, T) \\ p(T) = -\varphi_x(x^{u^*}(T)) \end{cases}$$

The solution to this problem is $p = -\tilde{p}$, where $\tilde{p}$ is the solution to (2.4).

Condition (2.5) becomes

$$G_u(\cdot, u^*, x^{u^*}) - f_u^*(\cdot, u^*, x^{u^*})p \in N_K(u^*),$$

and the Hamiltonian H is:

$$H(t, u, x, p) = -G(t, u, x) + f(t, u, x) \cdot p.$$

We, however, use both conventions (for the adjoint problem) in the next chapters.

In most situations (2.1) appears as a semilinear problem; that is, f has the following form,

$$f(t, u, x) = Ax + \tilde{f}(t, u, x),$$

where $A : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is a (particular) linear operator. Then

$$f_u = \tilde{f}_u, \quad f_x = A + \tilde{f}_x.$$

Several optimal control problems related to age-structured models, semilinear parabolic equations or to integroparabolic equations may be written in the abstract form (P1)–(2.1), where

$$G : [0, T] \times U \times X \to \mathbb{R},$$
$$\varphi : X \to \mathbb{R}$$

(U, X are appropriate real Hilbert spaces), $x_0 \in X$, and $K \subset L^2(0,T;U)$ is a closed convex subset. Here f has the above mentioned form, and A is a linear (possibly unbounded) operator, $A : D(A) \subset X \to X$ (see Chapters 4 and 5).

Here we have presented only a general scheme and not a rigorous proof of the maximum principle.

In the next sections we illustrate how this scheme works for significant examples of optimal control problems in life sciences and economics governed by ordinary differential systems. We deduce the maximum principle again for all these examples in a rigorous manner. We use the maximum principle to calculate or to approximate optimal control. Chapters 4 and 5 are devoted to control problems governed by partial differential equations. As announced the scheme is the same, but there are, of course, more technical difficulties.

2.2 Maximizing total consumption

We consider a mathematical model of a simplified economy. Let $x(t)$ be the rate of production at the moment $t \geq 0$ (the economical output). We have

$$x(t) = I(t) + C(t), \quad t \geq 0,$$

where

- $I(t)$ is the rate of investment at the moment t.
- $C(t)$ is the rate of consumption at the moment t.

Denote by $u(t) \in [0,1]$ the part of production $x(t)$ that is allocated to investment at moment t; that is,

$$I(t) = u(t)x(t).$$

We obtain that

$$C(t) = (1 - u(t))x(t), \quad t \geq 0.$$

We deal with the simple case when the production growth rate is proportional to the rate of investment. This means

$$x'(t) = \gamma u(t)x(t),$$

where $\gamma \in (0, +\infty)$.

We introduce a "utility" function $F(C)$, and we wish to find out the control that maximizes the welfare integral

$$\int_0^T e^{-\delta t} F(C(t))dt.$$

Here $T > 0$, and $\delta \geq 0$ is a discount rate (a measure of preference for earlier rather than later consumption).

We simplify our model by taking $F(C) = C$ and $\delta = 0$. The total consumption on the time interval $[0, T]$ is

$$\int_0^T C(t)dt = \int_0^T (1 - u(t))x(t)dt.$$

We therefore obtain the following optimal control problem (see [Bar94]),

$$\text{Maximize} \int_0^T (1 - u(t))x^u(t)dt, \qquad (\mathbf{P2})$$

subject to $u \in L^2(0, T)$, $0 \leq u(t) \leq 1$ a.e. $t \in (0, T)$, where x^u is the solution of

$$\begin{cases} x'(t) = \gamma u(t)x(t), & t \in (0, T) \\ x(0) = x_0 > 0. \end{cases} \qquad (2.6)$$

The problem seeks to find the control u that maximizes total consumption on the time interval $[0, T]$.

The solution x^u to (2.6) is given by

$$x^u(t) = x_0 \exp\left(\int_0^t \gamma u(s)ds\right), \quad t \in [0, T].$$

Problem (P2) is a particular case of (P1), for $m = 1$, $N = 1$,

$$G(t, u, x) = (1 - u)x,$$
$$\varphi(x) = 0,$$
$$f(t, u, x) = \gamma ux$$

and

$$K = \{w \in L^2(0, T);\ 0 \leq w(t) \leq 1 \text{ a.e. } t \in (0, T)\}.$$

Existence of an optimal pair for (P2)

Define

$$\Phi(u) = \int_0^T (1 - u(t))x^u(t)dt, \quad u \in K$$

and let

$$d = \sup_{u \in K} \Phi(u).$$

Because for any $u \in K$ we have that

$$0 < x^u(t) \leq x_0 e^{\gamma t}, \quad t \in [0, T],$$

then we get that

$$0 \leq \Phi(u) = \int_0^T (1 - u(t)) x^u(t) dt \leq x_0 T e^{\gamma T}.$$

In conclusion $d \in [0, +\infty)$.

So, for any $n \in I\!N^*$, there exists $u_n \in K$ such that

$$d - \frac{1}{n} < \Phi(u_n) \leq d. \tag{2.7}$$

K is a bounded subset of $L^2(0, T)$, therefore it follows that there exists a subsequence $\{u_{n_k}\}_{k \in I\!N^*}$ such that

$$u_{n_k} \longrightarrow u^* \quad \text{weakly in } L^2(0, T). \tag{2.8}$$

The limit u^* belongs to K because K is a closed convex subset of $L^2(0, T)$, and so it is weakly closed. The last convergence and the explicit formula for x^u imply that

$$x^{u_{n_k}} \longrightarrow x^{u^*} \quad \text{in } L^2(0, T). \tag{2.9}$$

By (2.7) we get that

$$d - \frac{1}{n_k} < \int_0^T (1 - u_{n_k}(t)) x^{u_{n_k}}(t) dt \leq d \quad \text{for any } k \in I\!N^*. \tag{2.10}$$

By (2.8) and (2.9) we obtain (we pass to the limit in (2.10)) that

$$d = \int_0^T (1 - u^*(t)) x^{u^*}(t) dt,$$

that is, (u^*, x^{u^*}) is an optimal pair (and u^* is an optimal control) for (P2).

In order to simplify the notations we denote $x^* := x^{u^*}$.

The maximum principle

For an arbitrary but fixed $v \in V = \{w \in L^2(0, T);\ u^* + \varepsilon w \in K$ for any $\varepsilon > 0$ sufficiently small$\}$ we denote by z the solution to

$$\begin{cases} z'(t) = \gamma u^*(t) z(t) + \gamma v(t) x^*(t), & t \in (0, T) \\ z(0) = 0. \end{cases} \tag{2.11}$$

z is given by

$$z(t) = \int_0^t \exp\{\int_s^t \gamma u^*(\tau)d\tau\}\gamma v(s)x^*(s)ds, \quad t \in [0,T]. \tag{2.12}$$

Inasmuch as

$$\int_0^T (1-u^*(t))x^*(t)dt \geq \int_0^T (1-u^*(t)-\varepsilon v(t))x^{u^*+\varepsilon v}(t)dt,$$

for any $\varepsilon > 0$ sufficiently small, we get that

$$\int_0^T [(1-u^*(t))\frac{x^{u^*+\varepsilon v}(t) - x^*(t)}{\varepsilon} - v(t)x^{u^*+\varepsilon v}(t)]dt \leq 0. \tag{2.13}$$

Let us prove that

$$x^{u^*+\varepsilon v} \longrightarrow x^* \quad \text{in } C([0,T])$$

and

$$\frac{x^{u^*+\varepsilon v} - x^*}{\varepsilon} \longrightarrow z \quad \text{in } C([0,T]),$$

as $\varepsilon \to 0+$.

Indeed, for any $\varepsilon > 0$ sufficiently small we have

$$x^{u^*+\varepsilon v}(t) = x_0 \exp\{\gamma \int_0^t (u^*(s)+\varepsilon v(s))ds\}$$

$$= x^{u^*}(t)\exp\{\varepsilon\gamma \int_0^t v(s)ds\}, \quad t \in [0,T],$$

which implies that

$$|x^{u^*+\varepsilon v}(t) - x^{u^*}(t)| = |x^{u^*}(t)| \cdot |\exp\{\varepsilon\gamma \int_0^t v(s)ds\} - 1|, \quad t \in [0,T].$$

Because

$$|\exp\{\varepsilon\gamma \int_0^t v(s)ds\} - 1| \longrightarrow 0,$$

uniformly on $[0,T]$, we may infer that

$$x^{u^*+\varepsilon v} \longrightarrow x^* \quad \text{in } C([0,T]).$$

For any $\varepsilon > 0$ sufficiently small we consider

$$w_\varepsilon(t) = \frac{x^{u^*+\varepsilon v} - x^*}{\varepsilon} - z(t), \quad t \in [0,T].$$

w_ε is the solution to

$$\begin{cases} w'(t) = \gamma u^*(t)w(t) + \gamma v(t)[x^{u^*+\varepsilon v}(t) - x^{u^*}(t)], & t \in (0,T) \\ w(0) = 0, \end{cases}$$

and is given by

$$w_\varepsilon(t) = \gamma \int_0^t \exp\{\gamma \int_s^t u^*(\tau)d\tau\} v(s)[x^{u^*+\varepsilon v}(s) - x^{u^*}(s)]ds, \quad t \in [0,T].$$

By taking into account the first convergence we deduce that

$$w_\varepsilon \longrightarrow 0 \quad \text{in } C([0,T]),$$

and consequently

$$\frac{x^{u^*+\varepsilon v} - x^*}{\varepsilon} \longrightarrow z \quad \text{in } C([0,T]).$$

By (2.13) we obtain now that

$$\int_0^T [(1 - u^*(t))z(t) - v(t)x^*(t)]dt \le 0. \tag{2.14}$$

Let us denote by p the solution to

$$\begin{cases} p'(t) = -\gamma u^*(t)p(t) + u^*(t) - 1, & t \in (0,T) \\ p(T) = 0. \end{cases} \tag{2.15}$$

p is given by

$$p(t) = -\int_t^T \exp\{\int_t^s \gamma u^*(\tau)d\tau\}(u^*(s) - 1)ds, \quad t \in [0,T].$$

