# THE MAXIMUM PRINCIPLE OF PONTRYAGIN

**Martin Brokate**
*Technische Universität München, Germany.*

**Keywords:** Optimal control, Maximum principle, Bang-bang control, Singular control, Adjoint function, Polyhedral constraints, Dynamic programming, Numerical solution, Shooting method, Multiple shooting method, Homotopy method

## Contents

1. Introduction
2. The Maximum Principle
2.1 Problem of Optimal Control
2.2 Statement of the Maximum Principle
2.3 Other Boundary Conditions
2.4 Degeneracy
2.5 Additional State and Control Constraints
2.6 Autonomous Problems
2.7 Problems with Free Final Time
3. Structure of Optimal Controls
3.1 No Explicit Control Constraints
3.2 Polyhedral Control Constraints
3.2.1 Optimal Bang-Bang Controls
3.2.2 Optimal Singular Controls
3.3 State Constraints
4. Relation to Dynamic Programming
5. Numerical Solution Based on the Maximum Principle
Glossary
Bibliography
Biographical Sketches



## Summary

This paper explains the maximum principle of Pontryagin, its setting and form, its consequences and its numerical evaluation. The maximum principle arises as a necessary condition of optimality when we want to optimize the performance of a dynamical system by the action of control functions, which influence directly the derivatives of the state of the system.

## 1. Introduction

A process, which evolves in time, may appear in many different forms. A mathematical description of its evolution, if it is done in continuous time, usually involves differential or integral equations. If the state of the process at each time $t$ is characterized by a vector $x(t) \in \mathbb{R}^n$, then the system of ordinary differential equations

$$\dot{x}(t) = f(t, x(t)) \tag{1}$$

constitutes the most common mathematical model for a deterministic evolution. Indeed, under natural assumptions concerning the right hand side $f$, the value of $x(t)$ is uniquely determined by an initial condition $x(t_0) = x_0$ for all times $t$. In a control system, we may influence the evolution by a control function $u = u(t)$ which enters the right hand side,

$$\dot{x}(t) = f(t, x(t)), u(t)). \tag{2}$$

In choosing the control, we may pursue different goals, like stabilization or optimization of certain performance parameters. The problem of optimal control arises if we want to choose $u$ such that a prescribed cost functional $J$ is optimized. Any control that achieves that goal is called optimal control.

If we minimize a functional $J$ without any constraints, then the condition

$$DJ=0 \tag{3}$$

must hold for the derivative of $J$ at the optimal solution, if such a solution exists. Accordingly, (3) is called a necessary condition of optimality. If there are constraints, condition (3) has to be replaced by a more general condition involving so-called Lagrange multipliers. When we evaluate this condition for the problem of optimal control, we get a set of equations and inequalities that are called the maximum principle, usually referred to as the maximum principle of Pontryagin. Sometimes, this necessary condition is also sufficient for optimality by itself (if the overall optimization is convex), or in combination with an additional condition on the second derivative.

This general approach also works for evolutions described by partial differential equations, integral equations or other functional-differential equations like delay differential equations. See *Optimization and Control of Distributed Processes*. It is not restricted to evolution problems where the independent variable $t$ has the meaning of time, so it can also be applied to partial differential equations of elliptic type which describe e.g. equilibrium situations in space. The following exposition is concerned, however, only with the case of ordinary differential equations.

The mathematical roots of the maximum principle lie in the calculus of variations, which has been under development for several hundred years. Its basic form has been developed as part of the early history of control theory during the period 1945-1960. Later, it has been extended to optimal control problems for many kinds of dynamical systems.

