Dynamic optimization

Dynamic optimization problems are considered, where the decision variables \(x(t)\) are no longer elements of the Euclidean space \(\mathbb{R}^n\) but are elements of an infinite–dimensional (normed) function space \((X, \| \cdot \|_X)\). Herein, the goal is to minimize (or to maximize) an objective functional, also referred to as cost functional or performance index \(J(\cdot) : X \rightarrow \mathbb{R}\) with respect to \(x(t)\).

4.1 Problem statement and preliminaries

The general formulation of the cost functional is the so–called Bolza form

\[
J(u) = \varphi(t_e, x(t_e)) + \int_{t_0}^{t_e} l(t, x(t), u(t))dt. \tag{4.1}
\]

Herein, \(t_e\) and \(x(t_e)\) are the terminal time or end–time, respectively, and terminal state, \(l : [t_0, t_e] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}\) is the running cost or Lagrangian or Lagrangian density, respectively, and \(\varphi : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}\) is the terminal cost. The trajectory \(x(t)\) is governed by the dynamic system

\[
\dot{x} = f(t, x, u), \quad t > t_0, \quad x(t_0) = x_0 \in \mathbb{R}^n \tag{4.2}
\]

with the vector \(u(t) \in U\) representing the control or input or manipulated variable, respectively. The space \(U\) refers to the control space, which is further precised below. It is assumed that there is at least a local unique solution to (4.2), which can be addressed by exploiting the local (or global) Lipschitz continuity of \(f(t, x, u)\). There are two important special cases of the Bolza form. If \(\varphi(t_e, x(t_e)) = 0\) the so–called Lagrange form arises, i.e.

\[
J(u) = \int_{t_0}^{t_e} l(t, x(t), u(t))dt. \tag{4.3}
\]

The so–called Mayer form is obtained when \(l(t, x, u) = 0\), i.e.,

\[
J(u) = \varphi(t_e, x(t_e)). \tag{4.4}
\]

Remark 4.1

It is possible to convert these two special cases into each other by simple transformations. Starting with the Mayer form (4.4) we obtain

\[
\varphi(t_e, x(t_e)) = \varphi(t_0, x(t_0)) + \int_{t_0}^{t_e} \frac{d}{dt} \varphi(t, x(t))dt = \varphi(t_0, x(t_0)) + \int_{t_0}^{t_e} \frac{\partial}{\partial t} \varphi(t, x(t)) + \frac{\partial}{\partial x} \varphi(t, x(t)) f(t, x(t), u(t))dt, \tag{4.5}
\]
Additional constraints may be imposed. It is thereby assumed that the initial time \( t_0 \) and the initial state \( x_0 \) are fixed. The trajectory \( x(t) \) is called admissible, if it fulfills all constraints in the interval \( t \in [t_0, t_e] \) and the set of all admissible trajectories is, as before, called \( X_{\text{ad}} \). For proper classification let \( X_{\text{ta}} = [t_0, \infty) \times \mathbb{R}^n \) denote the so–called target set and let \( t_e \) be the smallest time such that \( (t_e, x(t_e)) \in X_{\text{ta}} \). Let \( X_{\text{ta}} = [t_0, \infty) \times \{ x_e \} \) with \( x_e \) a fixed point in \( \mathbb{R}^n \), then this refers to a free–time, fixed–endpoint problem. If \( X_{\text{ta}} = \{ t_e \} \times \mathbb{R}^n \), then this defines a fixed–time, free–endpoint problem. The most restrictive case is obviously given by the fixed–time, fixed–endpoint problem with \( X_{\text{ta}} = \{ t_e \} \times \{ x_e \} \). The least restrictive free–time, free–endpoint problem is defined by means of \( X_{\text{ta}} = [t_0, \infty) \times \mathbb{R}^n \). All other cases may be similarly included in the formulation. In addition, path constraints

\[
\psi(t, x(t), u(t)) = 0, \quad t \in [t_0, t_e],
\]

and inequality constraints can be specified, i.e.

\[
\psi(t, x(t), u(t)) \leq 0, \quad t \in [t_0, t_e].
\]

The latter can be often reduced to input constraints

\[
u^r_j \leq u_j(t) \leq u^s_j, \quad j \in I_u \subseteq [1, \ldots, m]
\]

or state constraints

\[
x^r_j \leq x_j(t) \leq x^s_j, \quad j \in I_x \subseteq [1, \ldots, n]
\]

Moreover, so–called isoperimetric constraints may arise, which are given in the form

\[
\int_{t_0}^{t_e} \psi_k(t, x(t), u(t)) \, dt = a_k, \quad k = 1, \ldots, r < n
\]

During the subsequent analysis we make use of the space \( C^k([t_0, t_e], \mathbb{R}^n) \) of \( k \)-times continuously differentiable functions mapping the interval \([t_0, t_e]\) to \( \mathbb{R}^n \) as well as the space \( \dot{C}^k([t_0, t_e], \mathbb{R}^n) \) of piecewise \( k \)-times continuously differentiable functions. Note that a real–valued function \( f(t), t \in [t_0, t_e] \) is called piecewise \( k \)-times continuously differentiable if there is a finite (irreducible) partition \( t_0 = \tau_0 < \tau_1 < \cdots < \tau_N < \tau_{N+1} = t_e \) such that \( f(t) \in C^k((\tau_j, \tau_{j+1}), \mathbb{R}^n) \) for each \( j = 0, 1, \ldots, N \). If \( n = 1 \), then we often write \( C^k([t_0, t_e]) \) and \( \dot{C}^k([t_0, t_e]) \) to refer to \( C^k([t_0, t_e], \mathbb{R}) \) and \( \dot{C}^k([t_0, t_e], \mathbb{R}) \).

Similar to static constrained optimization problems dynamic constrained optimization problems lead to the definition of a feasible control and feasible pair [6].
Definition 4.1: Feasible control and feasible pair

An admissible control $u(t) \in U$ is said to be feasible if (i) the corresponding solution $x(t) = x(t; x_0, u(t))$ to (4.2) is defined on $t \in [t_0, t_e]$ and (ii) $u(t)$ and $x(t; x_0, u(t))$ satisfy all constraints for $t \in [t_0, t_e]$. In this case the pair $(u(t), x(t))$ is called a feasible pair. The set of feasible controls $U_{fe}$ is defined as the set $\{u \in U : u(t) \text{ is feasible}\}$.

It is furthermore required to define what is meant when referring to a global and local minimum of a cost functional $J(u)$.

Definition 4.2: Global and local minimizer

Let $u^*(t) \in U_{fe}$. Then $u^*(t)$ is a global minimizer of the cost functional $J(u)$ if

$$J(u^*) \leq J(u), \quad \forall u \in U_{fe}. \quad (4.12)$$

Moreover, $u^*(t)$ is a local minimizer of the cost functional $J(u)$ if

$$\exists \rho > 0 \text{ such that } J(u^*) \leq J(u), \quad \forall u \in U_{fe} \cap B_\rho(u^*) \quad (4.13)$$

with $B_\rho(u^*) = \{u \in U : \|u - u^*\| \leq \rho\}$.

This implies that the analysis of the local behavior in the neighborhood of $u^*(t)$ requires the definition of a norm $\|\cdot\|$. Given the space $C^0([t_0, t_e], \mathbb{R}^n)$ a commonly used norm is the maximum norm defined as

$$\|u\|_\infty = \max_{t \in [t_0, t_e]} \|u(t)\|_{\mathbb{R}^m} \quad (4.14)$$

with $\cdot_{\mathbb{R}^m}$ denoting the standard Euclidean norm on $\mathbb{R}^m$. For the space $C^1([t_0, t_e], \mathbb{R}^m)$ this extends to

$$\|u\|_{1,\infty} = \max_{t \in [t_0, t_e]} \|u(t)\|_{\mathbb{R}^m} + \max_{t \in [t_0, t_e]} \|\dot{u}(t)\|_{\mathbb{R}^m}. \quad (4.15)$$

Local minimizers according to Definition 4.2 can be further subdivided depending on the choice of the norm $\|\cdot\|$ in (4.13). If $\|\cdot\| = \|\cdot\|_\infty$, then one refers to the local minimizer also as strong local minimizer. If $\|\cdot\| = \|\cdot\|_{1,\infty}$, then the local minimizer is also called weak local minimizer.

Remark 4.2

While all norms defined on finite-dimensional spaces such as $\mathbb{R}^m$ are equivalent this no longer holds for infinite-dimensional function spaces. Hence, according to (4.13) $u^*(t)$ may be a local minimizer with respect to one norm but not with respect to another norm on $U_{fe}$.

4.2 Calculus of variations

To deduce optimality conditions we subsequently apply calculus of variations. Contrary to the determination of extrema of functions, calculus of variations enables us to determine the extrema of functionals, which map an element of a function space to the underlying field ($\mathbb{R}$ or $\mathbb{C}$). We have already seen the scalar product as an example of a functional. For the present considerations we focus
on integrals with respect to a single independent coordinate \( t \) with integral kernel in \( x(t) \) and \( \dot{x}(t) \), i.e.
\[
J(x) = \int_{t_0}^{t_e} l(t, x(t), \dot{x}(t)) \, dt.
\] (4.16)

For the sake of simplicity \( x(t) \) and its derivative \( \dot{x}(t) \) are used as dependent coordinates in the definition of the functional. We now seek for the function \( x(t), \ t \in [t_0, t_e] \) for which \( J(x) \) attains an extremum.

4.2.1 Preliminaries

In the following, selected results from calculus of variations are summarized, which are utilized throughout this chapter.

4.2.1.1 First variation or Gâteaux derivative

Searching for the zeros of the so-called (first) variation of a functional, also called Gâteaux derivative, enables us to generalize the necessary condition for an extremum \( (\nabla f)(x) = 0 \) given a function \( f(x) \) to the case of functionals [16, 20, 6, 15].

**Definition 4.3: Variation of a functional (Gâteaux derivative)**

Let \( J(x) \) be a functional defined on a linear space \( X \). The first variation of \( J \) at \( x \in X \) in the direction \( \xi \in X \), also called Gâteaux derivative with respect to \( \xi \) at \( x \), is defined as
\[
\delta J(x, \xi) = \lim_{\eta \to 0} \frac{J(x + \eta \xi) - J(x)}{\eta} = \frac{\partial}{\partial \eta} J(x + \eta \xi) \bigg|_{\eta=0}.
\] (4.17)

If \( \delta J(x, \xi) \) exists for all \( \xi \in X \), then \( J(x) \) is said to be Gâteaux differentiable at \( x \).

Note that this implies the existence of the Gâteaux derivative provided that \( J(x) \) is defined and \( J(x + \eta \xi) \) is differentiable with respect to \( \eta \) at \( \eta = 0 \). The first variation or Gâteaux derivative is a linear operation on the functional \( J \) fulfilling the

(i) additivity property, i.e.
\[
\delta (J_1 + J_2)(x, \xi) = \delta J_1(x, \xi) + \delta J_2(x, \xi)
\] (4.18a)

(ii) and the homogeneity property, i.e.
\[
\delta J(x, \alpha \xi) = \alpha \delta J(x, \xi).
\] (4.18b)

In addition, the Gâteaux derivatives satisfies

(iii) the product rule, i.e.,
\[
\delta (J_1J_2)(x, \xi) = \delta J_1(x, \xi)J_2(x, \xi) + J_1(x, \xi)\delta J_2(x, \xi)
\] (4.18c)
(iv) and the quotient rule

\[ \delta \left( \frac{J_1}{J_2} \right)(x, \xi) = \frac{\delta J_1(x, \xi)J_2(x, \xi) - J_1(x, \xi)\delta J_2(x, \xi)}{(J_2(x, \xi))^2}. \]  

(4.18d)

Since \( \delta J(x, \xi_1) + \delta J(x, \xi_2) \) is not necessarily equal to \( \delta J(x, \xi_1 + \xi_2) \) the Gâteaux derivative may in general not define a linear operator on \( X \). It is moreover important to note that the Gâteaux derivative, when non–vanishing, is independent of the norm on \( X \) and hence holds for any norm on \( X \).

**Example 4.1.** Consider the functional (4.16) with \( x(t) \in C^1([t_0, t_e], \mathbb{R}^n) \). With (4.17) the Gâteaux derivative of \( J(x) \) follows as

\[
\delta J(x, \xi) = \frac{\partial}{\partial \eta} \int_{t_0}^{t_e} l(t, x(t), \dot{x}(t) + \eta \xi(t)) dt \bigg|_{\eta=0} = \int_{t_0}^{t_e} \left\{ (\nabla_x l)^T(t, x(t), \dot{x}(t)) \xi(t) + (\nabla_x \dot{l})^T(t, x(t), \dot{x}(t)) \xi(t) \right\} dt
\]

for all \( \xi(t) \in C^1([t_0, t_e], \mathbb{R}^n) \). Hence, \( J(x) \) is Gâteaux differentiable at any \( x(t) \in C^1([t_0, t_e], \mathbb{R}^n) \).

The following example illustrates that the Gâteaux derivative must not necessarily exist.

**Example 4.2.** Consider the functional

\[ J(x) = \int_{t_0}^{t_e} |x(t)| dt \]

for \( x(t) \in C^1([t_0, t_e], \mathbb{R}) \). The value of \( J(x) \) is finite given the interval \( t \in [t_0, t_e] \) with \( 0 \leq t_0 < t_e < \infty \).

From (4.17), the Gâteaux derivative reads

\[
\delta J(x, \xi) = \lim_{\eta \to 0} \frac{1}{\eta} \left( \int_{t_0}^{t_e} |x(t) + \eta \xi(t)| dt - \int_{t_0}^{t_e} |x(t)| dt \right).
\]

Consider now the particular point \( x(t) = x_0(t) = 0, \xi(t) = \xi_0(t) = t \), which implies

\[
\delta J(x_0, \xi_0) = \lim_{\eta \to 0} \frac{1}{\eta} \int_{t_0}^{t_e} t dt = \int_{t_0}^{t_e} \frac{(t_e - t_0)^2}{2}, \quad \eta \to 0^+.
\]

Obviously in the direction \( \xi_0(t) = t \) the Gâteaux derivative at \( x_0(t) = 0 \) does not exist.

The first variation or Gâteaux derivative, respectively, enables us to formulate a geometric first order necessary optimality condition. For this, we first provide an exclusion criteria.

**Lemma 4.1: Exclusion of minima**

Let \( J \) be a functional defined on a normed linear space \((X, \| \cdot \|_X)\). Suppose that at a point \( \bar{x} \in X \) there exists a direction \( \xi \in X \) such that \( \delta J(\bar{x}, \xi) < 0 \). Then \( \bar{x} \in X \) cannot be a local minimizer for \( J \) (in the sense of the norm \( \| \cdot \|_X \)).
The proof of Lemma 4.1 can be found in [6] and makes use of the following remark.

**Remark 4.3**
A direction $\xi \in X$ such that $\delta J(\bar{x}, \bar{\xi}) < 0$ defines a *descent direction* for $J$ at $\bar{x} \in X$. In other words, $\delta J(\bar{x}, \bar{\xi}) < 0$ generalizes the algebraic condition $(\nabla f)^T (\bar{x}) \xi < 0$ for $\xi$ to be a descent direction at $\bar{x}$ given $f : \mathbb{R}^n \to \mathbb{R}$.

To properly address the minimization of a functional over a subset of a normed linear space so-called *admissible directions* need to be defined [6].

**Definition 4.4: Admissible directions**
Let $J$ be a functional defined on a subset $X_{ad}$ of a normed linear space $(X, \| \cdot \|_X)$ and let $\bar{x} \in X_{ad}$. Then, a direction $\xi \in X_{ad}, \xi \neq 0$ is said to be *admissible* (or $X_{ad}$-admissible) for $J$ at $\bar{x}$ if

(i) $\delta J(\bar{x}, \xi)$ exists and

(ii) $\bar{x} + \eta \xi \in X_{ad}$ for all sufficiently small $\eta$, i.e., $\exists \rho > 0$ such that $\forall \eta \in B_{\rho}(0)$ we have $\bar{x} + \eta \xi \in X_{ad}$.

With these preparations, a first order necessary optimality conditions is summarized in the theorem below.

**Theorem 4.1: First order necessary optimality condition**
Let $J$ be a functional defined on a subset $X_{ad}$ of a normed linear space $(X, \| \cdot \|_X)$ and suppose that $x^* \in X_{ad}$ is a local minimizer of $J$. Then

$$\delta J(x^*, \xi) = 0 \quad (4.19)$$

for all $X_{ad}$-admissible directions $\xi$ at the point $x^*$.

**Proof.** The proof of Theorem 4.1 follows by contradiction. Suppose there is a $X_{ad}$-admissible direction $\xi$ at $x^*$ with $\delta J(x^*, \xi) < 0$. Then by Lemma 4.1 the point $x^*$ cannot be a local minimizer for $J(u)$. To exclude $\delta J(x^*, \xi) > 0$ let $-\xi$ be a $X_{ad}$-admissible direction. Then $\delta J(x^*, -\xi) = -\delta J(x^*, \xi) < 0$ so that Lemma 4.1 excludes $x^*$ from the set of local minimizers. Hence, we must have $\delta J(x^*, \xi) = 0$ for any $X_{ad}$-admissible direction $\xi$ at $x^*$.

**4.2.1.2 Euler–Lagrange equations**
Consider the Lagrange form (4.16) with fixed initial and end point, i.e.,

$$J(x) = \int_{t_0}^{t_e} l(t, x(t), \dot{x}(t)) \, dt, \quad x(t_0) = x_0, \quad x(t_e) = x_e. \quad (4.20)$$

We subsequently seek for a local minimum $x^*(t) \in C^1([t_0, t_e], \mathbb{R}^n)$ of (4.20) such that $J(x) \geq J(x^*)$ with $x^*(t_0) = x_0$ and $x^*(t_e) = x_e$. Starting from the unknown $x^*(t)$ introduce so-called admissible functions defined as

$$x(t) = x^*(t) + \eta \xi(t), \quad \xi(t_0) = 0, \quad \xi(t_e) = 0. \quad (4.21)$$
The function \( \xi(t) \) fulfilling \( \xi(t_0) = 0 \) and \( \xi(t_e) = 0 \) is also called *admissible variation*. With this, \( J(x) \geq J(x^*) \) implies

\[
J(x) = J(x^* + \eta \xi) = J(x^*) + (\nabla_x J)^T(x^*) \eta \xi + (\nabla_x J)^T(x^*) \eta \xi + o^2(\eta) \geq J(x^*).
\]

By assumption \( J(x) \) attains a local extremum at \( x(t) = x^*(t) \) for \( \eta = 0 \) so that we have as a necessary optimality condition

\[
(\nabla_x J)^T(x^*) \eta \xi + (\nabla_x J)^T(x^*) \eta \xi = 0
\]

and hence

\[
\delta J(x^*, \xi) = 0
\]

since \( \eta \) is arbitrary. In other words, the *first variation* of \( J \) at \( x^*(t) \) has to evaluate to zero. The explicit evaluation of \( \delta J(x^*, \xi) \) is already presented in Example 4.1, i.e.,

\[
\delta J(x^*, \xi) = \int_{t_0}^{t_e} T(t, x^*(t), \dot{x}^*(t)) \xi(t) \, dt
\]

for all \( \xi(t) \in C^0([t_0, t_e], \mathbb{R}^n) \), then \( f(t) = 0, t \in [t_0, t_e] \) (almost everywhere possibly excluding a set of measure zero).

Thus, for arbitrary \( \xi(t) \in C^1([t_0, t_e], \mathbb{R}^n) \) the equation

\[
\delta J(x^*, \xi) = \int_{t_0}^{t_e} \left\{ (\nabla_x l)^T(t, x^*(t), \dot{x}^*(t)) - \frac{\partial}{\partial t} (\nabla_x l)^T(t, x^*(t), \dot{x}^*(t)) \right\} \xi(t) \, dt = 0
\]

yields

\[
(\nabla_x l)(t, x^*(t), \dot{x}^*(t)) - \frac{\partial}{\partial t} (\nabla_x l)(t, x^*(t), \dot{x}^*(t)) = 0.
\]

These equations are the so-called *Euler–Lagrange equations*, which impose a necessary optimality condition for \( x^*(t) \) to be the local minimizer of (4.20).

**Theorem 4.2: Euler–Lagrange equations**

Consider the functional

\[
J(x) = \int_{t_0}^{t_e} l(t, x(t), \dot{x}(t)) \, dt, \quad x(t_0) = x_0, \ x(t_e) = x_e
\]

(4.23)
for \( x(t) \in C^1([t_0, t_e], \mathbb{R}^n) \) and continuously differentiable Lagrangian density \( l: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \). Suppose that \( x^*(t) \) is a local minimizer of \( J(x) \) with \( x^*(t_0) = x_0 \) and \( x^*(t_e) = x_e \). Then \( x^*(t) \) fulfills the Euler–Lagrange equations

\[
\frac{\partial}{\partial t} (\nabla_\dot{x} l(t, x^*(t), \dot{x}^*(t))) - (\nabla_x l)(t, x^*(t), \dot{x}^*(t)) = 0
\]

(4.24)

for all \( t \in [t_0, t_e] \).

The solution of the Euler–Lagrange equations can be formulated using so–called first integrals in some special cases:

(i) The Lagrangian density is independent of \( t \), i.e., \( l = l(x(t), \dot{x}(t)) \). With the so–called Hamilton function

\[
H(x, \dot{x}) = (\nabla_\dot{x} l)^T(x, \dot{x}) \dot{x} - l(x, \dot{x})
\]

(4.25)

the Euler–Lagrange equations (4.24) reduce to

\[
\frac{d}{dt} H(x, \dot{x}) = \frac{d}{dt} (\nabla_\dot{x} l)^T(x, \dot{x}) \dot{x} + (\nabla_\dot{x} l)^T(x, \dot{x}) \ddot{x} - (\nabla_x l)^T(x, \dot{x}) \dot{x} - (\nabla_\dot{x} l)^T(x, \dot{x}) \ddot{x}
\]

= \left( \frac{d}{dt} (\nabla_\dot{x} l)^T(x, \dot{x}) - (\nabla_x l)^T(x, \dot{x}) \right) \ddot{x}.

(4.26)

Thus, either \( \dot{x}(t) = 0 \) and \( x(t) = c \) or the Hamiltonian \( H(x, \dot{x}) \) must remain constant along a local minimizer \( x^*(t) \) so that \( H(x, \dot{x}) \) is an invariant of the Euler–Lagrange equations.

(ii) The Lagrangian density is independent of \( x(t) \), i.e., \( l = l(t, \dot{x}(t)) \). Then the evaluation of the Euler–Lagrange equations (4.24) yields

\[
\frac{\partial}{\partial t} (\nabla_\dot{x} l)^T(t, \dot{x}) = 0.
\]

(4.27)

Thus, \( \frac{\partial}{\partial x_j} l(t, \dot{x}), \ j = 1, \ldots, n \) is an invariant of the Euler–Lagrange equations.

