The SILP-Relaxation Method in Optimal Control I: General Boundary Conditions

H. RUDOLPH

The paper deals with the relaxation method, a measure-theoretical approach for the treatment of classical control problems. After establishing a linear program over measure spaces as a substitute for a control problem, it is possible to use methods of semi-infinite linear programming (SILP), especially the semi-infinite simplex method (SISM), in order to estimate the solution of the linear program. In this way, approximations for the global optimal solution of the control problem can be obtained; a lower bound for the corresponding optimal value can also be found, by means of duality theory.

Key words: Optimal control, semi-infinite programs
AMS subject classification: 49M39

Introduction

The measure-theoretical approach to classical control problems, basing on ideas of YOUNG, GAMKRELIDZE, VARGA, and developed by RUBIO and VINTER, gives the possibility to develop numerical methods for solving nonlinear control problems by means of linear programs in finite dimensions. Recently the author pointed out how to use SILP-methods, for instance SISM, to get estimations for the optimal solutions of the control problem and the corresponding LP in measure space. Up to now all the methods listed above work only in the case of fixed boundary values; the aim of this note is to show what can be achieved in more general situations.

1. The Control Problem

We shall study a control problem of the following type:
Minimize the integral
\[ \int_0^T r(t, z, u) \, dt \]  
with respect to the state equation
\[ \dot{z} = g(t, z, u) \]  
and control resp. state constraints
\[ u(t) \in U, \quad x(t) \in X. \]

The boundary conditions are given in one of the forms (1.2b) or (1.2b)′:
\[ z_k(0) = \xi_k^0, \quad k \in K, \quad z_l(T) = \xi_l^T, \quad l \in L, \]  
resp.
\[ z(T) = Cx(0). \]
Here $\mathcal{K}, \mathcal{L}$ are subsets of $\{1, \ldots, n\}$ and $C$ is a regular $(n, n)$ matrix. If in case (1.2b) $\mathcal{K} = \mathcal{L} = \{1, \ldots, n\}$, we have a fixed end point problem as studied in [8]. The case (1.2b)' will be called a coupled boundary value problem.

The following analytical assumptions about data and solution functions of the control problem (1.1)–(1.3) have to be fulfilled:

1. $r(\cdot, \xi, v)$ is summable over $[0, T]$ for all $\xi \in X, v \in U$
2. $g(\cdot, \cdot, \cdot)$ is continuous over $X \times U$ for all $t \in [0, T]$
3. $z(\cdot) = [z_1(\cdot), \ldots, z_n(\cdot)]^T$ is an $n$-vector of absolutely continuous state functions
4. $u(\cdot) = [u_1(\cdot), \ldots, u_m(\cdot)]^T$ is an $m$-vector of bounded measurable control functions.

The state equation (1.2a) is to be satisfied almost everywhere over $[0, T]$ in the sense of Carathéodory.

Problem (1.1)–(1.3) is assumed to be consistent, that is, there exists at least one admissible pair (admissible process) $(x, u)$, which satisfies all the constraints in (1.1)–(1.3).

From this follows that initial and end point coordinates, $\xi^0, \xi^T$ are projections of vectors $\xi^0, \xi^T \in X$. We denote by $W$ the set of all admissible pairs, and will use the notation $Z = [0, T] \times X \times U, J = [0, T] \times X$.

2. Variational Description and the Relaxed Problem

In this section our goal is to give a relaxed formulation of the control problem by using an imbedding process: instead of admissible pairs $p = (x, u)$ we shall consider corresponding measures $\mu$ on the space $Z$, where the correspondence $(x, u) \mapsto \mu$ is not injective, the set of admissible measures $\mu$, which will be given by linear constraints, is a larger one then the set $W$. This goal can be achieved by a two-step procedure: variational description and relaxation. First let us give the variational description. Let $\varphi = \varphi(t, \xi), (t, \xi) \in J$, and for $v \in \mathbb{R}^m$ let

$$\varphi^v(t, \xi, v) = \varphi_t(t, \xi) + \nabla \varphi(t, \xi)v(t, \xi),$$

where $\nabla \varphi$ denotes the gradient of $\varphi$ with respect to the argument $\xi \in \mathbb{R}^n$. Then the variational description is implicit in the following theorem:

**Theorem.** A pair $p = (x, u)$ is admissible in (1.1)–(1.3) iff

$$\int_0^T \varphi^v(t, z(t), u(t)) dt = \Delta \varphi \quad \forall \varphi \in C^1(J), z(t) \in X, u(t) \in U \text{ a.e. on } [0, T],$$

where

$$\Delta \varphi = \varphi(T, z(T)) - \varphi(0, z(0)) = \Delta \varphi(0, T, \xi^0, \xi^T)$$

in the case of boundary conditions (1.2b) and

$$\Delta \varphi = \varphi(T, Cz(0)) - \varphi(0, z(0)) = \Delta \varphi(0, T, z(0))$$

in the case of boundary conditions (1.2b)'.

**Proof.** (i) Let $(x, u)$ be an admissible process in (1.1)–(1.3). Then for any $\varphi \in C^1(J)$ there holds

$$\int_0^T \varphi^v(t, z(t), u(t)) dt = \int_0^T [\varphi_t(t, z(t)) + \nabla \varphi(t, z(t))g(t, z(t), u(t))] dt$$

$$= \int_0^T \frac{d}{dt} \varphi(t, z(t)) dt = \varphi(T, z(T)) - \varphi(0, z(0)) = \Delta \varphi$$
in both cases of boundary conditions; the conditions (1.3) are satisfied by feasibility of \((z, u)\).

(ii) Let for any pair \((z, u)\) satisfying (1.3) the infinite set of integral relations be satisfied for all \(\varphi \in C^1(J)\). We choose \(\varphi_j(t, \xi) = \varphi(t, \xi_j), \varphi(\cdot) \in C^\infty(0, T), j \in \{1, \ldots, n\}\). Then we have

\[
\varphi_j^e(t, z(t), u(t)) = \dot{\varphi}(t)z_j(t) + \varphi(t)g_j(t, z(t), u(t)),
\]

and by partial integration there follows

\[
(*) \quad \int_0^T \varphi_j^e(t, z(t), u(t)) dt = \int_0^T \psi(t)[\dot{z}_j(t) - g_j(t, z(t), u(t))] dt.
\]

(a) By choosing \(\psi(t) \in C^\infty(0, T)\) we get \([\psi(t)z_j(t)]^T = 0 = \Delta \varphi_j\) and therefore

\[
\int_0^T \psi(t)[\dot{z}_j(t) - g_j(t, z(t), u(t))] dt = 0 \quad \forall \psi \in C^\infty_0.
\]

From this we obtain \(\dot{z}(t) = g(t, z(t), u(t))\) a.e., such that state equation (1.2a) is satisfied and the integral on the right-hand side in (*) is zero.

(b) The second part of the proof deals with the boundary conditions. In case of boundary conditions (1.2b) we obtain from (*)

\[
\text{for } k \in \mathcal{K}, \psi(0) = 1, \psi(T) = 0 : \quad z_k(0) = \xi^0_k,
\]

and

\[
\text{for } l \in \mathcal{L}, \psi(0) = 0, \psi(T) = 1 : \quad z_l(T) = \xi^T_l,
\]

but in case of boundary conditions (1.2b)' we get from (*) for \(j = 1, \ldots, n\) by choosing

\[
\psi(T) = 1, \psi(0) = 0 : \quad z_j(T) = (Cz(0))_j.
\]

Hence the proof is complete, and the control problem (1.1)-(1.3), with the notation

\[
P = \{p = (z, u) : z(t) \in X, u(t) \in U \text{ a.e. on } [0, T]\}
\]

is equivalent to the following variational description:

\[
\begin{align*}
\int_0^T r(t, z, u) dt & \longrightarrow \text{min} \quad (2.1) \\
\int_0^T \varphi^e(t, z, u) dt & = \Delta \varphi, \varphi \in C^1(J) \quad (2.2) \\
p = (z, u) & \in P. \quad (2.3)
\end{align*}
\]

