Course on Optimal Control
SOD311 Ensta Paris Tech
and Optimization master, U. Paris-Saclay
Part I: the Pontryagin approach
Version of Aug. 21, 2019

J. Frédéric Bonnans

Centre de Mathématiques Appliquées, Inria, Ecole Polytechnique, 91128 Palaiseau, France
(Frederic.Bonnans@inria.fr).

\footnote{Updates and additional material on the page
http://www.cmap.polytechnique.fr/~bonnans/notes/oc/oc.html}
Preface

These are lecture notes for the Optimal control course for students of both Ensta Paris Tech and the Optimization master, U. Paris-Saclay. We consider optimization problems for control systems modeled with ordinary differential equations. These lecture notes deal only with the Pontryagin approach, in which we mainly discuss necessary conditions for a trajectory to be optimal.

The notes are organized as follows. We give several examples of optimal control problems in chapter 1. Chapter 2 recalls general results in convex analysis and differential calculus (with a special focus on the notion of reduction) and obtains first order optimality conditions.

Chapter 3 states Pontryagin’s principle for a problem with control and initial-final constraints, as well as its extension for problems with decision variables (not depending on time), and variable horizon. It presents an application to minimal time problems, including geodesics in a Riemaniann setting, and a discussion of qualification conditions. The proof of Pontryagin’s principle is based on Ekeland’s principle.

Chapter 4 is devoted to the study of the shooting algorithm, that sometimes allows to reduce an optimal control problem to the computation of a zero of a finite dimensional shooting function. This assumes that Pontryagin’s principle allows to express the control as a function of the state and costate. We analyze problems that are unconstrained or with initial-final constraints. It is shown how the local well-posedness of the shooting function allows to obtain error estimates for discretized problems. Finally we show that this well-posedness can be characterized by some second order optimality conditions.

Chapter 5 deals with problems with running state constraints (constraints for each time). We state and prove a version of Pontryagin’s principle for this class. Junction conditions are analyzed, based on the notion of order of the constraints. As an illustration we discuss problems of elastic string and beam with obstacle.

Some useful references related to this course are: [15] for optimization theory (convex analysis, duality, Lagrange multipliers), [14] for an introduction to automatic control and optimal control, and [5], [6] for algorithmic aspects. Concerning applications: in spatial mechanics and quantum mechanics in the second case) [16], [17], in medicine, part IV of [27], in ecology [24], in chemistry [6], and on bioprocesses, e.g. [4]. We also mention the more advanced book [38], and [2] on the shooting approach.
## Chapter 1. Motivating examples

1. Simple linear quadratic problems
   1.1. Quadratic regulator
   1.2. Fuller’s problem

2. Population dynamics
   2.1. Fish harvesting
   2.2. Cell population control
   2.3. Predator-Prey

3. Bioreactors
   3.1. Chemostat
   3.2. Bio-gaz production

4. Flight mechanics
   4.1. The Goddard problem
   4.2. Space trajectories

## Chapter 2. First steps

1. Calculus in $\mathbb{R}^n$ and in Banach spaces
   1.1. Some notations
   1.2. Normal cones and separation of convex sets
      1.2.1. Preliminary
      1.2.2. Minimization over a convex set
      1.2.3. Normal cones
      1.2.4. Separation of sets
   1.3. Taylor expansions

2. Lagrange multipliers
   2.1. Implicit function theorem
   2.2. On Newton’s method
   2.3. Reduced equation
   2.4. Reduced cost
      2.4.1. First order analysis
      2.4.2. Computation of the derivative of the reduced cost
      2.4.3. Reduction Lagrangian
      2.4.4. Second order analysis
   2.5. Equality constrained optimization
      2.5.1. First order optimality conditions
      2.5.2. Second order optimality conditions
      2.5.3. Link with the reduction approach
   2.6. Calculus over $L^\infty$ spaces

3. Optimal control setting
   3.1. Weak derivatives
   3.2. Controlled dynamical system and associated cost
   3.3. Derivative of the cost function
   3.4. Optimality conditions with control constraints

4. Examples
4.1. Multiple integrators
4.1.1. Example: double integrator, quadratic energy
4.2. Control constraints

Chapter 3. Pontryagin’s principle
1. Pontryagin’s minimum principle
1.1. The easy Pontryagin’s minimum principle
1.1.1. Main result
1.1.2. A related nonlinear expansion of the state
1.2. Pontryagin’s principle with two point constraints
1.2.1. Main result
1.2.2. Specific structures
1.2.3. Link with Lagrange multipliers; convex problems
1.3. Variations of the pre-Hamiltonian
1.3.1. Pontryagin extremal over (0, T)
1.3.2. Constant pre-Hamiltonian for autonomous problem
1.3.3. Variation of the pre-Hamiltonian for non autonomous problem
2. Extensions
2.1. Decision variables
2.1.1. A design problem
2.2. Variable horizon
2.3. Varying initial and final time
2.4. Interior point constraints
2.4.1. A general setting
2.4.2. Variable durations
2.4.3. Variable dynamics structure
2.5. Minimal time problems, geodesics
2.5.1. General results
2.5.2. Linear dynamics
2.5.3. Geometric optics and Riemaniann geometry
3. Qualification conditions
3.1. Lagrange multipliers
3.1.1. General framework
3.1.2. Specific structures
3.2. Pontryagin multipliers
4. Proof of Pontryagin’s principle
4.1. Ekeland’s principle
4.2. Ekeland’s metric and the penalized problem
4.3. Proof of theorem 3.13

Chapter 4. Shooting algorithm, second order conditions
1. Shooting algorithm
1.1. Unconstrained problems
1.2. Problems with initial-final state constraints
1.3. Time discretization
2. Second-order optimality conditions
2.1. Expansion of the reduced cost
2.2. More on the Hessian of the reduced cost
2.3. Case of initial-final equality constraints
2.4. Links with the shooting algorithm
3. Control constraints
4. Notes
Motivating examples

Contents

1. Simple linear quadratic problems 9
   1.1. Quadratic regulator 9
   1.2. Fuller’s problem 10

2. Population dynamics 10
   2.1. Fish harvesting 10
   2.2. Cell population control 11
   2.3. Predator-Prey 12

3. Bioreactors 12
   3.1. Chemostat 12
   3.2. Bio-gaz production 13

4. Flight mechanics 13
   4.1. The Goddard problem 13
   4.2. Space trajectories 13

We present in this chapter some typical examples of optimal control problems arising in the computation of regulators, for population dynamics problem in ecology or biology, bioprocesses, and flight mechanics. The examples were solved numerically using the software bocop.org; some of them are taken from the collection [8].

1. Simple linear quadratic problems

In this section we present two academic examples, who may seem close one to each other. It appears, however, that the solutions have completely different behavior.

1.1. Quadratic regulator. Consider the problem:

\[
\begin{align*}
\min \quad & \frac{1}{2} \int_0^T (x_1^2(t) + x_2^2(t))\,dt; \\
\text{s.t.} \quad & \dot{x}_1(t) = x_2(t); \quad \dot{x}_2(t) = u(t), \quad t \in (0, T); \quad x_1(0) = 0, \quad x_2(0) = 1,
\end{align*}
\]

for \( t \in [0, T] \), \( T = 5 \), subject to the control and state bound constraint:

\[
-1 \leq u(t) \leq 1, \quad x_2(t) \geq -0.2, \quad t \in (0, T).
\]
We write the problem in the *Mayer form* (i.e., with a final cost only), defining an extra state variable with dynamics

\begin{equation}
\dot{x}_3(t) = \frac{1}{2}(x_1^2(t) + x_2^2(t)), \quad x_3(0) = 0.
\end{equation}

In this new setting the cost function can be expressed as $x_3(T)$. The optimal trajectory, displayed in figure 1, appears to have three arcs (connected by two junction points): lower control bound, active state constraint, and a final unconstrained arc; the control is discontinuous at the junction points.