If we multiply the differential equation in (2.15) by z and integrate over $[0,T]$ we get that

$$\int_0^T p'(t)z(t)dt = -\int_0^T \gamma u^*(t)p(t)z(t)dt + \int_0^T (u^*(t) - 1)z(t)dt.$$

If we integrate by parts it follows by ((2.11) and (2.15)) that

$$-\int_0^T p(t)z'(t)dt = -\int_0^T \gamma u^*(t)p(t)z(t)dt + \int_0^T (u^*(t) - 1)z(t)dt.$$

We again use (2.11) to obtain

$$\begin{aligned} &-\int_0^T \gamma u^*(t)z(t)p(t)dt - \int_0^T \gamma v(t)x^*(t)p(t)dt \\ &= -\int_0^T \gamma u^*(t)p(t)z(t)dt + \int_0^T (u^*(t) - 1)z(t)dt, \end{aligned}$$

which implies

$$\int_0^T (1-u^*(t))z(t)dt = \int_0^T \gamma v(t)x^*(t)p(t)dt.$$

This last relation and (2.14) imply that

$$\int_0^T x^*(t)(\gamma p(t)-1)v(t)dt \le 0, \tag{2.16}$$

for any $v \in V$. This is equivalent to

$$(\gamma p - 1)x^* \in N_K(u^*).$$

If we take into account the structure of $N_K(u^*)$ we may conclude that

$$u^*(t) = \begin{cases} 0 \ \text{ if } \gamma p(t) - 1 < 0 \\ \\ 1 \ \text{ if } \gamma p(t) - 1 > 0, \end{cases} \tag{2.17}$$

a.e. $t \in (0,T)$.

Let us give a direct proof of (2.17) starting from (2.16).

Denote by

$$A = \{t \in (0,T);\ \gamma p(t) - 1 < 0\}.$$

We prove that $u^*(t) = 0$ a.e. on A.

Assume by contradiction that there exists $\tilde{A} \subset A$, with $meas(\tilde{A}) > 0$ ($meas$ denotes the Lebesgue measure; see Appendix A.1.1) such that $u^*(t) > 0$ a.e. in $\tilde{A}$. We can choose $v \in L^2(0,T)$ such that $v(t) < 0$ a.e. in $\tilde{A}$, $v(t) = 0$ a.e. in $(0,T) \setminus \tilde{A}$ and $0 \le u^*(t) + \varepsilon v(t) \le 1$ a.e. in $(0,T)$. It follows that

$$\int_0^T x^*(t)(\gamma p(t)-1)v(t)dt = \int_{\tilde{A}} x^*(t)(\gamma p(t)-1)v(t)dt > 0,$$

because $v(t) < 0$, $\gamma p(t) - 1 < 0$, $x^*(t) > 0$ on $\tilde{A}$, and $meas(\tilde{A}) > 0$. This is, of course, in contradiction to (2.16).

In the same manner it follows that

$$u^*(t) = 1 \ \text{ a.e. } t \in \{s \in (0,T);\ \gamma p(s) - 1 > 0\}.$$

The conclusion follows.

Remark 2.1. Equations (2.6), (2.15), and (2.17) represent the maximum principle and (2.15) and (2.17) are the first-order necessary optimality conditions for (P2).

Calculation of the optimal control u^*

Our next goal is to use Pontryagin's principle in order to get more information on the optimal control u^*. We show that for our particular problem we are able to calculate it exactly.

Let $(T-\eta, T]$ $(\eta > 0)$ be a maximal interval where the continuous function p satisfies $\gamma p(t) < 1$. By (2.17) and (2.15) we see that

$$p'(t) = -1, \quad t \in [T-\eta, T],$$

which implies that

$$p(t) = T - t \quad t \in [T-\eta, T].$$

Therefore, if $\gamma T > 1$ we have

$$p(t) = T - t \quad t \in [T - \frac{1}{\gamma}, T]$$

and

$$u^*(t) = 0 \ \text{ a.e. } t \in (T - \frac{1}{\gamma}, T).$$

Because $p(T - (\frac{1}{\gamma})) = \frac{1}{\gamma}$, we see that $p'(t) \leq 0$ on a maximal interval $(T - (1/\gamma) - \delta, T - (1/\gamma)]$ $(\delta > 0)$, and therefore $\gamma p(t) > 1$ on this interval. It also follows that

$$\begin{cases} p'(t) = -\gamma p(t) \\ \\ u^*(t) = 1 \end{cases} \quad \text{on } (T - \frac{1}{\gamma} - \delta, T - \frac{1}{\gamma}).$$

Consequently

$$p(t) = \frac{1}{\gamma} \exp\{\gamma(T - \frac{1}{\gamma} - t)\} \quad t \in [T - \frac{1}{\gamma} - \delta, T - \frac{1}{\gamma}].$$

This implies that $\delta = T - (1/\gamma)$ and that $u^*(t) = 1$ a.e. $t \in [0, T - (1/\gamma))$.

The conclusion is that

- If $\gamma T > 1$, then

$$u^*(t) = \begin{cases} 1 & \text{if } t \in [0, T - \frac{1}{\gamma}) \\ \\ 0 & \text{if } t \in [T - \frac{1}{\gamma}, T]; \end{cases} \tag{2.18}$$

- If $\gamma T \leq 1$, then

$$u^*(t) = 0, \quad t \in [0, T]. \tag{2.19}$$

This means that if the time interval is sufficiently long, then for a certain interval of time the rate of investment should be maximal. After that we do not invest any more (we just put everything for consumption).

A control u^* that takes values in a finite set $\{\alpha_1,\ \alpha_2, \ldots, \alpha_k\}$, and $(u^*)^{-1}(\alpha_i)$ is a measurable set for any $i \in \{1,\ 2, \ldots,\ k\}$ is called a bang-bang control.

If there exist $t_0 < t_1 < \cdots < t_k$ such that u^* is constant on any interval (t_{i-1}, t_i) $(i = \overline{1,k})$, then u^* is a bang-bang control on (t_0, t_k) and $t_1,\ t_2, \ldots,\ t_{k-1}$ are called switching points.

Remark 2.2. (i) The optimal control in our example is a bang-bang control and has at most one switching point, namely $T - (1/\gamma)$.

(ii) For our example we were able to calculate the optimal control. The form of the optimal control is given by (2.18) and (2.19). This is, of course, a fortunate situation.

(iii) After identifying $\mathcal{L}$, G, φ, f, and K we were able to write Pontryagin's principle formally. What we have done in this section was to prove it and use it in order to calculate the optimal control.

2.3 Maximizing the total population in a predator–prey system

The following Lotka–Volterra system,

$$\begin{cases} x'(t) = r_1 x(t) - \mu_1 x(t) y(t), & t \in (0, T) \\ y'(t) = -r_2 y(t) + \mu_2 x(t) y(t), & t \in (0, T) \end{cases}$$

$(T > 0)$ describes the dynamics of a predator–prey system on the time interval $(0, T)$. Here $x(t)$ represents the density of the prey population at moment t, and $y(t)$ the density of predators at moment t.

- $r_1 > 0$ is the intrinsic growth rate of prey in the absence of predators.
- $r_2 > 0$ is the decay rate of the predator population in the absence of prey.
- μ_1 and μ_2 are positive constants; $\mu_1 y(t)$ is the additional mortality rate of prey at moment t, due to predation (it is proportional to the predator population density); and $\mu_2 x(t)$ is the additional growth rate of prey at moment t, due to the presence of prey (it is proportional to the prey population density).

A more general model for the predator–prey system has been presented in Section 1.7.

If the prey are partially separated from predators then the functional response to predation changes and the system becomes

$$\begin{cases} x'(t) = r_1 x(t) - \mu_1 u(t) x(t) y(t), & t \in (0, T) \\ y'(t) = -r_2 y(t) + \mu_2 u(t) x(t) y(t), & t \in (0, T), \end{cases} \tag{2.20}$$

where $1 - u(t)$ represents the segregation rate at moment t $(0 \le u(t) \le 1)$.

Let the initial conditions be

$$\begin{cases} x(0) = x_0 > 0 \\ y(0) = y_0 > 0. \end{cases} \tag{2.21}$$

We are interested in maximizing the total number of individuals of both populations at moment $T > 0$. The problem may be reformulated (see [Y82] and [Bar94]):

$$\text{Maximize}\{x^u(T) + y^u(T)\}, \tag{\textbf{P3}}$$

subject to $u \in L^2(0,T)$, $0 \leq u(t) \leq 1$ a.e. $t \in (0,T)$, where (x^u, y^u) is the solution to (2.20) and (2.21).

Problem (P3) is a particular case of (P1), for $m = 1$, $N = 2$,

$$G(t,u,(x,y)) = 0,$$
$$\varphi(x,y) = x + y,$$
$$f(t,u,(x,y)) = \begin{pmatrix} r_1 x - \mu_1 uxy \\ -r_2 y + \mu_2 uxy \end{pmatrix},$$

and

$$K = \{w \in L^2(0,T);\ 0 \leq w(t) \leq 1 \text{ a.e. } t \in (0,T)\}.$$

Existence of an optimal pair for (P3)

Define

$$\Phi(u) = x^u(T) + y^u(T), \quad u \in K,$$

and let

$$d = \sup_{u \in K} \Phi(u).$$

It is obvious that $d \in [0, +\infty)$. For any $n \in \mathbb{N}^*$, there exists $u_n \in K$ such that

$$d - \frac{1}{n} < \Phi(u_n) \leq d.$$

Because

$$x^{u_n}(t) = x_0 \exp\{\int_0^t (r_1 - \mu_1 u(s) y^{u_n}(s))ds\} > 0,$$
$$y^{u_n}(t) = y_0 \exp\{\int_0^t (-r_2 + \mu_2 u(s) x^{u_n}(s))ds\} > 0,$$

for $t \in [0,T]$, we get that $x^{u_n}(t), y^{u_n}(t) > 0$ for any $t \in [0,T]$, and so

$$0 \leq (x^{u_n})'(t) \leq r_1 x^{u_n}(t) \quad \text{a.e. } t \in (0,T).$$

This implies that

$$0 \le x^{u_n}(t) \le x_0 e^{r_1 T}, \quad t \in [0,T],$$

and that $\{(x^{u_n})'\}_n$ is bounded in $L^\infty(0,T)$.