The second step - the relaxation - consists in introducing a new kind of variables, and so a metamorphosis of the whole problem will be obtained (cf. [8]). For given \(p = (z, u)\) we consider the mapping

\[
A_p : C(Z) \rightarrow \mathbb{R}^1,
\]

given by the formula

\[
A_p f = \int_0^T f(t, z(t), u(t)) dt.
\]

Obviously the mapping \(A_p\) is well-defined, linear, continuous and nonnegative: for every \(f \geq 0\) there holds \(A_p f \geq 0\). Since \(A_p\) is a linear continuous functional over \(C(Z)\), by Riesz' theorem there exists a regular Borel measure \(\mu_p\) on \(Z\), such that

\[
A_p f = \langle f, \mu_p \rangle = \int_Z f d\mu_p.
\]
Then (2.1)-(2.3) can be rewritten equivalently as

$$\langle r, \mu \rangle \rightarrow \min_{\mu, \rho \in \mathcal{P}} \quad (2.4)$$

$$\langle \varphi^0, \mu \rangle = \Delta \varphi, \quad \varphi \in C^1(J) \quad (2.5)$$

$$\mu \geq 0 \quad (2.6)$$

(where the inequality (2.6) denotes the nonnegativity of the measure in the sense mentioned above and can be written as an cone ordering in the space $M(Z)$ of all Radon measures on $Z$, the dual space to $C(Z)$). If we now omit the correspondence $\rho \rightarrow \mu$, and minimize over arbitrary $\mu$ satisfying constraints (2.5), (2.6) and $\text{supp} \mu \subseteq Z$, we get the so-called relaxed problem

\[
\int_Z r \, d\mu \rightarrow \min
\]

$$\int_Z \varphi^0 \, d\mu = \Delta \varphi, \quad \varphi \in C^1(J) \quad (2.7)$$

$$\mu \geq 0 \quad (2.8)$$

This problem can be understood as a nonclassic linearization of the control problem (1.1)-(1.3); in fact it is an infinite-dimensional linear program over the spaces $M(Z)C^1(J)$ in the sense of [1, 4], hence it is possible to use LP-techniques (cf. [10]) in order to study the relaxed problem. As pointed out in [11], there is another interesting fact, which gives some insight in the nature of the relaxed problem: If we construct the dual problem to the control problem (1.1)-(1.3) by means of KLOTZLER's duality theory (see [6]) and on the other hand the dual LP to (2.7)-(2.9), we obtain by each method the same dual. That is, the LP (2.7)-(2.9) can also be interpreted as an convexification of the nonlinear (and in general non-convex) control problem. If we denote the control problem by (P), its KLOTZLER dual by (P)$^D$, the relaxed problem by (R) and its LP-dual by (R)$^D$, the connections between the four problems are shown in the following scheme:

\[
(P) \quad \text{C} \quad (R) \\
\text{K-dual} \quad \text{Relaxation} \quad \text{LP-dual} \\
(P)^D \equiv (R)^D
\]

Since we have the imbedding $(P) \subseteq (R)$, we get

$$\text{Min}(R) \leq \text{Inf}(P) \quad (2.10)$$

and by duality theory

$$\text{Sup}(R)^D = \text{Min}(R) \leq \text{Inf}(P). \quad (2.11)$$

There may be a gap between the values of the problems (R) and (P) (for an example see [5]), but by means of approximation schemes, which are closely related to the concept of asymptotic consistency of DUFFIN [2], RUBIO has proved the following theorem about the relations between problem (R) and the control problem (P) in the case of fixed boundary values.
Approximation Theorem (Rubio [8]). There exists a sequence of pairs $p^j = (x^j, u^j)$ with

$$z^j = g(t, x^j, u^j), \quad x^j(0) = \xi^0, \quad \lim_{j \to \infty} x^j(T) = \xi^T,$$

$$u^j(t) \in U, \quad \lim_{j \to \infty} \rho(x^j(t), X) = 0, \quad t \in [0, T]$$

such that

$$\lim_{j \to \infty} \int_0^T r(t, z^j, u^j) dt = \text{Min}(R) \leq \text{Inf}(P),$$

i.e., $p^j = (x^j, u^j)$ is an almost feasible (asymptotic feasible) global minimising sequence for (P) (here $\rho(x, X)$ is the distance between vector $x$ and set $X$).