**Example 4.3 (Brachistochrone problem).** The so–called brachistochrone problem traces back to Johann Bernoulli and refers to finding the path between two points in a vertical plane such that a particle sliding without friction and initial speed \( v_0 \) along this path takes minimal time to travel from the initial to the end point (cf. Figure 4.1).
Since the particle slides without friction energy is conserved so that the sum of kinetic and potential energy equals zero at any instance of time. With the mass \( m \) of the particle this implies

\[
\frac{1}{2} m (v^2(x) - v_0^2) + mg(y(x) - y_0) = 0
\]

and hence

\[
v(x) = \sqrt{\frac{v_0^2}{2} - 2g(y(x) - y_0)}.
\]

The objective function is the traveling time from the initial to the final point

\[
J(y) = \int_{x_0}^{x_1} dt = \int_{x_0}^{x_1} \frac{ds}{v(x)}
\]

with \( s \) denoting the Jordan length of \( y(x) \). Taking into account an infinitesimal element along the path, it follows that

\[
(ds)^2 = (dy)^2 + (dx)^2
\]

so that

\[
ds = \sqrt{(dy)^2 + (dx)^2} = dx \sqrt{1 + (y'(x))^2}.
\]

Substitution into \( J(y) \) results in

\[
J(y) = \int_{x_0}^{x_1} \sqrt{\frac{1 + (y'(x))^2}{\frac{v_0^2}{2} - 2g(y(x) - y_0)} \sqrt{\frac{1}{1 + (y'(x))^2}}} dx = l(y(x), y'(x))
\]

with the conditions that

\[
y(x_0) = y_0, \quad y(x_1) = y_1.
\]

Obviously, since \( J(y) \) does not explicitly depend on the independent coordinate \( x \) we know from the two special cases discussed above that the Hamiltonian function

\[
H(y, y') = \frac{\partial}{\partial y'} l(y, y') y' - l(y, y')
\]

\[
= \frac{(y'(x))^2}{\sqrt{\frac{v_0^2}{2} - 2g(y(x) - y_0)} \sqrt{1 + (y'(x))^2}} - \sqrt{\frac{1 + (y'(x))^2}{\frac{v_0^2}{2} - 2g(y(x) - y_0)}}
\]

is an invariant of the respective Euler–Lagrange equations (4.24) so that \( H(y, y') = c \) with constant \( c \). In order to solve this (differential) equation for \( y(x) \) it is convenient to introduce the substitution

\[
y(x) = y_0 - \bar{y}(x) + \frac{v_0^2}{2g} \Rightarrow y'(x) = -\bar{y}'(x) \land \bar{y}(x_0) = \frac{v_0^2}{2g}
\]

so that

\[
H(\bar{y}, \bar{y}') = \frac{1}{\sqrt{2g}} \left( \frac{(\bar{y}'(x))^2}{\sqrt{\bar{y}(x)} \sqrt{1 + (\bar{y}'(x))^2}} - \sqrt{\frac{1 + (\bar{y}'(x))^2}{\sqrt{\bar{y}(x)}}} \right) = c.
\]
The latter equation implies

\[
\frac{(y'(x))^2}{\sqrt{y(x)} \sqrt{1 + (y'(x))^2}} - \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{y(x)}} = \sqrt{2}gc
\]

and hence

\[
y(x)(1 + (y'(x))^2) = \frac{1}{2g}c^2
\]

With \(a = 1/(2gc^2)\) the differential equation allows for a formal integration in \(x\) to obtain

\[
x - x_0 = \int_{\theta_0}^{\theta} \frac{s}{a - s} \, ds.
\]

The integral on the right–hand side can be solved by making use of the substitutions

\[
s = a \sin^2 \left(\frac{\theta}{2}\right),
\]

which provides

\[
x - x_0 = \frac{a}{2} \left(\theta - \theta_0 + \sin(\theta_0) - \sin(\theta)\right).
\]

As a result, the solution to the brachistochrone problem follows in parametrized form

\[
x(\theta) = x_0 + \frac{a}{2} \left(\theta - \theta_0 + \sin(\theta_0) - \sin(\theta)\right)
\]

\[
y(\theta) = y_0 + \frac{v_0^2}{2g} - \frac{a}{2} \left(1 - \cos(\theta)\right)
\]

for \(\theta \in [\theta_0, \theta_1]\), where \(\theta_0\) and \(\theta_1\) have to be determined from the algebraic equations

\[
\frac{a}{2} \left(1 - \cos(\theta_0)\right) = \bar{y}(x_0) = \frac{v_0^2}{2g}
\]

\[
\frac{a}{2} \left(1 - \cos(\theta_1)\right) = \bar{y}(x_1) = y_0 - y_1 + \frac{v_0^2}{2g}.
\]

It should be pointed out that since Theorem 4.2 only provides a first order necessary optimality condition the solution determined above is only a candidate for a local minimizer and the actual verification requires further analysis.

This problem is moreover suitable to deduce a numerical solution procedure by evaluating the Euler–Lagrange equations for the brachistochrone problem, which yields

\[
\frac{y''(x)}{(y'(x))^2 + 1} \left(\frac{v_0^2 - 2g(y(x) - y_0)}{(y'(x))^2 + 1} + 1\right) = 0, \quad x \in (x_0, x_1)
\]

subject to

\[
y(x_0) = y_0, \quad y(x_1) = y_1.
\]
Equations (4.29) obviously define a boundary-value problem, which can be solved either analytically or numerically. Subsequently, a numerical solution approach using the function `bvp4c` of MATLAB is briefly summarized. This numerical solver relies on the formulation of the boundary-value problem as a system of coupled first-order ordinary differential equations. For this, introduce the state variables $z_1(x) = y(x)$ and $z_2(x) = y'(x)$ so that (4.29) reads

$$
\frac{dz_1}{dx} z_2 = \begin{bmatrix}
\frac{z_2^2}{v_0^2 - 2g(z_1 - y_0)} \\
g \left( \frac{z_2^2}{v_0^2 - 2g(z_1 - y_0)} + 1 \right)
\end{bmatrix}, \quad x \in (x_0, x_1)
$$

(4.30)

$z_1(x_0) = y_0, \quad z_2(x_1) = y_1.$

The following MATLAB function computes a solution to this boundary-value problem when varying $x_1 \in \{1, 2, 5, 7\}$. The respective numerical results are depicted in Figure 4.2.

```matlab
def brachistochrone_main()
    %System parameters
    p.g = 9.81;
    p.v0 = 0.5;
    p.x0 = 0.0;
    p.y0 = 0.0;
    p.y1 = -1.0;

    %Vary x1
    x1 = [1.0,2.0,5.0,7.0];
    col = {'b','r','g','k'};
    figure(1); hold on;
    for j=1:length(x1)
        p.x1 = x1(j);
        if j==1
            % Two runs with the solution of the first run as initial value
            disp('-> Run 1');
            sol = brachistochrone_bvp4c(p,1e-2);
        else
            %for the second run to improve the result
            disp('-> Run 2');
            sol = brachistochrone_bvp4c(sol,1e-2);
        end
    end
end
```

**Figure 4.2:** Numerical solution using the function `bvp4c` from MATLAB.

4.2 Calculus of variations
disp('-> Run 2');
sol = brachistochrone_bvp4c(p,1e-5,sol);
else
    % Use the previous solution as initial value mapped to the new % interval [x0,x1]
    ini.x = linspace(p.x0,p.x1,length(sol.x));
    ini.z = sol.z;
    disp('-> Run 1');
    sol = brachistochrone_bvp4c(p,1e-3,ini);
    disp('-> Run 2');
    sol = brachistochrone_bvp4c(p,1e-5,sol);
end
figure(1);
plot(sol.x,sol.z(1,:),strcat(col{j},'-'));
end

% --------------------------------------------------------
% SUBFUNCTIONS

function out = brachistochrone_bvp4c(p,reltol,varargin)
%
% Initialize
if nargin==2
    xini = linspace(p.x0,p.x1,21);
    a1 = (p.y0-p.y1)/(p.x0-p.x1);
    a0 = p.y0 - a1*p.x0;
    yini = a1*xini+a0;
    ypini = a1*ones(size(xini));
    solinit.x = xini;
    solinit.y = [yini;ypini];
    figure(11); plot(xini,yini); drawnow;
elseif nargin==3
    solinit.x = varargin{1}.x;
    solinit.y = varargin{1}.z;
end
%
%Solve BVP
options = bvpset('RelTol',reltol,'Stats','on');
options = bvpset(options,'FJacobian',@(x,z)jacelsys(x,z,p),...  
        'BCJacobian',@(za,zb)jacbcs(za,zb,p));
sol = bvp4c(@(x,z)elsys(x,z,p),@(za,zb)bcs(za,zb,p),solinit,options);
out.sol = sol;
out.x = sol.x;
out.z = sol.y;

function dzdx = elsys(x,z,p)
dzdx = [z(2);
        p.g*(1+z(2)^2)/(p.v0^2-2.0*p.g*(z(1)-p.y0))];

function out = bcs(za,zb,p)
\[
\text{out} = [\text{za}(1) - p.y0; \\
\text{zb}(1) - p.y1];
\]

function \text{out} = \text{jacelsys}(x,z,p)
\[
\text{out} = [0.0,1.0; \\
2.0*p.g^2*(1+z(2)^2)/(p.v0^2-2.0*p.g*(z(1)-p.y0))^2, \\
2.0*p.g*z(2)/(p.v0^2-2.0*p.g*(z(1)-p.y0))];
\]

function \[\text{dbcdza},\text{dbcdzb}\] = \text{jacbcs}(x,z,p)
\[
\text{dbcdza} = \text{zeros}(2,2); \\
\text{dbcdza}(1,1) = 1.0; \\
\text{dbcdzb} = \text{zeros}(2,2); \\
\text{dbcdzb}(2,1) = 1.0;
\]

It can be furthermore shown similar to the analysis of extrema of functions that if \(x^*(t)\) is a local minimizer, then the second variation of \(J(x)\) at \(x^*(t)\) needs to be positive semi–definite, i.e.,
\[
\delta^2 J(x^*, \xi) = \left. \frac{d^2 J(x^* + \eta \xi)}{d\eta^2} \right|_{\eta=0} \geq 0.
\]

This leads to the so–called Legendre condition \[7\].

**Theorem 4.3: Legendre condition**

Consider the functional
\[
J(x) = \int_{t_0}^{t_e} l(t, x(t), \dot{x}(t))dt, \quad x(t_0) = x_0, \ x(t_e) = x_e
\]

for \(x(t) \in C^1([t_0, t_e], \mathbb{R}^n)\) and twice continuously differentiable Lagrangian density \(l: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}\). Suppose that \(x^*(t)\) is a local minimizer of \(J(x)\) with \(x^*(t_0) = x_0\) and \(x^*(t_e) = x_e\). Then \(x^*(t)\) solves the Euler–Lagrange equations and satisfies the so–called Legendre condition
\[
(\nabla^2_X l)(t, x^*(t), \dot{x}^*(t)) \geq 0
\]

for all \(t \in [t_0, t_e]\).

### 4.2.1.3 Euler–Lagrange equations for problems with free end–point

We conclude this section by addressing a Bolza problem with a free end–point using calculus of variations, i.e.,

**Theorem 4.4: Euler–Lagrange equations for problems with free end–point and terminal cost**

Consider the linear functional
\[
J(t_e, x) = \varphi(t_e, x(t_e)) + \int_{t_0}^{t_e} l(t, x(t), \dot{x}(t))dt
\]
on the subset
\[
X_{\text{ad}} = \{(t_e, x(t)) : t_e \in (t_0, \bar{t}), \ x(t) \in C^1([t_0, \bar{t}], \mathbb{R}^n), \ x(t_0) = x_0\}
\]
for sufficiently large \( \bar{t} \gg t_e \) with \( l \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n, \varphi \in \mathbb{R} \times \mathbb{R}^n \) being continuously differentiable. Suppose that \((t^*_e, x^*(t))\) denotes a local minimum of \( J(t, x) \) on \( X_{ad} \). Then \( x^*(t) \) solves the Euler–Lagrange equations (4.24) on the interval \( t \in [t_0, t^*_e] \) and satisfies both the initial condition \( x^*(t_0) = x_0 \) and the transversality conditions

\[
\left[(\nabla_x l)(t, x(t), \dot{x}(t)) + (\nabla_x \varphi)(t, x(t))\right]_{t = t^*_e, x = x^*} = 0 \quad (4.34a)
\]

\[
\left[l(t, x(t), \dot{x}(t)) - (\dot{x})^T (\nabla_x l)(t, x(t), \dot{x}(t)) + \frac{\partial}{\partial t} \varphi(t, x(t))\right]_{t = t^*_e, x = x^*} = 0. \quad (4.34b)
\]

For the proof of Theorem 4.4 the definition of the Gâteaux derivative (4.17) can be directly modified according to

\[
\delta J(t_e, x, \tau, \xi) = \lim_{\eta \to 0} \frac{J(t_e + \eta \tau, x + \eta \xi) - J(t_e, x)}{\eta} = \frac{\partial}{\partial \eta} J(t_e + \eta \tau, x + \eta \xi) \bigg|_{\eta = 0}. \quad (4.35)
\]

Proof. If the end–point \( t^*_e \) is fixed, then Theorem 4.2 implies that \( x^*(t) \in X_{ad} \) must be a solution to the Euler–Lagrange equations (4.24) in the interval \( t \in [t_0, t^*_e] \) provided that \( x^*(t) \) is such that the admissible direction \( \xi(t) \) satisfies \( \xi(t_0) = \xi(t^*_e) = 0 \).

Taking into account the first order necessary optimality condition of Theorem 4.1 the application of (4.35) to (4.33) results in

\[
\frac{\partial}{\partial \eta} J(t^*_e + \eta \tau, x^* + \eta \xi) \\
= \frac{\partial}{\partial \eta} \varphi(t^*_e + \eta \tau, x^*(t^*_e + \eta \tau) + \eta \xi(t^*_e + \eta \tau)) \\
+ \frac{\partial}{\partial \eta} \int_{t_0}^{t^*_e + \eta \tau} l(t, x^*(t) + \eta \xi(t), \dot{x}^*(t) + \eta \dot{\xi}(t)) dt \\
= \left[ \frac{\partial}{\partial \tau} \varphi + \frac{\partial}{\partial x} \varphi \left[ \dot{x} + \eta \dot{\xi} \right] + \frac{\partial}{\partial x} \varphi \right]_{t = t^*_e + \eta \tau, x = x^* + \eta \xi} \\
+ \tau \left[l(t, x^*(t) + \eta \xi(t), \dot{x}^*(t) + \eta \dot{\xi}(t))\right]_{t = t^*_e + \eta \tau} + \int_{t_0}^{t^*_e + \eta \tau} \left[\frac{\partial}{\partial x} l(t, x^* + \eta \xi, \dot{x}^* + \eta \dot{\xi})\right]_{x = x^* + \eta \xi, \dot{x} = x^* + \eta \dot{\xi}} dt \\
= \left[ \frac{\partial}{\partial \tau} \varphi + \frac{\partial}{\partial x} \varphi \left[ \dot{x} + \eta \dot{\xi} \right] + \frac{\partial}{\partial x} \varphi \right]_{t = t^*_e + \tau, x = x^* + \eta \xi} \\
+ \tau \left[l(t, x^*(t) + \eta \xi(t), \dot{x}^*(t) + \eta \dot{\xi}(t))\right]_{t = t^*_e + \eta \tau} + \int_{t_0}^{t^*_e + \eta \tau} \left[\frac{\partial}{\partial x} l(t, x^* + \eta \xi, \dot{x}^* + \eta \dot{\xi})\right]_{x = x^* + \eta \xi, \dot{x} = x^* + \eta \dot{\xi}} dt.
\]

Evaluation of the limit as \( \eta \to 0 \) yields

\[
\delta J(t_e, x, \tau, \xi) = \left. \frac{\partial}{\partial \eta} J(t^*_e + \eta \tau, x^* + \eta \xi) \right|_{\eta = 0} \\
= \left[ \frac{\partial}{\partial \tau} \varphi + \frac{\partial}{\partial x} \varphi \left[ \dot{x} + \xi \right] + \tau l(t^*_e, x^*(t^*_e), \dot{x}^*(t^*_e)) + \frac{\partial}{\partial x} l(t^*_e, x^*(t^*_e), \dot{x}^*(t^*_e)) \right]_{t = t^*_e, x = x^*} \\
- \int_{t_0}^{t^*_e} \left[\frac{\partial}{\partial x} l(t, x^*(t), \dot{x}^*(t)) \right]_{x = x^*, \dot{x} = \dot{x}^*} dt.
\]
Since the initial value \(x(t_0) = x_0\) is fixed by assumption any admissible direction \(ξ(t)\) has to satisfy \(ξ(t_0) = 0\). Recalling also that the optimal solution \(x^*(t)\) has to fulfill the Euler–Lagrange equations (4.24) in the interval \(t \in [t_0, t_e]\), the Gâteaux derivative \(δJ(t_e, x, τ, ξ)\) reduces to

\[
δJ(t_e, x, τ, ξ) = \tau \left[ \frac{∂}{∂t} \varphi + \left( \frac{∂}{∂x} \varphi \right) \dot{x} + l \right]_{t=t_e, x=x^*} + \left[ \frac{∂}{∂x} \varphi + \frac{∂}{∂\dot{x}} l \right]_{t=t_e, x=x^*} ξ(t_e^*)
\]

\[
= \tau \left[ \frac{∂}{∂t} \varphi - \left( \frac{∂}{∂x} l \right) \dot{x} + l \right]_{t=t_e, x=x^*} + \left[ \frac{∂}{∂x} \varphi + \frac{∂}{∂\dot{x}} l \right]_{t=t_e, x=x^*} \left( τ \dot{x}^*(t_e^*) + ξ(t_e^*) \right)
\]

(4.36)

If both the end–time \(t_e\) and end–point \(x(t_e)\) are free, then \(τ\) and \(ξ(t_e^*)\) can be chosen independently. Thus, \(δJ(t_e, x, τ, ξ) = 0\) if the transversality conditions (4.34) are fulfilled.

This result can be further generalized according to the listing below.

(i) For fixed end–time \(t_e = t_e^*\) we have \(τ = 0\) such that the term \((*)\) vanishes in (4.36) and the transversality conditions reduce to (4.34a).

(a) If there is a component \(x_j(t), j \in \{1, \ldots, n\}\) of \(x(t)\) with fixed end–value \(x_j(t_e^*) = x_j^*(t_e^*) = x_{j,e}\), then \(ξ_j(t_e^*) = 0\) and the transversality condition (4.34a) vanishes for the \(j\)-th component.

(b) If there is a component \(x_j(t), j \in \{1, \ldots, n\}\) of \(x(t)\) with free end–value \(x_j(t_e^*)\), then \(ξ_j(t_e^*) \neq 0\) and the transversality condition (4.34a) for the \(j\)-th component is

\[
\left[ \frac{∂}{∂x_j} \varphi + \frac{∂}{∂\dot{x}_j} l \right]_{t=t_e, x=x^*} = 0.
\]

(ii) For free end–time \(t_e\) we have \(τ \neq 0\) and the term \((*)\) in (4.36) remains so that the transversality condition (4.34b) has to hold.

(a) If there is a component \(x_j(t), j \in \{1, \ldots, n\}\) of \(x(t)\) with fixed end–value \(x_j(t_e^*) = x_j^*(t_e^*) = x_{j,e}\), then an admissible direction \((τ, ξ_j(t))\) has to fulfill

\[
x_{j,e} = x_j^*(t_e^* + ητ) + η ξ_j(t_e^* + ητ) = \frac{∂}{∂η} x_{j,e} \bigg|_{η=0} = τ \dot{x}_j^*(t_e^*) + ξ_j(t_e^*) = 0.
\]

Hence, the corresponding entry in \((**\*)\) vanishes and there is no transversality condition (4.34a) for this component.

(b) If there is a component \(x_j(t), j \in \{1, \ldots, n\}\) of \(x(t)\) with free end–value \(x_j(t_e^*)\), then the transversality condition for this component is

\[
\left[ \frac{∂}{∂x_j} \varphi + \frac{∂}{∂\dot{x}_j} l \right]_{t=t_e, x=x^*} = 0.
\]
Remark 4.4: Natural boundary conditions

If $t_e$ is free and the terminal cost $\varphi(t_e, x(t_e)) = 0$, then the transversality conditions yield the so-called natural boundary conditions [20, 6]:

(i) If $t_e$ is free, then (4.34b) reduces to

\[
\left[ \frac{\partial}{\partial \dot{x}} l \right]_{t = t_e, x = x^*} = H(t_e, x^*(t_e), \dot{x}^*(t_e)) = 0.
\]

(ii) If $x(t_e)$ is free, then

\[
\left( \frac{\partial}{\partial \dot{x}} l \right)_{t = t_e, x = x^*} = 0.
\]

Example 4.4. Consider the functional

\[
J = p(x(t_e) - a)^2 + \int_0^{t_e} (\dot{x}(t))^2 \, dt
\]

for $x(t) \in X_{ad} = \{ x : x(t) \in C^1([0, t_e], \mathbb{R}), x(0) = x_0, t_e \text{ fixed} \}$. According to Theorem 4.4 a candidate $x(t)$ for a local minimizer has to solve the Euler–Lagrange equations (4.24), which for $l(\dot{x}(t)) = \dot{x}(t)^2$ read

\[
2\ddot{x}(t) = 0.
\]

This implies the solution

\[
x(t) = c_1 t + c_0
\]

with $c_0 = x_0$ to fulfill the initial condition $x(0) = x_0$. For the determination of $c_1$, the transversality conditions (4.34) have to be taken into account, which in view of case (i)(b) discussed above, $\varphi(x(t_e)) = p(x(t_e) - a)^2$ and $n = 1$ reduce to

\[
2p(x(t_e) - a) + 2\dot{x}(t_e) = 0 = 2p(c_1 t_e + x_0 - a) + 2c_1
\]

so that $c_1 = \frac{p(a - x_0)}{pt_e + 1}$. The unique solution to the Euler–Lagrange equations fulfilling the transversality conditions is given by

\[
x(t) = \frac{p(a - x_0)}{pt_e + 1} t + x_0.
\]

If $p \gg 1$ more weight is put on the terminal constraint than on the integral and $x(t_e)$ approaches $a$. In the limit $p \to \infty$ we obtain $x(t_e) = a$.