On the other hand we get that

$$0 \le y^{u_n}(t) \le y_0 \exp\{(-r_2 + \mu_2 x_0 e^{r_1 T})T\}, \quad t \in [0,T],$$

and as a consequence $\{(y^{u_n})'\}_n$ is bounded in $L^\infty(0,T)$. It follows that $\{x^{u_n}\}_n$ and $\{y^{u_n}\}_n$ are bounded in $C([0,T])$, and uniformly equicontinuous. By Arzelà's theorem, and by taking into account that $\{u_n\}_n$ is bounded in $L^2(0,T)$ we get that on a subsequence we have

$$\begin{aligned} u_{n_k} &\longrightarrow u^* \text{ weakly in } L^2(0,T) \\ x^{u_{n_k}} &\longrightarrow x^* \text{ in } C([0,T]) \\ y^{u_{n_k}} &\longrightarrow y^* \text{ in } C([0,T]) \end{aligned} \tag{2.22}$$

($u^* \in K$ because K is a closed convex subset of $L^2(0,T)$, and consequently weakly closed).

Inasmuch as

$$x^{u_{n_k}}(t) = x_0 + \int_0^t [r_1 x^{u_{n_k}}(s) - \mu_1 u_{n_k}(s) x^{u_{n_k}}(s) y^{u_{n_k}}(s)]ds,$$

$$y^{u_{n_k}}(t) = y_0 + \int_0^t [-r_2 y^{u_{n_k}}(s) + \mu_2 u_{n_k}(s) x^{u_{n_k}}(s) y^{u_{n_k}}(s)]ds,$$

for any $t \in [0,T]$, and by taking into account (2.22) we get that

$$x^*(t) = x_0 + \int_0^t [r_1 x^*(s) - \mu_1 u^*(s) x^*(s) y^*(s)]ds,$$

$$y^*(t) = y_0 + \int_0^t [-r_2 y^*(s) + \mu_2 u^*(s) x^*(s) y^*(s)]ds,$$

for any $t \in [0,T]$, which means that (x^*, y^*) is the solution to (2.20) and (2.21) corresponding to u^* (i.e., $x^* = x^{u^*}$ and $y^* = y^{u^*}$). On the other hand by

$$d - \frac{1}{n_k} < x^{u_{n_k}}(T) + y^{u_{n_k}}(T) \le d \quad \text{for any } k \in \mathbb{N}^*,$$

and by using the convergences in (2.22) we may pass to the limit and obtain that

$$d = x^{u^*}(T) + y^{u^*}(T);$$

that is, u^* is an optimal control for (P3); $((u^*, (x^*, y^*))$ is an optimal pair for (P3); i.e., u^* is an optimal control and $x^* = x^{u^*}$, $y^* = y^{u^*}$).

The maximum principle for (P3)

For an arbitrary but fixed $v \in V = \{w \in L^2(0,T);\ u^* + \varepsilon w \in K$ for any $\varepsilon > 0$ sufficiently small$\}$ we consider (z_1, z_2) the solution to

$$\begin{cases} z_1' = r_1 z_1 - \mu_1 u^* z_1 y^* - \mu_1 u^* x^* z_2 - \mu_1 v x^* y^*, \quad t \in (0,T) \\ z_2' = -r_2 z_2 + \mu_2 u^* z_1 y^* + \mu_2 u^* x^* z_2 + \mu_2 v x^* y^*, \ t \in (0,T) \\ z_1(0) = z_2(0) = 0. \end{cases} \tag{2.23}$$

Because

$$x^*(T) + y^*(T) \geq x^{u^*+\varepsilon v}(T) + y^{u^*+\varepsilon v}(T),$$

we get that

$$\frac{x^{u^*+\varepsilon v}(T) - x^*(T)}{\varepsilon} + \frac{y^{u^*+\varepsilon v}(T) - y^*(T)}{\varepsilon} \leq 0, \tag{2.24}$$

for any $\varepsilon > 0$ sufficiently small.
For $\varepsilon > 0$ sufficiently small we have that $x^{u^*+\varepsilon v}$ satisfies

$$(x^{u^*+\varepsilon v})'(t) \leq r_1 x^{u^*+\varepsilon v}(t) \ \text{ a.e. } t \in (0,T),$$

and consequently it follows that there exists $M \in (0, +\infty)$ such that

$$0 \leq x^{u^*+\varepsilon v}(t) \leq M \ \text{ for any } t \in [0,T],$$

for any $\varepsilon > 0$ sufficiently small. On the other hand

$$(y^{u^*+\varepsilon v})'(t) \leq (-r_2 + M\mu_2) y^{u^*+\varepsilon v}(t) \ \text{ a.e. } t \in (0,T),$$

and this implies that $\{y^{u^*+\varepsilon v}\}$ is bounded in $C([0,T])$ (for $\varepsilon > 0$ sufficiently small). It follows that both sequences $\{x^{u^*+\varepsilon v}\}$ and $\{y^{u^*+\varepsilon v}\}$ are uniformly bounded and uniformly equicontinuous on $[0,T]$. By Arzelà's theorem it follows that on a sequence $\varepsilon_n \searrow 0$ we have that

$$\begin{aligned} x^{u^*+\varepsilon_n v} &\longrightarrow \tilde{x} \ \text{ in } C([0,T]), \\ y^{u^*+\varepsilon_n v} &\longrightarrow \tilde{y} \ \text{ in } C([0,T]). \end{aligned} \tag{2.25}$$

Because

$$x^{u^*+\varepsilon_n v} = x_0 + \int_0^t [r_1 x^{u^*+\varepsilon_n v}(s) - \mu_1 (u^*(s) + \varepsilon_n v(s)) x^{u^*+\varepsilon_n v}(s) y^{u^*+\varepsilon_n v}(s)] ds$$

and

$$y^{u^*+\varepsilon_n v} = y_0 + \int_0^t [-r_2 y^{u^*+\varepsilon_n v}(s) + \mu_2 (u^*(s) + \varepsilon_n v(s)) x^{u^*+\varepsilon_n v}(s) y^{u^*+\varepsilon_n v}(s)] ds,$$

for any $t \in [0, T]$, we pass to the limit (and use (2.25)), and we get

$$\tilde{x}(t) = x_0 + \int_0^t [r_1 \tilde{x}(s) - \mu_1 u^*(s)\tilde{x}(s)\tilde{y}(s)]ds,$$

and

$$\tilde{y}(t) = y_0 + \int_0^t [-r_2 \tilde{y}(s) + \mu_2 u^*(s)\tilde{x}(s)\tilde{y}(s)]ds,$$

for any $t \in [0, T]$, which means that $(\tilde{x}, \tilde{y})$ is the solution to (2.20) corresponding to u^*; that is, $\tilde{x} = x^{u^*}$, $\tilde{y} = y^{u^*}$.

Define now

$$\alpha_n(t) = \frac{1}{\varepsilon_n} \left[x^{u^*+\varepsilon_n v}(t) - x^*(t) \right] - z_1(t), \quad t \in [0, T],$$

$$\beta_n(t) = \frac{1}{\varepsilon_n} \left[y^{u^*+\varepsilon_n v}(t) - y^*(t) \right] - z_2(t), \quad t \in [0, T].$$

(α_n, β_n) is the solution to

$$\begin{cases} \alpha_n' = r_1 \alpha_n - \mu_1 u^* \alpha_n y^* - \mu_1 u^* x^* \beta_n + f_{1n}(t), & t \in (0, T) \\ \beta_n' = -r_2 \beta_n + \mu_2 u^* \alpha_n y^* + \mu_2 u^* x^* \beta_n + f_{2n}(t), & t \in (0, T) \\ \alpha_n(0) = \beta_n(0) = 0 \end{cases}$$

and $f_{1n} \longrightarrow 0$, $f_{2n} \longrightarrow 0$ in $L^\infty(0, T)$.

This yields

$$\begin{aligned} \alpha_n(t)^2 + \beta_n(t)^2 &\le c \int_0^t [\alpha_n(s)^2 + \beta_n(s)^2] ds \\ &\quad + 2 \int_0^t [f_{1n}(s)\alpha_n(s) + f_{2n}(s)\beta_n(s)] ds \\ &\le (c+1) \int_0^t [\alpha_n(s)^2 + \beta_n(s)^2] ds \\ &\quad + \int_0^T [f_{1n}(t)^2 + f_{1n}(t)^2] dt, \end{aligned}$$

$t \in [0, T]$, where $c > 0$ is a constant independent of n. By Bellman's lemma (see Appendix A.2) we conclude that

$$0 \le \alpha_n(t)^2 + \beta_n(t)^2 \le e^{(c+1)t} \int_0^T [f_{1n}(t)^2 + f_{1n}(t)^2] dt,$$

for any $t \in [0, T]$. We pass to the limit and conclude that

$$\alpha_n \longrightarrow 0, \ \ \beta_n \longrightarrow 0 \ \text{ in } C([0, T]).$$

This implies that

$$\frac{1}{\varepsilon_n}[x^{u^*+\varepsilon_n v} - x^*] \longrightarrow z_1 \text{ in } C([0,T]),$$

and

$$\frac{1}{\varepsilon_n}[y^{u^*+\varepsilon_n v} - y^*] \longrightarrow z_2 \text{ in } C([0,T]).$$

If we again use (2.24) we may infer that

$$z_1(T) + z_2(T) \le 0. \tag{2.26}$$

Let (p_1, p_2) be the solution to

$$\begin{cases} p_1' = -r_1 p_1 + \mu_1 u^* y^* p_1 - \mu_2 u^* y^* p_2,\ t \in (0,T) \\ p_2' = r_2 p_2 + \mu_1 u^* x^* p_1 - \mu_2 u^* x^* p_2, \quad t \in (0,T) \\ p_1(T) = p_2(T) = 1. \end{cases} \tag{2.27}$$

