Chapter 2
Optimal Control Problem

Optimal control of any process can be achieved either in open or closed loop. In the following two chapters we concentrate mainly on the first class. The first chapter is devoted to definition of open-loop optimal control (dynamic optimisation) problems. Next, the chapter is concerned with practical ways (techniques) that can be used to solve such problems. It also discusses closed-loop implementation, handling of disturbances, and nonlinear state-feedback control.

We introduce three basic parts of an optimal control problem (OCP): an objective functional, constraint functions, and a process model and their common mathematical forms. The objective functional, optimisation criterion, or performance index represents mathematical expression of phenomenon whose minimum (or maximum) we want to attain. The constraint functions of various types determine a search space of decision (optimisation) variables whose time evolutions or values are searched for. The process model function ties inputs, states, and outputs of the process together and determines a search domain for the optimisation procedure in a similar way as the constraint functions do.

2.1 Objective Functional

Objective functional expresses costs or benefits of a process one wants to either avoid or reach. Dynamic optimisation objective functional can be in general defined in three different forms which can be converted easily from one to another [1]. Mathematical representations of these forms follow together with the year of approximate first appearance in the literature.

Lagrange form (1780)

$$\mathscr{J} = \int_{t_0}^{t_\mathrm{f}} \mathscr{F}\left(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{p}, t\right) \mathrm{d}t, \tag{2.1a}$$

R. Paulen and M. Fikar, *Optimal Operation of Batch Membrane Processes*,
Advances in Industrial Control, DOI 10.1007/978-3-319-20475-8_2

Mayer form (1890)

$$\mathscr{J} = \mathscr{G} = (\boldsymbol{x}(t_{\mathrm{f}}), \boldsymbol{p}, t_{\mathrm{f}})\,, \tag{2.1b}$$

Bolza form (1900)

$$\mathscr{J} = \mathscr{G}\,(\boldsymbol{x}(t_{\mathrm{f}}), \boldsymbol{p}, t_{\mathrm{f}}) + \int_{t_0}^{t_{\mathrm{f}}} \mathscr{F}\,(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{p}, t)\,\mathrm{d}t. \tag{2.1c}$$

Here t represents independent time variable, indices $\square_0$ and $\square_{\mathrm{f}}$ indicate initial and final process state, respectively, $\boldsymbol{x}(t) \in \mathbb{R}^{n_x}$ is a vector of state variables, $\boldsymbol{u}(t) \in \mathbb{R}^{n_u}$ denotes a vector of control variables which are going to be optimised.[1] Time-independent decision variables, traditionally called parameters, are denoted by $\boldsymbol{p} \in \mathbb{R}^{n_p}$. $\mathscr{J}$, $\mathscr{G}(\cdot)$ and $\mathscr{F}(\cdot)$ are then real-valued functions (functionals) since $\mathscr{G} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \times \mathbb{R} \to \mathbb{R}$ and $\mathscr{F} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \times [t_0, t_{\mathrm{f}}] \to \mathbb{R}$.

2.1.1 Typical Optimal Control Tasks

One chooses an approach of solving OCP based on the type (structure) of this problem. There are several types of optimal control tasks which are distinguished, e.g. by fixed/free initial/terminal time/state. Initial time is usually fixed since it is either the time when we start to observe the process or the time when some previous process (whose duration may have been optimised) ends. Many applications consider the initial state to be a free parameter which is chosen such that the designed process is optimal. However, this is omitted in Fig. 2.1 where we show schematic representation of the most common optimal control tasks.

Using different forms of objective functionals, one can consider different control problems. Historically the first form of functional, a Lagrange form, can be used, according to the mathematical form of $\mathscr{F}(\cdot)$ in (2.1a), for the following ones:

1. Minimum control effort problem

$$\mathscr{F} \equiv \|\boldsymbol{u}\|. \tag{2.2a}$$

   With $\|\cdot\|$ being appropriate norm, this form of a functional represents an approach of direct minimisation of process expenses since the goal is to minimise the control effort (e.g. energy) needed to transfer the state $\boldsymbol{x}(t)$ from a specified point $\boldsymbol{x}_0$ to the required $\boldsymbol{x}_{\mathrm{f}}$ in given or unspecified time t_{f}. This is visualised in Fig. 2.1a, b.
2. Minimum time problem

$$\mathscr{F} \equiv 1. \tag{2.2b}$$

[1]Explicit time-dependency of state and control variables will be usually omitted in this work for purpose of better readability of the text. Hence it holds that $\boldsymbol{x}(t) \equiv \boldsymbol{x}$ and $\boldsymbol{u}(t) \equiv \boldsymbol{u}$.