4.2.1.4 Piecewise continuous functions

The theory developed above can be extended and refined by including piecewise $C^1$–functions, i.e. searching for $x^*(t) \in \hat{C}([t_0, t_e], \mathbb{R}^n)$, as possible extrema. With this, one addresses the question whether these so–called cornered trajectories might yield improved results. Moreover, one might wonder if problems not admitting a solution in the class of $C^1$–functions have an extrema in the extended class of $\hat{C}^1$–functions. In the following, these results, in particular the Weierstrass–Erdmann (corner) conditions, will not be developed but the reader is referred to [6, 15, 7] and the references therein.
4.2.2 Problems with constraints

The Euler–Lagrange equations introduced in Theorem 4.2 provide a necessary optimality condition when the initial and terminal points are fixed but the curves are unconstrained otherwise. In the presence of constraints it is convenient to introduce Lagrange multipliers to derive the corresponding necessary optimality conditions.

4.2.2.1 Equality constraints

Making use of Theorem 4.1 it follows that given a functional $J$ defined on a subset $X_{\text{ad}}$ of a normed linear space $(X, \| \cdot \|_X)$ a local minimizer $x^* \in X_{\text{ad}}$ of $J$ can be characterized by

$$\delta J(x^*, \xi) = 0$$

for all $X_{\text{ad}}$–admissible directions $\xi$ at the point $x^*$. It should be noted that subsets $X_{\text{ad}}$ may exist such that the set of $X_{\text{ad}}$–admissible directions $\xi$ is empty, possibly at every point in $X_{\text{ad}}$. An example is given below [6].

**Example 4.5.** Consider the set

$$X_{\text{ad}} = \{ x(t) \in C^1([t_0, t_e], \mathbb{R}^2) : \sqrt{(x_1(t))^2 + (x_2(t))^2} - \sqrt{2} = 0, \forall t \in [t_0, t_e] \}.$$  

Obviously, $X_{\text{ad}}$ is the set of continuously differentiable curves lying on a cylinder of radius $\sqrt{2}$ whose axis is time $t$ centered at $x_1(t) = x_2(t) = 0$. Let $x^*(t) \in C^1([t_0, t_e], \mathbb{R}^2)$ with $x_1^*(t) = x_2^*(t) = 1$ so that $x^*(t) \in X_{\text{ad}}$. However, for every non–zero direction $\xi(t) \in C^1([t_0, t_e], \mathbb{R}^2)$ and for every $\eta \neq 0$ we have $x^*(t) + \eta \xi(t) \notin X_{\text{ad}}$. Thus, the set of $X_{\text{ad}}$–admissible directions is empty for any functional $J : C^1([t_0, t_e], \mathbb{R}^2) \rightarrow \mathbb{R}$.

The idea behind the introduction of Lagrange multipliers is to characterize the local minimizers or extremals of a functional $J$ defined in a normed linear space $(X, \| \cdot \|_X)$, when it is restricted to one of more level sets of other such functionals.

**Example 4.6.** Consider the previous example. Then $X_{\text{ad}}$ can be also considered as the intersection of the $0$–level sets of the family of functionals

$$G_t(x) = \sqrt{(x_1(t))^2 + (x_2(t))^2} - \sqrt{2}$$

for $t \in [t_0, t_e]$, i.e.,

$$X_{\text{ad}} = \bigcap_{t \in [t_0, t_e]} \Gamma_t(0)$$

for $\Gamma_{G}(s) = \{ x(t) \in C^1([t_0, t_e], \mathbb{R}^2) : G(x) = s \}$. This implies an uncountable number of functionals, which also illustrates why problems having path constraints are rather hard to solve in general.

The existence of a Lagrange multiplier is guaranteed by the following theorem, whose proof can be deduced from the exposition, e.g., in [24, 6].

**Theorem 4.5: Existence of Lagrange multipliers (single equality constraint)**

Let $J$ and $G$ be functionals defined in a neighborhood of $x^*$ in a normed linear space $(X, \| \cdot \|_X)$ having continuous Gâteaux derivatives in this neighborhood. Let $G(x^*) = s$ and suppose that
\( x^* \) is a (local) extremum for \( J \) constrained to \( \Gamma(s) = \{ x \in X : G(x) = s \} \). Suppose further that \( \delta G(x^*, \xi) \neq 0 \) for some direction \( \xi \in X \). Then there exists a scalar \( \lambda \in \mathbb{R} \) such that
\[
\delta J(x^*, \xi) + \lambda \delta G(x^*, \xi) = 0, \quad \forall \xi \in X. \tag{4.37}
\]

As in Chapter 3, the parameter \( \lambda \) is called Lagrange multiplier. Condition (4.37) implies that the directional derivatives of \( J \) are proportional to those of \( G \). In other words, the level sets of both \( J \) and \( G \) share a common tangent plane \( T_{x^*} \Gamma \) at \( x^* \), i.e., they meet tangential.

The extension to the case of multiple equality constraints is given below.

**Theorem 4.6: Existence of Lagrange multipliers (multiple equality constraints)**

Let \( J \) and \( G_j, j = 1, \ldots, p \) be functionals defined in a neighborhood of \( x^* \) in a normed linear space \( (X, \| \cdot \|_X) \) having continuous Gâteaux derivatives in this neighborhood. Let \( G_j(x^*) = s_j \) and suppose that \( x^* \) is a (local) extremum for \( J \) constrained to \( \Gamma(s) = \{ x \in X : G_j(x) = s_j, j = 1, \ldots, p \} \).

Suppose further that
\[
\begin{vmatrix}
\delta G_1(x^*, \xi_1) & \cdots & \delta G_1(x^*, \xi_p) \\
\vdots & \ddots & \vdots \\
\delta G_p(x^*, \xi_1) & \cdots & \delta G_p(x^*, \xi_p)
\end{vmatrix} \neq 0 \tag{4.38}
\]

for \( p \) independent directions \( \xi_j \in X, j = 1, \ldots, p \). Then there exists a vector \( \lambda \in \mathbb{R}^p \) such that
\[
\delta J(x^*, \xi) + [\delta G_1(x^*, \xi) \cdots \delta G_p(x^*, \xi)] \lambda = 0, \quad \forall \xi \in X. \tag{4.39}
\]

**Remark 4.5**

If \( x^* \in X_{\text{ad}} \) with \( X_{\text{ad}} \) a subset of a normed linear space \( (X, \| \cdot \|_X) \) and the \( X_{\text{ad}} \)-admissible directions form a linear subspace of \( X \), i.e., for all \( \eta_1, \eta_2 \in \mathbb{R} \) given \( \xi_1, \xi_2 \in X_{\text{ad}} \) we have \( \eta_1 \xi_1 + \eta_2 \xi_2 \in X_{\text{ad}} \), then the conclusions of Theorems 4.5 and 4.6 remain valid, when restricting the continuity of \( J \) to \( X_{\text{ad}} \) and considering \( X_{\text{ad}} \)-admissible directions only.

4.2.2.2 Inequality constraints

Similar to the previous section, Lagrange multipliers can be used to treat variational problems involving inequality constraints or mixed equality and inequality constraints.

**Theorem 4.7: Existence of Lagrange multipliers (multiple inequality constraints)**

Let \( J \) and \( G_j, j = 1, \ldots, p \) be functionals defined in a neighborhood of \( x^* \) in a normed linear space \( (X, \| \cdot \|_X) \) having continuous Gâteaux derivatives in this neighborhood. Suppose that \( x^* \) is a (local) minimizer for \( J \) constrained to \( \Gamma(s) = \{ x \in X : G_j(x) \leq s_j, j = 1, \ldots, p \} \).

Suppose further that \( q \leq p \) constraints are active, say \( G_j, j = 1, \ldots, q \) for simplicity, and satisfy
\[
\begin{vmatrix}
\delta G_1(x^*, \xi_1) & \cdots & \delta G_1(x^*, \xi_q) \\
\vdots & \ddots & \vdots \\
\delta G_q(x^*, \xi_1) & \cdots & \delta G_q(x^*, \xi_q)
\end{vmatrix} \neq 0 \tag{4.40}
\]
Recalling from (4.11) isoperimetric constraints defined according to 4.2.2.3 Isoperimetric constraints theorem below gives a characterization of the (local) minimizer using Lagrange multipliers [6].

\[ t \] involves constraints given as integrals of a function over parts or all of the horizon \( t \in [t_0, t_e] \). The theorem below gives a characterization of the (local) minimizer using Lagrange multipliers [6].

\[ \text{Theorem 4.8: First–order necessary optimality condition for problems with isoperimetric constraints} \]

Consider the functional

\[ J(x) = \int_{t_0}^{t_e} l(t, x(t), \dot{x}(t)) dt \]  

(4.42a)

on \( X_{ad} = \{x(t) \in C^1([t_0, t_e], \mathbb{R}^n) : x(t_0) = x_0, x(t_e) = x_e\} \) subject to the isoperimetric constraint

\[ G_k(x) = \int_{t_0}^{t_e} \psi_k(t, x(t), \dot{x}(t)) dt = a_k, \quad k = 1, \ldots, r < n \]  

(4.42b)

with Lagrangian density \( l : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) and \( \psi_k : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, k = 1, \ldots, r \) being continuously differentiable. Suppose that \( x^*(t) \in X_{ad} \) is a (local) minimizer for this problem and

\[ \det \begin{bmatrix} \delta G_1(x^*, \xi_1) & \cdots & \delta G_p(x^*, \xi_1) \\ \vdots & & \vdots \\ \delta G_1(x^*, \xi_r) & \cdots & \delta G_p(x^*, \xi_r) \end{bmatrix} \neq 0 \]  

(4.43)

for \( r \) independent directions \( \xi_j \in X, j = 1, \ldots, r \). Then there exists a vector \( \lambda^* \in \mathbb{R}^p \) such that \( x^*(t) \) is a solution to the Euler–Lagrange equations

\[ \frac{\partial}{\partial t} \left( \nabla_x L(t, x^*(t), \dot{x}^*(t), \lambda^*) - (\nabla_x L(t, x^*(t), \dot{x}^*(t), \lambda^*) \right) = 0, \]  

(4.44a)

where

\[ L(t, x(t), \dot{x}(t), \lambda) = l(t, x(t), \dot{x}(t)) + \lambda^T \psi(t, x(t), \dot{x}(t)) \]  

(4.44b)

with \( \psi = [\psi_1, \ldots, \psi_r]^T \).
It can be similarly shown that if \( L \) does not depend on \( t \), then the Hamilton function
\[
H(x, \dot{x}) = (\nabla_{\dot{x}} L)(x, \dot{x}, \lambda) \dot{x} - L(x, \dot{x}, \lambda)
\]
is constant along any (local) minimizer \( x^*(t) \). It is hence an invariant of the Euler–Lagrange equations introduced in the theorem above involving the Lagrange multiplier.

These results provide a fundamental tool to analyze optimal control problems by making use of variational calculus.

### 4.3 Unconstrained optimal control

Subsequently, unconstrained dynamic optimization problems in Bolza form are considered given by
\[
\begin{align*}
\min_u J(u) &= \varphi(t_e, x(t_e)) + \int_{t_0}^{t_e} l(t, x(t), u(t)) \, dt \\
\text{subject to} \quad &\dot{x} = f(t, x, u), \quad t > t_0, \quad x(t_0) = x_0 \in \mathbb{R}^n
\end{align*}
\]
with free or fixed endtime \( t_e \) and endpoint \( x_e \), respectively.

We will thereby primarily focus on the optimal open–loop control, where we seek \( u(t) = u^*(t) \) as a function of time for a specified initial state and in case final state.

#### 4.3.1 Existence of an optimal control

It can be shown that if \( g(t, x) = f(t, x(t), u(t)) \) is piecewise continuous in \( t \) and locally Lipschitz continuous in \( x \) according to Remark 2.3, then there exists a \( \delta > 0 \) such that the initial value problem
\[
\begin{align*}
\dot{x} &= g(t, x), \\
x(t_0) &= x_0
\end{align*}
\]
has a unique solution for \( t \in [t_0, t_0 + \delta] \) (see, e.g., [11, 23]). Global existence can be achieved by either restricting \( g(t, x) \) to the very restrictive set of global Lipschitz continuous functions or by imposing further information about the solution of the system. In particular, assume that \( g(t, x) \) is piecewise continuous in \( t \) and locally Lipschitz continuous in \( x \) for all \( x \in \mathbb{V} \subset \mathbb{R}^n \) and let \( W \) be a compact subset of \( \mathbb{V} \) so that \( x_0 \in W \). Suppose that the solution \( x(t) \) to (4.47) lies entirely in \( W \), then there is a unique solution for all \( t \geq t_0 \). Due to the assumption of piecewise continuity in \( t \) obviously input trajectories \( u(t) \in C^0([t_0, t_e], \mathbb{R}^m) \) are admissible so that \( x(t) \in C^1([t_0, t_e], \mathbb{R}^n) \) with the cornering points at the discontinuities of \( u(t) \).

Besides the situations where there does not exist a feasible control or feasible pair, respectively, according to Definition 4.1, the non–existence of a solution to an optimal control problem results from the failure of the set \( U_{fe} \) of feasible controls to be compact.

**Remark 4.6: Compact set**

A set \( V \) in a normed linear space \((X, \|\cdot\|)\) is said to be compact if every sequence in \( V \) contains a convergent subsequence with its limit point in \( V \). The set \( V \) is called relatively compact if its closure \( \overline{V} \) (add to \( V \) all limit points of sequences in \( V \)) is compact.
In view of the previous discussion, issues arise when the solution to (4.46b) becomes unbounded in the interval \( t \in [t_0, t_e] \), \( t_e < \infty \) so that the cost functional \( J(u) \) approaches infinity. This corresponds to a so-called finite escape time. Hence, one typically requires the solutions to (4.46b) to be bounded, i.e.,

\[
\| x(t; x_0, u(t)) \| \leq \alpha, \quad t \geq t_0
\]

for finite \( \alpha > 0 \). One particular class of systems not exhibiting finite escape behavior is given by systems that are affine in \( x \), i.e., \( \dot{x} = A(t, u)x + b(t, u) \), \( t > t_0 \) with \( x(t_0) = x_0 \). In addition, if the optimization interval \( [t_0, t_e] \) is unbounded, i.e., \( t_e = \infty \), then the set of feasible controls is itself unbounded and hence not compact. Hence, operations should be restricted to a compact finite optimization interval \( [t_0, T] \) with \( T \) sufficiently large so that the set of feasible controls \( U_{t_0} = U_{t_0}([t_0, t_e]) \neq \emptyset \) for some \( t_e \in (t_0, T] \).

Example 4.7. Consider a point mass \( m \) that is accelerated by a force \( u(t) \) with \( 0 \leq u(t) \leq 1 \) in the interval \( t \in [t_0, t_e] \), i.e.,

\[
\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u, \quad t > t_0, \quad x(t_0) = x_0 = \begin{bmatrix} x_0 \\ 0 \end{bmatrix}.
\]

It is desired to determine the input \( u(t) \) so that starting from \( x_1(t_0) = x_0 \) the mass reaches the point \( x_1(t_e) = x_e \) at time \( t = t_e \) while minimizing the cost functional

\[
J(u) = \int_{t_0}^{t_e} u^2(t) dt.
\]

Given \( x_{1,e} > x_{1,0} \) it follows immediately that \( u(t) \equiv 0 \) is infeasible so that \( J(u) > 0 \) for any feasible control \( u(t) \). For the sequence of constant admissible controls \( u_k(t) = 1/k \), \( k \geq 1 \) and \( t \geq t_0 \) the solution to the differential equation is obtained as

\[
x_1 = \frac{1}{2mk}(t - t_0)^2 + x_0
\]

\[
x_2 = \frac{1}{mk}(t - t_0).
\]

The value \( x_1(t_e) = x_e \) is hence reached at time

\[
t_{e,k} = t_0 + \sqrt{2mk(x_e - x_0)}.
\]

With this, the value of the cost functional in the interval \( t \in [t_0, t_{e,k}] \) follows as

\[
J(u_k) = \int_{t_0}^{t_{e,k}} \frac{1}{k^2} dt = \frac{2m}{k^3}(x_e - x_0).
\]

Thus, for \( k \to \infty \) we have \( J(u_k) \to 0 \) while \( t_{e,k} \to \infty \). This implies \( \inf_k J(u_k) = 0 \) so that the problem does not have a minimum.

Similarly examples can be constructed, where an optimal control does not exist although the time horizon is finite and the solution to the dynamic system remains bounded.

The discussion reveals that additional conditions have to be imposed on the class of admissible controls. On the one hand the input may be required to fulfill an additional Lipschitz condition, i.e.,

\[
\| u(t) - u(s) \| \leq L_u |t - s|, \quad L_u \in (0, \infty)
\]  

(4.48)
for all \( t, s \in [t_0, t_e] \). On the other hand the class of admissible controls may be restricted to the set of piecewise constant inputs with at most a finite number of points of discontinuity.

### 4.3.2 Application of variational calculus

In the following, necessary optimality conditions are derived for optimal control problems of the form (4.46) making use of variational calculus as introduced in the sections before.

#### 4.3.2.1 Fixed–time, free–endpoint problems

In the optimal control problem with fixed–time and free–endpoint we seek the input \( u(t) \in C([t_0, t_e], \mathbb{R}^m) \) minimizing the cost functional

\[
J(u) = \int_{t_0}^{t_e} l(t, x(t), u(t)) \, dt
\]

subject to the equality constraint given by the dynamic system

\[
\dot{x} = f(t, x, u), \quad t > t_0, \quad x(t_0) = x_0 \in \mathbb{R}^n
\]

for fixed endtime \( t_e \). Here it is obvious that a variation of the state trajectory \( x(t) \) in terms of \( x(t) = x^*(t) + \eta \xi(t) \) does not explicitly relate to a variation in the input \( u(t) \) due to the implicit coupling given by the differential equation (4.49b). Hence to deduce first order optimality conditions we proceed by introducing a one–parameter family of comparison trajectories in \( u(t) \), i.e., \( u(t) = u^*(t) + \eta \omega(t) \) with \( \omega(t) \in C([t_0, t_e], \mathbb{R}^m) \). These considerations will allow to prove the following result.

**Theorem 4.9: First order necessary condition (fixed–time, free–endpoint problem)**

Consider the minimization problem

\[
\min_u J(u) = \int_{t_0}^{t_e} l(t, x(t), u(t)) \, dt
\]

subject to

\[
\dot{x} = f(t, x, u), \quad t > t_0, \quad x(t_0) = x_0 \in \mathbb{R}^n
\]

for fixed terminal time \( t_e > t_0 \). Assume that \( l \) and \( f \) are continuous in \((t, x, u)\) and continuously differentiable with respect to \( x, u \) for all \((t, x, u) \in [t_0, t_e] \times \mathbb{R}^n \times \mathbb{R}^m \).

Suppose that \( u^*(t) \in C([t_0, t_e], \mathbb{R}^m) \) is a (local) minimizer of (4.50) with \( x^*(t) \in C^1([t_0, t_e], \mathbb{R}^n) \) denoting the corresponding solution of the initial–value problem (4.50b). Then there is a \( \lambda^*(t) \in C^1([t_0, t_e], \mathbb{R}^n) \) such that the triple \((u^*(t), x^*(t), \lambda^*(t))\) satisfies

\[
\dot{x}^* = f(t, x^*, u^*), \quad x^*(t_0) = x_0
\]

\[
\dot{\lambda}^* = -(\nabla_x l)(t, x^*, u^*) - (\nabla_x f)^T(t, x^*, u^*) \lambda^*(t), \quad \lambda^*(t_e) = 0
\]

\[
0 = (\nabla_u l)(t, x^*, u^*) + (\nabla_u f)^T(t, x^*, u^*) \lambda^*(t)
\]

for \( t \in [t_0, t_e] \). These equations are called Euler–Lagrange equations of the optimal control problem (4.50) and \( \lambda^*(t) \) is referred to as the adjoint state or co–state.

**Proof.** Consider the one–parameter family of comparison functions \( u(t) = u(t; \eta) = u^*(t) + \eta \omega(t) \) with \( \omega(t) \in C([t_0, t_e], \mathbb{R}^m) \) and the scalar parameter \( \eta \). Due to the assumption of continuity and
Taking into account the Taylor series, then
\[ \delta \]
the integral kernels are continuous so that
\[ \text{The initial condition } x(t; \eta) \text{ exists and is differentiable in } \eta \text{ for all } \eta \in \mathcal{D}_0(0) \text{ for all } t \in [t_0, t_e] \) (see the discussion in Section 4.3.1). In addition, \( \eta = 0 \) implies \( x(t; 0) = x^*(t) \) for \( t \in [t_0, t_e] \).

To proceed, recall from Section 4.2.2 that equality constraints can be included in the formalism by introducing a Lagrange multiplier \( \lambda(t) \). Hence, substitution of \( u(t; \eta) \) into the cost functional \( J(u) \) resulting from (4.50)

\[
J(u(t; \eta)) = \int_{t_0}^{t_e} \left\{ l(t, x(t; \eta), u(t; \eta)) + \lambda^T(t) \left[ f(t, x(t; \eta), u(t; \eta)) - \dot{x}(t; \eta) \right]\right\} dt
\]

\[ = \int_{t_0}^{t_e} \left\{ l(t, x(t; \eta), u(t; \eta)) + \lambda^T(t) f(t, x(t; \eta), u(t; \eta)) + \dot{\lambda}^T(t)x(t; \eta) \right\} dt
\]

\[ - \left[ \lambda^T(t)x(t; \eta) \right]_{t=t_0}^{t=t_e}
\]

for any \( \lambda(t) \in C^1([t_0, t_e], \mathbb{R}^n) \) and each \( \eta \in \mathcal{D}_0(0) \). The first order necessary optimality condition introduced in Theorem 4.1 hence imposes

\[
\delta J(u^*, \omega) = \frac{\partial}{\partial \eta} J(u(t; \eta)) \Bigg|_{\eta=0} = 0
\]

so that

\[
0 = \int_{t_0}^{t_e} \left\{ (\nabla_x l)^T(t, x^*, u^*) + \lambda^T(t)(\nabla_x f)(t, x^*, u^*) + \dot{\lambda}^T(t) \right\} \xi(t) dt
\]

\[ + \int_{t_0}^{t_e} \left\{ (\nabla_u l)^T(t, x^*, u^*) + \lambda^T(t)(\nabla_u f)(t, x^*, u^*) \right\} \omega(t) dt - \lambda^T(t_0) \xi(t_0) = \lambda^T(t_0) \xi(t_0)
\]

for any \( \omega(t) \in C([t_0, t_e], \mathbb{R}^m) \) and any \( \lambda(t) \in C^1([t_0, t_e], \mathbb{R}^n) \). Herein,

\[
\xi(t) = \left( \frac{\partial}{\partial \eta} x \right)(t; 0).
\]

The integral kernels are continuous so that \( \delta J(u^*, \omega) \) exists.