1.2. Fuller’s problem. Here is a variant of the previous problem, due to [23]:

\begin{equation}
\min \int_0^T x^2(t) \, dt; \quad \ddot{x}(t) = u(t) \in [-10^{-2}, 10^{-2}]; \quad x(0) = x_1; \quad \dot{x}(0) = x_2 \text{ given.}
\end{equation}

For large enough $T$, the solution is as follows. There exists $\tau \in (0, T)$ such that: (i) over $(0, \tau)$, the control has values $\pm 1$, the arc lengths converging geometrically to 0, and (ii) over $(\tau, T)$, $x(t) = u(t) = 0$. These switches are not easy to reproduce numerically. Those in figure 2 were obtained with 1000 time steps, $T = 3.5$, $x(0) = 0, \dot{x}(0) = 1, x(T) = \dot{x}(T) = 0$.

2. Population dynamics

We next discuss examples of population dynamics problems, that are widely used in ecology and biology.

2.1. Fish harvesting. The state equation is

\begin{equation}
\dot{x}(t) = B(x(t))x(t) - au(t)x(t)
\end{equation}

Here $x(t)$ is the size of the fish population, $B(\cdot)$ is the birth rate, $a > 0$ is the harvesting efficiency, $u(t)$ is the harvesting effort; the cost function is

\begin{equation}
\int_0^T (c(u(t)) - u(t)x(t)) \, dt
\end{equation}
where \( c(u) \) is the harvesting cost, with constraint \( 0 \leq u(t) \leq 1 \) and \( x(T) \geq x_T \).

Classical birth rate models are the exponential, logistics and Gomperz ones, with positive parameters \( a, b \):

\[
(1.7) \quad B_E(x) := ax; \quad B_L(x) := a(1 - x/b); \quad B_G(x) := a \log(b/x).
\]

In the uncontrolled case, that is, when \( u(t) = 0 \) for all \( t \), for the logistics and Gomperz models, \( x(t) \to b \) in a monotone way, with an exponential speed of convergence: for some \( C > 0 \) and \( d > 0 \) depending on the initial state \( x(0) \):

\[
(1.8) \quad |x(t) - b| \leq Ce^{-dt}.
\]

One can consider the optimization of a steady-state solution: compute \( u \) solution of

\[
(1.9) \quad \text{Min}_u c(u) - ax; \quad B(x)x = aux; \quad u \in [0,u_M].
\]

About this and related optimal control problems, see [24, Ch. 4].

2.2. Cell population control. Similar state equation

\[
(1.10) \quad \dot{x}(t) = B(x(t))x(t) - aw(t)x(t)
\]

are used for modelling cell populations. Here \( w(t) \geq 0 \) is the drug concentration, subject to the upper bound \( w(t) \leq w_M \), and possibly satisfying the dynamics

\[
(1.11) \quad \dot{w}(t) = -bw(t) + v(t),
\]

with \( b > 0 \) elimination rate and \( v(t) \) drug injection rate. One calls (1.10) the pharmacodynamics equation (modelling the action of the drug) and (1.11) the pharmacokinetics equation (which models the dynamics of the drug itself). We want to minimize \( x(T) \), with possibly constraints on time integrals of \( w(t) \):

\[
(1.12) \quad \int_{\alpha_i}^{\beta_i} w(t)dt \leq \gamma_i,
\]

with \( 0 \leq \alpha_i < \beta_i \leq T \) and \( \gamma_i > 0 \), for \( i = 1 \) to \( N \). Extensions of such models are applied to the problem of optimizing cancer treatments [39].
2.3. Predator-Prey. A simple extension to the controlled setting of the Lokta-Volterra system is, see [24, Ch. 5] or [25]:

\[
\begin{align*}
\dot{x} &= \alpha x - \beta xy - a_1 u, \\
\dot{y} &= \delta xy - \gamma y - a_2 u,
\end{align*}
\]

with positive parameters \(\alpha, \beta, \delta, a_1, a_2\). The uncontrolled dynamics (with \(u(t) = 0\) for all \(t\)) has a unique nonzero equilibrium point

\[
\hat{x} = \frac{\gamma}{\delta}; \quad \hat{y} = \frac{\alpha}{\beta}.
\]

The uncontrolled trajectories are a closed loop around this point. The objective is to reach \((\hat{x}, \hat{y})\) by minimizing the sum of final time and \(c\) times the integral of the control. For the numerical illustration the constants are \(\alpha = \beta = \gamma = \delta = 1; b_1 = 0.9, b_2 = 1, c = 1;\) see figure 3.

3. Bioreactors

3.1. Chemostat. The simplest bioreactor model is the chemostat one:

\[
\begin{align*}
\dot{x}(t) &= (\mu(s(t)) - u(t)/v)x(t), \\
\dot{s}(t) &= -\frac{1}{\gamma}\mu(s(t))x(t) + \frac{u(t)}{v}(s_{in} - s(t)).
\end{align*}
\]

The state variables are the biomass concentration \(x\) and substrate \(s\). We can take the growth function \(\mu\) of Monod type: \(\mu(s) = \mu_{max}s/(k+s)\), with \(\mu_{max} > 0\) and \(k > 0\). The yield coefficient is \(\gamma > 0\). The control is the flow \(u \in [0, u_M]\). The volume \(v > 0\) is either constant, of for batch processes a function of time, subject to the dynamics

\[
\dot{v}(t) = u(t)
\]

The problem is usually to reach a given state in minimal time. See the analysis of the solution in [33].
3.2. Bio-gaz production. The state equation is

\[
\begin{align*}
\dot{y}(t) &= \nu(t)y(t)/(1 + y(t)) - (r + u(t))y(t), \\
\dot{x}(t) &= (\mu(s(t)) - \beta u(t))x(t), \\
\dot{s}(t) &= \beta u(\delta y(t) - s(t)) - \mu(s(t))x(t).
\end{align*}
\]

(1.17)

Here \(y\) is the mass of algae, \(x\) is the biomass concentration and \(s\) is the substrate; \(\beta > 0\) is the volume ratio between the algae tanks and the bioreactor. To function to be maximized is

\[
\int_0^T \mu(s(t))x(t)\,dt.
\]

(1.18)

Given a light model \(\nu(t)\) that is periodic with period of a day, the solution tends to have a periodic behavior when optimizing with a large horizon; see [3].

4. Flight mechanics

4.1. The Goddard problem. This is the problem of a vertical ascent of a rocket, see [40], State variables: position \(h\), speed \(v\), mass \(m\). Control variable \(u\): normalized thrust (maximal thrust denoted by \(T_{\text{max}}\)) Cost: opposite of final mass (minimize fuel consumption) \(T\): free final time, \(D(h, v)\) is the drag function. The state equation is

\[
\begin{align*}
\dot{h} &= v, \\
\dot{v} &= -1/h^2 + 1/m(T_{\text{max}} u - D(h, v)), \\
\dot{m} &= -b T_{\text{max}} u, \\
0 &\leq u \leq 1, \quad D(h, v) \leq D_{\text{max}}, \\
\end{align*}
\]

(1.19)

We take the simplest model of drag function: \(c\rho(h)v^2\), with \(c > 0\) the aerodynamic coefficient, and \(\rho\) volumic mass expressed as function of altitude, in the exponential volumic mass model: \(\rho(h) = \alpha e^{-\beta h}\) for some positive \(\alpha, \beta\). The optimal solution has four arcs, see figure 4. For a three dimensional extension of this problem, see [13].