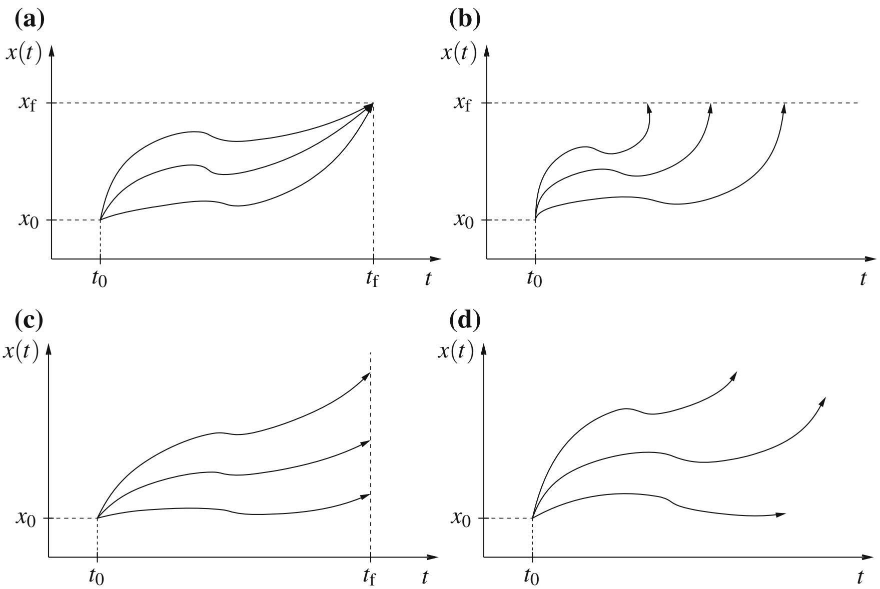

Fig. 2.1 Typical optimal control tasks. **a** Fixed terminal time and terminal state. **b** Free terminal time, fixed terminal state. **c** Fixed terminal time, free terminal state. **d** Free terminal time and terminal state

represents historically the first optimal control problem treated (brachistochrone problem proposed by Johan Bernoulli). For the optimisation of chemical processes, use of this functional is one possibility to indirectly (but globally) minimise operational process expenses such as electricity needed to power the pumps and heaters, amount of heating and cooling media used, and so on. This problem is to find such control that drives the process from a given initial state $\boldsymbol{x}_0$ to the prescribed (fixed) final state $\boldsymbol{x}_\mathrm{f}$ in minimum time possible. This can be schematically represented by Fig. 2.1b as finding a time-optimal transition trajectory between $\boldsymbol{x}_0$ and $\boldsymbol{x}_\mathrm{f}$.

3. LQ (linear-quadratic) problem

$$\mathscr{F} \equiv \frac{1}{2}(\boldsymbol{x}^T\mathbf{Q}\boldsymbol{x} + \boldsymbol{u}^T\mathbf{R}\boldsymbol{u}). \tag{2.2c}$$

This problem considers a linear process model (to be defined later herein) and a quadratic criterion such that $\mathbf{Q} \in \mathbb{S}^{n_x}$ and $\mathbf{R} \in \mathbb{S}^{n_u}$ are symmetric positive semi-definite and positive definite matrices, respectively, which weight the steering from a point $\boldsymbol{x}_0$ to a point as close as possible to the origin (zero) and the control effort used for this task. Terminal time is usually specified for this kind of problem and it is then desired to choose one of the admissible state trajectories (shown in Fig. 2.1c) which minimises the functional. This type of functional is often used in closed-loop optimal control since it yields Lyapunov function whose properties guarantee the stability of the closed-loop control system.

Mayer form allows for optimisation of some criterion evaluated at the final point and can be used as a criterion for many optimal control tasks.