Since the effect of the variation on \( u^*(t) \) on the solution in terms of \( \xi(t) \) is difficult to determine, \( \lambda(t) = \lambda^*(t) \) is chosen so that

\[
\dot{\lambda}^*(t) = -(\nabla_x l)(t, x^*, u^*) - (\nabla_x f)^T(t, x^*, u^*) \lambda^*(t).
\]

Taking into account the Taylor series, then \( x(t; \eta) \) approximately satisfies

\[
x(t; \eta) \approx x(t; 0) + \left( \frac{\partial}{\partial \eta} x \right)(t; 0) \eta = x^*(t) + \xi(t) \eta.
\]

The initial condition \( x^*(t_0) = x_0 = x(t_0; 0) \) hence implies \( \xi(t_0) = 0 \). Recalling that \( x(t_e) \) is free we have

\[
\lambda^*(t_e) = 0
\]

in (4.52). The adjoint differential equation is linear in \( \lambda^*(t) \) and its solution exists and is unique in the interval \( t \in [t_0, t_e] \) by the continuity and differentiability assumption imposed in \( l \) and \( f \).
With this, (4.52) reduces to
\[0 = \int_{t_0}^{t_e} \left\{ (\nabla_u l)^T(t, x^*, u^*) + \lambda^T(t)(\nabla_u f)(t, x^*, u^*) \right\} \omega(t) dt\]
so that the fundamental lemma of variational calculus provided in Lemma 4.2 yields
\[(\nabla_u l)^T(t, x^*, u^*) + \lambda^T(t)(\nabla_u f)(t, x^*, u^*) = 0\]
for all \(t \in [t_0, t_e]\).

Some comments are in order [24, 6, 15, 7].

(i) The optimality conditions (4.51) consist of 2n differential equations (4.51a), (4.51b) in \(x^*(t), \lambda^*(t)\) and m algebraic equations (4.51c). Since the initial state at \(t = t_0\) in \(x^*(t)\) and the final state at \(t = t_e\) in \(\lambda^*(t)\) are provided, (4.51) corresponds to a so-called two–point boundary–value problem.

(ii) The Euler–Lagrange equations (4.51) can be re–written using the Hamiltonian function
\[
H(t, x, u, \lambda) = l(t, x, u) + \lambda^T f(t, x, u) 
\]
(4.53)
as
\[
\dot{x}^* = (\nabla H)(t, x^*, u^*, \lambda^*), \quad x^*(t_0) = x_0 \quad \\text{(4.54a)} \\
\dot{\lambda}^* = -(\nabla_x H)(t, x^*, u^*, \lambda^*), \quad \lambda^*(t_e) = 0 \quad \\text{(4.54b)} \\
0 = (\nabla_u H)(t, x^*, u^*, \lambda^*). \quad \\text{(4.54c)}
\]
for \(t \in [t_0, t_e]\). The last condition (4.54c) illustrates that for the triple \((u^*(t), x^*(t), \lambda^*(t))\) to be a local minimizer of \(l\) the input \(u^*(t)\) must necessarily be a stationary point of the Hamiltonian function for each \(t \in [t_0, t_e]\).

(iii) The variation of the Hamiltonian function along an optimal trajectory in view of (4.54) results in
\[
\frac{d}{dt} H = \frac{\partial}{\partial t} H + (\nabla H)^T \dot{x}^* + (\nabla u H)^T \dot{u}^* + f^T \dot{\lambda}^* \\
= \frac{\partial}{\partial t} H + (\nabla u H)^T \dot{u}^* + f^T \left[ (\nabla H) + \dot{\lambda}^* \right] = \frac{\partial}{\partial t} H.
\]
If neither \(f\) nor \(l\) depend explicitly on time, then the Hamiltonian function \(H\) is constant along an optimal trajectory and is an invariant of the two–point boundary–value problem (4.51). The Hamiltonian function is in this case also called a first integral of (4.51).

(iv) The Euler–Lagrange equations (4.51) or (4.54), respectively, are necessary for both a minimization and a maximization problem. For a (local) minimizer \(u^*(t)\) the Legendre condition introduced in Theorem 4.3 implies that necessarily
\[(\nabla^2_H)(t, x^*, u^*, \lambda^*) \succeq 0, \quad (4.55)\]
i.e., the Hessian matrix of the Hamiltonian function must be positive semi–definite. For a (local) maximizer \(u^*(t)\) it is obvious that \((\nabla^2_H)(t, x^*, u^*, \lambda^*)\) must be negative semi–definite.
(v) If the cost functional includes a term involving a terminal cost, i.e.,

\[ J(u) = \phi(t_e, x(t_e)) + \int_{t_0}^{t_e} l(t, x(t), u(t)) dt \]

it is an easy exercise to show that the optimal solution \((u^*(t), x^*(t), \lambda^*(t))\) must still satisfy (4.51) or (4.54), respectively, but with the terminal condition \(\lambda^*(t_e) = 0\) replaced by

\[ \lambda^*(t_e) = (\nabla_x \phi)(t_e, x^*(t_e)). \] (4.56)

(vi) In some situations it is convenient to express \(u^*(t)\) in terms of \(x^*(t)\) and \(\lambda^*(t)\) using (4.54c) and then substitute this expression into (4.54a), (4.54b) to obtain a two–point boundary–value problem in \(x^*(t)\) and \(\lambda^*(t)\) alone.

(vii) The adjoint state \(\lambda^*(t)\) can be interpreted in the sense that \(\lambda^*(t_0)\) corresponds to the sensitivity of the cost functional (4.50a) to changes in the initial condition \(x_0\).

Moreover, note the following remark referring to the case, where \(u^*(t)\) is sought for in the class of piecewise continuous functions.

**Remark 4.7**

Theorem 4.9 relies on the assumption that \(u^*(t)\) is continuous, i.e., \(u^*(t) \in C([t_0, t_e], \mathbb{R}^m)\). There are examples, where no solution of the Euler–Lagrange equations (4.51) can be found in this function class. Hence, one tries to seek for minimizers in the extended class of only piecewise continuous functions so that \(u^*(t) \in \dot{C}([t_0, t_e], \mathbb{R}^m)\). As is noted in Section 4.3.1 given \(u(t) \in \dot{C}([t_0, t_e], \mathbb{R}^m)\) the corresponding solutions to the differential equation (4.49b) are piecewise continuously differentiable, i.e., \(x(t) \in \dot{C}^1([t_0, t_e], \mathbb{R}^n)\) with the cornering points at the discontinuities of \(u(t)\). Referring by \(u^*(t) \in \dot{C}([t_0, t_e], \mathbb{R}^m)\) to the optimal input with \(x^*(t)\) and \(\lambda^*(t)\) the corresponding state and adjoint state of the optimization problem (4.49a), then at each cornering point \(c \in [t_0, t_e]\) the conditions

\[ x^*(c^-) = x^*(c^+) \] (4.57a)
\[ \lambda^*(c^-) = \lambda^*(c^+) \] (4.57b)
\[ H(c^-, x^*(c^-), u^*(c^-), \lambda^*(c^-)) = H(c^+, x^*(c^+), u^*(c^+), \lambda^*(c^+)) \] (4.57c)

have to hold, where \(c^-\) and \(c^+\) denote the left and right limit point.
4.3.2.2 Free–time, fixed–endpoint / terminal constraints problems

Differing from the previous section we subsequently seek the input $u(t) \in C([t_0, t_e], \mathbb{R}^m)$ solving an optimal control problem with free terminal time and fixed endpoint.

**Theorem 4.10: First order necessary optimality condition (free–time, terminal constraints problem)**

Consider the minimization problem

$$\min_{u} J(u) = \varphi(t_e, x(t_e)) + \int_{t_0}^{t_e} l(t, x(t), u(t))dt$$

subject to

$$\dot{x} = f(t, x, u), \quad t > t_0, \quad x(t_0) = x_0 \in \mathbb{R}^n$$
$$G_k(t_e, u) = \psi_k(t_e, x(t_e)) = 0, \quad k = 1, \ldots, p$$

for fixed initial time $t_0$ and free terminal time $t_e \ll T$. Assume that $l$ and $f$ are continuous in $(t, x, u)$ and continuously differentiable with respect to $x, u$ for all $(t, x, u) \in [t_0, T] \times \mathbb{R}^n \times \mathbb{R}^m$. In addition, assume that $\varphi$ and $\psi_k, k = 1, \ldots, p$ are continuously differentiable with respect to $t_e$.

Suppose that $(u^*(t), t^*_e) \in C([t_0, t_e], \mathbb{R}^m) \times [t_0, T]$ is a (local) minimizer of (4.58) with $x^*(t) \in C^1([t_0, T], \mathbb{R}^n)$ denoting the corresponding solution of the initial–value problem (4.58b). Suppose further that the **regularity condition**

$$\det \begin{bmatrix} \delta G_1(t^*_e, u^*(t), \bar{\tau}_1, \bar{\omega}_1(t)) & \cdots & \delta G_1(t^*_e, u^*(t), \bar{\tau}_p, \bar{\omega}_p(t)) \\ \vdots & \ddots & \vdots \\ \delta G_p(t^*_e, u^*(t), \bar{\tau}_1, \bar{\omega}_1(t)) & \cdots & \delta G_p(t^*_e, u^*(t), \bar{\tau}_p, \bar{\omega}_p(t)) \end{bmatrix} \neq 0$$

4.3 Unconstrained optimal control
holds for \( p \) independent directions \((\omega_k(t), \bar{r}_k) \in C([t_0, t_e], \mathbb{R}^m) \times \{t_0, T\}, k = 1, \ldots, p\). Then there is a \( \lambda^*(t) \in C^1([t_0, t_e], \mathbb{R}^m) \) and a \( \mu^* \in \mathbb{R}^p \) such that the tuple \((u^*(t), x^*(t), t_e^*, \lambda^*(t), \mu^*)\) satisfies the Euler–Lagrange equations

\[
\dot{x}^* = (\nabla_x H)(t, x^*, u^*, \lambda^*), \quad x^*(t_0) = x_0 \quad \text{(4.60a)}
\]

\[
\dot{\lambda}^* = -(\nabla_{x^*} H)(t, x^*, u^*, \lambda^*), \quad \lambda^*(t_e^*) = (\nabla_{x^*} \psi)(t_e^*, x^*(t_e^*), \mu^*) \quad \text{(4.60b)}
\]

\[
0 = (\nabla_u H)(t, x^*, u^*, \lambda^*) \quad \text{(4.60c)}
\]

for \( t \in [t_0, T] \) and the transversality conditions

\[
\psi(t_e^*, x^*(t_e^*)) = [\psi_1(t_e^*, x^*(t_e^*)) \cdots \psi_p(t_e^*, x^*(t_e^*))]^T = 0 \quad \text{(4.61a)}
\]

\[
\left(\frac{\partial}{\partial t_e}\nabla \psi\right)(t_e^*, x^*(t_e^*), \mu^*) + H(t_e^*, x^*(t_e^*), u^*(t_e^*), \lambda^*(t_e^*)) = 0. \quad \text{(4.61b)}
\]

with the Hamiltonian function

\[
H(t, x, u, \lambda) = l(t, x, u) + \lambda^T(t) f(t, x, u) \quad \text{(4.62)}
\]

and

\[
\phi(t_e, x_e, \mu) = \psi(t_e, x_e) + \mu^T \psi(t_e, x_e). \quad \text{(4.63)}
\]

The proof of this result in principle combines the procedure used to verify Theorem 4.4 and Theorem 4.9 above.

**Proof.** Consider the one–parameter family of comparison functions \( u(t) = u(t; \eta) = u^*(t) + \eta \omega(t) \) with \( \omega(t) \in C([t_0, T], \mathbb{R}^m) \) and the scalar parameter \( \eta \). Due to the assumption of continuity and differentiability of \( f \) there exists a \( \eta > 0 \) such that the solution \( x(t; \eta) \) of (4.58b) associated with \( u(t; \eta) \) exists, is unique and is differentiable in \( \eta \) for all \( \eta \in \mathcal{B}_0(0) \) for all \( t \in [t_0, T] \) (see the discussion in Section 4.3.1). In addition, \( \eta = 0 \) implies \( u(t; 0) = u^*(t) \) and \( x(t; 0) = x^*(t) \) for \( t \in [t_0, t_e^*] \).

The cost functional (4.58a) extended by the equality constraints (4.58b), (4.58c) and the terminal time \( t_e = t_e^* + \eta \tau \) reads

\[
j(t_e, u(t; \eta)) = \phi(t_e, x(t_e; \eta)) + \mu^T \psi(t_e, x(t_e; \eta)) \\
\quad + \int_{t_0}^{t_e} \left[ l(t, x(t; \eta), u(t; \eta)) + \lambda^T(t) \left[ f(t, x(t; \eta), u(t; \eta)) - \dot{x}(t; \eta) \right] \right] dt \\
= \phi(t_e, x(t_e; \eta)) + \mu^T \psi(t_e, x(t_e; \eta)) - \left[ \lambda^T(t) x(t; \eta) \right]_{t=t_0}^{t=t_e} \\
\quad + \int_{t_0}^{t_e} \left[ l(t, x(t; \eta), u(t; \eta)) + \lambda^T(t) f(t, x(t; \eta), u(t; \eta)) + \dot{\lambda}^T(t) x(t; \eta) \right] dt
\]

The first order necessary optimality condition of Theorem 4.1 implies for the local minimizer that

\[
\delta j(t_e^*, u^*, \tau, \omega) = \frac{\partial}{\partial \eta} j(t_e, u(t; \eta)) \bigg|_{\eta=0} = 0.
\]

The Gâteaux derivative at \((u^*(t), t_e^*)\) in any direction \((\omega(t), \tau) \in C([t_0, T], \mathbb{R}^m) \times [0, T] \) evaluates\(^1\) to

\[
0 = \delta j(t_e^*, u^*, \tau, \omega)
\]

\(^1\)Remember that the upper integration limit depends on \( \eta \) since \( t_e = t_e^* + \eta \tau \).
The transversality condition (4.61b) follows from lines 1 and 2 of (4.64) and the remaining equations. Some comments to Theorem 4.10 are in order.

\[\left\{-\lambda^T(t_\tau^*) \dot{x}(t_\tau^*;0) - \lambda^T(t_\tau^*) \dot{x}(t_\tau^*;0) + \left[\frac{\partial}{\partial t_e} \psi\right](t_\tau^*, x(t_\tau^*)) \right\} \tau\]

\[+ \mu^T \left[ \left(\frac{\partial}{\partial t_e} \psi\right)(t_\tau^*, x(t_\tau^*)) + \left(\frac{\partial}{\partial t_e} \psi\right)(t_\tau^*, x(t_\tau^*)) \dot{x}(t_\tau^*) \right] + l(t_\tau^*, x(t_\tau^*), u(t_\tau^*)) \]

\[+ \lambda^T(t_\tau^*) f(t_\tau^*, x(t_\tau^*), u(t_\tau^*)) + \lambda^T(t_\tau^*) \dot{x}(t_\tau^*) \right\} \tau\]

\[+ \{\nabla_x \psi\}^T(t_\tau^*, x(t_\tau^*)) + \mu^T (\nabla_x \psi)(t_\tau^*, x(t_\tau^*)) - \lambda^T(t_\tau^*) \} \xi(t_\rho) + \{\lambda^T(t_\rho)\} \xi(t_\rho)\]

\[= \int_{t_\rho}^{t_\tau} \{\nabla_x \psi\}(t, x(t;0), u(t;0)) + \lambda^T(t) (\nabla_x f)(t, x(t;0), u(t;0)) + \lambda^T(t) \right\} \xi(t) dt\]

with \(\xi(t) = \left(\frac{\partial}{\partial \eta} x\right)(t_\rho;0)\). Since the initial value is fixed we have, as in the proof of Theorem 4.9, that \(x(t_\rho;\eta) = x_0\) and \(\xi(t_\rho) = 0\). Noting that \(x(t_\rho;0) = x^*(t_\rho)\) and \(u(t_\rho;0) = u^*(t_\rho)\) the expression reduces to

\[0 = \left\{\frac{\partial}{\partial t_e} \psi\right\}(t_\rho^*, x^*(t_\rho^*)) + \mu^T \left[ \left(\frac{\partial}{\partial t_e} \psi\right)(t_\rho^*, x^*(t_\rho^*)) + l(t_\rho^*, x^*(t_\rho^*), u^*(t_\rho^*)) \right] \]

\[+ \lambda^T(t_\rho^*) f(t_\rho^*, x^*(t_\rho^*), u^*(t_\rho^*)) \}

\[+ \{\nabla_x \psi\}^T(t_\rho^*, x^*(t_\rho^*)) + \mu^T (\nabla_x \psi)(t_\rho^*, x^*(t_\rho^*)) - \lambda^T(t_\rho^*) \} \xi(t_\rho) + \{\lambda^T(t_\rho)\} \xi(t_\rho)\]

\[= \int_{t_\rho}^{t_\tau} \{\nabla_x \psi\}(t, x^*(t;0), u^*(t;0)) + \lambda^T(t) (\nabla_x f)(t, x^*(t;0), u^*(t;0)) + \lambda^T(t) \right\} \xi(t) dt\]

Since the effect on \(u^*(t)\) on the solution in terms of \(\xi(t)\) is hard to determine the adjoint state \(\lambda^*(t)\) is chosen so that the first integral arising in (4.64) vanishes, i.e.,

\[\dot{\lambda}^*(t) = -\{\nabla_x \psi\}^T(t, x^*(t), u^*) - (\nabla_x f)^T(t, x^*, u^*) \lambda^*(t)\]

with the terminal condition obtained from setting the argument in front of \(\{\lambda^T(t_\rho)\} \xi(t_\rho)\) to zero, i.e.,

\[\lambda^*(t_\rho) = (\nabla_x \psi)(t_\rho^*, x^*(t_\rho^*)) + (\nabla_x \psi)^T(t_\rho^*, x^*(t_\rho^*)) \mu^*\]

The transversality condition (4.61b) follows from lines 1 and 2 of (4.64) and the remaining equations are obtained by making use of the fundamental lemma of variational calculus.

Some comments to Theorem 4.10 are in order.

(i) The optimality conditions (4.60) consist of \(2n\) differential equations (4.60a), (4.60b) in \(x^*(t)\), \(\lambda^*(t)\) and \(m\) algebraic equations (4.60c). Since the initial state at \(t = t_\rho\) in \(x^*(t)\) and the final state at \(t = t_\rho^*\) in \(\lambda^*(t)\) are provided we have to deal with a two-point boundary-value problem. The Lagrange multiplier \(\mu\) and — in case — the terminal time \(t_\rho^*\) follow from the transversality conditions (4.61) so that a complete set of equations for the \(2n + m + p + 1\) unknowns is available.

(ii) For fixed end-time \(t_\rho = t_\rho^*\) we have \(\tau = 0\) such that the term \(\{\cdot\} \tau\) vanishes in (4.64) and the transversality conditions reduce to (4.61a).
(a) If there is a component \(x_j(t), j \in \{1, \ldots, n\}\) of \(x(t)\) with fixed end-value \(x_j(t_e^*) = x_{j,e} = x_j^*(t_e^*) = 0\) and there is no terminal condition for the corresponding adjoint state \(\lambda_j^*(t_e^*)\).

(b) If there is a component \(x_j(t), j \in \{1, \ldots, n\}\) of \(x(t)\) with free end-value \(x_j(t_e^*)\), then \(\xi_j(t_e^*) \neq 0\). In this case the terminal condition for the corresponding adjoint state \(\lambda_j^*(t_e^*)\) is determined from (4.65).

(iii) For free end-time \(t_e\) we have \(\tau \neq 0\) the transversality condition (4.61b) has to hold.

(a) If there is a component \(x_j(t), j \in \{1, \ldots, n\}\) of \(x(t)\) with fixed end-value \(x_j(t_e^*) = x_{j,e} = x_j^*(t_e^*)\), then an admissible direction \(\tau, \xi_j(t)\) has to fulfill

\[ x_j(t_e = x_j^*(t_e + \eta \tau) + \eta \xi_j(t_e + \eta \tau) \Rightarrow \frac{\partial}{\partial \eta} x_j \bigg|_{\eta=0} = \tau \dot{x}_j(t_e^*) + \xi_j(t_e^*) = 0. \]

Hence, there is no terminal condition for the corresponding adjoint state \(\lambda_j^*(t_e^*)\).

(b) If there is a component \(x_j(t), j \in \{1, \ldots, n\}\) of \(x(t)\) with free end-value \(x_j(t_e^*)\), then the terminal condition is determined from (4.65).

(iv) If the equality constraints in Theorem 4.10 are replaced by inequality constraints

\[ G_k(t_e, u) = \psi_k(t_e, x(t_e)) \leq 0 \quad k = 1, \ldots, p, \]

then only (4.61a) has to be replaced by

\[ \psi_k(t_e, x(t_e)) \leq 0 \quad k = 1, \ldots, p \] (4.67a)

\[ \mu^* \geq 0 \] (4.67b)

\[ \psi^T(t_e, x(t_e)) \mu^* = 0, \] (4.67c)

which needs to be interpreted in the sense of the complementary slackness condition.

Remark 4.8: Reachability condition

Theorem 4.10 relies on the verification of the regularity condition (4.59), which is difficult in general. As is elaborated, e.g., in [6], this condition can be interpreted as a reachability condition. Hence, if it does not hold, then it is may not be possible to find a control \(u^*(t)\) so that the terminal conditions are fulfilled at the terminal time.
Example 4.8. Consider a particle of mass \( m \) moving in the \((x, y)\)-plane subject to a thrust force of magnitude \( ma(t) \) with the thrust acceleration \( a(t) \) a known function of time \([5, 4]\). The goal is to steer the point mass in minimal time \( t \in [0, t_e] \) from the initial position \((x_0, y_0)\) to a prescribed final position \((x_e, y_e)\). Let \( u(t) \) denote the input corresponding to the angle between the direction of thrust and the \( x \)-axis. The optimization problem hence reads

\[
\min_{u(t)} t_e \tag{4.68}
\]

subject to

\[
\begin{bmatrix}
\dot{x} \\
\dot{v} \\
\dot{y} \\
\dot{w}
\end{bmatrix} = \begin{bmatrix}
v \\
a \cos(u) \\
w \\
a \sin(u)
\end{bmatrix}, \quad t > 0,
\]

\[
\begin{bmatrix}
x(0) \\
v(0) \\
y(0) \\
w(0)
\end{bmatrix} = \begin{bmatrix}
x_0 \\
v_0 \\
y_0 \\
w_0
\end{bmatrix}
\]

with the terminal constraint

\[
\psi_1(t_e, x(t_e)) = x(t_e) - x_e = 0
\]

\[
\psi_2(t_e, y(t_e)) = y(t_e) - y_e = 0.
\]

Note that no conditions are imposed on the terminal velocities. Let \( x(t) = [x(t), v(t), y(t), w(t)]^T \), then the Hamiltonian function \( H \) and the function \( \phi \) follow as

\[
H(x, u, \lambda) = \lambda_1 v + \lambda_2 a \cos(u) + \lambda_3 w + \lambda_4 a \sin(u)
\]

\[
\phi(t_e, x(t_e), \mu) = t_e + \mu_1 (x(t_e) - x_e) + \mu_2 (y(t_e) - y_e).
\]

With these preparations the Euler–Lagrange equations (4.60) for a minimizer candidate \((t_e, u(t))\) are obtained as (4.69) and

\[
\dot{\lambda} = -(\nabla_x H)(x, u, \lambda) = \begin{bmatrix}
0 \\
\lambda_1 \\
0 \\
\lambda_3
\end{bmatrix}, \quad \lambda(t_e) = \begin{bmatrix}
\mu_1 \\
0 \\
\mu_2 \\
0
\end{bmatrix}
\]

\[
0 = (\nabla_u H)(x, u, \lambda) = -\lambda_2 a \sin(u) + \lambda_4 a \cos(u).
\]

The differential equations for the adjoint state \( \lambda(t) \) directly admit a solution given by

\[
\lambda(t) = \begin{bmatrix}
\mu_1 \\
(t_e - t) \mu_1 \\
\mu_2 \\
(t_e - t) \mu_2
\end{bmatrix}
\]

The homogeneous equation moreover implies

\[
\tan(u) = \frac{\lambda_4 a}{\lambda_2 a} = \frac{\mu_2}{\mu_1}
\]

so that

\[
u(t) = u = \arctan\left(\frac{\mu_2}{\mu_1}\right), \quad |u| < \frac{\pi}{2}.
\]
(a) Initial velocities \((v_0, w_0) = (0, 0)\).