The global optimality of the minimizing sequence seems to be one of the greatest advantages of the relaxation approach; most of the numerical methods based on first or higher order necessary conditions like Pontryagin's maximum principle gave only local solutions for the control problem (P).

3. The Numerical Method

We now describe our numerical method, which can be called a semi-infinite approach to problem (R). As mentioned above, the problem (2.7)-(2.9) resp. (R) is an infinite-dimensional linear program with a variable $\mu$, varying in a measure space. Rubio in his approximation scheme has discretized the equations (2.8) by choosing a grid in the set $Z = [0, T] \times X \times U$ and a finite set of functions $\varphi_i$ instead of the infinite set $\varphi \in C$; the result of this discretization procedure was a finite-dimensional linear program with number of equations $M \approx 25$ and number of variables $N \approx 1000$ even in small academic examples. Unfortunately in this approach there is no evident possibility to get a lower bound for the optimal value $\text{Min}(R)$ of problem (2.7)-(2.9) resp. (R) in terms of the values $\text{Min}(R)_{M,N}$; the accuracy of the approximation can be estimated only theoretically in proving convergence, with some unknown constants present. For the semi-infinite approach we choose a finite set of functions $\varphi_i$, $i = 1, \ldots, M$, and consider the semi-infinite LP

$$\begin{align*}
\langle \varphi_i, \mu \rangle &= \Delta \varphi_i, \quad i = 1, \ldots, M \\
\mu &\geq 0.
\end{align*}$$

\begin{align*}
\text{Theorem.} \quad \text{There is an optimal solution } \mu^* = \sum_{k=1}^M \alpha_k \varepsilon_{t_k} \text{ of } (R)_M \text{ with }
\end{align*}\
\begin{align*}
\alpha_k &\geq 0, \quad \sum_{k=1}^M \alpha_k = T, \quad \varepsilon_k = (t_k, \xi_k, \nu_k) \in Z
\end{align*}\

($\varepsilon_z$ is the Dirac measure concentrated in $z \in Z$).

\begin{align*}
\text{Proof.} \quad \text{By consistency of (1.1)-(1.3) we get the consistency of (R) and hence of (R)_M. It is well known that the feasible region is weak* compact and the objective functional is continuous; from this the existence of } \mu^* \text{ follows. The structure of } \mu^* \text{ as a measure with finite support is an obvious consequence of convexity and Carathéodory's theorem.}
\end{align*}\

The numerical procedure consists of two main procedures:

1. Numerical estimation of $\mu^*$ in \( (R)_M \)

10*
The numerical estimation of \(\mu^*\) resp. \((r, \mu^*)\) is possible by the semi-infinite simplex-method SISM (see [10]); the "exact" determination of \((r, \mu^*)\) needs much computational effort, because SISM has in general only linear convergence rate. However in the \(k\)-th step of the algorithm there can be obtained a dual (lower) bound

\[
de_k = (r, \mu^k) - \sum_{i=1}^{M} \eta_i^k \left( \langle \varphi_i, \mu^k \rangle - \Delta \varphi_i \right) \leq (r, \mu^*) \tag{3.3}
\]

with

\[
\eta_i^k : \text{ vector of actual LAGRANGE-multipliers}
\]
\[
\mu_i^k : \text{ measure to exchange in the basic feasible solution,}
\]
so by (2.11), (3.3) we obtain a lower bound for the global minimum in problem (P).