4.2. Space trajectories. There is a large literature on the subject; let us just mention the case of space launchers [30], and atmosphere reentry [11].
Figure 4. Four arcs: full thrust, maximum drag, unconstrained arc, zero thrust. Discontinuity of control at all junction points.
### Contents

<table>
<thead>
<tr>
<th></th>
<th>Calculus in $\mathbb{R}^n$ and in Banach spaces</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Some notations</td>
<td>15</td>
</tr>
<tr>
<td>1.2</td>
<td>Normal cones and separation of convex sets</td>
<td>17</td>
</tr>
<tr>
<td>1.3</td>
<td>Taylor expansions</td>
<td>18</td>
</tr>
<tr>
<td>2.1</td>
<td>Lagrange multipliers</td>
<td>20</td>
</tr>
<tr>
<td>2.2</td>
<td>Implicit function theorem</td>
<td>20</td>
</tr>
<tr>
<td>2.3</td>
<td>On Newton’s method</td>
<td>21</td>
</tr>
<tr>
<td>2.4</td>
<td>Reduced equation</td>
<td>21</td>
</tr>
<tr>
<td>2.5</td>
<td>Reduced cost</td>
<td>22</td>
</tr>
<tr>
<td>2.6</td>
<td>Equality constrained optimization</td>
<td>24</td>
</tr>
<tr>
<td>3.1</td>
<td>Calculus over $L^\infty$ spaces</td>
<td>26</td>
</tr>
<tr>
<td>3.2</td>
<td>Optimal control setting</td>
<td>27</td>
</tr>
<tr>
<td>3.3</td>
<td>Weak derivatives</td>
<td>27</td>
</tr>
<tr>
<td>3.4</td>
<td>Controlled dynamical system and associated cost</td>
<td>28</td>
</tr>
<tr>
<td>3.5</td>
<td>Derivative of the cost function</td>
<td>30</td>
</tr>
<tr>
<td>4.1</td>
<td>Optimality conditions with control constraints</td>
<td>31</td>
</tr>
<tr>
<td>4.2</td>
<td>Examples</td>
<td>33</td>
</tr>
<tr>
<td>4.3</td>
<td>Multiple integrators</td>
<td>33</td>
</tr>
<tr>
<td>4.4</td>
<td>Control constraints</td>
<td>33</td>
</tr>
</tbody>
</table>

1. Calculus in $\mathbb{R}^n$ and in Banach spaces

1.1. Some notations. We denote by $\mathbb{R}^n$ the Euclidean space of dimension $n$, whose elements are identified with vertical vectors, with norm $|x| := (\sum_{i=1}^n x_i^2)^{1/2}$ and scalar product $x \cdot y := \sum_{i=1}^n x_i y_i$, for all $x, y$ in $\mathbb{R}^n$. The dual space $\mathbb{R}^n^*$ is identified with the set of $n$ dimensional horizontal vectors. Other useful norms in $\mathbb{R}^n$ are the $\ell^s$ norm $|x|_s := (\sum_{i=1}^n |x_i|^s)^{1/s}$, for $s \in [1, \infty]$, and the uniform norm $|x|_\infty := \max\{|x_i|, 1 \leq i \leq n\}$. Note that norms in $\mathbb{R}^n$ are denoted with a single bar. Let $A, B$ be $n$ dimensional symmetric matrices. We write $A \succeq B$ if $A - B$ is semi positive definite, and $A \succ B$ if $A - B$ is positive definite.

Let $X$ be a Banach space (a normed vector space that is complete, i.e., every Cauchy sequence has a limit), with norm denoted by $\|x\|_X$ or $\|x\|$ if there is no ambiguity. If $Y$ is another Banach space, we denote by $L(X,Y)$ the set of linear continuous mappings $X \to Y$. Endowed with the
norm
\[ \|A\| := \sup\{\|Ax\|; \|x\|_X \leq 1\}, \]

\(L(X,Y)\) is a Banach space. Note that a linear mapping \(A : X \to Y\) is continuous iff the above r.h.s. is finite.

**Exercice 2.1.** Prove that, similarly, if \(Z\) is a third Banach space and \(a : X \times Y \to Z\) is bilinear, then it is continuous iff
\[ \|a\| := \sup\{\|a(x,y)\|_{Z}; \|x\|_X \leq 1, \|y\|_Y \leq 1\}, \]
is finite, and that the set of continuous bilinear forms endowed with this norm is a Banach space.

The set of *linear continuous forms* over \(X\) (linear and continuous applications \(X \to \mathbb{R}\)) is denoted by \(X^*\), and the action of \(x^* \in X^*\) over \(x \in X\) is denoted by \(\langle x^*,x \rangle_X\). The space \(X^*\) is a Banach space, endowed with the *dual norm*
\[ \|x^*\|_{X^*} := \sup\{\langle x^*,x \rangle_X; \|x\|_X \leq 1\}. \]
If \(A \in L(X,Y)\) with \(X\) and \(Y\) Banach spaces, we denote by \(A^\dagger \in L(Y^*,X^*)\) the *transpose operator* defined, for all \(y^* \in Y^*\), by
\[ \langle A^\dagger y^*,x \rangle_X = \langle y^*,Ax \rangle_Y, \quad \text{for all } x \in X. \]
That \(A^\dagger\) is continuous follows from the relations
\[ \|A^\dagger y^*\|_{X^*} = \sup_{\|x\| \leq 1} \langle A^\dagger y^*,x \rangle_X \leq \|A\|\|y^*\|_{Y^*}. \]
We say that the Banach space \(X\) is a *Hilbert space* if there exists a symmetric continuous bilinear form \(a(\cdot,\cdot)\) over \(X\) such that
\[ \|x\|^2 = a(x,x), \quad \text{for all } x \in X. \]
A Hilbert space \(H\) is endowed with the *scalar product*
\[ \langle x,x' \rangle_X := a(x,x'), \quad \text{for all } x, x' \text{ in } X. \]
By the *Fréchet-Riesz representation theorem*, if \(X\) is a Hilbert space, we have that
\[ \text{Given } x^* \in X^*, \text{ there exists } x' \in X \text{ such that } \langle x^*,x \rangle_X = \langle x',x \rangle_X, \quad \text{for all } x \in X. \]

**Exercice 2.2.** Check that \(x'\) is uniquely defined, and prove the Fréchet-Riesz representation theorem by showing that we can take for \(x'\) the unique solution of the optimization problem below:
\[ \text{Min}_{x \in X} \frac{1}{2}\|x\|_X^2 - \langle x^*,x \rangle_X. \]

In the sequel \(X\) is a Banach space. If \(A\) and \(B\) are subsets of \(X\), their *Minkowski sum* and *difference* are
\[ \begin{cases} A + B := \{a + b; \ a \in A, \ b \in B\}; \\ A - B := \{a - b; \ a \in A, \ b \in B\}. \end{cases} \]
If \(E \subset \mathbb{R}\) we define the product
\[ EA := \{ea; \ e \in E, a \in A\}. \]
If \(f : X \to Y, A \subset X\) and \(B \subset Y\), we set
\[ \begin{cases} f(A) := \{f(a); \ a \in A\}; \\ f^{-1}(B) := \{x \in \mathbb{R}^n; \ f(x) \in B\}. \end{cases} \]
The *closure* of \(A \subset X\) is the intersection of closed sets containing \(A\), and is denoted by \(\text{cl}(A)\). The *interior* and *boundary* of \(A\) are denoted by
\[ \begin{cases} \text{int}(A) := \{x \in X; \ y \in A \text{ if } y \text{ is close enough to } x\}; \\ \partial A := \text{cl}(A) \setminus \text{int}(A). \end{cases} \]
The segment \([x, y]\), where \(x\) and \(y\) belong to \(X\) is
\[
(2.14) \quad [x, y] := \{\alpha x + (1 - \alpha)y; \ \alpha \in [0, 1]\}.
\]
We say that \(A \subset X\) is convex if
\[
(2.15) \quad [x, y] \subset A, \quad \text{for all } x, y \in A.
\]
By \(\{0\}_X\) we denote the set having for unique element the zero of \(X\). We say that a subset \(C\) of a vector space \(Y\) is a cone if \(\alpha y \in C\), for all \(\alpha > 0\) and \(y \in C\). The positive (resp. negative) cone of \(\mathbb{R}^n\) is the set of element of \(\mathbb{R}^n\) whose all coordinates are nonnegative (resp. nonpositive), and is denoted by \(\mathbb{R}^n_+\) (resp. \(\mathbb{R}^n_-\)). More generally, if \(X\) is a space of real valued functions (perhaps defined only a.e.), we call positive (resp. negative) cone of \(X\) and denote by \(X_+\) (resp. \(X_-\)) the set of nonnegative (resp. nonpositive) (perhaps only a.e.) functions of \(X\). We then denote by \(X^*_+\) the dual positive cone defined by
\[
(2.16) \quad X^*_+ := \{x^* \in X^*; \ (x^*, x) \geq 0, \ \text{for all } x \in X_+\}.
\]