## 2. The Maximum Principle

### 2.8 Problem of Optimal Control

We consider the following problem of optimal control for a system of ordinary differential equations. We want to minimize the *cost functional*

$$J(x,u)=\int_0^T L(t,x(t),u(t))dt+L_f(x(T)) \tag{4}$$

where the *state function* $x:[0,T]\to\mathbb{R}^n$ solves the differential equation

$$\dot{x}(t)=f(t,x(t)),u(t)), \tag{5}$$

subject to the *initial condition*

$$x(0)=x_0\ . \tag{6}$$

The *control function* $u$: $[0,T]\to\mathbb{R}^m$ has to obey a control constraint

$$u(t)\in U_{ad}. \tag{7}$$

The functions $L$, $L_f$, $f$, the set $U_{ad}\subset\mathbb{R}^m$ of admissible values of the control function, the final time $T$ and the initial state $x_0$ are given. We want to determine an optimal control $u_*$ and the corresponding state $x_*$ such that

$$J(x_*,u_*)\le J(x,u) \tag{8}$$

for all pairs $(x, u)$ of state and control functions which satisfy (5)-(7).

## 2.9 Statement of the Maximum Principle

Let $(x_*, u_*)$ be a solution of the optimal control problem (4)-(7). The maximum principle asserts that the maximum condition

$$f(t,x_*(t),u_*(t))^T p(t)-L(t,x_*(t),u_*(t))=\max_{v\in U_{ad}}[f(t,x_*(t),v)^T p(t)-L(t,x_*(t),v)] \tag{9}$$

holds; here, $p:[0,T]\to\mathbb{R}^n$ is the *adjoint function,* which solves the *adjoint equation*

$$\dot{p}(t)=-\partial_x f(t,x_*(t),u_*(t))^T p(t)+\partial_x L(t,x_*(t),u_*(t))\ , \tag{10}$$

with the boundary condition

$$p(T)=-\partial_x L_f(x_*(T)). \tag{11}$$

In addition, for the function

$$h(t)=f(t,x_*(t),u_*(t))^T p(t)-L(t,x_*(t),u_*(t)) \tag{12}$$

it holds that

$$h'(t) = \partial_t f(t, x_*(t), u_*(t))^T p(t) - \partial_t L(t, x_*(t), u_*(t)). \tag{13}$$

In this manner, the original problem is reduced to a boundary value problem coupled to the finite-dimensional optimization problem

$$\text{Maximize } p^T f(t, x, u) - L(t, x, u)\,, \text{ subject to } u \in U_{ad}\,, \tag{14}$$

where $t \in \mathbb{R}$, $x, p \in \mathbb{R}^n$ are fixed, and an optimal vector $u \in \mathbb{R}^m$ has to be found. In many cases, problem (14) is very simple and can be solved explicitly. We can substitute its solution

$$u = u(t, x, p) \tag{15}$$

into the differential equations. We then have to determine the optimal state function $x_* = x_*(t)$ and the corresponding adjoint function $p = p(t)$ as the solution of the boundary value problem

$$\dot{x} = f(t, x, u(t, x, p))\,, \qquad x(0) = x_0, \tag{16}$$

$$\dot{p} = \partial_x f(t, x, u(t, x, p))^T p - \partial_x L(t, x, u(t, x, p))\,, \quad p(T) = -\partial_x L_f(x(T)). \tag{17}$$

Once we have accomplished that, we get the optimal control

$$u_*(t) = u(t, x_*(t), p(t)) \tag{18}$$

by substituting the solution of (16), (17) into (15).

If the approach outlined above works, we have effectively reduced the optimal control problem to the two-point boundary value problem (16), (17). That problem is solved by appropriate numerical methods, see below.

## 2.10 Other Boundary Conditions

Often, the state function has to satisfy a boundary condition at $t = T$,

$$\psi_f(x(T)) = 0\,, \tag{19}$$

where $\psi_f : \mathbb{R}^n \to \mathbb{R}^k$ is a prescribed function. In that case, the boundary condition (11) for the adjoint function changes to

$$p(T) = -\partial_x L_f(x_*(T)) + \partial_x \psi_f(x_*(T))^T \alpha\,. \tag{20}$$