By multiplying the first equation in (2.27) by z_1 and the second one by z_2 and integrating over $[0,T]$ we get that

$$\begin{aligned} &\int_0^T [p_1'(t)z_1(t) + p_2'(t)z_2(t)]dt \\ &= \int_0^T [-r_1 p_1(t)z_1(t) + \mu_1 u^*(t)y^*(t)p_1(t)z_1(t) - \mu_2 u^*(t)y^*(t)p_2(t)z_1(t) \\ &\quad + \mu_1 u^*(t)x^*(t)p_1(t)z_2(t) - \mu_2 u^*(t)x^*(t)p_2(t)z_2(t) + r_2 p_2(t)z_2(t)]dt. \end{aligned}$$

If we integrate by parts and use (2.23) we get after some calculation that

$$\begin{aligned} &p_1(T)z_1(T) + p_2(T)z_2(T) - p_1(0)z_1(0) - p_2(0)z_2(0) \\ &= \int_0^T x^*(t)y^*(t)v(t)[\mu_2 p_2(t) - \mu_1 p_1(t)]dt, \end{aligned}$$

and consequently by (2.23) and (2.26) we get that

$$z_1(T) + z_2(T) = \int_0^T x^*(t)y^*(t)v(t)[\mu_2 p_2(t) - \mu_1 p_1(t)]dt \le 0,$$

for any $v \in V$. This implies (as in the previous section) that

$$u^*(t) = \begin{cases} 0 \text{ if } x^*(t)y^*(t)[\mu_2 p_2(t) - \mu_1 p_1(t)] < 0 \\ 1 \text{ if } x^*(t)y^*(t)[\mu_2 p_2(t) - \mu_1 p_1(t)] > 0 \end{cases}$$

a.e. on $(0,T)$. Because $x_0,\ y_0 > 0$, and x^* and y^* are positive functions, we may conclude that

$$u^*(t) = \begin{cases} 0 \text{ if } \mu_2 p_2(t) - \mu_1 p_1(t) < 0 \\ 1 \text{ if } \mu_2 p_2(t) - \mu_1 p_1(t) > 0 \end{cases} \tag{2.28}$$

a.e. on $(0,T)$.

Equations (2.27) and (2.28) are the first-order necessary optimality conditions, and (2.20)–(2.21), (2.27)–(2.28) represent the maximum principle for (P3).

The structure of the optimal control u^* for (P3)

Our next goal is to obtain more information about the structure of the optimal control u^*.

- If $\mu_2 < \mu_1$, then $\mu_2 p_2(T) - \mu_1 p_1(T) = \mu_2 - \mu_1 < 0$, and then we may choose a maximal interval $(T-\eta, T]$ $(\eta > 0)$ where $\mu_2 p_2(t) - \mu_1 p_1(t) < 0$. By (2.28) we have $u^*(t) = 0$ on $(T-\eta, T]$ and consequently

$$p_1'(t) = -r_1 p_1(t), \quad p_2'(t) = r_2 p_2(t) \ \text{ a.e. } t \in (T-\eta, T).$$

  This yields

$$p_1(t) = \exp\{-r_1(t-T)\}, \quad p_2(t) = \exp\{r_2(t-T)\}, \quad t \in [T-\eta, T].$$

  The function $t \mapsto \mu_2 \exp\{r_2(t-T)\} - \mu_1 \exp\{-r_1(t-T)\}$ is increasing on $[T-\eta, T]$, and this implies that $T - \eta = 0$ and

$$\mu_2 p_2(t) - \mu_1 p_1(t) < 0, \quad t \in (0, T),$$

  and so $u^*(t) = 0$ a.e. on $(0, T)$.
- If $\mu_2 = \mu_1$, then (p_1, p_2) is the solution to

$$\begin{cases} p_1' = -r_1 p_1 - \mu_1 u^* y^* (p_2 - p_1), \, t \in (0, T) \\ p_2' = r_2 p_2 - \mu_1 u^* x^* (p_2 - p_1), \quad t \in (0, T) \\ p_1(T) = p_2(T) = 1. \end{cases}$$

  In conclusion

$$p_2(t) - p_1(t) = -\int_t^T [r_2 p_2(s) + r_1 p_1(s)] \exp\{\mu_1 \int_s^t u^*(\tau)[y^*(\tau) - x^*(\tau)] d\tau\} ds,$$

  $t \in [0, T]$. So, $p_2(t) - p_1(t) < 0$ on a maximal interval $(T-\eta, T]$ $(\eta > 0)$ and, in the same manner as in the previous case, it follows that $u^*(t) = 0$ a.e. on $(0, T)$.
- If $\mu_2 > \mu_1$, then there exists a maximal interval $(T-\eta, T]$ $(\eta > 0)$ such that

$$\mu_2 p_2(t) - \mu_1 p_1(t) > 0, \quad t \in (T-\eta, T].$$

  By (2.28) we have $u^*(t) = 1$ on $(T-\eta, T]$. We intend to prove that $T-\eta$ is a switching point for the optimal control u^*. Indeed, by (2.27) we get that

$$\mu_2 p_2(t) - \mu_1 p_1(t) = -\int_t^{T-\eta} [r_2 \mu_2 p_2(s) + r_1 \mu_1 p_1(s)] \cdot \exp\{\int_s^t u^*(\tau)[\mu_2 x^*(\tau) - \mu_1 y^*(\tau)] d\tau\} \, ds, \tag{2.29}$$

  $t \in [0, T-\eta]$. On the other hand (p_1, p_2) is a solution to

$$\begin{cases} p_1' = -p_1(r_1 - \mu_1 y^*) - \mu_2 y^* p_2, \ t \in (T-\eta, T) \\ p_2' = -p_2(\mu_2 x^* - r_2) + \mu_1 x^* p_1, \ t \in (T-\eta, T) \\ p_1(T) = p_2(T) = 1. \end{cases} \tag{2.30}$$

Because $\mu_2 p_2(t) - \mu_1 p_1(t) > 0$, for any $t \in (T-\eta, T]$, then we get that

$$p_1(t) \geq \exp\{r_1(T-t)\} \geq 1, \quad t \in [T-\eta, T].$$

Using the fact that $\mu_2 p_2(T-\eta) - \mu_1 p_1(T-\eta) = 0$ and (2.29) we obtain that $p_2(T-\eta) > 0$ and consequently $\mu_2 p_2(t) - \mu_1 p_1(t) < 0$ in a maximal interval $(T-\eta-\varepsilon, T-\eta]$ $(\varepsilon > 0)$. This implies that $u^*(t) = 0$ on $(T-\eta-\varepsilon, T-\eta]$. On this interval we have

$$p_1(t) = p_1(T-\eta)\exp\{r_1(T-\eta-t)\},$$
$$p_2(t) = p_2(T-\eta)\exp\{r_2(t-T+\eta)\},$$

and in conclusion $\mu_2 p_2 - \mu_1 p_1$ is increasing on $(T-\eta-\varepsilon, T-\eta)$. Hence

$$\mu_2 p_2(t) - \mu_1 p_1(t) < 0, \quad t \in (T-\eta-\varepsilon, T-\eta),$$

and consequently $T - \eta - \varepsilon = 0$. The conclusion is that

$$u^*(t) = \begin{cases} 0, & t \in [0, T-\eta] \\ 1, & t \in (T-\eta, T] \end{cases} \tag{2.31}$$

a.e. on $(0, T)$.

So, we have a bang-bang optimal control with at most one switching point. We can determine the switching point $T - \eta$, either by taking into account (2.30) and $\mu_2 p_2(T-\eta) - \mu_1 p_1(T-\eta) = 0$, or by finding $T - \eta \in [0, T]$, which maximizes $\Phi(u^*)$, where u^* is given by (2.31).

Approximating the optimal control for (P3)

In order to approximate the optimal control u^* we have to find η from formula (2.31). A simple idea is to try τ ($T - \eta$ in (2.31)) as switching point for the control of the elements of a grid defined on $[0, L]$ (we put L instead of T) and to get the one that provides the maximum value for $\Phi(u)$. Here

$$u(t) = \begin{cases} 0, & t \in [0, \tau] \\ 1, & t \in (\tau, L]. \end{cases}$$

Here is the algorithm.

Algorithm 2.1

```
/* Build the grid */
tspan = 0:h1:L ;
/* Try the grid points */
m = length(tspan) ;
```

for i = 1 to m
  τ = tspan(i) ;
  /* **S1 :** Build the corresponding control u_τ */

$$u_\tau(t) = \begin{cases} 0, & t \in [0, \tau] \\ 1, & t \in (\tau, L]. \end{cases}$$

  /* **S2 :** Compute the state $[x, y]$, the corresponding solution of system (2.20) corresponding to $u := u_\tau$, with the initial conditions */

$$x(0) = x_0, \quad y(0) = y_0.$$

  /* **S3 :** Compute the corresponding value of the cost functional Φ */
  fiu(i) = x(L) + y(L) ;
end–for
  /* **S4 :** Find the maximal value of vector *fiu* */

Here is the corresponding program.

```
% file ppp1.m
% predator–prey model with bang-bang optimal control
clear
global r1 r2 mu1 mu2
global tsw
disp('get model parameters') ;
r1 = input('r1 : ') ;
mu1 = input('mu1 : ') ;
r2 = input('r2 : ') ;
mu2 = input('mu2 : ') ;
disp('get data') ;
L = input('final time : ') ;
h = input('grid step : ') ;
h1 = input('switch step : ') ;
x0 = input('x(0) : ') ;
y0 = input('y(0) : ') ;
lw = input('LineWidth : ') ; % for graphs ( plot )
tt = 0:h:L ; % ODE integration grid
n = length(tt) ;
tspan = 0:h1:L ; % switching points grid
m = length(tspan) ;
for i = 1:m
    i
    tsw = tspan(i) ; % tsw stands for switching point τ
    [t q] = ode45('bp2',tt,[x0 ; y0]) ;
    k = length(t) ;
    fiu(i) = q(k,1) + q(k,2) ; % store cost functional value
    clear t q % clear memory to avoid garbage for the next iteration
```

```
end
w = fiu' ;
save cont.txt w -ascii
disp('FILE MADE') ;
[vmax,j] = max(fiu) ; % maximal value and corresponding index
j
a1 = ['max = ', num2str(vmax)] ;
disp(a1) ;
a2 = ['switch = ', num2str(tspan(j))] ;
disp(a2) ;
plot(tspan,fiu,'LineWidth',lw) ; grid
xlabel('\bf u switch','FontSize',16)
ylabel('\bf \Phi(u_{\tau})','FontSize',16)
figure(2)
bar(fiu)
title('\bf \Phi(u_{\tau})','FontSize',16)
```

We have used a vector, namely *tt*, for *ode45* and another one, namely *tspan*, for the switching points grid to get a faster program.