(b) Initial velocities \((v_0, w_0) = (0, 1)\).

(c) Initial velocities \((v_0, w_0) = (-1, 0)\).

(d) Initial velocities \((v_0, w_0) = (-1, 1)\).

**Figure 4.3:** Time optimal control of the point mass for different initial velocities \(v_0\) and \(w_0\) with \(a = 1\).

*Obviously, the candidate solution is constant on the whole interval \([0, t_e]\). For constant thrust acceleration \(a(t) = a\) the corresponding states are obtained analytically by solving (4.69), which results in*

\[
\begin{align*}
x(t) &= x_0 + t v_0 + \frac{t^2}{2} a \cos(u) \\
v(t) &= v_0 + t a \cos(u) \\
y(t) &= y_0 + t w_0 + \frac{t^2}{2} a \sin(u) \\
w(t) &= w_0 + t a \sin(u).
\end{align*}
\]

*With the previous results, we have herein*

\[
\begin{align*}
\cos(u) &= \frac{1}{\sqrt{1 + \left(\frac{\mu_2}{\mu_1}\right)^2}}, & \sin(u) &= \frac{\mu_2}{\mu_1} \frac{1}{\sqrt{1 + \left(\frac{\mu_2}{\mu_1}\right)^2}}.
\end{align*}
\]
Since the terminal time is free the transversality condition (4.61b) has to be fulfilled, i.e.,

\[
0 = \left( \frac{\partial}{\partial t_e} \phi \right) (t_e, x(t_e), \mu) + H(x(t_e), u, \lambda(t_e)) = 1 + \mu_1 v(t_e) + \mu_2 w(t_e).
\] (4.71)

The remaining three unknowns \( t_e, \mu_1 \) and \( \mu_2 \) can be thus determined as the solution of the nonlinear system of algebraic equations

\[
\begin{bmatrix}
    x_0 + t_e v_0 + \frac{t_e^2}{2} a - \frac{1}{\sqrt{1 + \left( \frac{a}{\mu_1} \right)^2}} - x_e \\
    y_0 + t_e w_0 + \frac{1}{\sqrt{1 + \left( \frac{a}{\mu_1} \right)^2}} - y_e \\
    1 + \mu_1 \left( v_0 + t_e a - \frac{1}{\sqrt{1 + \left( \frac{a}{\mu_1} \right)^2}} \right) + \mu_2 \left( w_0 + t_e a \right) - \frac{1}{\sqrt{1 + \left( \frac{a}{\mu_1} \right)^2}}
\end{bmatrix} = 0
\] (4.72)

comprised of the two terminal constraints (4.70) and the transversality condition (4.71). Numerical techniques for the determination of the zero of (4.72) have already been introduced in Section 2.2.

For the numerical results depicted in Figure 4.3 the function \texttt{fsolve} of MATLAB is used.

Example 4.9 (Linear–quadratic regulation). Consider the linear time–varying MIMO system

\[
\dot{x} = A(t) x + B(t) u(t), \quad t > 0, \quad x(t_0) = x_0
\]

with \( x(t) \in \mathbb{R}^n \) and \( u(t) \in \mathbb{R}^m \). Determine the input \( u^*(t) \) to minimize the quadratic cost functional

\[
J(u) = \frac{1}{2} x^T(t_e) S_e x(t_e) + \frac{1}{2} \int_{t_0}^{t_e} \left[ u^T(t) R(t) u(t) + x^T(t) Q(t) x(t) \right] dt
\]

for given endtime \( t_e \). Herein, \( S_e \) and \( Q(t) \) are assumed positive semi–definite and \( R(t) \) is assumed positive definite for all \( t \in [t_0, t_e] \). With the Hamiltonian function

\[
H(t, x, u, \lambda) = \frac{1}{2} \left[ u^T(t) R(t) u(t) + x^T(t) Q(t) x(t) \right] + \lambda^T(t) \left[ A(t) x + B(t) u(t) \right]
\]

and

\[
\phi(t_e, x(t_e)) = \frac{1}{2} x^T(t_e) S_e x(t_e)
\]

the Euler–LaGrange equations (4.60) read

\[
\begin{align*}
\dot{x} &= (V A) H(t, x, u, \lambda) = A(t) x + B(t) u, \quad x(t_0) = x_0 \\
\dot{\lambda} &= -(V x) H(t, x, u, \lambda) = -Q(t) x - A^T(t) \lambda, \quad \lambda(t_e) = S_e x(t_e) \\
0 &= R(t) u + B^T(t) \lambda.
\end{align*}
\]

The candidate for the optimal control follows from the last equation, i.e.,

\[
u = -R^{-1}(t) B^T(t) \lambda,
\] (4.73)

and relies on the computation of the adjoint state \( \lambda(t) \). Substitution of \( u(t) \) into the Euler–LaGrange equations yields the two–point boundary–value problem

\[
\begin{bmatrix}
    \dot{x} \\
    \dot{\lambda}
\end{bmatrix}
= \begin{bmatrix}
    A(t) & -B(t) R^{-1}(t) B^T(t) \\
    -Q(t) & -A^T(t)
\end{bmatrix}
\begin{bmatrix}
    x \\
    \lambda
\end{bmatrix}, \quad \begin{bmatrix}
    x(t_0) \\
    \lambda(t_e)
\end{bmatrix} = \begin{bmatrix}
    x_0 \\
    S_e x(t_e)
\end{bmatrix}.
\] (4.74)
To solve this problem, we use a sweep method by assuming that \( \lambda(t) = P(t)x(t) \) so that \( \lambda(t_0) = P(t_0)x_0 \), i.e., the terminal condition \( \lambda(t_e) = S_e x(t_e) \) is swept back in time. Substitution of this relation provides
\[
P(t) = -P(t)A(t) - A^T(t)P(t) + P(t)B(t)R^{-1}(t)B^T(t)P(t) - Q(t), \quad P(t_e) = S_e. \tag{4.75}
\]

Hence, \( P(t) \) has to solve this matrix Riccati differential equation. Obviously, the differential equation is nonlinear in \( P(t) \) and has to be solved backward in time with the terminal condition \( P(t_e) = S_e \). Once \( P(t) \) is computed, the optimal trajectories \( x^*(t) \) and \( \lambda^*(t) \) can be determined forward in time by solving (4.74) with the initial value \( \lambda(t_0) = P(t_0)x_0 \). The optimal control hence follows from (4.73) in the form
\[
u^*(t) = -R^{-1}(t)B^T(t)P(t)x^*(t). \tag{4.76}
\]

Moreover note that since
\[
(V^2_H(t, x^*, u^*, \lambda^*)) = R(t)
\]
with \( R(t) \) positive definite by assumption the determined optimal control \( u^*(t) \) is a minimizer.

The corresponding optimal feedback controller minimizing the quadratic cost functional is similarly obtained as
\[
u = -K(t)x, \quad K(t) = R^{-1}(t)B^T(t)P(t) \tag{4.77}
\]
with \( P(t) \) solving (4.75) and \( x(t) \) the system state at time \( t \).

For the special case of an infinite optimization horizon with \( t_e \to \infty \), the terminal cost is meaningless since \( \lim_{t \to \infty} x(t) = 0 \) needs to hold to ensure the existence of the integral term in \( J(u) \). Hence, we have to impose \( S_e = 0 \). For the sake of simplicity consider the time–invariant setting with \( A(t) = A, B(t) = B, Q(t) = Q, \) and \( R(t) = R \) given \( Q \) positive semi–definite and \( R \) positive definite as before. Let the linear time–invariant system be stabilizable (there exists a matrix \( K \) such that the eigenvalues of \( A - BK \) have negative real part) and let the pair \( (A, C) \) be detectable, where \( C \) originates from the Cholesky decomposition of \( Q \), i.e., \( Q = C^T C \). Then the matrix Riccati differential equation (4.75) reduces to the algebraic Riccati equation
\[
PA + A^T P - PBR^{-1}B^T P + Q = 0. \tag{4.78}
\]

In view of the assumptions above it can be shown that there exists a unique positive semi–definite solution \( P \) to (4.78). In addition, it can be shown that the static feedback law
\[
u = -Kx, \quad K = R^{-1}B^T P \tag{4.79}
\]
with \( P \) the unique positive semi–definite solution to (4.78) under the imposed assumptions implies asymptotic stability of the closed–loop control, i.e., the eigenvalues of the matrix \( A - BR^{-1}B^T P \) have strictly negative real part.

For further details the reader is, e.g., referred to [14, 18].
**Exercise 4.17.** Determine the solution of the optimal control problem

$$\min_u J(u) = \int_0^1 \left( \frac{1}{2} u^2(t) + \frac{a}{2} x^2(t) \right) dt$$

subject to

$$\dot{x} = u, \quad t > 0, \quad x(0) = 1$$

and $x(1) = 0$.

**Solution 4.17.** The solution is given by

$$x^*(t) = -\frac{\sinh(\sqrt{a}(t-1))}{\sinh(\sqrt{a})}, \quad u^*(t) = -\frac{\sqrt{a}\cosh(\sqrt{a}(t-1))}{\sinh(\sqrt{a})}, \quad \lambda^*(t) = -u^*(t).$$

A graphical illustration when varying the parameter $a$ is given below.

---

### 4.4 Input constrained optimal control

In the previous section it was assumed that no constraints are imposed on the optimization problem in addition to a possibly fixed endtime and/or endpoint. This assumption will be weakened by introducing *Pontryagin's maximum principle*.

#### 4.4.1 Pontryagin's maximum principle

Subsequently, we consider in a first step dynamic optimization problems comprised of the cost functional

$$J(u) = \int_{t_0}^{t_e} l(x(t), u(t)) dt$$

with free endtime $t_e$ and fixed terminal value $x(t_e) = x_e$ subject to the *time invariant or autonomous*\(^2\) dynamic system

$$\dot{x} = f(x, u), \quad t > t_0, \quad x(t_0) = x_0.$$

---

\(^2\)Both $f$ and $l$ do not explicitly depend on time $t$. 

128  Chapter 4  Dynamic optimization
Differing from the previous section, \textit{admissible controls} are taken in the class of piecewise continuous functions
\[
u \in \mathcal{U} = \{ u \in \mathcal{C}^0_0([t_0, T], \mathbb{R}^m) : u(t) \in U, t_0 \leq t \leq t_e \}.
\]
for \( T \gg t_e \) sufficiently large and \( U \) the non-empty set of input constraints.

Some ideas underlying Pontryagin’s maximum principle\(^3\) can be motivated by properly reformulating the optimization problem in an extended state space. For this, consider
\[
x_{n+1}(t) = \int_{t_0}^{t} l(x(s), u(s)) \, ds.
\]
Hence, the optimization problem can be considered as finding an admissible control \( u(t) \in \mathcal{U} \) and an
\[
\begin{align*}
x_0 & \quad \text{at the point } (x_1, x_2) = (x_0, 0) \\
x_e & \quad \text{at the point } (x_1, x_2) = (x_e, 0)
\end{align*}
\]
Figure 4.4: Geometric interpretation of the reformulated optimization problem.

endtime \( t_e \) so that the solution of the extended system
\[
\dot{\bar{x}} = \begin{bmatrix} \dot{x} \\ \dot{x}_{n+1} \end{bmatrix} = \begin{bmatrix} f(x, u) \\ l(x, u) \end{bmatrix}, \quad t > t_0, \quad \bar{x}(t_0) = \begin{bmatrix} x_0 \\ 0 \end{bmatrix}
\]
(4.80)

terminates at the point \( \bar{x}^T(t_e) = [x_1^T(x_e), x_{n+1}(t_e)] \) with \( x_{n+1}(t_e) \) taking the smallest possible value. To illustrate this fact consider Figure 4.4. Along the red line passing through the point \( (x^T_e, 0) \) one finds all terminal points of solution trajectories of the extended system for \( t = t_e \) with different values of the cost functional \( x_{n+1}(t_e) \). No other trajectory can intersect with this line at a point below \( (x^T_e, x^*_{n+1}(t_e)) \).

These geometric observations form the basis for the derivation of the Pontryagin maximum principle. Due to the necessary technical efforts, in the following only results are stated without providing any proof. For this, the reader is referred to the available literature [19, 10].

\(^3\)As the name refers to the maximum principle was originally defined for maximization problems [19]. In view of the consideration of minimization problems one should hence refer to it as minimum principle. However, the name maximum principle is so commonly used in dynamic optimization and optimal control that we will not make this distinction.
Theorem 4.11: Pontryagin maximum principle for autonomous systems

Consider the optimal control problem

$$\min_{u \in \mathcal{U}} f(u) = \int_{t_0}^{t_e} l(x(t), u(t)) dt$$

subject to

$$\dot{x} = f(x, u), \quad t > t_0, \quad x(t_0) = x_0, \quad x(t_e) = x_e, \quad (4.81b)$$

with $x_0, x_e \in \mathbb{R}^n$ for fixed initial time $t_0$ and free terminal time $t_e \ll T$. Assume that $l$ and $f$ are continuous in $(x, u)$ and continuously differentiable with respect to $x$ for all $(x, u) \in \mathbb{R}^n \times \mathbb{R}^m$.

Suppose that $(u^*(t), t_e^*) \in \mathcal{U} \times [t_0, T)$ is a minimizer of $(4.81)$ with $x^*(t)$ the corresponding solution to $(4.81b)$. Then there exists a non–zero $\tilde{\lambda}^*(t) \in \dot{C}^1([t_0, t_e^*], \mathbb{R}^{n+1})$, $\tilde{x}^*(t) = [\tilde{x}^1_1(t), \ldots, \tilde{x}^n_{n+1}(t)]^T$ such that the canonical equations

$$\begin{align*}
\tilde{x}^* &= (\nabla_x H)(x^*, u^*, \tilde{\lambda}^*) = \left[ \frac{f(x^*, u^*)}{l(x^*, u^*)} \right], \\
\tilde{x}^*(t_0) &= [x_0^0, 0] \quad \tilde{x}^*(t_e) = x_e \quad (4.82a) \\
\tilde{\lambda}^* &= -(\nabla_x H)(x^*, u^*, \tilde{\lambda}^*) = \left[ \frac{(-\nabla_x H)(x^*, u^*, \tilde{\lambda}^*)}{0} \right] \\
&= (4.82b)
\end{align*}$$

are fulfilled with the Hamiltonian function

$$H(x, u, \tilde{\lambda}) = \tilde{\lambda}^T f(x, u) = [\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n] f(x, u) + \tilde{\lambda}_{n+1} l(x, u) \quad (4.82c)$$

of the extended system and:

(i) The optimal control $u^*(t)$ minimizes the function $H(x^*(t), u(t), \tilde{\lambda}^*(t))$ for all $t \in [t_0, t_e]$ in the set of input constraints $U$, i.e.,

$$H(x^*(t), \nu, \tilde{\lambda}^*) \geq H(x^*, u^*, \tilde{\lambda}^*), \quad \forall \nu \in U \quad (4.83)$$

(ii) For any $t \in [t_0, t_e]$ the relations

$$\begin{align*}
\tilde{\lambda}^*_{n+1} &= \text{const.} \geq 0 \\
H(x^*, u^*, \tilde{\lambda}^*) &= \text{const.} \geq 0
\end{align*} \quad (4.84a) \quad (4.84b)$$

hold.

(iii) For free endtime $t_e$ the transversality condition

$$H(x^*(t_e^*), u^*(t_e^*), \tilde{\lambda}^*(t_e^*)) = 0 \quad (4.85)$$

holds.

According to Theorem 4.11 $2n + m + 3$ equations are available for the computation of the $2n + m + 3$ unknowns $(\tilde{x}^*, u^*, \tilde{\lambda}^*, t_e^*)$. These are given by the $2n + 2$ differential equations $(4.82a)$ and $(4.82b)$ for the extended state $\tilde{x}^*(t)$ and the adjoint state $\tilde{\lambda}^*(t)$, $m$ algebraic equations arising from $(4.83)$, namely that for all $t \in [t_0, t_e]$ the input $u(t) = u^*(t)$ minimizes the function $H(x^*(t), u(t), \tilde{\lambda}^*(t))$ in the set of input constraints $U$, and the transversality condition $(4.85)$. These are completed with $n + 1$
With these considerations the general procedure for the application of the maximum principle in the normal case arises, where $H$ is independent of $l$ and hence of the cost functional so that the optimization problem is obviously ill-posed. Here, the adjoint variables $\lambda_j^*(t)$, $j = 1, \ldots, n$ are uniquely determined.

Remark 4.9: Connection to first and second order necessary optimality conditions
The necessary conditions for $u(t)$ being a minimizer of $H(x(t), u(t), \lambda(t))$ for $\lambda^{n+1}(t) = 1$ according to (4.83) directly coincide with the necessary first and second order conditions (4.54c) and (4.55), i.e.,

\[
\begin{aligned}
\nabla_u H(x^*, u^*, \lambda^*) &= 0, \\
(\nabla^2_u H)(x^*, u^*, \lambda^*) &\geq 0, \\
\forall & t \in [t_0, t_e].
\end{aligned}
\]

However, Pontryagin’s maximum principle is more general. The first condition $(\nabla_u H)(x^*, u^*, \lambda^*) = 0$ in general does not hold if the minimum is located at the boundary of the set $U$ of input constraints. Moreover, in Theorem 4.11 it is required that $f$ and $l$ are only continuous in $u$ while the derivation of the Euler–Lagrange equations relies on their continuous differentiability in $u$.

With these considerations the general procedure for the application of the maximum principle in the normal case with $\lambda^{n+1}(t) = 1$ so that $\lambda(t)$ is used for $\lambda(t)$ is given by the following steps:

(i) Set up the Hamiltonian function $H(x, u, \lambda) = \sum_{j=1}^n \dot{\lambda}_j f_j(x, u) + l(x, u)$.

(ii) Solve the minimization problem

$$H(x, v, \lambda) \geq H(x, u, \lambda), \quad \forall v \in U$$

or equivalently

$$u = \arg\min_v \{H(x, v, \lambda) : v \in U, \ t \in [t_0, t_e]\}.$$ 

depending on $(x, \lambda)$. This yields $u = k(x, \lambda)$.

(iii) Substitution of $u = k(x, \lambda)$ into (4.82a), (4.82b) results in the boundary–value problem

$$\dot{x} = (\nabla_A H)(x, k(x, \lambda), \lambda), \quad \ddot{x}(t_0) = \begin{bmatrix} x_0 \\ 0 \end{bmatrix}, \quad x(t_e) = x_e$$

$$\dot{\lambda} = -(\nabla_x H)(x, k(x, \lambda), \lambda)$$

with the transversality condition (4.85) if $t_e$ is free.

(iv) The (numerical) solution of the boundary–value problem yields $(x^*(t), \lambda^*(t))$ and the optimal control $u^*(t) = k(x^*(t), \lambda^*(t))$. 

4.4 Input constrained optimal control
Example 4.10 (Normal case). Consider the optimal control problem

\[
\min_{u \in U} J(u) = \int_{t_0}^{t_e} \frac{1}{2} u^2(t) \, dt
\]

subject to

\[
\dot{x} = u - x, \quad t > 0, \quad x(0) = 1, \quad x(1) = 0
\]

with the input constrained to \(u(t) \in [-0.6, 0]\) for \(t \in [0, 1]\). The Hamiltonian function of the extended system reads

\[
H(x, u, \lambda) = \lambda_1 (u - x) + \frac{1}{2} \lambda_2 u^2
\]

so that (4.82a), (4.82b) evaluate to

\[
\dot{x} = u - x, \quad x(0) = 1, \quad x(t_e) = 0
\]

\[
\dot{\lambda}_1 = -\left(\frac{\partial}{\partial x} H\right)(x, u, \lambda) = \lambda_1
\]

\[
\dot{\lambda}_2 = 0
\]

This implies the solutions for the optimal adjoint states in the form

\[
\lambda_1^* = C_1 e^{t}, \quad \lambda_2^* = C_2
\]

for constants \(C_1\) and \(C_2 \geq 0\). Since the problem is normal choose \(C_2 = 1\). By (4.83) the optimal solution \(u^*(t)\) has to satisfy the inequality

\[
H(x^*, v, \lambda^*) \geq H(x^*, u^*, \lambda^*), \quad \forall v \in [-0.6, 0], \quad \forall t \in [0, 1]
\]

or in other words

\[
u^* = \arg \min_{v} \{H(x^*, v, \lambda^*) : v \in [-0.6, 0], \ t \in [0,1]\}.
\]

With

\[
\left(\frac{\partial}{\partial u} H\right)(x^*, u^*, \lambda^*) = \lambda_1^* + \lambda_2^* u^* = 0
\]

the optimal control follows as

\[
u^* = \begin{cases} 
0, & \lambda_1^* \leq 0 \\
-\lambda_1^*, & \lambda_1^* \in (0,0.6) \\
-0.6, & \lambda_1^* \geq 0.6
\end{cases}
\]

in view of the input constraint. Taking \(C_1 \leq 0\) implies \(\lambda_1^*(t) = C_1 \exp(t) \leq 0\) so that \(u^*(t) = 0\) for \(t \in [0,1]\) and thus \(x^*(t) = \exp(-t)\). This solution is infeasible since \(x^*(1) = \exp(-1) \neq 0\). Hence, we have to consider \(C_1 > 0\).