1.2. Normal cones and separation of convex sets.

1.2.1. Preliminary. Again in the sequel \(X\) is a Banach space. The orthogonal space to \(E \subset X\) is the closed vector space of \(X^*\) defined by
\[
(2.17) \quad E^\perp := \{x^* \in X^*; \ (x^*, x) = 0, \ \text{for all } x \in E\}.
\]

Remark 2.3. When \(X\) is a Hilbert space and \(E \subset X\) we have another notion of primal orthogonal space which is (we remind that \((\cdot, \cdot)_X\) denotes the scalar product in \(X\)):
\[
(2.18) \quad \{x' \in X; \ (x', x)_X = 0, \ \text{for all } x \in E\}.
\]

There should be no confusion in our applications.

1.2.2. Minimization over a convex set. Let \(K\) be a convex subset of a Banach space \(X\), and let \(f : X \to \mathbb{R}\) be differentiable. We say that \(f\) attain a local minimum over \(K\) at \(\bar{x} \in K\) if \(f(\bar{x}) \leq f(x)\), for all \(x \in K\) sufficiently close to \(\bar{x}\).

Lemma 2.4. Let \(f\) attain a local minimum over \(K\) at \(\bar{x}\). Then
\[
(2.19) \quad Df(\bar{x})(x - \bar{x}) \geq 0, \quad \text{for all } x \in K.
\]
Conversely, if \(f\) is convex, \((2.19)\) implies that \(f\) attains its (global) minimum over \(K\) at \(\bar{x}\).

Proof. Let \(x \in K\). For \(t \in [0, 1]\), \(x_t := (1 - t)\bar{x} + tx\) belongs to \(K\), and \(x_t \to \bar{x}\) when \(t \downarrow 0\), so that \(f(\bar{x}) \leq f(x_t)\), for small enough \(t\). Therefore the conclusion follows from
\[
(2.20) \quad 0 \leq \lim_{t \downarrow 0} \frac{1}{t}(f(x_t) - f(\bar{x})) = Df(\bar{x})(x - \bar{x}).
\]
Conversely, if \(f\) is convex, and \(x \in K\), \(f(x) - f(\bar{x}) \geq Df(\bar{x})(x - \bar{x})\), so that \((2.19)\) implies \(f(x) - f(\bar{x}) \geq 0\).

1.2.3. Normal cones.

Definition 2.5. Let \(K\) be a closed convex subset of \(X\) and \(\bar{x} \in K\). The normal cone to \(K\) at \(\bar{x}\) is the set
\[
(2.21) \quad N_K(\bar{x}) := \{x^* \in X^*; \ (x^*, x' - \bar{x}) \leq 0, \ \text{for all } x' \in K\}.
\]
Its elements are called normal directions. This is a closed and convex cone. When \(X = \mathbb{R}^n\) we identify \(X\) with its dual and see normal directions as elements of \(\mathbb{R}^n\).

Exercise 2.6. Check that (i) If \(K\) is a singleton, the normal cone is \(X^*\), (ii) If \(x \in \text{int}(K)\), the normal cone is the singleton \(\{0\}\). (iii) If \(K\) is a vector subspace, the normal cone coincides with the orthogonal space. (iv) If \(K\) has a smooth boundary, the normal cone at \(x \in K\) is of the form \(\mathbb{R}_+y\), where \(y\) is the outward normal at \(x\) to \(K\).
EXERCISE 2.7 (Product form). Let $X = X' \times X''$, where $X'$ and $X''$ are Banach spaces, and $K = K' \times K''$, with $K' \subset X'$ and $K'' \subset X''$ closed and convex. Let $\bar{x} = (\bar{x}', \bar{x}'') \in K$ with $\bar{x}' \in K'$ and $\bar{x}'' \in K''$. Check that
\begin{equation}
N_K(\bar{x}) = N_{K'}(\bar{x}') \times N_{K''}(\bar{x}'').
\end{equation}

EXERCISE 2.8. Let $K = \{0\} \times \mathbb{R}^m$. Check that, for $\bar{x} \in K$, we have
\begin{equation}
N_K(\bar{x}) = \mathbb{R}^n \times \{ \lambda \in \mathbb{R}^m ; \; \lambda_i x_{n+i} = 0, \; i = 1, \ldots, m \}.
\end{equation}

DEFINITION 2.9. Let $E \subset X$. Its negative polar cone is
\begin{equation}
E^- := \{ x^* \in X^* ; \; \langle x^*, x \rangle \leq 0, \; \text{for all } x \in E \}.
\end{equation}
Its positive polar cone is $E^+ := -E^-$. By 'polar cone' we mean the negative one.

We see that the polar cone is a closed convex cone, intersection of closed half spaces.

**Lemma 2.10.** Let $C \subset X$ be a closed convex cone, and $x \in C$. Then
\begin{equation}
N_C(x) = C^- \cap x^\perp.
\end{equation}

**Proof.** Let $x^* \in C^- \cap x^\perp$ and $y \in C$. Then $\langle x^*, y \rangle \leq 0$ and $\langle x^*, x \rangle = 0$, so that $\langle x^*, y-x \rangle \leq 0$, proving that $x^* \in N_C(x)$.

Conversely, let $x^* \in N_C(x)$. For $y \in C$, since $C$ is a convex cone, $z := x + y = 2(1/2 x + 1/2 y) \in C$, and therefore $0 \geq \langle x^*, z-x \rangle = \langle x^*, y \rangle$, proving that $x^* \in C^-$. Since $0 \in C$ we also have $0 \geq \langle x^*, 0-x \rangle = -\langle x^*, x \rangle$, and the converse inequality holds since $x^* \in C^-$. The conclusion follows. \qed

EXERCISE 2.11. Let $K := \mathbb{R}^n$. Deduce from the previous lemma that if $x \in K$, then $\lambda \in N_K(x)$ if $\lambda \in \mathbb{R}^n_+$ and $\lambda_i x_i = 0$, $i = 1$ to $n$.

1.2.4. Separation of sets.

DEFINITION 2.12. Let $A$ and $B$ be two convex subsets of a Banach space $X$. Let $x^* \in X^*$, $x^* \neq 0$. We say that $x^*$ separates $A$ and $B$ if
\begin{equation}
\langle x^*, a \rangle \leq \langle x^*, b \rangle, \quad \text{for all } a \in A \text{ and } b \in B.
\end{equation}

We start with the following geometric form of the Hahn-Banach theorem.

**Theorem 2.13.** Let $A$ and $B$ be two convex subsets of a Banach space $X$, with empty intersection. If $A-B$ has a nonempty interior, then there exists $x^*$ separating $A$ and $B$.