1. Minimum time problem
$$\mathscr{G} \equiv t_{\mathrm{f}}. \tag{2.3a}$$
This represents the very same problem as discussed before (2.2b) and thus an example of how considered functional forms are closely connected and interchangeable. Indeed, the correct approach of interchanging those forms is to introduce an additional differential equation $\dot{x}_{n_x+1} = 1$ and to state the criterion in Mayer form as $\equiv x_{n_x+1}(t_{\mathrm{f}})$, which is equivalent to (2.3a).
2. Terminal control problem
$$\mathscr{G} \equiv \|\boldsymbol{x}(t_{\mathrm{f}}) - \boldsymbol{x}_{\mathrm{f}}\|. \tag{2.3b}$$
This form of a functional represents a direct expression of goals of the process itself which are to be maximised or minimised. If speaking about chemical processes, this might represent minimisation of the difference between a desired and an achieved purity or quantity of product or the same on the side of some unwanted by-products. Graphically, this OCP can be represented by Fig. 2.1c, d.

Bolza form of an objective functional combines Lagrange and Mayer form and thus stands for the most general form. It is considered in our further discussions on optimal control since it comprises all possible objective functional forms.

2.2 Constraints

As stated above, constraint functions determine the optimisation search space and can take various forms in general. Different equality and inequality constraints may be considered to bound the values of decision as well as state variables to respect safety or environmental safeguards, to state certain setpoints in control loop, and so on. The following list covers all types of considered constraints:

- Infinite dimensional equality constraint
$$h(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t) = 0, \qquad \forall t \in [t_{c,0}, t_{c,\mathrm{f}}],\ [t_{c,0}, t_{c,\mathrm{f}}] \subseteq [t_0, t_{\mathrm{f}}], \tag{2.4a}$$
is typically present in chemical processes. An example of using such kind of constraints is a separation of a gas mixture where sum of mole fractions of all components, $\mathbf{1}^T \boldsymbol{y}$, must be equal to one at each time instance of the process.
- Infinite dimensional inequality constraint
$$g(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t) \leq 0, \qquad \forall t \in [t_{c,0}, t_{c,\mathrm{f}}],\ [t_{c,0}, t_{c,\mathrm{f}}] \subseteq [t_0, t_{\mathrm{f}}], \tag{2.4b}$$

is also of great practical interest since it may express a limit of some resource that may not be overrun. Traditionally, such kind of constraints are present as so-called *box constraints* of a form $\boldsymbol{x} \in [\boldsymbol{x}_{\min}, \boldsymbol{x}_{\max}]$, where $\boldsymbol{x}_{\min}$ and $\boldsymbol{x}_{\max}$ denote lower and upper bounds on $\boldsymbol{x}$ respectively.

- Point equality constraint

$$h(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t_c) = 0, \quad t_c \in [t_0, t_\mathrm{f}], \tag{2.4c}$$

is used to determine a specific value (or set of values) of state/decision variables at distinct time points. This constraint is commonly used when we solve the fixed terminal point problem, e.g. concentration of product is restricted to some value at the end of the reaction.

- Point inequality constraint

$$g(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t_c) \leq 0, \quad t_c \in [t_0, t_\mathrm{f}], \tag{2.4d}$$

may be used in a similar way as the previous type, to ensure that some quantity does not exceed the specified limit at some distinct time point. For example, we may express the required purity of the product of a separation by this kind of constraint.

We note that all these constraints can be expressed in the standard canonical form, an equivalent to the functional form (2.1c)

$$\mathscr{J}_c = \mathscr{G}_c\left(\boldsymbol{x}(t_c), \boldsymbol{p}, t_c\right) + \int_{t_{c,0}}^{t_{c,\mathrm{f}}} \mathscr{F}_c\left(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t\right) \mathrm{d}t, \tag{2.5}$$

where $c = 1, \ldots, n_c$ and n_c is the number of constraints. Constraints (2.4) can be rewritten in the canonical form such that for:

- Infinite dimensional equality constraint

$$\mathscr{G}_c = 0, \qquad \mathscr{F}_c = \omega\left(h(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t)\right)^2, \qquad \mathscr{J}_c = 0. \tag{2.6a}$$

where ω is zero if t runs outside the interval $[t_{c,0}, t_{c,f}]$. Otherwise, it is an empirical positive and adjustable weighting factor which is used, for example, to improve the numerical accuracy of the optimisation solver.