Here is the function file *bp2.m* for the right-hand side of the differential system.

```
function out1 = bp2(t,q)
global r1 r2 mu1 mu2
global tsw
if t > tsw
    u = 1 ;
else
    u = 0 ;
end
out1 = [ r1*q(1) – mu1*u*q(1)*q(2) ; mu2*u*q(1)*q(2) – r2*q(2) ] ;
```

For a numerical test we have used $r_1 = 0.07$, $\mu_1 = 1$, $r_2 = 0.6$, $\mu_2 = 2$, $L = 50$, $h = 0.1$, $h1 = 1$, $x(0) = 0.04$, $y(0) = 0.02$, and $lw = 5$. The graph of the corresponding function $\tau \mapsto \Phi(u_\tau)$, where τ is the switching point of u_τ, can be seen in Figures 2.1 and 2.2.
We have obtained a global maximum on $[0, L]$ for $\tau^* = 15$, and the maximal value of the cost functional is 1.5006. The program that uses the switch point of the optimal control in order to plot the graphs for the corresponding state components is *ppp2.m*:

```
% file ppp2.m
% predator–prey model with bang-bang optimal control
% makes graphs by using the switching point obtained by ppp1.m
clear
global r1 r2 mu1 mu2
global tsw
```

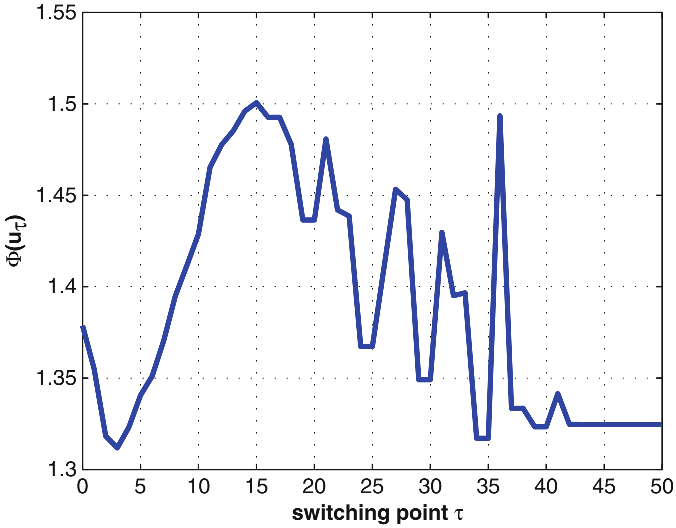

Fig. 2.1. The dependence of the cost function with respect to the switching point

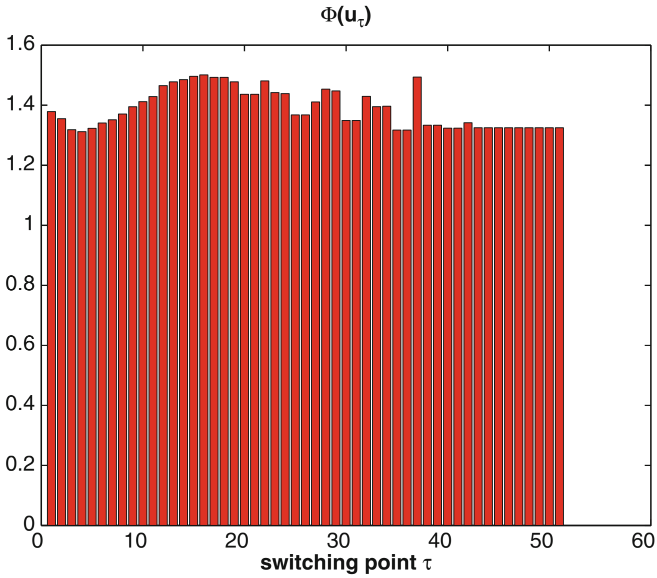

Fig. 2.2. Another representation of the dependence of the cost function with respect to the switching point τ

```
... read parameters and data as in ppp1.m (except h) ...
tspan = 0:h1:L ;
% graph of the control
m = length(tspan) ;
for i = 1:m
    if tspan(i) > tsw
        z(i) = 1 ;
    else
        z(i) = 0 ;
    end
end
plot(tspan,z,'rs') ; grid
axis([0 L -0.2 1.2])
xlabel('\bf t','Fontsize',16)
ylabel('\bf u(t)','Fontsize',16)
[t q] = ode45('bp2',[0 L],[x0 ; y0]) ;
% predator-prey populations graph
figure(2)
plot(t,q(:,1),'*',t,q(:,2),'ro') ; grid
xlabel('\bf t','FontSize',16)
legend('prey','predator',0)
% xOy graph
figure(3)
plot(q(:,1),q(:,2),'LineWidth',lw) ; grid
xlabel('\bf x','Fontsize',16)
ylabel('\bf y','Fontsize',16)
```

2.4 Insulin treatment model

We consider a model for insulin treatment for patients with diabetes. The main problem for such a patient is to keep the blood glucose level close to a convenient value and to avoid large variations of it. In practice insulin injections are used. An optimal control problem with impulsive controls is considered to maintain a steady state of the blood glucose level. This problem does not fit in the framework of Problem (P1) from Section 2.1 mainly because the control considered here is of impulsive type. We do, however, obtain first-order necessary optimality conditions which are used to write a program.

In the case of diabetes the pancreas (the beta cells) is not able to provide enough insulin to metabolize glucose. Blood glucose concentration increases when glucose is administrated in mammals whereas insulin accelerates the removal of glucose from the plasma. Therefore blood sugar decays to a normal value of 0.8–1.2 g/l. Let us denote by $I(t)$ the insulin concentration, and by $G(t)$ the glucose concentration at moment $t \in [0, L]$ $(L > 0)$.

We now consider diabetic patients who are not able to produce enough insulin. The insulin is supplied by injections. The glucose concentration can be easily determined (measured). A corresponding simplified model for dynamics of the insulin–glucose system is the following one (see [Che86, Chapter 6]):

$$\begin{cases} I'(t) = dI(t), & t \in (0, L) \\ G'(t) = bI(t) + aG(t), & t \in (0, L) \\ I(0) = I_0, \quad G(0) = G_0, & \end{cases} \tag{2.32}$$

where $d < 0$ ($|d|$ is the decay rate of insulin), a is the growth rate of glucose ($a \neq d$), b is a negative constant that can be measured, I_0 is the initial concentration of insulin (injected), and G_0 is the initial concentration of glucose. The numerical tests show that model (2.32) works well only for $I(t)$ and $G(t)$ between appropriate limits. For I_0 and G_0 outside the usual medical limits it is possible to obtain negative values for $I(t)$ and $G(t)$ and therefore the model fails. A more accurate model is, however, indicated at the end of this subsection. The reaction between $I(t)$ and $G(t)$ in (2.32) is a local linearization of the full model presented later (see (2.41)).

The first program plots the graphs of insulin concentration and of glucose concentration.

```
% file dbt1.m
% blood insulin–glucose system
% y(1) = insulin concentration
% y(2) = glucose concentration
clear
global a b d
L = input('final time : ') ;
h = input('h : ') ;
I0 = input('I(0) : ') ;
G0 = input('G(0) : ') ;
a = 0.0343 ;
b = −0.05 ;
d = −0.5 ;
tspan = 0:h:L ;
[t y] = ode45('hum1',tspan,[I0 ; G0]) ;
plot(t,y(:,1),'∗') ; grid
xlabel('\bf t','FontSize',16)
ylabel('\bf I(t)','FontSize',16)
figure(2)
plot(t,y(:,2),'r∗') ; grid
xlabel('\bf t','FontSize',16)
ylabel('\bf G(t)','FontSize',16)
```

We also have

```
function out1 = hum1(t,y)
global a b d
out1 = [d*y(1) ; b*y(1) + a*y(2)] ;
```

The numerical test is done for $L = 10$, $h = 0.01$, $I0 = 15$, and $G0 = 2$. The evolution of insulin and glucose concentration are presented in Figures 2.3 and 2.4, respectively. Notice that $I(t)$ decays to zero (the effect of the decay rate) and $G(t)$ reaches a convenient level. The insulin has a good effect because the glucose level at the beginning was $G_0 = 2$, and reaches approximatively, the value 0.8 at the moment $t = 6$. After $t = 7$ the insulin effect almost vanishes and the glucose level increases slowly.