**Proof.** (a) The result is obtained in [18, Ch. 1, Thm 1.6], assuming that either $A$ or $B$ is open.

(b) Set $E := \text{int}(B-A)$. Since $A \cap B = \emptyset$, $E$ does not contain zero. By step (a), there exists $x^*$ separating 0 and $E$. Now let $a \in A$ and $b \in B$, and set $e := b-a$. Then $e$ is the limit of a sequence $e_k$ in $E$ (take $e_k := (1-1/k)e + (1/k)\bar{e}$, with $\bar{e} \in E$). Therefore, $\langle x^*, b-a \rangle = \lim_k \langle x^*, e_k \rangle \geq 0$, as was to be proved. \qed

**Remark 2.14.** Let $X$ be finite dimensional, and $A$, $B$ be convex subsets with empty intersection. Then there exists $x^*$ separating $A$ and $B$. See (the stronger result) [36, Thm 11.3].

1.3. Taylor expansions. Let $f : X \to Y$, where $X$ and $Y$ are Banach spaces. We call directional derivatives of $f$ at $\bar{x} \in X$ in direction $h \in X$ the following amount, if it exists:
\begin{equation}
f'(\bar{x}, h) := \lim_{t \downarrow 0} \frac{f(\bar{x} + th) - f(\bar{x})}{t}.
\end{equation}

We say that $f$ is differentiable, or Fréchet differentiable, at $x \in X$, if there exists a linear continuous mapping $X \to Y$, denoted by $Df(x)$ or $f'(x)$, and called derivative of $f$ at $x$, such that
\begin{equation}
\|f(x+h) - f(x) - f'(x)h\|_Y = o(\|h\|_X).
\end{equation}
In this case
\[(2.29)\quad f'(\bar{x}, h) = Df(\bar{x})h, \quad \text{for all } h \in X.\]

When \(X = \mathbb{R}^n\) and \(Y = \mathbb{R}^p\), we identify \(f'(x)\) with the usual \(p \times n\) Jacobian matrix:
\[(2.30)\quad f'_{ij}(x) = \frac{\partial f_i(x)}{\partial x_j}, \quad 1 \leq i \leq p; \quad 1 \leq j \leq n.\]

When \(Y = \mathbb{R}\), and \(X\) is a Hilbert space, we denote by \(\nabla f(x)\) and call gradient of \(f\) at \(x\) the element of \(X\) associated with \(f'(x)\) by the Fréchet-Riesz representation theorem, characterized by
\[(2.31)\quad f'(x)h = (\nabla f(x), h)_X, \quad \text{for all } h \in X.\]

When \(X = \mathbb{R}^n\), \(\nabla f(x)\) is the element of \(\mathbb{R}^n\) whose coordinates are equal to the partial derivatives of \(f\) at \(x\), and \(f'(x) = \nabla f(x)^\dagger\).

If \(Z\) is another Banach space and \(f : X \times Y \to Z\), we denote by e.g. \(D_x f(x, y)\) or \(f_x(y, x)\) its partial derivatives. If in addition \(Z = \mathbb{R}\), we denote by \(\nabla_x f(x, y)\) its partial gradient, and then, if \(X = \mathbb{R}^n\) and \(Y = \mathbb{R}^q\), \(f_{x,y}(x, y)\) is identified with the \(n \times q\) matrix with general term \(\partial f(x, y)/\partial x_i \partial y_j\).

We recall the Taylor expansion with integral term: if \(f : X \to Y\) is \((n+1)\) times continuously differentiable, with \(n \geq 0\), then
\[(2.32)\quad f(x + h) = f(x) + \cdots + \frac{1}{n!} D^n f(x)(h)^n + \int_0^1 \frac{(1-t)^n}{n!} D^{n+1} f(x + th)(h)^{n+1} dt.\]

For \(n = 0\) and \(1\), this gives
\[(2.33)\quad \begin{cases} f(x + h) = f(x) + \int_0^1 Df(x + th)h dt, \\ f(x + h) = f(x) + Df(x)h + \int_0^1 (1-t)D^2 f(x + th)(h)^2 dt. \end{cases}\]

The Taylor expansion (2.32) may be expressed as
\[(2.34)\quad f(x + h) = f(x) + \cdots + \frac{1}{(n+1)!} D^{n+1} f(x)(h)^{n+1} + r_n(x, h)\]

where the remainder satisfies
\[(2.35)\quad \begin{cases} (i) \quad r_n(x, h) = a(x, h)(h)^{n+1}; \\ (ii) \quad a(x, h) = \int_0^1 \frac{(1-t)^n}{n!} (D^{n+1} f(x + th) - D^{n+1} f(x)) dt. \end{cases}\]

The above relation uses the identity
\[(2.36)\quad \int_0^1 \frac{(1-t)^n}{n!} dt = \frac{1}{(n+1)!}\]

Since \(f : X \to Y\) is \((n+1)\) times continuously differentiable if follows that
\[(2.37)\quad \|r_n(x, h)\|_Y = o(\|h\|_X^{n+1}) \quad \text{when } h \to 0, \text{ for any given } x \in X.\]

Sometimes we need more precise estimates of the remainder, as the one below:

**Lemma 2.15.** Let \(f : X \to Y\) be \((n+1)\) times continuously differentiable, \(D^{n+1} f(x)\) being Lipschitz with constant \(L\). Then
\[(2.38)\quad \|r_n(x, h)\|_Y \leq \frac{L}{(n+2)!} \|h\|_X^{n+2}.\]
Proof. By (2.35), we have that
\[
(2.39) \quad \|r_n(x,h)\|_Y \leq L\|h\|_X^{n+2} \int_0^1 \frac{t(1-t)^n}{n!} dt.
\]
Integrating by parts, we see that the integral is equal to \(((n+1)!)^{-1} \int_0^1 (1-t)^{n+1} dt = 1/(n+2)!\)
The conclusion follows. \(\square\)

Some more accurate estimates of the remainder are based on the following notion.

**Definition 2.16.** (i) Let \(X\) and \(Y\) be Banach spaces, \(E \subset X\) and \(f : E \to Y\). The *modulus of continuity of \(f\), over \(E\), at \(\bar{x} \in E\) is the nondecreasing function \(\omega_{f,\bar{x}} : \mathbb{R}_+ \to \mathbb{R}_+ \cup \{+\infty\}\) defined by
\[
(2.40) \quad \omega_{f,\bar{x},E}(\varepsilon) := \sup\{\|f(x) - f(\bar{x})\| : \|x - \bar{x}\| \leq \varepsilon, x \in E\}.
\]
(ii) The *modulus of continuity of \(f\) over \(E\) is
\[
(2.41) \quad \omega_{f,E}(\varepsilon) := \sup\{\|f(x') - f(x)\| : \|x' - x\| \leq \varepsilon, x',x \in E\}.
\]

The moduli of continuity are nondecreasing functions with value 0 at 0, and the restriction of \(f\) to \(E\) is continuous at \(\bar{x}\) iff \(\omega_{f,\bar{x},E}(\varepsilon)\) is continuous at 0, i.e., if \(\omega_{f,\bar{x},E}(\varepsilon) \downarrow 0\) when \(\varepsilon \downarrow 0\).

**Definition 2.17.** We say that \(f\) is *uniformly continuous over \(E\)* if \(\omega_{f,E}(\varepsilon) \downarrow 0\) when \(\varepsilon \downarrow 0\).

We recall the link between continuity and uniform continuity.

**Lemma 2.18.** Let \(E\) be a compact subset (any open covering contains a finite covering) of a Banach space \(X\). Then a continuous function over \(E\) is uniformly continuous over \(E\).