Here $\alpha \in \mathbb{R}^k$ is an additional unknown vector that makes up for the additional $k$ scalar conditions to be satisfied in (19). In the extreme case, which occurs if we want to steer the state to a particular value $x_T \in \mathbb{R}^n$,

$$x(T) = x_T \,, \tag{21}$$

the value of $p(T)$ is completely unrestricted. A more general situation arises with a *multi-point boundary condition* like

$$\psi\,(x(0), x(t_1), \ldots,\; x(t_j), x(T\,)) = 0 \,, \tag{22}$$

which interconnects the values of the state variable for $t = 0, t_1, \ldots, t_j, T$. The adjoint will then in general have jumps at the interior points $t_i$ with a structure corresponding to (20). For example, if a single condition

$$\psi_i\,(x(t_i)) = 0 \,, \qquad \psi_i : \mathbb{R}^n \to \mathbb{R} \,, \tag{23}$$

is prescribed, the corresponding jump condition becomes

$$p(t_i^+) - p(t_i^-) = \alpha \partial_x \psi_i (x_*(t_i)) \tag{24}$$

with a single degree of freedom $\alpha \in \mathbb{R}$. Another particular case of (23) is the periodicity boundary condition

$$x(0) = x(T) \,, \tag{25}$$

which occurs in *optimal periodic control*.

In all these cases, however, the maximum principle may have to be modified if degeneracy occurs, as outlined in the next section.

-
-
-



**Bibliography**

Berkovitz L.D. (1974). *Optimal Control Theory*, 304 pp. New York: Springer. [Treats existence theory and the maximum principle for optimal control problems for ordinary differential equations]

Carlson D.A. and Haurie A. (1987). *Infinite Horizon Optimal Control*, Lecture Notes in Economics and Mathematical Sciences, 254 pp. Berlin: Springer-Verlag. [Treats optimal control problems where the final time T is infinite, with applications in economics]

Clark C.W. (1990). *Mathematical Bioeconomics*: *The Optimal Management of Renewable Resources*, 386 pp. New York: Wiley. [Treats the maximum principle and its applications in biology, environmental science and economics]

Clark C.W. (1985). *Bioeconomic Modeling and Fisheries Management*. New York: Wiley. [Treats optimal resource management in fisheries models]

Feichtinger G. and Hartl R.F. (1986). *Optimal Control of Economic Processes*. Berlin: de Gruyter. In German. [Treats applications of the maximum principle in the economic sciences]

Fattorini H.O. (1999). *Infinite-Dimensional Optimization and Control Theory*, 798 pp. Encyclopedia of Mathematics and its Applications, 62. Cambridge: Cambridge University Press. [Treats the theory of optimal control with emphasis on optimality conditions, partial differential equations and relaxed solutions]

Fleming W.H. and Rishel R.W. (1975). *Deterministic and Stochastic Optimal Control*, 222 pp. Berlin: Springer-Verlag. [Classical introductory textbook, includes also the stochastic maximum principle]

Sontag E.D. (1990). *Mathematical Control Theory*: *Deterministic Finite Dimensional Systems*, 396 pp. New York: Springer-Verlag. [Standard textbook on all fundamental aspects of control theory for ordinary differential equations]

Stoer J. and Bulirsch R. (1993). *Introduction to Numerical Analysis*, 660 pp. New York: Springer-Verlag. [Includes a discussion of the multiple shooting method]

Zabczyk J. (1992). *Mathematical Control Theory: an Introduction*, 260 pp. Boston: Birkhäuser. [Introductory textbook for control theory with emphasis on ordinary differential equations]

**Biographical Sketch**

**Martin Brokate** is a Full Professor in the Faculty of Mathematics of the Technical University of Munich, Germany. He received his diploma and his doctoral degree in Mathematics from the Free University of Berlin, and his habilitation in Mathematics from the University of Augsburg. He has done extensive research in mathematical analysis and optimization, particularly in the area of optimal control theory and of the mathematical modeling of hysteresis phenomena. He has published more than forty papers. He is author of the book Optimal Control of OrdinaryDifferential Equations with Nonlinearities of Hysteresis Type and, together with J. Sprekels, author of the book Hysteresis and Phase Transitions.