- Infinite dimensional inequality constraint

$$\mathscr{G}_c = 0, \qquad \mathscr{F}_c = \omega \max\left(0, g(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t)\right), \qquad \mathscr{J}_c = 0, \tag{2.6b}$$

- Point equality constraint

$$\mathscr{G}_c = h\left(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t_c\right), \qquad \mathscr{F}_c = 0, \qquad \mathscr{J}_c = 0, \tag{2.6c}$$

- Point inequality constraint

$$\mathscr{G}_c = g\left(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t_c\right), \qquad \mathscr{F}_c = 0, \qquad \mathscr{J}_c \leq 0. \tag{2.6d}$$

The presented canonical forms of constraint functions can be adjoined to the cost functional $\mathscr{J}$ using a vector of Lagrange multipliers $\boldsymbol{\nu} \in \mathbb{R}^{n_c}$ to form an augmented functional $\bar{\mathscr{J}}$

$$\bar{\mathscr{J}} = \mathscr{J} + \sum_{c=1}^{n_c} \nu_c \mathscr{J}_c. \tag{2.7}$$

Equivalently, this functional can be written as

$$\bar{\mathscr{J}} = \bar{\mathscr{G}} + \int_{t_0}^{t_\mathrm{f}} \bar{\mathscr{F}} \mathrm{d}t, \tag{2.8}$$

with

$$\bar{\mathscr{G}} = \mathscr{G} + \sum_{c=1}^{n_c} \nu_c \mathscr{G}_c, \quad \text{and} \quad \bar{\mathscr{F}} = \mathscr{F} + \sum_{c=1}^{n_c} \nu_c \mathscr{F}_c. \tag{2.9}$$

2.3 Process Model

In principle, the process model represents an additional set of equality constraints since in general it consists of a set of algebraic, differential, and/or functional equations which, if satisfied, give input–output or inner mathematical description of phenomena taking place in the observed system.

While studying dynamic optimisation of processes running in continuous time, continuous time-dependent (dynamical) models are involved. The simplest type of the model which will guarantee these properties is a model described by a set of ordinary differential equations (ODEs)

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{f}\left(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t\right), \qquad \forall t \in [t_0, t_\mathrm{f}]. \tag{2.10a}$$

Here the vector function $\boldsymbol{f}(\cdot)$ is such that $\boldsymbol{f} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \times [t_0, t_\mathrm{f}] \to \mathbb{R}^{n_x}$. If the solution $\boldsymbol{x}(t)$ to (2.10a) additionally satisfies the initial condition

$$\boldsymbol{x}(t_0, \boldsymbol{p}) = \boldsymbol{x}_0(\boldsymbol{p}), \tag{2.10b}$$

where $\boldsymbol{x}_0(\cdot)$ is such that $\boldsymbol{x}_0 : \mathbb{R}^{n_p} \to \mathbb{R}^{n_x}$, it is a solution to initial value problem (IVP) defined by (2.10).

Hence, Eq. (2.10a) is written in the form we classify as a non-autonomous differential equation. However, the majority of physical processes are described by autonomous ODEs in the following form:

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{f}\left(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}\right), \qquad \forall t \in [t_0, t_\mathrm{f}]. \tag{2.11}$$

This means that dynamics of the process are affected just by dynamics of process states and inputs, and by parameters, not by the time itself. Simply spoken, for the distinct process state, an action (control) performed at time τ has the same effect as the same action carried out at time $\tau + \Delta\tau$ where $\Delta\tau$ is arbitrary.

It is often the case that the structure of the process model affects the complexity of optimal control problem dramatically. This is why a particular type of the process model may decide about the optimisation strategy used for solving the OCP.

2.3.1 Linear Time-Invariant System

A common approach to represent linear time-invariant systems is the state-space form