**Proof.** We must prove that if \(x_k, x'_k\) in \(E\) satisfy \(\|x'_k - x_k\| \to 0\), then \(\|f(x'_k) - f(x_k)\| \to 0\). If this does not hold, extracting a subsequence we may assume that \(\liminf_k \|f(x'_k) - f(x_k)\| > 0\). Since any sequence in a compact metric space has a convergent subsequence, extracting if necessary a subsequence, we may assume that \((x'_k, x_k) \to (x',x)\). Since \(\|x'_k - x_k\| \to 0\) we must have \(x' = x\). A compact set being closed, \(x \in E\). From the continuity of \(f\) at \(x\), it follows that \(\|f(x'_k) - f(x_k)\| \to 0\). We have obtained a contradiction. The conclusion follows. \(\square\)

2. Lagrange multipliers

2.1. Implicit function theorem. Let \(U, Y\) and \(Z\) be Banach spaces, and \(\Psi : U \times Y \to Z\) be of class \(C^1\). Let \((\bar{x}, \bar{y}) \in U \times Y\) satisfy
\[
(2.42) \quad \Psi(\bar{u}, \bar{y}) = 0; \quad D\Psi(\bar{u}, \bar{y}) \text{ is invertible.}
\]
We recall that \(A \in L(U,Y)\) is said to be *invertible* if it is a bijection. It can then be proved that its inverse \(A^{-1}\) (obviously linear) is continuous (as a consequence of the celebrated open mapping theorem, see [18, Corollary 2.7]).

**Theorem 2.19.** [IFT] If (2.42) holds, there exist an open neighbourhood \(V\) of \(\bar{u}\) and a \(C^1\) mapping \(\varphi : V \to Y\) and \(\gamma > 0\) such that the equation
\[
(2.43) \quad \Psi(u,y) = 0; \quad u \in V; \quad \|y - \bar{y}\|_Y \leq \gamma
\]
holds iff \(y = \varphi(u)\).

Given \(u \in V\), we have that \(\Psi(u, \varphi(u)) = 0\). The derivative of this function is also equal to zero, i.e.
\[
(2.44) \quad \Psi_u(u, \varphi(u)) + \Psi_y(u, \varphi(u)) D\varphi(u) = 0.
\]
Taking \(u = \bar{u}\), we find that
\[
(2.45) \quad \Psi_u(\bar{u}, \bar{y}) + \Psi_y(\bar{u}, \bar{y}) D\varphi(\bar{u}) = 0.
\]
Since \(\Psi_y(\bar{u}, \bar{y})\) is invertible, we deduce that
\[
(2.46) \quad D\varphi(\bar{u}) = -\Psi_y(\bar{u}, \bar{y})^{-1}\Psi_u(\bar{u}, \bar{y}).
\]
From this expression, and since the inverse mapping \( A \rightarrow A^{-1} \) is of class \( C^\infty \) over the (open) set of invertible mappings \( Y \rightarrow Z \), we easily deduce that:

**Corollary 2.20.** Under the hypotheses of theorem 2.19, if \( \Psi \) is of class \( C^p \), \( p \geq 1 \), then \( \varphi \) is also of class \( C^p \). If \( D^p \Psi \) is Lipschitz, then so is \( D^p \varphi \).

### 2.2. On Newton’s method

Let \( X \) and \( Y \) be Banach spaces and \( F : X \rightarrow Y \) be of class \( C^1 \). We say that \( \bar{x} \in X \) is a zero of \( F \) if \( F(\bar{x}) = 0 \), and that it is a regular zero if in addition \( DF(\bar{x}) \) is invertible. Newton’s method for finding a zero of \( F \) consists in computing the Newton sequence \( x^k \) in \( X \) such that

\[
F(x^k) + DF(x^k)(x^{k+1} - x^k) = 0, \quad k \in \mathbb{N},
\]

starting from some point \( x^0 \in X \). The sequence is well-defined as long as \( DF(x^k) \) is invertible. It is easily checked that the set of linear invertible elements of \( L(X,Y) \) is open. Since \( DF(x) \) is a continuous function of \( x \) it follows that \( DF(x) \) is invertible in a vicinity of \( \bar{x} \). It can be proved that:

**Theorem 2.21.** if \( x^0 \) is close enough to the regular zero \( \bar{x} \), the Newton sequence is well-defined and \( x^k \rightarrow \bar{x} \) superlinearly, i.e.,

\[
\frac{\|x^{k+1} - \bar{x}\|}{\|x^k - \bar{x}\|} \rightarrow 0.
\]

If in addition \( x \mapsto DF(x) \) is Lipschitz near \( \bar{x} \), then we have quadratic convergence, i.e.

\[
\|x^{k+1} - \bar{x}\| = O \left( \|x^k - \bar{x}\|^2 \right).
\]

**Proof.** See e.g. [9, Ch. 6]. \( \square \)

### 2.3. Reduced equation

Consider a system with two (vector) equations and unknowns, such that the second variable can be eliminated from the second equation. We will compare the original system of two equations, with the reduced one obtained after elimination of the second variable. Namely, let \( U, Y, W, Z \) be Banach spaces, \( \Phi : U \times Y \rightarrow W \) and \( \Psi : U \times Y \rightarrow Z \) be of class \( C^1 \). Consider the equations, in \( W \times Z \):

\[
(2.50) \quad \Phi(u, y) = 0; \quad \Psi(u, y) = 0.
\]

Let \((\bar{u}, \bar{y}) \in U \times Y\) satisfy

\[
(2.51) \quad (\bar{u}, \bar{y}) \text{ is a zero of } (\Phi, \Psi), \text{ and } D_y \Psi(\bar{u}, \bar{y}) \text{ is invertible.}
\]

By the IFT, in the vicinity of \((\bar{u}, \bar{y})\), we have that

\[
(2.52) \begin{cases} 
\text{The relation } \Psi(u, y) = 0 \text{ is equivalent to } y = \chi(u), \\
\text{for some } C^1 \text{ function } \chi : V \rightarrow Y, \text{ where } V \text{ is a neighbourhood of } \bar{u}, \\
\text{such that whenever } y = \chi(u): \\
D_u \Psi(u, y) + D_y \Psi(u, y) D\chi(u) = 0.
\end{cases}
\]

We will call \( \Psi(u, y) = 0 \) the state equation. Locally, (2.50) is equivalent to the reduced equation

\[
(2.53) \quad \Xi(u) = 0, \quad \text{where } \Xi(u) := \Phi(u, \chi(u)).
\]

In addition

\[
(2.54) \quad D\Xi(\bar{u}) := D_u \Phi(\bar{u}, \bar{y}) + D_y \Phi(\bar{u}, \bar{y}) D\chi(\bar{u}) = D_u \Phi(\bar{u}, \bar{y}) - D_y \Phi(\bar{u}, \bar{y}) D\Psi(\bar{u}, \bar{y})^{-1} D_u \Psi(\bar{u}, \bar{y}).
\]

Given \((v, w) \in U \times W\), one easily checks that

\[
(2.55) \begin{cases} 
D\Xi(\bar{u}) v = w \text{ iff there exists } z \in Y \text{ such that } \\
D\Phi(\bar{u}, \bar{y})(v, z) = w; \quad D\Psi(\bar{u}, \bar{y})(v, z) = 0.
\end{cases}
\]

**Lemma 2.22.** The point \( \bar{u} \) is a regular zero of \( \Xi \) iff \((\bar{u}, \bar{y})\) is a regular zero of \((\Phi, \Psi)\).
Proof. Consider the system

\[(2.56)\]
\[
\begin{aligned}
D_u \Phi(\bar{u}, \bar{y}) v + D_y \Phi(\bar{u}, \bar{y}) z &= a; \\
D_u \Psi(\bar{u}, \bar{y}) v + D_y \Psi(\bar{u}, \bar{y}) z &= b.
\end{aligned}
\]