For \(C_1 > 0\) every optimal control is piecewise continuous and takes the values \(-C_1 \exp(t)\) and \(-0.6\) with at most 1 corner point. To illustrate this note that \(\lambda_1^*(t) = C_1 \exp(t)\) is strictly monotonically
increasing such that \( u^* (t) \) decreases monotonically. Starting with \( u_{(1)} (t) = - C_1 \exp (t) \) for \( t \in [0, c] \) the solution of the differential equation

\[
\dot{x}^*_{(1)} = u^*_{(1)} - x^*_{(1)}, \quad x^*_{(1)} (0) = 1
\]

in this time interval is given by

\[
x^*_{(1)} = e^{-t} \left( 1 + \frac{C_1}{2} \right) - \frac{C_1}{2} e^t.
\]

Similarly for \( t \in (c, 1] \) with \( u^*_{(2)} (t) = -0.6 \) the solution is obtained from

\[
\dot{x}^*_{(2)} = u^*_{(2)} - x^*_{(2)}, \quad x^*_{(2)} (1) = 0
\]

as

\[
x^*_{(2)} = \frac{3}{5} (e^{1-t} - 1).
\]

For the determination of the constant \( C_1 \) recall from (4.84b) that the Hamiltonian function \( H(x^*, u^*, \hat{\lambda}^*) \) has to be constant for all \( t \in [0, 1] \). Hence we have

\[
\lambda_1^* (u^*_{(1)} - x^*_{(1)}) + \frac{1}{2} \lambda_2^* (u^*_{(1)})^2 = \lambda_1^* (u^*_{(2)} - x^*_{(2)}) + \frac{1}{2} \lambda_2^* (u^*_{(2)})^2 \Rightarrow - C_1 \left( 1 + \frac{C_1}{2} \right) = - \frac{3}{5} e C_1 + \frac{9}{50},
\]

which yields the two solutions

\[
C_{1,1} = 0.435721, \quad C_{1,2} = 0.826218.
\]

The switching time \( t_s \) is deduced from the continuity condition

\[
x^*_{(1)} (t_s) = x^*_{(2)} (t_s) \Rightarrow - \frac{C_1}{2} e^{2t_s} + \frac{3}{5} e^{t_s} + 1 - \frac{3}{5} e + \frac{C_1}{2} = 0.
\]

Depending on the determined values of \( C_1 \) the solution of this quadratic equation in \( \exp (t_s) \) yields the switching times

\[
t_{s,1} = 0.319929, \quad t_{s,2} = -0.319929.
\]

Since \( t \in [0, 1] \) the only possible value is \( t_{s,1} = 0.319929 \) so that \( C_1 = C_{1,1} = 0.435721 \). The optimal control hence follows as

\[
u^* = \begin{cases} 
-0.435721 e^{t'}, & t \in [0, 0.319929] \\
-0.6, & t > 0.319929
\end{cases}
\]

and is shown in Figure 4.5.
Example 4.11 (Abnormal case). Consider the optimal control problem

\[
\min_{u \in \mathcal{U}} J(u) = \int_{t_0}^{t_e} l(x(t), u(t)) dt
\]

subject to

\[
\dot{x} = u, \quad t > 0, \quad x(0) = 0, \quad x(1) = 1
\]

with the input constrained to \( u(t) \in [0,1] \) for \( t \in [0,1] \).

There is only the single \( u^*(t) = 1 \) that transfers the state \( x(t) = t \) from the initial state \( x(0) = 0 \) to the terminal state \( x(1) = 1 \). This optimal control is, however, independent of the cost functional in \( l(x(t), u(t)) \).

In the following, Theorem 4.11 is extended by replacing the terminal constraint \( x(t_e) = x_e \) with the target set condition \( x(t_e) \in X_{ta} \subset \mathbb{R}^n \). Herein, \( X_{ta} \) is assumed to be a smooth manifold of dimension \( n - p \leq n \). Recall from Section 3.1.1 that an \( (n-p) \)-dimensional manifold is typically defined in terms of the set

\[ X_{ta} = \{ x \in \mathbb{R}^n : g_j(x) = 0, \ j = 1, \ldots, p \} . \]

The corresponding tangent space \( \mathcal{T}_{x^*} X_{ta} \) at the point \( x = x^* \in X_{ta} \) is hence defined as

\[ \mathcal{T}_{x^*} X_{ta} = \{ d \in \mathbb{R}^n : (\nabla_x g_j)^T(x^*) d = 0, \ j = 1, \ldots, p \} . \quad (4.87) \]

If the functions \( g_j(x), \ j = 1, \ldots, p \) are linearly independent, then the set of functions satisfies the constraint qualification condition (cf. also Remark 3.2)

\[ \text{rank}(\nabla_x g)(x^*) = \text{rank}[ (\nabla_x g_1)(x^*), \ldots, (\nabla_x g_p)(x^*) ] = p . \quad (4.88) \]

With this, the following extension of Pontryagin’s maximum principle can be verified.
Theorem 4.12: Pontryagin maximum principle for autonomous systems with target set condition
Consider the optimal control problem
\[
\min_{u \in U} J(u) = \int_{t_0}^{t_e} l(x(t), u(t)) dt
\]
subject to
\[
\dot{x} = f(x, u), \quad t > t_0, \quad x(t_0) = x_0, \quad x(t_e) \in X_{ia},
\]
with \(x_0 \in \mathbb{R}^n\) for fixed initial time \(t_0\), free terminal time \(t_e \ll T\) and \(X_{ia}\) a smooth manifold of dimension \((n-p)\). Assume that \(l\) and \(f\) are continuous in \((x, u)\) and continuously differentiable with respect to \(x\) for all \((x, u) \in \mathbb{R}^n \times \mathbb{R}^m\).

Suppose that \((u^*(t), t_e^*) \in \mathcal{U} \times [t_0, T)\) is a minimizer of (4.89) with \(x^*(t)\) the corresponding solution to (4.89b). Then there exists a non-zero \(\tilde{\lambda}^*(t) \in \mathcal{C}^1([t_0, t_e^*], \mathbb{R}^{n+1}), \tilde{\lambda}^*(t) = [\tilde{\lambda}_1^*(t), \ldots, \tilde{\lambda}_{n+1}^*(t)]^T\) such that conditions (4.82a)–(4.85) of Theorem 4.11 are fulfilled.

Moreover, \(\lambda^*(t_e^*) = [\tilde{\lambda}_1^*(t_e^*), \ldots, \tilde{\lambda}_{n+1}^*(t_e^*)]^T\) is orthogonal to the tangent space \(\mathcal{T}_{x^*(t_e^*)}X_{ia}\), i.e., the transversality conditions
\[
(\lambda^*(t_e^*))^T d = 0, \quad \forall d \in \mathcal{T}_{x^*(t_e^*)}X_{ia}
\]
hold.

In view of (4.87), (4.88) this obviously implies that \(\lambda^*(t_e^*)\) is a linear combination of the individual \((\nabla x g_j)(x^*(t_e^*))\) and hence admits a representation in the form
\[
\lambda^*(t_e^*) = \sum_{j=1}^{p} \mu_j (\nabla x g_j)(x^*(t_e^*))
\]
with the Lagrange multiplier \(\mu = [\mu_1, \ldots, \mu_p]^T \in \mathbb{R}^p\).

Finally we point our attention to the case of non–autonomous nonlinear dynamic systems depending explicitly on time. The procedure is basically identical to the time–invariant case and again relies on the introduction of an extended system as in (4.80).

Theorem 4.13: Pontryagin maximum principle for non–autonomous systems
Consider the optimal control problem
\[
\min_{u \in U} J(u) = \int_{t_0}^{t_e} l(t, x(t), u(t)) dt
\]
subject to
\[
\dot{x} = f(t, x, u), \quad t > t_0, \quad x(t_0) = x_0, \quad x(t_e) = x_e,
\]
with \(x_0, x_e \in \mathbb{R}^n\) for fixed initial time \(t_0\) and free terminal time \(t_e \ll T\). Assume that \(l\) and \(f\) are continuous in \((t, x, u)\) and continuously differentiable with respect to \((t, x)\) for all \((t, x, u) \in [t_0, T] \times \mathbb{R}^n \times \mathbb{R}^m\).
Suppose that \((u^*(t), t^*_e) \in \mathcal{U} \times \{t_0, T\}\) is a minimizing solution to (4.92). Then there exists a non-zero \(\tilde{\lambda}^n(t) \in \tilde{C}^1([t_0, t^*_e], \mathbb{R}^{n+1})\), \(\tilde{\lambda}^n(t) = [\tilde{\lambda}^1(t), \ldots, \tilde{\lambda}^n(t)]^T\) such that the canonical equations

\[
\dot{x}^* = (\nabla_x H)(t, x^*, u^*, \tilde{\lambda}^n) = \begin{bmatrix} 0 & f(t, x^*, u^*) \\ l(t, x^*, u^*) \end{bmatrix}, \quad \tilde{\lambda}^n(t_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad x^*(t_0) = x_e
\]

are satisfied with the Hamiltonian function

\[
H(t, x, u, \tilde{\lambda}) = \tilde{\lambda}^T f(t, x, u) = [\tilde{\lambda}_1 \ldots \tilde{\lambda}_n] f(t, x, u) + \tilde{\lambda}_{n+1} l(t, x, u)
\]

of the extended system and:

(i) The optimal control \(u^*(t)\) minimizes the function \(H(t, x^*(t), u(t), \tilde{\lambda}^n(t))\) for all \(t \in [t_0, t_e]\) in the set of input constraints \(U\), i.e.,

\[
H(t, x^*, v, \tilde{\lambda}^n) \geq H(t, x^*, u^*, \tilde{\lambda}^n), \quad \forall v \in U
\]

(ii) For any \(t \in [t_0, t_e]\) the relations

\[
\tilde{\lambda}_{n+1} = \text{const.} \geq 0
\]

\[
\left(\frac{\partial}{\partial t} H\right)(t, x^*, u^*, \tilde{\lambda}^n) = \left(\tilde{\lambda}^n\right)^T \left(\frac{\partial}{\partial t} f\right)(t, x^*, u^*)
\]

hold.

(iii) For free endtime \(t_e\) the transversality condition

\[
H(t_e^*, x^*(t_e^*), u^*(t_e^*), \tilde{\lambda}^n(t_e^*)) = 0
\]

holds.

This result can be deduced from Theorem 4.11 by first introducing the auxiliary variable \(x_{n+1}(t) = 1, x_{n+1}(t_0) = t_0\) so that (4.92b) reads

\[
\frac{d}{dt} \begin{bmatrix} x_{n+1} \\ \frac{f(x_{n+1}, x, u)}{1} \end{bmatrix}, \quad t > t_0, \quad \begin{bmatrix} x(t_0) \\ x_{n+1}(t_0) \end{bmatrix} = \begin{bmatrix} x_0 \\ t_0 \end{bmatrix}
\]

and secondly applying Theorem 4.11 to this extended but autonomous system. This also implies the assumption that \(f\) and \(l\) have to be continuously differentiable in \(t\). Theorem 4.12 including a target set condition can be similarly extended to the non-autonomous case.

Further extensions to problems involving inequality constraints can be, e.g., found in [10, 6] and the references therein.
4.4.2 Application to nonlinear affine input systems

Subsequently, Pontryagin's maximum principle is applied to autonomous nonlinear systems that are affine in the input, i.e.

\[ \dot{x} = f(x, u) = f_0(x) + \sum_{j=1}^{m} f_j(x)u_j, \quad t > t_0, \quad x(t_0) = x_0 \]  

(4.97a)

with the input constraints

\[ u \in U = [u^-, u^+] \]  

(4.97b)

or equivalently \( u_j \in [u^-_j, u^+_j] \), \( j = 1, \ldots, m \). For this particular class of nonlinear systems different cost functionals are studied that rather commonly arise in applications.

4.4.2.1 Cost functionals minimizing consumption

Optimal control in view of minimizing consumption can be addressed by cost functionals of the form

\[ J(u) = \int_{t_0}^{t_e} \left( l_0(x(t)) + \sum_{j=1}^{m} r_j |u_j(t)| \right) dt, \quad r_j > 0. \]  

(4.98)

The corresponding Hamiltonian function for the normal case with \( \bar{\lambda}_{n+1}(t) = 1 \) reads

\[ H(x, u, \lambda) = l_0(x) + \sum_{j=1}^{m} r_j |u_j| + \lambda^T \left( f_0(x) + \sum_{j=1}^{m} f_j(x)u_j \right) \]  

(4.99)

Since the term \( l_0(x) + \lambda^T f_0(x) \) is independent of \( u(t) \) it can be neglected when addressing the minimization problem \( H(x^*, v, \lambda^*) \geq H(x^*, u^*, \lambda^*) \), \( \forall v \in U \) arising in the theorems introduced above. In view of the affine input structure the problem can be split into \( m \) independent problems of the form

\[ \min_{u_j \in [u^-_j, u^+_j]} H_j(u_j) = r_j |u_j| + q_j(x, \lambda)u_j, \quad q_j(x, \lambda) = \lambda^T f_j(x). \]  

(4.100)

Noting that minimization has to be achieved for fixed \( (x, \lambda) \), i.e., actually for \( (x^*(t), \lambda^*(t)) \) at each \( t \in [t_0, t_e] \) according to Pontryagin's maximum principle, we have to distinguish between the 4 cases shown in Figure 4.6. The optimal control can be hence deduced as

\[ u^*_j = \begin{cases} u^-_j, & q_j(x, \lambda) > r_j \\ 0, & q_j(x, \lambda) \in (-r_j, r_j), \quad j = 1, \ldots, m. \\ u^+_j, & q_j(x, \lambda) < -r_j \end{cases} \]  

(4.101)

When \( q_j(x(t), \lambda(t)) = \pm r_j \) on a subset \( I_s \subset [t_0, t_e] \) the optimal control \( u^*_j(t) \) can no longer be determined uniquely from (4.100). This case is referred to as the singular case and is considered in Section 4.4.2.4.

4.4.2.2 Cost functionals addressing energy optimality

Optimal control in view of minimizing energy can be addressed by cost functionals of the form
\[ J(u) = \int_{t_0}^{t_f} \left( l_0(x(t)) + \frac{1}{2} \sum_{j=1}^{m} r_j(u_j(t))^2 \right) dt, \quad r_j > 0. \] (4.102)

The corresponding Hamiltonian function for the normal case with \( \bar{\lambda}_{N+1}(t) = 1 \) reads

\[ H(x, u, \lambda) = l_0(x) + \frac{1}{2} \sum_{j=1}^{m} r_j(u_j)^2 + \lambda^T \left( f_0(x) + \sum_{j=1}^{m} f_j(x)u_j \right) \] (4.103)

Proceeding as before the optimal control can be determined for each component \( u_j(t) \) by solving the minimization problem

\[ \min_{u_j \in [u_j^-, u_j^+]} H_j(u_j) = \frac{1}{2} r_j(u_j)^2 + q_j(x, \lambda)u_j, \quad q_j(x, \lambda) = \lambda^T f_j(x). \] (4.104)

Without constraints the minimizer follows as

\[ u_j^0 = -\frac{q_j(x, \lambda)}{r_j} \] (4.105)
Hence, if in the constrained case \( u_0^j \in [u^-_j, u^+_j] \), then \( u^*_j = u_0^j \). Moreover, if \( u_0^j \not\in [u^-_j, u^+_j] \), then the minimizer is defined as a boundary value of the admissible interval. As a result, the optimal control is given by

\[
\begin{align*}
    u^*_j &= \begin{cases} 
        u^-_j, & u_0^j \leq u^-_j \\
        u_0^j, & u_0^j \in (u^-_j, u^+_j), \\
        u^+_j, & u_0^j \geq u^+_j
    \end{cases} \quad j = 1, \ldots, m.
\end{align*}
\] (4.106)

### 4.4.2.3 Cost functionals addressing time optimality

In order to impose time optimality the cost functional reads

\[
  J(u) = \int_{t_0}^{t_e} 1 \, dt = t_e - t_0
\] (4.107)

so that the Hamiltonian function (for the normal case) evaluates to

\[
  H(x, u, \lambda) = 1 + \lambda^T \left[ f_0(x) + \sum_{j=1}^{m} f_j(x) u_j \right]
\] (4.108)

Neglecting terms unaffected by \( u_j \) reduces the respective minimization problem to

\[
  \min_{u_j \in [u^-_j, u^+_j]} H_j(u_j) = q_j(x, \lambda) u_j, \quad q_j(x, \lambda) = \lambda^T f_j(x),
\] (4.109)

whose solution can be immediately determined as

\[
  u^*_j = \begin{cases} 
        u^-_j, & q_j(x, \lambda) > 0 \\
        u^+_j, & q_j(x, \lambda) < 0
    \end{cases} \quad j = 1, \ldots, m.
\] (4.110)

Since \( u^*_j(t) \) only switches between minimal and maximal values this type of optimal control is often referred to as bang–bang control.

Similar to the situation in Section 4.4.2.1 the singular case arises when \( q_j(x(t), \lambda(t)) = 0 \) for \( t \in I_s \subset [t_0, t_e] \). Here, the Hamiltonian function is independent of \( u_j(t) \) so that minimality is trivially ensured without providing information about the minimizer \( u^*_j(t) \). To avoid the singular case, the cost functional (4.107) is often extended by a so-called regularization term

\[
  \int_{t_0}^{t_e} \left( 1 + \frac{1}{2} \sum_{j=1}^{m} r_j(u_j(t))^2 \right) dt
\] (4.111)

with \( r_j > 0 \) but sufficiently small to achieve near time–optimality. Obviously, the cost functional (4.111) corresponds to (4.102) so that the optimal control \( u^*_j(t), j = 1, \ldots, m \) is determined by (4.106).

**Example 4.12.** Consider the linear time–optimal control problem (4.107) for the double integrator

\[
\begin{align*}
    \frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} x_2 \\ u \end{bmatrix}, \quad t > 0, \quad x(0) = x_0, \quad x(t_e) = 0
\end{align*}
\] (4.112)
with $t_e$ free and the input constraint

$$u \in [-1, 1] \quad \forall \, t \in [0, t_e].$$

The Hamiltonian function is given by $H(x, u, \lambda) = 1 + \lambda_1 x_2 + \lambda_2 u$ so that the adjoint equations follow from (4.82b) as

$$\lambda_1^* = 0$$

$$\dot{\lambda}_2^* = -\lambda_1^*.$$

These result in

$$\lambda_1^* = C_1, \quad \lambda_2^* = -C_1 \, t + C_2$$

with the constants of integration $C_1, C_2$. The solution of the minimization problem (4.109) hence yields that the optimal control $u^*(t)$ must satisfy

$$u^*(t) = \begin{cases} -1, & \lambda_2^*(t) > 0 \\ +1, & \lambda_2^*(t) < 0 \end{cases}.$$

Note that the singular case cannot arise since $\lambda_2^*(t) = 0$ for some subinterval $I_s \subset [0, t_e]$ implies $\lambda_2^*(t) = 0$ and $\lambda_1^*(t) = 0$ for all $t \in [0, t_e]$. This contradicts the transversality condition (4.85) for the free terminal time

$$H(x^*(t^*_e), u^*(t^*_e), \lambda^*(t^*_e)) = 1 + \lambda_1^*(t_e)x_2^*(t_e) + \lambda_2^*(t_e)u^*(t_e) = 0.$$

In particular, due to $\lambda_2^*(t)$ being affine in $t$ the situation $\lambda_2^*(t) = 0$ can only arise at a single discrete instance of time $t = t_s$ referring to a switch in $u^*(t)$ between $-1$ and $+1$. Thus, at most 4 switching sequences may arise, i.e., $[+1, [-1, [-1, +1], \text{and } [+1, -1]$. Taking into account (4.112) with piecewise constant $u(t) = u^*(t)$ yields that trajectories $x^*(t)$ denote parabolas in the $(x_1, x_2)$–plane. Indeed it is an easy exercise to show that given $u(t) = u = \pm 1$ we have

$$x_1(t) = \frac{1}{2}ut^2 + x_{2,0}t + x_{1,0}, \quad x_2(t) = ut + x_{2,0} \quad (4.113)$$

and moreover

$$x_1 = \frac{x_2^2}{2u} + x_{1,0} - \frac{x_{2,0}^2}{2u}. \quad (4.114)$$

The optimal control problem is set up to ensure the transfer from an arbitrary initial state $x(0) = x_0$ to the zero state $x(t_e) = 0$ in minimal time. Taking into account (4.114) it becomes apparent that the origin can only be reached along the switching curve defined by

$$x_1 = \frac{x_2^2}{2}, \quad \text{for } u = +1$$

$$x_1 = -\frac{x_2^2}{2}, \quad \text{for } u = -1.$$

Introducing the curve $S(x_2) = -\frac{1}{2}x_2|x_2|$ the switching curve is given by

$$x_1 = S(x_2). \quad (4.115)$$
In view of the possible switching sequences we have to distinguish between the following situations:

(i) If \( \mathbf{x}_0 \) lays on the switching curve, i.e., \( x_{1,0} = S(x_{2,0}) \), then the origin \( x(t_e) = 0 \) is directly reached without any switching along the switching curve (4.115) with \( u(t) = +1 \) or \( u(t) = -1 \). This scenario is shown in Figure 4.7(a).

(ii) If \( \mathbf{x}_0 \) does not lay on the switching curve, i.e., \( x_{1,0} > S(x_{2,0}) \) or \( x_{1,0} < S(x_{2,0}) \), then a single switch in \( u(t) \) is required to reach the switching curve (4.115) and travel along this path to the origin.

Figure 4.7: Switching curve and optimal response for the double integrator of Example 4.12.