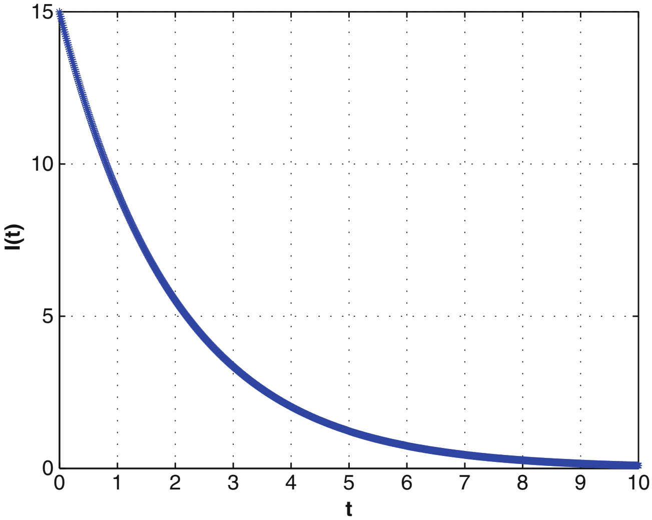

Fig. 2.3. Insulin dynamics

System (2.32) can also be integrated mathematically. We first consider the problem of insulin dynamics:

$$\begin{cases} I'(t) = dI(t), \, t \in (0, L) \\ I(0) = I_0 \end{cases}$$

which has a unique solution given by

$$I(t) = I_0 e^{dt}, \quad t \in [0, L]. \tag{2.33}$$

If we use the form of $I(t)$ given by (2.33), we obtain from (2.32) the following linear model for the glucose dynamics,

$$\begin{cases} G'(t) = bI_0 e^{dt} + aG(t), \quad t \in (0, L) \\ G(0) = G_0, \end{cases}$$

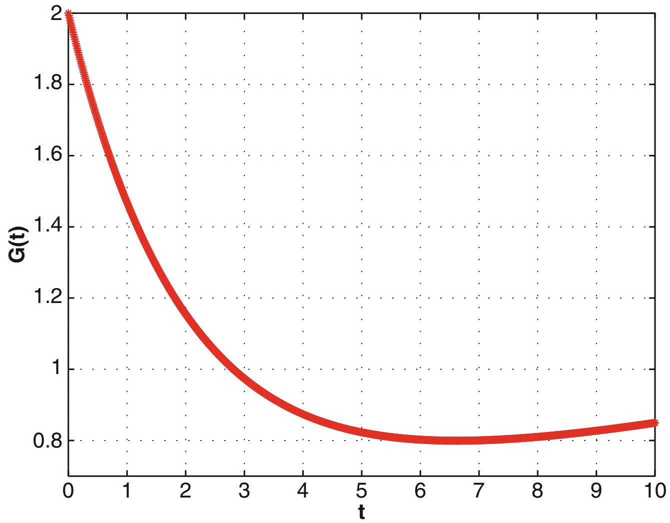

Fig. 2.4. Glucose dynamics

which gives the following formula for the glucose concentration,

$$G(t) = G_0 e^{at} + \frac{bI_0}{d-a}(e^{dt} - e^{at}), \quad t \in [0, L]. \tag{2.34}$$

Hence the solution of system (2.32) is given by formulae (2.33) and (2.34). The corresponding program is

```
% file dbt2.m
% blood insulin–glucose system
% y1(t) = insulin concentration
% y2(t) = glucose concentration
% mathematical integration
clear
L = input('final time : ') ;
h = input('h : ') ;
I0 = input('insulin(0) : ') ;
G0 = input('glucose(0) : ') ;
a = 0.0343 ;
b = −0.05 ;
d = −0.5 ;
temp = b*I0/(d − a) ;
t = 0:h:L ;
v = exp(d*t) ;
w = exp(a*t) ;
y1 = I0*v ;
```

```
y2 = G0*w + temp*(v - w) ;
% make figures as in previous program
% .....
```

The figures obtained are similar to the previous two figures.

We now consider an optimal control problem with impulsive control to obtain a scheme of insulin treatment providing good control of glycemia over some time interval. We denote by A the desired level of glucose. Assume that the patient gets m injections of insulin at moments

$$0 = t_1 < t_2 < \cdots < t_m = L,$$

with corresponding amounts $c_j = c(t_j)$, $j = 1, 2, \ldots, m$ and that the initial concentration of insulin is $I_0 = 0$. Usually the moments for injections are fixed and we have $t_{j+1} - t_j = h$ for $j = 1, \ldots, m-1$. The dynamics of the insulin–glucose system is then described by

$$\begin{cases} I'(t) = dI(t) + \displaystyle\sum_{j=1}^{m} c_j \delta_{t_j}, \\ G'(t) = bI(t) + aG(t), \\ I(0) = 0, \quad G(0) = G_0, \end{cases} \tag{2.35}$$

where δ_{t_j} is the Dirac mass at t_j. System (2.35) is equivalent to the following one

$$\begin{cases} I'(t) = dI(t), & t \in (t_j, t_{j+1}), j \in \{1, \ldots, m-1\} \\ I(0) = 0 \\ I(t_j+) = I(t_j-) + c_j, & j \in \{1, \ldots, m-1\} \\ G'(t) = bI(t) + aG(t), & t \in (0, L) \\ G(0) = G_0. \end{cases} \tag{2.36}$$

The solution of (2.35) in the sense of the theory of distributions (which is also the solution to (2.36)) is given by

$$\begin{cases} I(t) = \displaystyle\sum_{j=1}^{m} c_j H(t - t_j) e^{d(t-t_j)}, \\ G(t) = G_0 e^{at} + \dfrac{b}{d-a} S(t), \end{cases} \tag{2.37}$$

$t \in [0, L]$, where

$$S(t) = \sum_{j=1}^{m} c_j H(t - t_j) \left[e^{d(t-t_j)} - e^{a(t-t_j)} \right], \tag{2.38}$$

and H is the step (Heaviside) function (i.e., $H : I\!R \to I\!R$),

$$H(t) = \begin{cases} 1 & \text{if } t \geq 0 \\ 0 & \text{if } t < 0. \end{cases}$$

Therefore, function $t \mapsto H(t - t_j)$ in formulae (2.37) and (2.38), defined for $t \in [0, L]$, reads

$$H(t - t_j) = \begin{cases} 1 & \text{if } t \in [t_j, L] \\ 0 & \text{if } t \in [0, t_j). \end{cases}$$

The formula for G says that the effect of the insulin injection received at the moment $t = t_j$ is valid only for $t \geq t_j$. The effect vanishes after some time due to the exponential function with negative exponent.

Here is the optimal control problem (the insulin treatment) related to (2.35):

$$\text{Minimize } \Psi(c) = \frac{1}{2} \int_0^L [G(t) - A]^2 dt, \qquad (\mathbf{I})$$

subject to $c = (c_1, \ldots, c_m) \in I\!R^m$, where (I, G) is the solution to (2.35). Here the vector c is the control (which is in fact an impulsive control, a control that acts only at some discrete moments of time).

The functional Ψ is quadratic with respect to every c_j, thus it means that there exists at least an optimal control $c = (c_1, \ldots, c_m) \in I\!R^m$. The optimal control satisfies

$$\frac{\partial \Psi}{\partial c_j}(c) = 0, \quad j = 1, \ldots, m, \qquad (2.39)$$

a linear algebraic system with the unknowns c_j, $j = 1, \ldots, m$. We calculate the partial derivatives and use formula (2.39) to get the following algebraic linear system,

$$\sum_{i=1}^{m} q_{ij} c_i = B_j, \quad j \in \{1, \ldots, m\},$$

where

$$q_{ij} = \alpha \int_0^L H(t - t_i) H(t - t_j) e_i(t) e_j(t) dt, \qquad (2.40)$$

$$B_j = \int_0^L H(t - t_j) e_j(t)(A - G_0 e^{at}) dt = \int_{t_j}^L e_j(t)(A - G_0 e^{at}) dt,$$

$i, j \in \{1, \ldots, m\}$. We have denoted

$$\alpha = \frac{b}{d - a},$$

and

$$e_j(t) = e^{d(t - t_j)} - e^{a(t - t_j)}, \quad t \in [0, L],\ j \in \{1, \ldots, m\}.$$

If $i > j$, then $t_i > t_j$ and Formula (2.40) reads

$$q_{ij} = \alpha \int_{t_i}^{L} e_i(t)e_j(t)dt.$$

Our goal is to solve system (2.39). However if a certain component c_j is negative, this is meaningless from the medical point of view. If we introduce the restrictions $c_j \geq 0$, $j \in \{1, \ldots, m\}$, we get a mathematical programming problem which is more complicated. Another possibility is to introduce restrictions of the form $0 \leq c_j \leq \bar{c}$, $j \in \{1, \ldots, m\}$, and to use a projected gradient method (see Chapter 3). But this is more complicated also. To establish a treatment policy we can simply take $c_j := 0$ if $c_j < 0$. Then we have to add glucose, usually from food, or to replace the negative dose of the injection by $c_j = 0$, and consequently to obtain suboptimal control. For our numerical test made for medically appropriate values of $G(0)$ the solution was positive.

We return to the linear system. The algorithm to compute the transpose of matrix Q, that is, $Q^T = [q_{ij}]$, is:

```
for j = 1 to m
    for i = 1 to j
        compute q_ij = α ∫_{t_j}^L e_i(t)e_j(t)dt
    end–for
    for i = j+1 to m
        compute q_ij = α ∫_{t_i}^L e_i(t)e_j(t)dt
    end–for
end–for
```

Then we transpose the matrix $[q_{ij}]$ obtained above and we get Q. We leave it to the reader to write the corresponding program. The values of the system parameters are $a = 0.1$, $b = -0.05$, and $d = -0.5$. Below we give only the sequence to compute the matrix Q and the right-hand side B of the system $Qc = B$.

```
...
Q = zeros(m − 1) ;
for j = 1:m − 1
    tj = t(j) ;
    for i = 1:j
        ti = t(i) ;
        Q(i,j) = alf*quadl('fi1',tj,L) ;
    end
    for i = j+1:m − 1
        ti = t(i) ;
        Q(i,j) = alf*quadl('fi2',ti,L) ;
    end
end
```

```
Q = Q' ;
for j = 1:m - 1
    tj = t(j) ;
    B(j) = quadl('psi',t(j),L) ;
end
B = B' ;
% solve system Qc = B
c = Q\B ;
...
```

The function file $fi1.m$ computes the matrix components q_{ij} for $i \leq j$.

```
function y = fi1(t)
global ti tj
global a d
y = 0 ;
if t >= tj
     temp1 = exp(d*(t - tj)) - exp(a*(t - tj)) ;
     temp2 = exp(d*(t - ti)) - exp(a*(t - ti)) ;
     y = temp1 .* temp2 ;
end
```

The function file $fi2.m$ computes the matrix components q_{ij} for $i > j$. It is similar to $fi1.m$. There is only one difference. The statement

```
if t >= tj
```

is replaced by

```
if t >= ti
```

The function file $psi.m$ computes the right-hand side components B_j.

```
function y = psi(t)
global tj
global a d
global a1
global G0
y = 0 ;
if t >= tj
     temp1 = exp(d*(t - tj)) - exp(a*(t - tj)) ;
     temp2 = a1 - G0*exp(a*t) ;
     y = temp1 .* temp2 ;
end
```

We pass now to numerical examples.