2. Reconstruction of a corresponding to \(\mu^*\) process

The reconstruction of the process \((x^*, u^*)\) corresponding to \(\mu^*\) requires an integration of the state equations, using all the informations obtained from \(\mu^*\). Indeed, the points \(z_k\) in (3.2) have to be interpreted as support points for mean values of the integrals in (3.1) with weights \(\alpha_k\) in a quadrature formula. So we have used as an integration routine a RUNGE-KUTTA-FEHLBERG method for solving the initial value problems

\[
\dot{x}_k = g(t, \xi_k, u_k), \quad x_k(t_k) = x_{k-1}(t_k), \quad t \in [t_k, t_k + \alpha_k], \quad k = 1, \ldots, M \tag{3.4}
\]

by forward integration with \(t_0 = 0\), \(x_0(0) = \xi^0\) (see the example in Section 4, \(\xi^0\) was given, only \(\xi^T\) was not completely known). The integration routine makes use not only of the computed initial values \(z_{k-1}(t_k)\) in the \(k\)-th integration step, but also of the "mean directions" \(g(t_k, \xi_k, u_k)\). The computations have shown the efficiency of such an approach.

4. Application to a Minimum Fuel Space Travel Problem

We shall now show how the method described above works for a control problem known in the literature as minimum fuel space travel problem (see OBERLE [7]). We consider the planar Earth-Mars transfer by a low thrust ion rocket, following OBERLE. The problem is to pilot an ion rocket from the Earthian orbit to the Martian orbit, where both orbits are assumed to be circular and coplanar (HÖHMANN'S assumption) and the influences of the gravitational fields of Earth and Mars are neglected with respect to the gravitation of the Sun. The corresponding control problem is as follows:

\[
\begin{array}{ll}
\text{State variables} & \text{Control variables} \\
r & \text{distance from the sun} & \psi \in [-\pi, \pi] \\
w & \text{radial velocity} & \text{thrust angle} \\
v & \text{tangential velocity} & k \in [0, k_{\text{max}}] \\
m & \text{mass} & \text{thrust}
\end{array}
\]

The problem is to determine control functions \(\psi(t), k(t), t \in [0, T]\) \((T=5.1\text{ is given, i.e. we consider transfer in fixed time})\), such that

\[
F = m(0) - m(T) \rightarrow \text{min} \tag{4.1}
\]

subject to state equations

\[
\begin{align*}
\dot{r} &= w \\
\dot{w} &= \frac{v^2}{r} - \frac{\kappa}{r^2} + k \frac{c}{m} \sin \psi \\
\dot{v} &= -\frac{wv}{r} + k \frac{c}{m} \cos \psi \\
m &= -k,
\end{align*}
\tag{4.2}
\]
boundary conditions

\begin{align*}
    r(0) &= 1, \quad r(T) = r_f = 1.525 \\
    w(0) &= 0, \quad w(T) = 0 \\
    v(0) &= 1, \quad v(T) = \frac{1}{\sqrt{r(T)}} \\
    m(0) &= 1,
\end{align*}

and control restrictions

\[ \psi(t) \in [-\pi, \pi], \quad k(t) \in [0, k_{\text{max}}], \quad t \in [0, T]. \]  

(4.4)

Physically reasonable state constraints are for instance

\begin{align*}
    1 &\leq r(t) \leq r_f \\
    w(t) &\geq 0, \quad v(t) \geq 0 \\
    0 &\leq m(t) \leq 1.
\end{align*}

(4.5)

The parameters in the physical model are \( \kappa = 1, \quad c = 1.872, \quad k_{\text{max}} = 0.075 \).

Problem (4.1)–(4.5), rewritten with state variables \( z_1 = r, \quad z_2 = w, \quad z_3 = v, \quad z_4 = m \), control variables \( u_1 = k, \quad u_2 = \psi \), objective functional \( F = \int_0^T u_1(t) \, dt \) and normalized, such that the control interval is \([0, 1]\), then corresponds to the following semi-infinite relaxed problem:

\begin{align*}
    \int \varphi_i \, d\mu &\longrightarrow \text{min} \tag{4.6} \\
    \int \varphi_i^2 \, d\mu = \Delta \varphi_i, & \quad i \in I_\varphi \tag{4.7} \\
    \int \theta_j \, d\mu &= 0, \quad j \in J_\theta \tag{4.8} \\
    \int_0^1 \chi_k \, d\mu &\geq \tau \delta_k, \quad k \in K_X \tag{4.9} \\
    \mu &\geq 0 \\
\end{align*}