Since \(D_y \Psi(\bar{u}, \bar{y})\) is invertible, we can eliminate \(z\) from the second equation. Using (2.54), we see that the above system is equivalent to

\[(2.57)\]
\[
D \Xi(\bar{u}) v = a - D_y \Phi(\bar{u}, \bar{y}) D_y \Psi(\bar{u}, \bar{y})^{-1} b.
\]

The conclusion easily follows. \(\square\)

2.4. Reduced cost.

2.4.1. First order analysis. We next consider an optimization problem of the form

\[(2.58)\]
\[
\min_{u,y} f(u, y); \quad \Psi(u, y) = 0.
\]

Here \(U, Y, Z\) are Banach spaces, \(f : U \times Y \rightarrow \mathbb{R}\) and \(\Psi : U \times Y \rightarrow Z\) are of class \(C^1\). Let \((\bar{u}, \bar{y})\) be a zero of \(\Psi\), such that \(D_y \Psi(\bar{u}, \bar{y})\) is invertible. Then for \((u, y)\) close to \((\bar{u}, \bar{y})\), the consequence (2.52) of the IFT applies. Let \(F(u) := f(u, \chi(u))\) be the reduced cost. Given a neighbourhood \(V\) of \(\bar{u}\), we may define a (localized) reduced problem as

\[(2.59)\]
\[
\min_u F(u); \quad u \in V.
\]

2.4.2. Computation of the derivative of the reduced cost. In view of (2.52), the derivative of \(F\) at \(u\) close to \(\bar{u}\) is, writing \(y = \chi(u)\):

\[(2.60)\]
\[
D F(u) v = D_a f(u, y) v + D_y f(u, y) D_\chi(u) v \\
= D_a f(u, y) v - D_y f(u, y) D_y \Psi(u, y)^{-1} D_u \Psi(u, y) v.
\]

Therefore,

\[(2.61)\]
\[
D F(u) = D_a f(u, y) - D_y f(u, y) D_y \Psi(u, y)^{-1} D_u \Psi(u, y).
\]

However, we must pay attention to the fact that we usually represent the action of the linear form \(D F(u)\) over \(v\) with the dual notation \(\langle DF(u), v \rangle\). Note that the transpose of an invertible linear mapping is invertible, and that transposition and inversion commute\(^1\), which justifies the notation \(A^{-1}\). We have that

\[(2.62)\]
\[
\langle DF(u), v \rangle = \langle D_a f(u, y), v \rangle + \langle D_y f(u, y), D_\chi(u) v \rangle \\
= \langle D_a f(u, y) + D_\chi(u) D_y f(u, y), v \rangle,
\]

that is,

\[(2.63)\]
\[
DF(u) = D_a f(u, y) + D_\chi(u) D_y f(u, y) \\
= D_a f(u, y) - D_u \Psi(u, y) D_y \Psi(u, y)^{-1} D_y f(u, y).
\]

One must be careful of avoiding confusions between (2.61) and (2.63), which, despite an apparent difference, represent the same operator.

\(^1\)Indeed, if \(A \in L(X, Y)\) is invertible, solving \(A^* x^* = x^*\), with \(y^* \in Y^*\) and given \(x^* \in X^*\), amounts to \((y^*, Ax) = (x^*, x)\) for all \(x \in X\), or equivalently \((y^*, y) = (x^*, A^{-1} y) = ((A^{-1})^* x^*, y)\), for all \(y \in Y\). Since \(y \mapsto \langle (A^{-1})^* x^*, y \rangle\) is linear and continuous, the existence and uniqueness of \(y^*\) follows, and we have that \((A^{-1})^{-1} x^* = (A^{-1})^* x^*\).
2.4.3. Reduction Lagrangian. In practice, one obtains the expression of $DF(u)$ using the reduction Lagrangian, where $(u, y, p) \in U \times Y \times Z^*$:

\begin{equation}
L_R(u, y, p) := f(u, y) + \langle p, \Psi(u, y) \rangle_z,
\end{equation}

and the costate equation, where $y = \chi(u)$:

\begin{equation}
0 = D_y L_R(u, y, p) = D_y f(u, y) + D_y \Psi(u, y)^\dagger p = 0.
\end{equation}

Note the use of the dual notation. Given $(u, y) \in U \times Y$, $y$ being the state associated with $u$, since $D_y \Psi(u, y)$ (and hence, $D_y \Psi(u, y)^\dagger$) is invertible, the costate equation has a unique solution $p \in Z^*$, called the costate associated with $u$:

\begin{equation}
p = -D_y \Psi(u, y)^{-1} D_y f(u, y).
\end{equation}

Observe now that

\begin{equation}
D_u L_R(u, y, p) = D_u f(u, y) + D_u \Psi(u, y)^\dagger p,
\end{equation}

where the last equality uses (2.63), and is therefore equal to 0 if $u$ is a local solution. We say that $u \in U$ is a stationary point of $F(u)$ if $DF(u) = 0$. We have proved the following useful rule:

**Lemma 2.23.** (i) The derivative of the reduced cost $F$ is equal to the partial derivative w.r.t. the control of the reduction Lagrangian, computed at the control and associated state and costate. (ii) An element $u$ of $U$ is a stationary point of $F(u)$ iff there exists $p \in Z^*$ such that, $y$ being the associated state:

\begin{equation}
D_u L_R(u, y, p) = 0; \quad D_y L_R(u, y, p) = 0.
\end{equation}

2.4.4. Second order analysis. The costate approach is also useful for obtaining a simple expression of the Hessian (the second derivative seen as a bilinear form) of the reduced cost, assuming now that $f$ and $\Psi$ are $C^2$. The Hessian of the reduction Lagrangian (w.r.t. the control and costate variables) is the quadratic form $Q : U \times Y \to \mathbb{R}$ defined by

\begin{equation}
Q(v, z) := D^2 f(u, y)(v, z)^2 + \langle p, D^2 \Psi(u, y)(v, z) \rangle.
\end{equation}

Here we assume that $y$ and $p$ are the state and costate associated with the control $u$. We denote by $z[v]$ the solution of the linearized state equation

\begin{equation}
D_u \Psi(u, y)v + D_y \Psi(u, y)z = 0.
\end{equation}

Given $v \in U$ it has a unique solution denoted by $z[v]$ in $Y$, namely

\begin{equation}
z[v] := -D_y \Psi(u, y)^{-1} D_u \Psi(u, y)v.
\end{equation}

**Lemma 2.24.** The Hessian of the reduced cost satisfies

\begin{equation}
D^2 F(u)(v, v) = Q(v, z[v]), \quad \text{for all } v \in U.
\end{equation}

**Proof.** Given $\sigma \geq 0$, set $u_\sigma := u + \sigma v$. The associated state denoted by $y_\sigma$ satisfies

\begin{equation}
\| y_\sigma - y - \sigma z[v]\|_Y = O(\sigma^2).
\end{equation}

So we have that

\begin{equation}
F(u_\sigma) = f(u_\sigma, y_\sigma) = L(u_\sigma, y_\sigma, p) = L(u, y) + \sigma D_u L(u, y, p)v + \frac{1}{2} \sigma^2 Q(v, z[v]) + o(\sigma^2),
\end{equation}

where we have used the fact that $D_y L(u, y, p) = 0$. The result follows.

**Corollary 2.25.** Let $F$ have a local minimum at $u \in U$. Then $Q(v, z[v]) \geq 0$, for all $v \in U$. 23
Proof. From the Taylor expansion of $F$ it follows that, for $v \in U$ and $t \in \mathbb{R}$:
\begin{equation}
F(u + tv) = F(u) + tDF(u)v + \frac{1}{2}t^2D^2F(u)(v,v) + o(t^2),
\end{equation}
and since $DF(u) = 0$, the local optimality of $u$ implies
\begin{equation}
0 \leq \lim_{t \downarrow 0} \frac{F(u + tv) - F(u)}{t^2} = D^2F(u)(v,v).
\end{equation}
We conclude using lemma 2.24.