Figure 4.7(b) provides a graphical illustration of these two cases. As a result, the optimal control is

\[
u^*(t) = \begin{cases} 
+1, & \text{if } x_1 < S(x_2) \smallskip 
+1, & \text{if } x_1 = S(x_2) \text{ and } x_1 > 0 \smallskip 
-1, & \text{if } x_1 > S(x_2) \smallskip 
-1, & \text{if } x_1 = S(x_2) \text{ and } x_1 < 0 
\end{cases}
\]

With this, the optimal switching points \( t_s \) and the minimal endtime \( t_e^* \) can be determined taking into account (4.113), which provides

\[
t_s = \begin{cases} 
x_{2,0} + \sqrt{\frac{1}{2} x_{2,0}^2 + x_{1,0}}, & x_{1,0} > S(x_{2,0}) 
-x_{2,0} + \sqrt{\frac{1}{2} x_{2,0}^2 - x_{1,0}}, & x_{1,0} < S(x_{2,0}) 
\end{cases}
\]

and

\[
t_e^* = \begin{cases} 
x_{2,0} + \sqrt{2 x_{2,0}^2 + 4 x_{1,0}}, & x_{1,0} > S(x_{2,0}) 
x_{2,0}, & x_{1,0} = S(x_{2,0}) 
x_{2,0} - \sqrt{2 x_{2,0}^2 - 4 x_{1,0}}, & x_{1,0} < S(x_{2,0}) 
\end{cases}
\]

Solutions for the double integrator example when varying the initial condition \( \mathbf{x}(0) = \mathbf{x}_0 \) are shown in Figure 4.8.
Chapter 4  Dynamic optimization

4.4.2.4  Singular problems

As pointed out in Section 4.4.2.1 it may happen that the optimal control $u^*(t)$ cannot be determined from the minimization problem (4.83) for $t \in I_s \subset [t_0, t_e]$. In order to illustrate this and to discuss some measure to proceed in this situation we consider a particular scalar case given by

$$\min_{u \in \mathcal{U}} J(t) = \int_{t_0}^{t_e} \left( l_0(x(t)) + l_1(x(t))u(t) \right) dt$$

subject to

$$\dot{x} = f_0(t, x) + f_1(t, x)u, \quad t > t_0, \quad x(t_0) = x_0, \quad x(t_e) = x_e.$$  \hspace{1cm} (4.118a)

The set $\mathcal{U}$ thereby comprises all piecewise continuous functions $u(t)$ on the interval $[t_0, t_e]$ bounded according to $u(t) \in [u^-, u^+]$. Due to the affine input structure the corresponding Hamiltonian function is affine in $u(t)$, i.e.,

$$H(t, x, u, \tilde{\lambda}) = l_0(x) + \tilde{\lambda}^T f_0(t, x) + \left( \tilde{\lambda}^T f_1(t, x) + l_1(x) \right) u.$$  \hspace{1cm} (4.119)

Finding $u(t)$ minimizing $H(t, x, u, \tilde{\lambda})$ for given $(x(t), \tilde{\lambda}(t))$ hence requires to determine the zeros of $(\nabla_u H)(t, x, u, \tilde{\lambda})$ in $u$. To analyze this introduce the switching function

$$\zeta(t) = (\nabla_u H)(t, x, u, \tilde{\lambda}) = \tilde{\lambda}^T f_1(t, x) + l_1(x)$$  \hspace{1cm} (4.120)

with $\zeta^*(t) = \zeta(t)|_{x(t)=x^*(t),\tilde{\lambda}(t)=\tilde{\lambda}^*(t)}$. If $\zeta^*(t) = 0$ on a finite time interval $I_s \subset [t_0, t_e]$, then the minimization requirement (4.83) does not provide any information about $u^*(t)$ for $t \in I_s$. In other words the control does not affect the Hamiltonian function on $I_s$. This curve $\zeta^*(t)$ is called singular path or singular arc [18, 6].

---

Exercise 4.18. Verify (4.116) and (4.117) for the switching time and minimal endtime.
For the determination of an optimal control along a singular arc one proceeds by adding the requirement that in the interval \( I_s \) also successive derivatives of \( \zeta^*(t) \) must vanish. In particular, the smallest positive integer \( k \) such that

\[
\frac{d^k}{dt^k} \zeta(t) = 0, \quad \frac{\partial}{\partial u} \left[ \frac{d^k}{dt^k} \zeta(t) \right] \neq 0
\]  

(4.121)
can be shown to be even \( k = 2r \), provided it exists. The integer \( k \) is often called the degree (or the order) of the singularity. Along a singular arc the state \( x^*(t) \) and the adjoint state \( \lambda^*(t) \) are restricted to the manifold defined by

\[
\zeta^*(t) = \frac{d}{dt} \zeta^*(t) = \cdots = \frac{d^{k-1}}{dt^{k-1}} \zeta^*(t) = 0
\]  

(4.122)
together with the condition (4.95b) in the non–autonomous case governed by Theorem 4.13. The resulting manifold is also referred to as singular surface. In order to ensure a minimum along the singular arc the generalized Legendre (or Legendre–Clebsch) condition has to hold, which imposes

\[
(-1)^{\frac{k}{2}} \frac{\partial}{\partial u} \left[ \frac{d^k}{dt^k} \zeta^*(t) \right] \geq 0.
\]  

(4.123)

Similar to non–singular problems both the adjoint state \( \lambda(t) \) and the Hamiltonian function \( H \) must be continuous along an optimal trajectory. It should be noted that in general the solution to an optimal control problem may evolve a mixture of singular and non–singular arcs. Finding their proper sequence is a difficult task and may even be impossible for certain problems.

**Example 4.13.** Consider the minimization problem

\[
\min_{u \in U} J(u) = \int_0^2 \frac{1}{2} x_1^2(t) \, dt
\]  

(4.124a)
subject to

\[
\begin{bmatrix}
\frac{dx_1}{dt} \\
\frac{dx_2}{dt}
\end{bmatrix} = \begin{bmatrix} x_2 + u \\ -u \end{bmatrix}, \quad t > 0, \quad x(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad x(2) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]  

(4.124b)

with the input constraint

\[
\forall u \in [-10, 10], \quad \forall t \in [0, 2].
\]

The Hamiltonian function with \( \lambda_3 = 1 \) is given by

\[
H(x, u, \lambda) = \frac{1}{2} x_1^2 + \lambda_1 (x_2 + u) - \lambda_2 u.
\]

The adjoint states are hence governed by

\[
\begin{align*}
\dot{\lambda}_1^* &= -\left( \frac{\partial}{\partial x_1} H \right)(x^*, u^*, \lambda^*) = -x_1^* \\
\dot{\lambda}_2^* &= -\left( \frac{\partial}{\partial x_2} H \right)(x^*, u^*, \lambda^*) = -\lambda_1^*.
\end{align*}
\]
From Theorem 4.11 the optimal control must fulfill condition \((4.83)\), i.e.,

\[
H(x^*, v, \lambda^*) \geq H(x^*, u^*, \lambda^*), \quad \forall v \in [-10, 10].
\] (4.125)

Taking into account the input constraint, this implies

\[
u^* = \begin{cases} 
+10, & \lambda_1^* < \lambda_2^* \\
-10, & \lambda_1^* > \lambda_2^*. \\
\text{undefined}, & \lambda_1^* = \lambda_2^*
\end{cases}
\]

Thus, a singular arc occurs if \(\lambda_1^*(t) = \lambda_2^*(t)\) for some finite time interval \(I_s \subset [0, 2]\). To determine the order of the singular arc consider the following sequence for the switching function \(\zeta^*(t) = (\nabla_u H)(x^*(t), u^*(t), \lambda^*(t)) = \lambda_1^*(t) - \lambda_2^*(t)\)

\[
\zeta^*(t) = \dot{\lambda}_1^*(t) - \dot{\lambda}_2^*(t) = 0
\]

\[
\frac{d}{dt}\zeta^*(t) = \dot{\lambda}_1^*(t) - \dot{\lambda}_2^*(t) = \dot{\lambda}_1^*(t) - x_1^*(t)
\]

\[
\Rightarrow \frac{\partial}{\partial u}\left(\frac{d}{dt}\zeta^*(t)\right) = 0
\]

\[
\frac{d^2}{dt^2}\zeta^*(t) = \ddot{\lambda}_1^*(t) - \dot{x}_1^*(t) = -x_1^*(t) - x_2^*(t) - u^*(t)
\]

\[
\Rightarrow \frac{\partial}{\partial u}\left(\frac{d^2}{dt^2}\zeta^*(t)\right) = -1.
\]

This yields \(k = 2\). The final equation provides the optimal control along the singular arc

\[
u^*(t) = -x_1^*(t) - x_2^*(t)
\]

and the generalized Legendre-Clebsch condition \((4.123)\), i.e.,

\[
(-1)^1 \frac{\partial}{\partial u} \left[ \frac{d^2}{dt^2} \zeta^*(t) \right] = 1 \geq 0,
\]

is fulfilled implying a minimum along the singular arc.

To proceed further note that the Hamiltonian function by \((4.84b)\) must be constant for all \(t \in [0, 2]\) along an optimal trajectory so that

\[
H(x^*, u^*, \lambda^*) = \frac{1}{2}(x_1^*)^2 + \lambda_1^* x_2^* + (\lambda_1^* - \lambda_2^*) u^* = C
\] (4.126)

for some constant \(C\). In addition, along the singular arc \(x^*(t)\) and \(\lambda^*(t)\) are restricted to the singular manifold defined by \((4.122)\), which imposes \(\lambda_1^* = \lambda_2^* = x_1^*\). In view of \((4.126)\) this yields

\[
\frac{1}{2}(x_1^*)^2 + x_1^* x_2^* = C.
\]

The remaining steps can be summarized as follows:

(i) Starting with \(u^*(t) = 10\) solve the differential equations for \(x^*(t)\) and \(\lambda^*(t)\) with the initial condition \(x^*(0) = [1, 1]^T\). Denote the arising 4 constants of integration by \(P_1, ..., 4\) and refer to the solution as \(x^*_1(t)\).

(ii) Since we know that there is a singular arc, solve the differential equations for \(x^*(t)\) and \(\lambda^*(t)\) for \(u^*(t) = -10\) taking into account the terminal condition \(x^*(2) = 0\). Denote the arising 4 constants of integration by \(R_1, ..., 4\) and refer to the solution as \(x^*_3(t)\).
(iii) On the singular arc take into account \( u^*(t) = -x_1^*(t) - x_2^*(t) \) and solve the differential equations for \( x_i^*(t) \) and \( \bar{\lambda}^*(t) \). Denote the arising 4 constants of integration by \( Q_j \), \( j = 1, \ldots, 4 \) and refer to the solution as \( x_{(a)}^*(t) \).

(iv) The conditions determining the singular manifold provide 2 equations to determine 2 of the 4 constants \( Q_j \), \( j = 1, \ldots, 4 \). At the unknown switching times \( t_{s,1} \) onto the singular arc and \( t_{s,2} \) from the singular arc continuity implies \( x_{(1)}^*(t_{s,1}) = x_{(2)}^*(t_{s,1}) \) and \( x_{(2)}^*(t_{s,2}) = x_{(3)}^*(t_{s,2}) \) and hence 4 equations for 4 unknowns including \( t_{s,1} \) and \( t_{s,2} \). Numerically solving this system of coupled nonlinear algebraic equations provides \( t_{s,1} \approx 0.299 \), \( t_{s,2} = 1.927 \) and thus the optimal control

\[
\begin{align*}
u^* &= \begin{cases} 
  +10, & 0 \leq t \leq 0.299 \\
  -(x_1^* + x_2^*), & 0.299 < t < 1.927 \\
  -10, & 1.927 \leq t \leq 2 
\end{cases}
\end{align*}
\]

*aThe minimization problem (4.125) reduces to the analysis of \( (\bar{\lambda}_1^*(t) - \bar{\lambda}_2^*(t))u^*(t) \).

### 4.5 Numerical solution of optimal control problems

In the following, a brief introduction is given to techniques for the numerical solution of optimization problems. We thereby distinguish between

(i) *indirect methods*, where the optimal trajectory is determined by solving the necessary optimality conditions introduced in Sections 4.3 and 4.4, and

(ii) *direct methods*, where the infinite–dimensional optimization problem in the input \( u(t), \ t \in [t_0, t_e] \) is discretized in \( t \) to obtain a finite–dimensional static optimization problem as considered in Chapters 2 and 3.

*Dynamic programming* or *stochastic optimization* are not addressed so that the reader is referred to the respective literature. It is also assumed that the reader is familiar with numerical techniques for the solution of systems of differential equations (see, e.g., [21, 8]).

#### 4.5.1 Indirect methods

As outlined above, indirect methods directly approach the two–point boundary–value problem defined by the necessary optimality conditions. Techniques involve, e.g., *discretization methods*, *shooting methods* or *collocation methods*. In order to introduce the individual concepts the optimization problem

\[
\min_{u \in \mathcal{U}} J(u) = \varphi(t_e, x(t_e)) + \int_{t_0}^{t_e} l(t, x(t), u(t)) \, dt
\]

subject to

\[
\begin{align*}
\dot{x} &= f(t, x, u(t)), & t > 0, & x(0) = x_0 \\
\psi_k(t_e, x(t_e)) &= 0, & k = 1, \ldots, p
\end{align*}
\]

is considered for fixed endtime \( t_e \). Following Theorem 4.10 the Euler–Lagrange equations are given by

\[
\dot{x} = (\nabla_\lambda H)(t, x, u, \lambda), \quad x(t_0) = x_0
\]

(4.128a)
\[
\dot{\lambda} = -(\nabla_x H)(t, x, u, \lambda), \quad \lambda(t_e) = (\nabla_x \phi)(t_e, x(t_e), \mu) \tag{4.128b}
\]
\[
0 = (\nabla_u H)(t, x, u, \lambda) \tag{4.128c}
\]

for \( t \in [t_0, t_e] \) and the transversality condition yields
\[
\psi(t_e, x(t_e)) = [\psi_1(t_e, x(t_e)) \cdots \psi_p(t_e, x(t_e))]^T = 0 \tag{4.129}
\]

with the Hamiltonian function \( H(t, x, u, \lambda) = l(t, x, u) + \lambda^T f(t, x, u) \) and \( \phi(t_e, x(t_e), \mu) = \phi(t_e, x(t_e)) + \mu^T \psi(t_e, x(t_e)) \) for \( \mu \in \mathbb{R}^p \). Under the assumption that (4.128c) can be solved (at least locally) for \( u(t) \) so that \( u(t) = k(x(t), \lambda(t)) \) the substitution of this expression into (4.128a), (4.128b) reduces the problem formulation to a two–point boundary value problem

\[
\begin{align*}
\dot{x} &= (\nabla_\lambda H)(t, x, k(x, \lambda), \lambda), \quad x(t_0) = x_0, \quad \psi(t_e, x(t_e)) = 0 \tag{4.130a} \\
\dot{\lambda} &= -(\nabla_x H)(t, x, k(x, \lambda), \lambda), \quad \lambda(t_e) = (\nabla_x \phi)(t_e, x(t_e), \mu) \tag{4.130b}
\end{align*}
\]

for \( t \in [t_0, t_e] \) with a vector \( \mu \) of free parameters. Since the entries in \( \mu \) are constant they can be integrated into the formulation in terms of \( \dot{\mu} = 0 \) so that (4.130) reads
\[
\begin{align*}
\dot{x} &= (\nabla_\lambda H)(t, x, k(x, \lambda), \lambda), \quad x(t_0) = x_0, \quad \psi(t_e, x(t_e)) = 0 \tag{4.131a} \\
\dot{\lambda} &= -(\nabla_x H)(t, x, k(x, \lambda), \lambda), \quad \lambda(t_e) = (\nabla_x \phi)(t_e, x(t_e), \mu) \tag{4.131b} \\
\dot{\mu} &= 0. \tag{4.131c}
\end{align*}
\]

Introducing
\[
\begin{align*}
Z &= \begin{bmatrix} x \\ \lambda \\ \mu \end{bmatrix}, \quad F(t, Z) &= \begin{bmatrix} (\nabla_\lambda H)(t, x, k(x, \lambda), \lambda) \\ -(\nabla_x H)(t, x, k(x, \lambda), \lambda) \\ 0 \end{bmatrix}, \\
G(t, Z) &= \begin{bmatrix} \psi(t_e, x(t_e)) \\ \lambda(t_e) - (\nabla_x \phi)(t_e, x(t_e), \mu) \end{bmatrix}
\end{align*}
\]

with \( x(t) = MZ(t) \), then (4.131) can be re–formulated according to
\[
\dot{Z} = F(t, Z), \quad t \in (t_0, t_e), \quad MZ(t_0) = x_0, \quad G(t_e, Z(t_e)) = 0. \tag{4.132}
\]

Both problem formulations (4.130) and (4.132) are subsequently exploited depending on the used solution approach.

### 4.5.1.1 Discretization methods

Discretization or relaxation methods in principle make use of an appropriate finite difference approximation of the arising differentials in (4.131). For this, discretize the time–interval \([t_0, t_e]\) in \( N + 1 \) steps
\[
t_j = t_0 + j\delta t, \quad j = 0, 1, \ldots, N, \quad \delta t = \frac{t_e - t_0}{N}
\]

so that the solution can be approximated at the discretization points by \( \hat{z}^j = z(t_j^j), \ j = 0, \ldots, N \). Making, e.g., use of the trapezoidal rule, the respective discretization of (4.132) is obtained as
\[
\frac{z^{j+1} - z^j}{\delta t} = \frac{1}{2} \left[ F(t^j, z^{j+1}) + F(t^j, z^j) \right], \quad j = 0, \ldots, N - 1 \tag{4.133a}
\]
This algebraic system is implicit in the \((2n + p)(N + 1)\) unknowns \(\{z^j\}_{j=0}^N\) and is comprised of \((2n + p)(N + 1)\) nonlinear equations. Note that since \(\mu \in \mathbb{R}^p\) is constant the last \(p\) rows in (4.133a) reduce to \(\mu^{j+1} = \mu^j\). The numerical solution of (4.133) is equivalent to the computation of the zeros of the nonlinear algebraic system

\[
\begin{bmatrix}
M & 0 & 0 & \cdots & 0 & 0 \\
-E & E & 0 & \cdots & 0 & 0 \\
0 & -E & E & \cdots & 0 & 0 \\
& \vdots & & \ddots & \vdots & \vdots \\
0 & \cdots & -E & 0 & E & 0 \\
0 & 0 & 0 & \cdots & -E & E \\
\end{bmatrix}
\begin{bmatrix}
z^0 \\
z^1 \\
z^2 \\
\vdots \\
z^{N-1} \\
z^N \\
\end{bmatrix}
- \begin{bmatrix}
0 \\
F(t^1, z^1) + F(t^0, z^0)^{\delta t} \\
F(t^2, z^2) + F(t^1, z^1)^{\delta t} \\
\vdots \\
F(t^N, z^N) + F(t^{N-1}, z^{N-1})^\delta t \\
G(t^N, z^N) \\
\end{bmatrix}
= \begin{bmatrix}
x_0 \\
0 \\
0 \\
\vdots \\
0 \\
0 \\
\end{bmatrix}
\] (4.134)

making use of, e.g., Newton's method (see also Section 2.2.2.2). Note that in order to obtain a finite difference approximation of (4.132) other techniques can used with different error orders such as backward, forward or central differences

\[\dot{z}(t^j) = \frac{z^{j+1} - z^j}{\delta t} + o(\delta t), \quad \dot{z}(t^j) = \frac{z^{j+1} - z^j}{\delta t} + o(\delta t), \quad \dot{z}(t^j) = \frac{z^{j+1} - z^j}{2\delta t} + o((\delta t)^2)\]

or higher–order discretizations involving Runge–Kutta techniques. Some remarks are in order:

- Discretization methods in general lead to a numerically robust solution due to the simultaneous consideration of the differential equations and the boundary conditions.
- Convergence essentially relies on a proper selection of the initial guess of the adjoint variables.
- The special structure of the arising matrices, cf. (4.134), can be exploited for the numerical solution.
- The number of discretization steps \(N + 1\) inherently influences accuracy and computational burden.

### 4.5.1.2 Shooting method

The shooting method traces the solution of the boundary value problem (4.130) back to the solution of an initial value problem by guessing the values \(\lambda(t_0) = \lambda_0\) and \(\mu\), i.e.,

\[
\begin{align*}
\dot{x} &= (V_{x} H)(t, x, k(x, \lambda), \lambda), \\
\dot{\lambda} &= -(V_{x} H)(t, x, k(x, \lambda), \lambda),
\end{align*}
\] (4.135a, b)

Let \(x(t) = x(t; \lambda_0)\) and \(\lambda(t) = \lambda(t; \lambda_0)\) denote the solution to (4.135). Then the \((n + p)\) residual terms at the boundary \(t = t_e\) given by

\[
B(\lambda_0, \mu) = \left[\lambda(t_e; \lambda_0) - (V_{x} \psi)(t_e, x(t_e; \lambda_0)) + (V_{x} \psi)^T(t_e, x(t_e; \lambda_0)) \mu\right] = 0
\] (4.135c)

need to be fulfilled by properly determining \(\lambda_0 \in \mathbb{R}^n\) and \(\mu \in \mathbb{R}^p\). For this, Newton's method can be utilized to obtain an iterative numerical solution of (4.135c). Let \(\eta = [\lambda_0^T, \mu^T]^T\) denote the stacked vector of unknowns so that \(B(\eta) = B(\lambda_0, \mu)\) and let \(\eta^j\) denote the \(j–th\) Newton iterate, then

\[
(V_{\eta} B)(\eta^j)(\eta^{j+1} - \eta^j) = -B(\eta^j)
\] (4.136)
or in other words
\[ \eta^{j+1} = \eta^j - (\nabla \eta B)^{-1}(\nabla \eta^j) B(\eta^j), \quad \eta^0 = \eta_0 \] (4.137)
for a suitable starting value \( \eta_0 \). The iteration is stopped at \( j = j^* \), e.g., when \( \|\eta^{j+1} - \eta^j\| < \varepsilon \max(1, \|\eta^j\|) \) for some \( \varepsilon \ll 1 \). In addition note:

- The implementation effort for the shooting method is rather small.
- The shooting method essentially relies on a proper choice of the initial state \( \lambda_0 \).
- Since the canonical equations tend to be only weakly stable numerical issues may arise in the forward integration for large intervals \([t_0, t_e]\). This issue can be resolved by considering so-called multiple shooting methods.

**Remark 4.10: Multiple shooting**

In the multiple shooting method the interval \([t_0, t_e]\) is subdivided as in discretization methods and the (single) shooting method is applied in each interval. The solution of the boundary value problem is hence obtained by collecting the individual solutions for each interval and ensuring continuity at the borders between subinterval. With this, the residual terms also cover the conditions arising when connecting the intervals.