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{u}, \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \tag{2.12}$$

with state matrix $\mathbf{A} \in \mathbb{R}^{n_x \times n_x}$ and input matrix $\mathbf{B} \in \mathbb{R}^{n_x \times n_u}$. Solution of such a set of ODEs is usually easily found to be:

$$\boldsymbol{x}(t) = \mathrm{e}^{\mathbf{A}t}\boldsymbol{x}_0 + \int_{t_0}^{t} \mathrm{e}^{\mathbf{A}(t-\tau)}\mathbf{B}\boldsymbol{u}(\tau)\mathrm{d}\tau, \tag{2.13}$$

where matrix $\mathrm{e}^{\mathbf{A}t}$ is traditionally called fundamental matrix of a system. Existence of such explicit solution gives certain advantage and implies that it is somewhat easier to find optimal control of processes which can be modelled in the stated manner. This is why these kinds of models are widely used for representing process behaviour in optimal, robust, or feedback control. Moreover, Eq. (2.12) yields a set of convex equality constraints and therefore, provided that other considered constraints (2.4) are convex and the objective functional is of the form (2.2) or (2.3), the resulting optimal control problem is convex. In general, it is easier to solve convex rather than non-convex problems.

2.3.2 Input Affine System

In this work, a majority of the interest is devoted to systems of the form

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{a}(\boldsymbol{x}) + \mathbf{B}(\boldsymbol{x})\boldsymbol{u}, \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \tag{2.14}$$

where $\boldsymbol{a}(\cdot)$ and $\mathbf{B}(\cdot)$ represent non-linear vector and matrix functions, respectively, such that $\boldsymbol{a} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ and $\mathbf{B} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x \times n_u}$. Presence of non-linearities in such set of ODEs establishes the fact that in general, no analytical solution can be given and we usually rely on numerical techniques. These numerical techniques, described in detail in [2], include a use of Euler explicit/implicit, Runge-Kutta, Backward differentiation formula, and Adams-Moulton method.

2.4 Summary of Problem Definition

We have introduced various forms of three basic parts of DO problem (the objective functional (2.1), constraints (2.4), and the process model in Sect. 2.3). These can be summarised to general OCP of the following form[2]

$$\begin{aligned}
&\min_{\boldsymbol{u}(t),\boldsymbol{p}} \left\{ \mathscr{G}(\boldsymbol{x}(t_\mathrm{f}), \boldsymbol{p}) + \int_{t_0}^{t_\mathrm{f}} \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t)\,\mathrm{d}t \right\}, \\
&\text{s.t.} \quad \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}), \quad \forall t \in [t_0, t_\mathrm{f}], \\
&\qquad \boldsymbol{x}(t_0, \boldsymbol{p}) = \boldsymbol{x}_0(\boldsymbol{p}), \\
&\qquad \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t) = \boldsymbol{0}, \quad \forall t \in [t_0, t_\mathrm{f}], \\
&\qquad \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}, t) \le \boldsymbol{0}, \quad \forall t \in [t_0, t_\mathrm{f}], \\
&\qquad \boldsymbol{u}(t) \in [\boldsymbol{u}_\mathrm{min}(t), \boldsymbol{u}_\mathrm{max}(t)], \\
&\qquad \boldsymbol{p} \in [\boldsymbol{p}_\mathrm{min}, \boldsymbol{p}_\mathrm{max}].
\end{aligned} \tag{2.15}$$

In order to find a solution to this problem, one can exploit various techniques, stochastic (such as genetic algorithms, simulated annealing, etc.) or deterministic ones. Deterministic methods, in the scope of this work, are based on the following three principles [3–5]: variational calculus (developed by Euler and Lagrange), dynamic programming, and Pontryagin's minimum principle. The latest one gave rise to most popular numerical techniques such as control vector parameterisation, control vector iteration, boundary condition iteration, orthogonal collocation, and multiple shooting techniques.

[2]For lucidity of this definition and of the following text we use a dot over a variable in order to represent its derivative with respect to time.

References

1. Kirk DE (1970) Optimal control theory: an introduction. Prentice-Hall, London
2. Brenan KE, Campbell SE, Petzold LR (1989) Numerical solution of initial value problems in differential-algebraic equations. North-Holland, New York
3. Bellman R (1957) Dynamic programming. Princeton University Press, Princeton
4. Hull DG (2003) Optimal control theory for applications. Mechanical engineering series. Springer, New York
5. Pontryagin LS, Boltyanskii VG, Gamkrelidze RV, Mishchenko EF (1962) The mathematical theory of optimal processes. Wiley, New York

http://www.springer.com/978-3-319-20474-1

Optimal Operation of Batch Membrane Processes
Paulen, R.; Fikar, M.
2016, XXV, 158 p. 60 illus., 42 illus. in color. With online files/update., Hardcover
ISBN: 978-3-319-20474-1