Example 1. We take $L = 48$ (hours), $G(0) = 2$, $A = 1$, and $m = 9$ (number of injections). It follows that the interval between successive injections is $h = 6$ (hours). The insulin "shots" are represented in Figure 2.5.

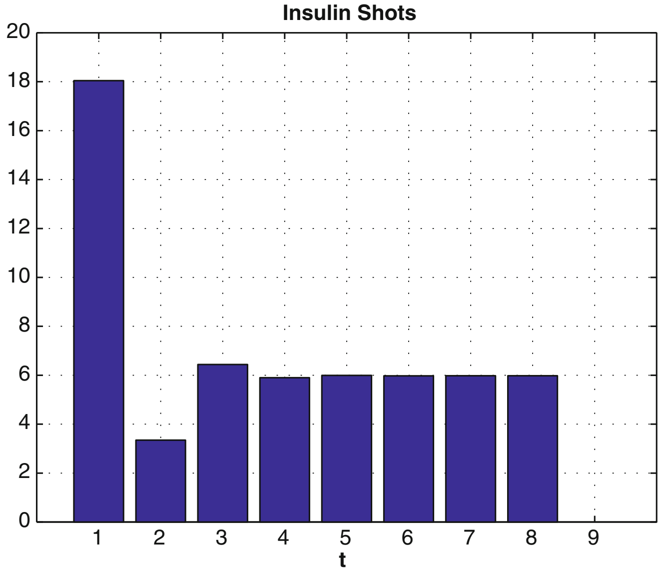

Fig. 2.5. The insulin doses for 48 h

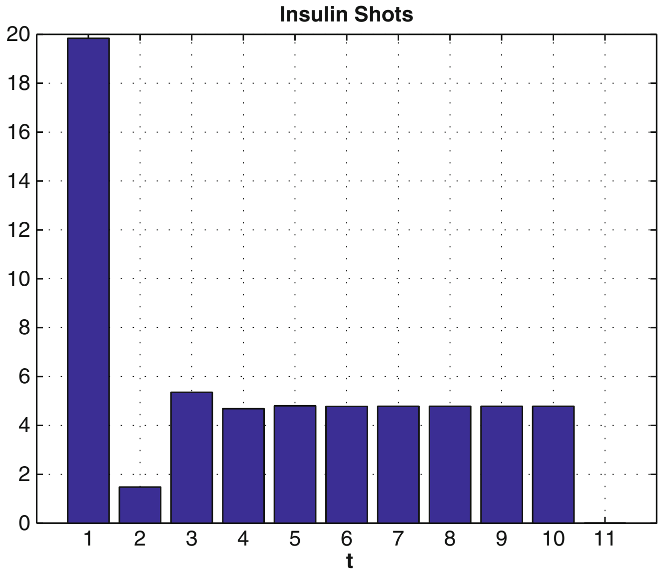

Fig. 2.6. The insulin doses for 60 h

Example 2. Another experiment was done with $L = 60$, $G(0) = 2$, $A = 0.8$, and $m = 11$ $(h = 6)$. The results are given in Figure 2.6.

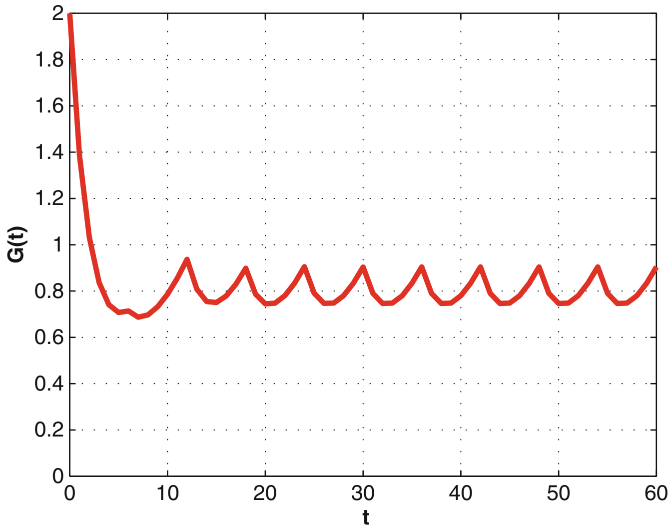

Fig. 2.7. The blood glucose concentration for the second numerical experiment

To complete our investigation we have also computed the glucose level given by formulae (2.37) and (2.38). The shape of the blood glucose concentration for the second numerical experiment is given in Figure 2.7. Let us remark that the glucose level decays from $G(0) = 2$ under the desired level $A = 0.8$ and then remains quite close to it. For the first numerical experiment the behavior of $G(t)$ is similar.

Remark 2.3. Equation (2.39) are the first-order optimality conditions for Problem (I).

We propose that the reader investigate in a similar manner the following optimal control problem.

$$\text{Minimize } \Psi(c) = \frac{1}{2}\int_0^L [G(t) - A]^2 dt, \tag{\textbf{I1}}$$

subject to $c = (c_1, c_2, \ldots, c_m) \in \mathbb{R}^m$, where (I, G) is the solution to the following more accurate model,

$$\begin{cases} I'(t) = dI(t) + \displaystyle\sum_{j=1}^{m} c_j \delta_{t_j}, \\ G'(t) = bI(t)G(t) + aG(t), \\ I(0) = 0, \quad G(0) = G_0. \end{cases} \tag{2.41}$$

It is also important to investigate both optimal control problems under the control constraints

$$c_j \geq 0, \quad j \in \{1, 2, \dots, m\}.$$

A better way to control the glucose level is, however, to act on insulin concentration (by injections) as well as on glucose concentration (by the food from usual meals).

2.5 Working examples

2.5.1 HIV treatment

We consider here a mathematical model that describes the interaction of the immune system with the HIV (human immunodeficiency virus) proposed in [KLS97]. Next we propose two optimal control problems based on chemotherapy which affects either the viral infectivity or the viral productivity.

The immune system is modeled in terms of the population of $CD4^+$ T cells (see [PKD93], [HNP95], and [PN02]). Let

$T(t)$ denote the concentration of uninfected $CD4^+$ T cells.
$T_i(t)$ denote the concentration of infected $CD4^+$ T cells.
$V(t)$ denote the concentration of free infectious virus particles

at moment t. The dynamics of the system is modeled by the following initial-value problem.

$$\begin{cases} T'(t) = \dfrac{s}{1+V(t)} - \mu_1 T(t) + rT(t)\left(1 - \dfrac{T(t)+T_i(t)}{T_{max}}\right) - k_1 V(t) T(t), \\ \\ T_i'(t) = k_1 V(t) T(t) - \mu_2 T_i(t), \\ \\ V'(t) = -k_1 V(t) T(t) - \mu_3 V(t) + N\mu_2 T_i(t), \\ \\ T(0) = T_0, \quad T_i(0) = T_{i0}, \quad V(0) = V_0, \end{cases} \tag{2.42}$$

$t \in (0, L)$ $(L > 0)$, where $s, k_1, r, N, \mu_1, \mu_2, \mu_3, T_{\max}$ are positive constants and $T_0, T_{i0}, V_0 \geq 0$ are the initial concentrations of $CD4^+$ T cells, infected $CD4^+$ T cells, and free infectious virus particles, respectively.

The term $s/(1+V)$ represents a source term; the dependence upon the viral concentration V models the fact that infection of precursors of T cells may occur, thus reducing the production of the uninfected T cells.

The term $-k_1VT$ in the first equation in (2.42) together with $+k_1VT$ in the second equation in (2.42) models is the infection of T cells due to the viral concentration V; the term $-k_1VT$ in the third equation in (2.42) models the binding of viruses to uninfected T cells, thus leading to infection.

μ_1, μ_2, μ_3 denote natural decay rates.

The term $N\mu_2 T_i$ in the third equation in (2.42) models the production of viruses during the decay of infected T cells.

The term
$$r\left(1-\frac{T(t)+T_i(t)}{T_{\max}}\right)$$
represents the production rate of T cells.

Chemotherapy by a drug may either:

- Affect the virus infectivity, so that the second equation in (2.42) is modified into the following (see [BKL97]),
$$T_i'(t) = u(t)k_1 V(t)T(t) - \mu_2 T_i(t), \quad t \in (0, L),$$
$u(t)$ being the control variable, that is, the strength of the chemotherapy. The first and third equations should be modified accordingly.
- Or reduce the viral production, which is most applicable to drugs such as protease inhibitors (see [KLS97]), thus modifying the third equation in (2.42) into
$$V'(t) = -k_1 V(t)T(t) - \mu_3 V(t) + u(t)N\mu_2 T_i(t), \quad t \in (0, L).$$
The second equation should be modified accordingly.
In either case the cost functional to maximize is
$$\int_0^L [aT(t) - \frac{1}{2}(1-u(t))^2]dt,$$

($a > 0$) subject to $u \in L^2(0, L)$, $0 \le u(t) \le 1$ a.e. $t \in (0, L)$, which means maximizing the number of uninfected T cells, while simultaneously minimizing the "cost" of the chemotherapy to the human body.
A greater or lower value for a corresponds to a lower or greater importance given to minimizing the "cost" of the chemotherapy to the human body.

We propose that the reader derive the first-order necessary conditions of optimality for both optimal control problems.