### 4.5.1.3 Collocation methods

Collocation methods make use of a solution ansatz in terms of a linear combination of suitable basis functions \( \theta_k(t) \), \( k = 0, \ldots, K \) fulfilling the boundary conditions in (4.130) at \( t = t_0 \) and \( t = t_e \), i.e.,
\[ x(t) \approx \hat{x}(t) = \sum_{k=0}^{K} a_k^x \theta_k(t), \quad \lambda(t) \approx \hat{\lambda}(t) = \sum_{k=0}^{K} a_k^\lambda \theta_k(t), \quad a_k^x, a_k^\lambda \in \mathbb{R}^n. \] (4.138)

The basis functions are thereby assumed to be linear independent and are often chosen as polynomials, Legendre polynomials, trigonometric functions or other families of functions. For the determination of the coefficients \( a_k^x, a_k^\lambda, k = 0, \ldots, K \) collocation conditions are imposed by requiring that the differential equations and boundary conditions (4.130) are satisfied pointwise at \( K + 1 \) distinct collocation points \( t^j \in [t_0, t_e], j = 0, \ldots, K \) with \( t^0 = t_0, t^K = t_e \) that need to be fixed properly. Let \( \hat{x}^j(t) = \hat{x}(t^j) \) and \( \hat{\lambda}^j(t) = \hat{\lambda}(t^j) \), then (4.130) evaluates to
\begin{align}
\dot{x}^j &= (\nabla \lambda H)(t^j, \dot{x}^j, k(\dot{x}^j, \dot{\lambda}^j), \dot{\lambda}^j), \quad (4.139a) \\
\dot{\lambda}^j &= - (\nabla x H)(t^j, \dot{x}^j, k(\dot{x}^j, \dot{\lambda}^j), \dot{\lambda}^j), \quad (4.139b) \\
\dot{x}^0 - x_0 &= 0, \quad (4.139c) \\
\psi(t_e, \dot{x}^K) &= 0, \quad (4.139d) \\
\hat{\lambda}^K - \{(\nabla x, \varphi)(t_e, \dot{x}^K) + (\nabla x, \psi) (t_e, \dot{x}^K) \mu \} &= 0 \quad (4.139e)
\end{align}

This nonlinear system of algebraic equations can be solved, e.g., by making use of Newton’s method.

Besides a global setup also local ansatz functions can be used in each subinterval \([t^j, t^{j+1}]\). As an example consider the MATLAB function bvp4c [22]. Here third order polynomials are used in each subinterval \([t^j, t^{j+1}]\) of a mesh \( t_0 = t^0 < t^1 < \cdots < t^K = t_e \) and the collocation conditions are determined by making the approximate solution fulfill the boundary conditions and the differential equations at both ends and the midpoint of each subinterval. This again results in nonlinear system
of algebraic equations, which can be solved iteratively. It is thereby crucial to recall that boundary value problems can have more than one solution so that a suitable guess for the solution has to be supplied for the iteration.

4.5.1.4 Extensions to free endtime problems

If the endtime $t_e$ is free, then the necessary optimality conditions of Theorem 4.10 include the transversality condition (4.61b), i.e.,

$$\left(\frac{\partial}{\partial t_e} \phi\right)(t_e^*, x^*(t_e^*), \mu^*) + H(t_e^*, x^*(t_e^*), u^*(t_e^*), \lambda^*(t_e^*)) = 0.$$

By introducing the time scaling

$$t = \nu \tau, \quad \tau \in [0, 1], \quad \nu > 0$$

(4.140)

the Euler–Lagrange equations (and similarly the canonical equations introduced by making use of Pontryagin's maximum principle in Section 4.4) can be transformed to the fixed time interval $[0, 1]$ in the new independent coordinate $\tau$. For this, note that differentiation with respect to $t$ is related to differentiation with respect to $\tau$ by

$$\frac{d}{dt} = \frac{1}{\nu} \frac{d}{d\tau}.$$  

(4.141)

As a result, the boundary value problem consisting of the differential equations for the state $x(t)$ and the adjoint state $\lambda(t)$, the algebraic equation $(\nabla_u H) = 0$, and the transversality conditions can be easily re–formulated on the fixed time interval $\tau \in [0, 1]$. Thereby the determination of the unknown optimal endtime $t_e^*$ reduces to the computation of the constant scaling factor $\nu > 0$.

4.5.2 Direct methods

*Direct methods* are based on the direct discretization of the (infinite–dimensional) optimal control problem so that methods of static optimization can be applied to the resulting finite–dimensional problem. Differing from indirect methods, which

- provide insight into the structure of the optimal solution,
- allow to determine a highly accurate or even exact solution,
- enable to utilize the adjoint variables for sensitivity analysis and controller design,

direct methods allow

- to avoid the determination of the canonical equations,
- a simpler incorporation of in particular state and path constraints,
- to compute the Lagrange multiplier in a post–processing step,
- often improved convergence behavior,
- the solution of optimal control problems for systems governed by ordinary differential equations, differential–algebraic equations and partial differential equations

at the cost of providing only a *suboptimal solution* due to discretization.
In order to illustrate so-called direct sequential methods and direct simultaneous methods we focus on the optimal control problem

$$\min_{u \in U} J(u) = \varphi(t_e, x(t_e)) + \int_{t_0}^{t_e} l(t, x(t), u(t)) \, dt \quad (4.142a)$$

subject to

$$\dot{x} = f(t, x, u), \quad t > 0, \quad x(t_0) = x_0 \quad (4.142b)$$

$$g_k(t_e, x(t_e)) = 0, \quad k = 1, \ldots, p \quad (4.142c)$$

$$h_l(t, x(t), u(t)) \leq 0, \quad t \in [t_0, t_e] \quad (4.142d)$$

involving both equality and inequality constraints. The endtime $t_e$ is assumed fixed since the procedure introduced in Section 4.5.1.4 can be used to reduce the case of free endtime to the determination of a single constant scaling parameter.

### 4.5.2.1 Control parametrization

In direct methods the time interval $[t_0, t_e]$ is discretized into $N+1$ stages

$$t_0 = t^0 < t^1 < \cdots < t^N = t_e \quad (4.143)$$

and the control inputs $u(t)$ are parametrized on each individual subinterval $[t^j, t^{j+1}]$ to obtain

$$u(t) = r^j(t, v^j), \quad t \in [t^j, t^{j+1}] \quad (4.144)$$

with $v^j \in \mathbb{R}^{mK}$, where $K$ refers to the order of approximation by functions that are piecewise constant, piecewise linear, piecewise cubic, etc. Examples are shown in Figure 4.9. In practice Lagrange polynomials are often employed for control parametrization. For further details the reader is referred to, e.g., [2, 6].
4.5.2.2 Direct sequential methods

In direct sequential methods\(^5\) the control variables are parametrized according to (4.144) so that (4.142) reduces to the static (finite–dimensional) optimization problem

\[
\min_v J(v) = \phi(t_e, x(t_e)) + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} l(t, x(t), r_j(t), v_j) dt
\]  

with \(v^T = [(v_0)^T, ..., (v_N)^T] \in \mathbb{R}^{mK N}\) subject to

\[
\begin{align*}
\dot{x} &= f(t, x, r_j(t), v_j), & j &= 0, 1, ..., N-1 \quad (4.145b) \\
x(t_0) &= x_0 \quad (4.145c) \\
x(t_j) &= x(t_{j+1}), & j &= 0, 1, ..., N-1 \quad (4.145d) \\
g(t, x(t), x(t)) &= 0 \quad (4.145e) \\
h(t, x(t), r_j(t), v_j) &\leq 0, & t &\in [t_j, t_{j+1}] \quad j = 0, 1, ..., N-1. \quad (4.145f)
\end{align*}
\]

The differential equations (4.145b) need to be solved (numerically) in each subinterval taking into account the initial condition (4.145c), the terminal condition (4.145e) and conditions (4.145d) ensuring continuity of the solution on the interval \([t_0, t_e]\).

The key issue in this setting is the consideration of the path constraint (4.145e), which has to be satisfied for all \(t \in [t_j, t_{j+1}]\). Since this would imply an infinite number of constraints one often replaces this condition by interior–point constraints

\[
h(t_j, x(t_j), r_j(t_j), v_j) \leq 0, & i = 0, ..., I_j, j = 0, 1, ..., N-1. \quad (4.146)
\]

so that (4.145f) has to hold only at a finite number of points interior to each subinterval \([t_j, t_{j+1}]\). Note also the following remarks:

- The control parametrization and the approximation of the path constraints yields \(N\) static optimal control problems in either the Mayer of the Lagrange form, which are subject to an initial value problem. For an efficient numerical solution of (4.145) with (4.146) methods of nonlinear optimization such as the SQP method can be applied to determine the decision variables.
- The accuracy of the solution of the differential equations only depends on the used numerical solver and is hence independent of the grid induced by (4.143). However, problems may arise when the system is unstable or the differential equation does not have a solution for certain decision variables.

4.5.2.3 Direct simultaneous methods

In direct simultaneous methods\(^6\) the infinite–dimensional optimal control problem (4.142) is discretized both in the control and the state variables. While different methods such as Lagrange polynomials or monomial basis functions are available for the parametrization of the state variables in a way similar to (4.144), subsequently numerical routines for the solution of ordinary differential equations are incorporated into the setup. Examples include explicit and implicit Euler or Runge–Kutta methods.

\(^5\)Direct sequential methods are also referred to as semi–discretization methods or control vector parametrization methods.

\(^6\)In the literature direct simultaneous methods are also referred to as full discretization methods.
as well as trapezoidal and Simpson’s rule. These schematically lead to an approximation of the solution to \( \dot{x}(t) = f(t, x(t), u(t)) \) by

\[
\alpha^j(x^{j+1}, u^{j+1}, x^j, u^j) = \beta^j(x^j, u^j), \quad j = 0, \ldots, N - 1
\]

with \( \alpha^j \) and \( \beta^j \) depending on the used approach. With this, the optimal control problem (4.142) is reduced to the static optimal control problem

\[
\min_{\hat{x}, \hat{u}} J(\cdot) = \varphi(t_e, x^N) + \sum_{j=0}^{N-1} \frac{t^{j+1} - t^j}{2} \left[ I(t^{j+1}, x^{j+1}, u^{j+1}) + I(t^j, x^j, u^j) \right]
\]

subject to

\[
\alpha^j(x^{j+1}, u^{j+1}, x^j, u^j) = \beta^j(x^j, u^j), \quad j = 0, 1, \ldots, N - 1 \tag{4.148b}
\]

\[
x^0 = x_0 \tag{4.148c}
\]

\[
g(t^N, x^N) = 0 \tag{4.148d}
\]

\[
h(t^j, x^j, u^j) \leq 0, \quad j = 0, 1, \ldots, N - 1 \tag{4.148e}
\]

on the grid imposed by (4.143). For its solution, the methods introduced in Section 3 can be applied to determine the \((N + 1)(n + m)\) decision variables

\[
\hat{x} = \begin{bmatrix} x^0 \\ \vdots \\ x^N \end{bmatrix}, \quad \hat{u} = \begin{bmatrix} u^0 \\ \vdots \\ u^N \end{bmatrix} \tag{4.149}
\]

It becomes apparent that typically a large-scale static optimization problem is obtained, whose numerical solution might require tailored techniques exploiting, e.g., certain sparsity or block structure properties. Some further remarks are in order:

- The differential equations are fulfilled at the converged solution \( \hat{x}^*, \hat{u}^* \) only.
- The inequality constraints are satisfied at the discretization points \( t^j, j = 0, \ldots, N \) only.
- The number of time intervals \( N + 1 \) not only influences the approximation of the optimal control but also the accuracy of approximation of the solution to the differential equations.

### 4.6 Benchmark example

For the illustration of the discussed numerical approaches for the solution of optimal control problems, we consider the example of an evasive maneuver of a ship. Parts of this example are borrowed from [9] and the equations of motion are derived in [3]. For further details the reader is also referred to [17].

The motion of the ship is governed by

\[
\dot{x} = \begin{bmatrix} x_2 \\ c_1 x_2 \\ c_3 v \sin(x_1 - x_3) \\ c_4 x_2 \\ v \cos(x_1 - x_3) \end{bmatrix} + \begin{bmatrix} 0 \\ c_2 \\ 0 \\ 0 \\ 0 \end{bmatrix} u, \quad t > 0, \quad x(0) = 0 \tag{4.150}
\]
with $x_1(t)$ the heading angle, $x_2(t)$ the yaw rate, $x_3(t)$ the drift angle, $(x_4(t), x_5(t))$ the ship position, and $u(t)$ the rudder angle. The ship velocity is $v$ and is assumed constant with $v = 4$ m/s. The remaining parameters are identified for a real ship and are given by $c_1 = -0.26 \text{ 1/s}$, $c_2 = 0.2 \text{ 1/s}$, $c_3 = -1.87 \text{ 1/(rad m)}$, and $c_4 = 0.6 \text{ 1/s}$. The rudder angle is assumed to be bounded according to

$$u \in [u^-, u^+] .$$

The considered cost functional is given in Bolza form

$$\min_{u \in [u^-, u^+]} J(u) = \frac{1}{2} \Delta x^T(t_e) S_e \Delta x(t_e) + \frac{1}{2} \int_0^{t_e} \{\Delta x^T(t) Q \Delta x(t) + r u^2(t)\} \, dt$$

with $\Delta x(t) = x(t) - x_e$ denoting the distance to a desired final ship position $x_e$. The terminal time $t_e$ is fixed to $t_e = 20$ s.

Subsequently, a sidestep of the ship is considered with

$$x_e = [0 \ 0 \ 0 \ 10 \text{ m free}]^T .$$

For the sake of simplicity the weighting matrices are chosen as $S_e = Q = E$.

**Remark 4.11**

The desired final position $x_e$ is set–up with the final value for the position $x_5(t_e)$ to be free. For the numerical evaluation this is included by neglecting the state $x_5(t)$ throughout the solution of the optimal control problem. Note that this is reasonable since $x_5(t)$ does not couple into the ordinary differential equations of the state variables $x_1(t)$ to $x_4(t)$.

In view of (4.150) the Hamiltonian function reads

$$H(x, u, \lambda) = \frac{1}{2} \Delta x^T Q \Delta x + \frac{r}{2} u^2 + \lambda^T [f_0(x) + f_1 u] .$$

This implies the adjoint system

$$\dot{\lambda} = - (\nabla_x H)(x, u, \lambda) = - Q(x - x_e) - (\nabla_x f_0)(x) \lambda , \quad \lambda(t_e) = S_e(x(t_e) - x_e) .$$

Since the system (4.150) is an affine input system and the cost functional contains a term related to energy optimality the results of Section 4.4.2.2 apply so that the optimal control follows from (4.106) in the form

$$u^*_j = \begin{cases} u^- , & u^0 \leq u^- \\ u^0 , & u^0 \in (u^-, u^+) \\ u^+ , & u^0 \geq u^+ \end{cases} , \quad u^0 = - \frac{1}{r} \lambda^T f_1 = - \frac{1}{r} c_2 \lambda_2 .$$

**Collocation method using bvp4c**

Indirect methods directly act on the canonical equations given by (4.150) and (4.154) with (4.155). Proceeding as in Section 4.5.1 a collocation approach is applied by making use of the MATLAB function bvp4c. The resulting code is split into two subparts with the first one initializing the problem, defining the parameters of the optimal control problem, calling the numerical solver, and finally simulating the original system (recall Remark 4.11).
function ship()

% System parameters
p.c1 = -0.26;
p.c2 = 0.2;
p.c3 = -1.87;
p.c4 = 0.6;
p.v = 3.5;

% Initial state
p.x0 = zeros(4,1);

% Optimization parameters
p.te = 20.0;
p.r = 10.0;
p.Q = diag([1,1,1,1]);
p.S = diag([1,1,1,1]);
p.xe = [0,0,0,10]';

% Scenarios
umax = [15,10,5];
umin = -umax;

for j=1:length(umin)

% Run 1
p.r=10.0;
bvp=ship_bvp4c(umin(j),umax(j),p,1e-4);

% Run 2 with result of run 1 as initial condition, adjusted RelTol and reduced weight r
ini.x = bvp.t;
ini.y = [bvp.x;bvp.l];
p.r=1.0;
bvp=ship_bvp4c(umin(j),umax(j),p,1e-6,ini);

% Run 3 with result of run 2 as initial condition, adjusted RelTol and reduced weight r
ini.x = bvp.t;
ini.y = [bvp.x;bvp.l];
p.r=0.1;
bvp=ship_bvp4c(umin(j),umax(j),p,1e-6,ini);

% Simulate extended system
tspan = linspace(0.0,p.te,201);
[t,x] = ode45(@(t,x)ship_ode(t,x,bvp,p),tspan,zeros(5,1));

end

%=================================================================

Chapter 4  Dynamic optimization
function out = ship_ode(t,x,inp,p);
%ODEs including x5 governing ship motion
u = interp1(inp.t,inp.u,t);
out = zeros(5,1);
out(1) = x(2);
out(2) = p.c1*x(2)+p.c2*u;
out(3) = p.c3*p.v*abs(x(3))*x(3)+p.c4*x(2);
out(4) = p.v*sin(x(1)-x(3));
out(5) = p.v*cos(x(1)-x(3));

It has to be pointed out that the use of bvp4c does for this example not allow to start with the small parameter \( r = 0.1 \) but requires warm-up steps by starting with a larger value of \( r \) (here \( r = 10 \)) and successively reducing \( r \). Thereby, it is crucial to properly assign starting conditions, which are herein chosen as the solution of the previous warm-up step. The problem setup for bvp4c is shown below.

function out = ship_bvp4c(umin,umax,p,reltol,varargin)

%Input constraints
p.umax = umax*pi/180;
p.umin = umin*pi/180;

%Calling bvp4c
opts = bvpset('Stats','on','RelTol',reltol);
if nargin==4
    ic = bvpinit(linspace(0,p.te,40),[p.x0;zeros(size(p.x0))]);
else
    ic = varargin{1};
end
sol = bvp4c(@ship_canon,@ship_bcs,ic,opts,p);
t = sol.x;
x = sol.y(1:4,:);
l = sol.y(5:end,:);

%Determine input
for j=1:length(sol.x)
    u(j) = ship_oc(x(:,j),l(:,j),p);
end

out.t=t;
out.x=x;
out.l=l;
out.u=u;

%=================================================================

%SUBFUNCTIONS

function out = ship_ode(t,x,u,p);

%SUBEQUATIONS

function out = ship_ode(t,x,u,p);
%ODEs governing ship motion
out = zeros(4,1);
out(1) = x(2);
out(2) = p.c1*x(2)+p.c2*u;
out(3) = p.c3*p.v*abs(x(3))*x(3)+p.c4*x(2);
out(4) = p.v*sin(x(1)-x(3));
%out(5) = p.v*cos(x(1)-x(3));

%=================================================================
function out = ship_fjac(x,p);
%Jacobian of the ship ODEs
out = zeros(4,4);
out = [0,1,0,0;
      0,p.c1,0,0;
      0,p.c4,2*p.c3*p.v*x(3)*sign(x(3)),0;
      p.v*cos(x(1)-x(3)),0,-p.v*cos(x(1)-x(3)),0];
%=================================================================

function out = ship_canon(t,z,p)
%Canoncial equations
x = z(1:4);
l = z(5:end);
u = ship_oc(x,l,p);
xdot = ship_ode(t,x,u,p);
ldot = -(p.Q*(x-p.xe) + ship_fjac(x,p)’*l);
out = [xdot;ldot];

%=================================================================

function out = ship_bcs(z0,ze,p)
%Boundary conditions
x0 = z0(1:4);
xe = ze(1:4);
le = ze(5:end);
out = [x0-p.x0;
       le-p.S*(xe-p.xe)];

%=================================================================

function out = ship_oc(x,l,p)
%Optimal control
f1 = [0,p.c2,0,0]’;
u = -l’*f1/p.r;
%Take into account input constraints
if u< p.umin
    u = p.umin;
elseif u> p.umax

156 Chapter 4 Dynamic optimization
Numerical results when varying the input constraints are shown in Figure 4.10. The maneuver is driven rather aggressively as can be seen from the rudder angle that almost approaches a bang–bang behavior switch between the maximal and minimal allowed input value. A relaxation is achieved by either increasing $r$ or the terminal time $t_e$.

![Figure 4.10: Sidestep of the ship when varying the input constraints: Collocation method using bvp4c.](image)

**Multiple shooting using ACADO**

As a second approach, the open–source code ACADO [13, 12] is utilized to compute the optimal solution of the maneuvering problem. Besides a C++ interface ACADO provides a MATLAB interface [1] that is used to compute the following results. Provided that the necessary libraries are available in the MATLAB search path the problem can be setup as follows.

```plaintext
clear;
BEGIN_ACADO;

acadoSet('problemname', 'ship');

%Optimization parameters
te = 20.0;
xe = [0,0,0,10];

%Initialize
DifferentialState x1 x2 x3 x4 L;
Control u;

% Set default objects
f = acado.DifferentialEquation();
f.linkMatlabODE('ship_ode');
```

%Optimal control problem
ocp = acado.OCP(0.0, te, 200);
ocp.minimizeMayerTerm(L + (x1-xe(1))*(x1-xe(1)) + (x2-xe(2))*(x2-xe(2)) + ... 
(x3-xe(3))*(x3-xe(3)) + (x4-xe(4))*(x4-xe(4)));

ocp.subjectTo( f );
ocp.subjectTo( 'AT_START', x1 == 0.0 );
ocp.subjectTo( 'AT_START', x2 == 0.0 );
ocp.subjectTo( 'AT_START', x3 == 0.0 );
ocp.subjectTo( 'AT_START', x4 == 0.0 );
ocp.subjectTo( 'AT_START', L == 0.0 );

%[-15 deg,+15 deg]
ocp.subjectTo( -0.261799 <= u <= 0.261799 );
%[-10 deg,+10 deg]
ocp.subjectTo( -0.174533 <= u <= 0.174533 );
%[-5 deg,+5 deg]
ocp.subjectTo( -0.087266 <= u <= 0.087266 );

algo = acado.OptimizationAlgorithm(ocp);
algo.set('INTEGRATOR_TOLERANCE',1e-5);
algo.set('KKT_TOLERANCE', 1e-6);
algo.initializeControls([0 0]);

END_ACADO; % Always end with "END_ACADO".
clear;
out = ship_RUN();
draw; % Graphical output needs to be supplied by the user

In the setup above, the user has to specify the system (4.150) as a MATLAB function provided below. Note that the system equations can be also provided inline. However, due to the unavailability of a function realizing the absolute value \( | \cdot | \) in the ACADO MATLAB–interface linking to the external function was chosen to overcome this restriction.

function [ dx ] = ship_ode( t,x,u,p,w )

%System parameters
c1 = -0.26;
c2 = 0.2;
c3 = -1.87;
c4 = 0.6;
v = 3.5;

%Optimization parameters
r = 0.1;
xe = [0,0,0,10]';
de = (x(1)-xe(1))*(x(1)-xe(1)) + (x(2)-xe(2))*(x(2)-xe(2)) + ... 
(x(3)-xe(3))*(x(3)-xe(3)) + (x(4)-xe(4))*(x(4)-xe(4));

%abs() seems not to work in ACADO
if x(3)<0
    h=-x(3)*x(3);
else
    h=x(3)*x(3);
end

dx(1) = x(2);
dx(2) = c1*x(2)+c2*u;
dx(3) = c3*v+h+c4*x(2);
dx(4) = v*sin(x(1)-x(3));
dx(5) = 0.5*(de +r*u*u);
end

Numerical results achieved using ACADO are shown in Figure 4.11. Both optimal control and the ship trajectory are almost identical to those obtained in Figure 4.10 when taking into account the different input constraints.

Figure 4.11: Sidestep of the ship when varying the input constraints: Multiple shooting using ACADO.
References