where \( |I_\varphi| \approx 15, \quad |J_\theta| \approx 12, \quad |K_X| \approx 10, \quad \tau \approx 0.75 \ldots 0.99 \). Here equations (4.7) with functions \( \varphi_i \) as monomials in \( t, \xi_1, \ldots, \xi_4 \) are connected with the boundary conditions, equations (4.8) with functions \( \theta_j \) as trigonometric polynomials in \( t \) with coefficients in \( \xi_j \) are corresponding to the state equations, and the third set of conditions, inequalities (4.9) with indicator functions \( \chi_k \), guarantees the uniform distribution of the measure \( \mu \) with respect to \( t \) over the interval \([0, 1]\) in a discrete manner (for details see RUBIO [8], Chapters 4, 5). We have one free boundary value, namely \( m(T) = z_4(1) \); this requires an additional variable for the unknown boundary value. Problem (4.6)–(4.10) corresponds to \((R)_M\) from Section 3; we will also denote it by \((R)_M\). The numerical model, corresponding to \((R)_M\), is as follows:

Let \( z = (t, r, w, v, m, k, \psi) = (t, z_1, \ldots, z_4, u_1, u_2) \in \mathbb{Z} \), and let \( \{z^l : l = 1, \ldots, N\} \subset \mathbb{Z} \) be a grid, \( \mu = \sum_{l=1}^N \alpha_l \varepsilon_l \), a discrete measure. The measure \( \mu \) is feasible in \((R)_M\) iff

\begin{align*}
    \sum_{l=1}^N \alpha_l \varphi_i^l (z^l) &= \Delta \varphi_i, & \quad i \in I_\varphi \\
    \sum_{l=1}^N \alpha_l \theta_j (z^l) &= 0, & \quad j \in J_\theta \\
    \sum_{l=1}^N \alpha_l \chi_k (z^l) &\geq \tau \delta_k, & \quad k \in K_X \\
    \alpha_l &\geq 0, & \quad l = 1, \ldots, N.
\end{align*}
Let $\bar{\mu} = \sum_{i=1}^{M} a_i \xi^*_i$ be a basic feasible solution of $(R)_M$. Then SISM generates $\mu^* = \sum_{i=1}^{M} a_i \xi^*_{i+1}$ as an $(\varepsilon)$-optimal solution of $(R)_M$. Let $I_M = \text{Inf}(R)_M$ and $S_M$ a dual bound (compare (3.3)), namely

$$S_M = \text{Inf} \left\{ (r, \mu) - \sum \eta_i ((a'_i, \mu) - b_i) \mid \mu \geq 0, \text{supp} \mu = Z \right\}$$

with some LAGRANGE multipliers $\{\eta_i\}$ from SISM; then we get

$$S_M \leq I_M \leq \text{Min}(R) \leq \text{Inf}(P) \quad (4.11)$$

such that $S_M$ is a lower bound for the global minimum in problem (P).

Results: With $N \approx 450, M \approx 15 \ldots 25, \delta = 0.05$ (end point accuracy) a computational test gave the following pictures:

The end point tolerances are

$$\max_{i=1,2,3} |z^*_i(T) - \xi^*_i| = 0.03 < \delta$$

(compare with the asymptotic feasibility of $z^*$!) and

$$m^*(0) - m^*(T) = 1.000 - 0.906 = 0.094.$$ 

A dual bound obtained after 21 steps of SISM was:

$$S_M = 1.000 - 0.908 = 0.092,$$

such that (4.11) gives

$$0.092 \leq \text{Min}(R)_M \leq \text{Inf}(P) \leq 0.096,$$

where the value on the right-hand side is taken from OBERLE [7]. It seems to be remarkable that regardless of the "great" end point tolerance the computed bound is (numerically) exact; its goodness is depending only of the chosen $M$, but not depending of $\delta$.

Acknowledgement: The author wishes to thank Prof. Dr. R. BULIRSCH and his group from TU Munich for support, suggestions and helpful discussions at his stay in Munich in March 1990.
References


Received 12. 03. 1991; in revised form 07. 06. 1991

Prof. Dr. Helmut Rudolph
Pädagogische Hochschule Güstrow
Institut für Mathematik
Goldbergerstr. 12
D(Ost)-2600 Güstrow