Hint. The first optimal control problem proposed here is a particular case of (P1) (Section 2.1), for
$$G(t, u, T, T_i, V) = aT - \frac{1}{2}(1-u)^2, \quad \varphi(T, T_i, V) = 0,$$
$$f(t, u, T, T_i, V) = \begin{pmatrix} \dfrac{s}{1+V} - \mu_1 T + rT(1 - \dfrac{T+T_i}{T_{\max}}) - k_1 uVT \\ k_1 uVT - \mu_2 T_i \\ -k_1 uVT - \mu_3 V + N\mu_2 T_i \end{pmatrix}$$
and
$$K = \{w \in L^2(0, L);\ 0 \le w(t) \le 1 \text{ a.e. } t \in (0, L)\}.$$

The second optimal control problem proposed here is a particular case of (P1), for the same G, φ, K (as for the previous proposed problem), and

$$f(t,u,T,T_i,V) = \begin{pmatrix} \dfrac{s}{1+V} - \mu_1 T + rT(1 - \dfrac{T+T_i}{T_{\max}}) - k_1 VT \\ \\ k_1 VT - \mu_2 u T_i \\ \\ -k_1 VT - \mu_3 V + uN\mu_2 T_i \end{pmatrix}.$$

2.5.2 The control of a SIR model

We describe here the dynamics of a disease (transmitted only by contact between infectious and susceptible individuals) in a biological population using the following standard SIR model with vital dynamics (see [Cap93]).

$$\begin{cases} S'(t) = mN - mS(t) - cS(t)I(t) - u(t)S(t), \\ \\ I'(t) = -mI(t) + cS(t)I(t) - dI(t), \\ \\ R'(t) = -mR(t) + u(t)S(t) + dI(t), \end{cases} \tag{2.43}$$

for $t \in (0, L)$, $L > 0$, together with the initial conditions

$$S(0) = S_0 > 0, \quad I(0) = I_0 > 0, \quad R(0) = R_0 \geq 0. \tag{2.44}$$

Here

$S(t)$ represents the density of susceptible individuals,
$I(t)$ represents the density of infectious individuals, and
$R(t)$ represents the density of recovered (and immune) individuals

at moment t. $N = S(t) + I(t) + R(t) = S_0 + I_0 + R_0 > 0$ is a constant that represents the density of total population which is assumed to be constant.

Here m, c, d are positive constants. The incidence of the disease is described by the term $cS(t)I(t)$. The constant d represents the rate at which the infectious individuals recover.

The control u represents the part of the susceptible population being vaccinated. The vaccinated individuals recover.

We propose that the reader investigate the following optimal control problem for the above-mentioned SIR model:

$$\text{Minimize} \int_0^L [I(t) + au(t)^2]dt,$$

($a > 0$) subject to $u \in L^2(0, L)$, $0 \leq u(t) \leq M$ ($M > 0$) a.e. $t \in (0, L)$, where (S, I, R) is the solution to (2.43) and (2.44).

This means we are interested in minimizing the infectious population while simultaneously minimizing the "cost" of vaccination. A greater or lower value for a means a greater or lower importance given to minimizing the cost of vaccination.
Derive the maximum principle.

Hint. This problem is a particular case of (P1) (Section 2.1), for $m = 1$, $N = 3$, $T := L$,

$$G(t,u,S,I,R) = I + au^2, \quad \varphi(S,I,R) = 0,$$
$$f(t,u,S,I,R) = \begin{pmatrix} mN - mS - cSI - uS \\ -mI + cSI - dI \\ -mR + uS + dI \end{pmatrix},$$

and

$$K = \{w \in L^2(0,L);\ 0 \le w(t) \le M \text{ a.e. } t \in (0,L)\}.$$

Another important optimal control problem related to the SIR model proposed to the reader is the following identification problem,

$$\text{Minimize } \int_0^L [I(t) - \tilde{I}(t)]^2 dt,$$

subject to $c \in [0,M]$ $(M > 0)$, where $\tilde{I} \in C([0,L])$, $\tilde{I}(t) \ge 0$ for any $t \in [0,L]$ is a known function and (S,I,R) is the solution to

$$\begin{cases} S'(t) = mN - mS(t) - cS(t)I(t), & t \in (0,L) \\ I'(t) = -mI(t) + cS(t)I(t) - dI(t), & t \in (0,L) \\ R'(t) = -mR(t) + dI(t), & t \in (0,L) \\ S(0) = S_0,\ I(0) = I_0,\ R(0) = R_0. \end{cases}$$

Here m, d, S_0, I_0, R_0 are given constants. The meaning of this problem is the following one. Knowing the number of infectious individuals at any moment we wish to determine the infectivity rate c.

Bibliographical Notes and Remarks

There is an extensive mathematical literature devoted to optimal control theory. This domain developed enormously after the pioneering work of Pontryagin and his collaborators. One of the main purposes when investigating an optimal control problem is to derive first-order necessary conditions of optimality (Pontryagin's principle). Here is a list of important monographs devoted to this subject: [LM67], [Kno81], [Bar93], [Bar94], and [Son98]. More applied optimal control problems can be found only in a few monographs; see [Kno81], [Che86], [Bar94], [Ani00], and [Tre05]. For applications in biology, with a few MATLAB programs we cite [LW07].

Exercises

2.1. Derive the maximum principle for the following problem:

$$\text{Maximize}\{x^u(T) + \gamma y^u(T)\},$$

subject to $u \in L^2(0,T)$, $0 \leq u(t) \leq 1$ a.e. $t \in (0,T)$, where (x^u, y^u) is the solution to the predator–prey system:

$$\begin{cases} x'(t) = r_1 x(t) - \mu_1 u(t)x(t)y(t), & t \in (0,T) \\ y'(t) = -r_2 y(t) + \mu_2 u(t)x(t)y(t), & t \in (0,T) \\ x(0) = x_0,\ y(0) = y_0. \end{cases}$$

Hint. Proceed as in Section 2.3. This problem is a particular case of (P1) (Section 2.1), for $m = 1$, $N = 2$,

$$G(t,u,x,y) = 0, \quad \varphi(x,y) = x + \gamma y,$$

$$f(t,u,x,y) = \begin{pmatrix} r_1 x - \mu_1 uxy \\ \\ -r_2 y + \mu_2 uxy \end{pmatrix},$$

and

$$K = \{w \in L^2(0,T);\ 0 \leq w(t) \leq 1 \text{ a.e. } t \in (0,T)\}.$$

2.2. Derive the maximum principle for the following problem,

$$\text{Maximize}\{x^u(T) + y^u(T)\},$$

subject to $u \in L^2(0,T)$, $0 \leq u(t) \leq 1$ a.e. $t \in (0,T)$, where (x^u, y^u) is the solution to the predator–prey system

$$\begin{cases} x'(t) = r_1 x(t) - kx(t)^2 - \mu_1 u(t)x(t)y(t), & t \in (0,T) \\ y'(t) = -r_2 y(t) + \mu_2 u(t)x(t)y(t), & t \in (0,T) \\ x(0) = x_0,\ y(0) = y_0. \end{cases}$$

Here $r_1, r_2, k, \mu_1, \mu_2$ are positive constants, and kx represents an additional mortality rate and is due to the overpopulation; kx^2 is a logistic term for the prey population.

Hint. Proceed as in Section 2.1. This problem is a particular case of (P1) (Section 2.1).

2.3. Obtain the maximum principle for the following optimal harvesting problem:

$$\text{Maximize} \int_0^T u(t)x^u(t)dt,$$

subject to $u \in L^2(0,T)$, $0 \le u(t) \le M$ $(M > 0)$ a.e. $t \in (0,T)$, where x^u is the solution to the following Malthusian model of population dynamics,

$$\begin{cases} x'(t) = r(t)x(t) - u(t)x(t), & t \in (0,T) \\ x(0) = x_0 > 0. \end{cases}$$

Here $x^u(t)$ represents the density of individuals of a population species at time t, $r \in C([0,T])$ gives the growth rate, and $u(t)$ is the harvesting effort (a control) and plays the role of an additional mortality rate. $\int_0^T u(t)x^u(t)dt$ represents the total harvested population on the time interval $[0,T]$.

Hint. Let u^* be an optimal control. Here are the first-order necessary optimality conditions:

$$\begin{cases} p'(t) = -r(t)p(t) + u^*(t)(1+p(t)), & t \in (0,T) \\ p(T) = 0, \end{cases}$$

$$u^*(t) = \begin{cases} 0 & \text{if } 1 + p(t) < 0 \\ M & \text{if } 1 + p(t) > 0. \end{cases}$$

2.4. Obtain the maximum principle for the following optimal harvesting problem:

$$\text{Maximize} \int_0^T u(t)x^u(t)dt,$$

subject to $u \in L^2(0,T)$, $0 \le u(t) \le M$ $(M > 0)$ a.e. $t \in (0,T)$, where x^u is the solution to the following logistic model of population dynamics,

$$\begin{cases} x'(t) = rx(t) - kx(t)^2 - u(t)x(t), & t \in (0,T) \\ x(0) = x_0 > 0. \end{cases}$$

Here r, k, x_0 are positive constants.

2.5. Derive the optimality conditions for the following problem,

$$\text{Maximize} \int_0^T u(t)x^u(t)dt - c\int_0^T u(t)^2 dt,$$

subject to $u \in L^2(0,T)$, $0 \le u(t) \le M$ $(M > 0)$ a.e. $t \in (0,T)$, where x^u is the solution to the following logistic model of population dynamics,

$$\begin{cases} x'(t) = rx(t) - kx(t)^2 - u(t)x(t), & t \in (0,T) \\ x(0) = x_0 > 0. \end{cases}$$

Here c, r, k, x_0 are positive constants. This problem seeks to maximize the harvest while minimizing effort.

http://www.springer.com/978-0-8176-8097-8

An Introduction to Optimal Control Problems in Life Sciences and Economics
From Mathematical Models to Numerical Simulation with MATLAB®

Aniţa, S.; Arnăutu, V.; Capasso, V.

2011, XII, 232 p., Hardcover
ISBN: 978-0-8176-8097-8
A product of Birkhäuser Basel