2.5. Equality constrained optimization.

2.5.1. First order optimality conditions. Consider next an 'abstract' optimization problem with equality constraints:
\begin{equation}
\min_{x} f(x); \quad g(x) = 0.
\end{equation}
Here $g : X \to Y$ (Banach spaces) and $f : X \to \mathbb{R}$ are of class $C^2$. This is obviously a generalization of the previous setting, but without the possibility of reduction to an unconstrained reduced problem. Let $\bar{x}$ be a local solution of this problem, in the sense that for some $\varepsilon > 0$:
\begin{equation}
f(\bar{x}) \leq f(x), \quad \text{whenever } g(x) = 0 \text{ and } ||x - \bar{x}||_X \leq \varepsilon.
\end{equation}
We assume that $\bar{x}$ satisfies the following qualification condition
\begin{equation}
The operator $Dg(\bar{x})$ is surjective.
\end{equation}
The Lagrangian of the problem is the function $L_g : X \times Y^* \to \mathbb{R}$ defined by
\begin{equation}
L_g(x, \lambda) := f(x) + \langle \lambda, g(x) \rangle_Y.
\end{equation}

Theorem 2.26. There exists a unique Lagrange multiplier $\lambda \in Y^*$ such that (dual notation)
\begin{equation}
0 = D_xL_g(\bar{x}, \lambda) := Df(\bar{x}) + Dg(\bar{x})^\dagger \lambda.
\end{equation}
The proof is based on the following lemmas below. The starting point is the concept of metric regularity.

Lemma 2.27. Let (2.79) hold. Then the following metric regularity property is satisfied: there exists $c_g > 0$ such that, if $x \in X$ is close enough to $\bar{x}$, there exists $x' \in X$ such that
\begin{equation}
||x' - x||_X \leq c_g ||g(x)||_Y \text{ and } g(x') = 0.
\end{equation}
Proof. See e.g. [9, Ch. 3].

Exercice 2.28. Take $X = Y = \mathbb{R}$ and $g(x) = x^2$. Show that the metric regularity property does not hold at $\bar{x} = 0$. Also, show that the problem of minimizing $f(x) = x$ under the constraint $g(x) = 0$ has solution $\bar{x} = 0$, with which no Lagrange multiplier is associated.

Corollary 2.29. If (2.79) holds, any $h \in \ker g'(\bar{x})$ is tangent to the manifold $g^{-1}(0)$ in the sense that, setting $x_\sigma := \bar{x} + \sigma h$, for $\sigma \in \mathbb{R}$:
\begin{equation}
\text{There exists } x'_\sigma \in U \text{ such that } g(x'_\sigma) = 0 \text{ and } ||x'_\sigma - x_\sigma|| = o(\sigma).
\end{equation}
Proof. That $h \in \ker g'(\bar{x})$ implies $||g(x_\sigma)|| = o(\sigma)$. We conclude with lemma 2.27.

Corollary 2.30. If (2.79) holds, then $f'(\bar{x}) \in (\ker g'(\bar{x}))^\perp$.

Proof. Otherwise there would exists $h \in \ker g'(\bar{x})$ such that $f'(\bar{x})h \neq 0$. Changing $h$ into $-h$ if necessary we may assume that $f'(\bar{x})h < 0$. Define $x'_\sigma$ as above. By the previous corollary,
\begin{equation}
\lim_{\sigma \downarrow 0} \frac{f(x'_\sigma) - f(\bar{x})}{\sigma} = f'(\bar{x})h < 0,
\end{equation}
contradicting the local optimality of $\bar{x}$. The conclusion follows.
Proof of theorem 2.26. By the above corollary, \( f'(\bar{x}) \in (\text{Ker } g'(\bar{x}))^\perp \). Since \( g'(\bar{x}) \) is surjective, its image is closed. Therefore, the orthogonal of its kernel coincides with the image of its transpose operator, see e.g. [18, Thm. 2.19]. So, there exists \( \lambda \in Y^* \) such that \( f'(\bar{x}) + g'(\bar{x})^\lambda = 0 \), as was to be proved.

Remark 2.31. In a finite dimensional setting, the orthogonal of the kernel of a linear operator always coincides with the image of its transpose. For a counterexample in a Banach space setting, let \( A \) be the injection from \( L^2(0,1) \) (identified with its dual) into \( L^1(0,1) \). Its image is dense, but not closed. Now \( A^\dagger : L^\infty(0,1) \to L^2(0,1) \) is defined, for all \( w \in L^\infty(0,1) \) and \( u \in L^1(0,1) \) by \( \langle A^\dagger w, u \rangle_{L^2(0,T)} = \langle w, Au \rangle_{L^1(0,T)} = \int_0^1 w(t)u(t)dt \). Since \( A \) is injective its kernel is reduced to \{0\} and the orthogonal of the kernel is \( L^2(0,1) \). On the other hand the image of its transpose is \( L^\infty(0,1) \).

2.5.2. Second order optimality conditions. We next obtain second order necessary conditions as an easy consequence of the previous results:

Proposition 2.32. Let \( \bar{x} \) be a qualified local solution of (2.77), with associated Lagrange multiplier \( \lambda \). Then
\[
D^2_{xx}L_g(\bar{x}, \lambda)(h,h) \geq 0, \quad \text{for all } h \in \text{Ker } g'(\bar{x}).
\]

Proof. Set as before \( x_\sigma := \bar{x} + \sigma h \) and let \( x'_\sigma \) be given by corollary 2.29. Since \( g(x'_\sigma) = 0 \), and \( D_xL_g(\bar{x}, \lambda) = 0 \) by (2.81), we have for \( |\sigma| \) small enough
\[
0 \leq f(x'_\sigma) - f(\bar{x}) = L_g(x'_\sigma, \lambda) - L_g(\bar{x}, \lambda) = \frac{1}{2}\sigma^2 D^2_{xx}L_g(\bar{x}, \lambda)(h,h) + o(\sigma^2).
\]

Dividing by \( \sigma^2 \) and making \( \sigma \to 0 \) we obtain the conclusion. \( \square \)

Exercise 2.33. Let \( X \) be a Hilbert space, \( x_0 \in X \), \( M \in L(X) \) symmetric and invertible, and for \( x \in X \), \( g(x) := \frac{1}{2}\langle x, Mx \rangle_X - \frac{1}{2} \). Write the first and second order optimality conditions for the problem of projecting \( x_0 \) over \( g^{-1}(0) \) (we do not discuss the existence of a solution).

Hint: Set \( f(x) := \frac{1}{2}\|x - x_0\|^2 \). Check that (i) \( \nabla g(x) = Mx \) (deduce that any feasible point is qualified), (ii) if \( \bar{x} \) is solution, then \( \bar{x} - x_0 + \lambda M\bar{x} = 0 \), (iii) the Hessian of Lagrangian is \( I + \lambda M \).

2.5.3. Link with the reduction approach. We next establish the relation with the 'non reduced' formulation, assuming that \( x = (u, y) \), i.e.
\[
\begin{align*}
\min_{u,y} \Phi(u, y); \quad \Psi(u, y) = 0; \quad g(u, y) = 0.
\end{align*}
\]
We assume that \( D_y\Psi(\bar{u}, \bar{y}) \) is invertible so that 'locally' \( \Psi(u, y) = 0 \) iff \( y = \varphi(u) \) where \( \varphi \) is given by the IFT. So the corresponding reduced problem is
\[
\min_u F(u); \quad G(u) = 0,
\]
where
\[
F(u) := \Phi(u, \varphi(u)); \quad G(u) := g(u, \varphi(u));
\]
The Lagrangian functions associated with the original and reduced problems are resp.
\[
\begin{align*}
\begin{cases}
L(u, y, p, \lambda) := & 
\Phi(u, y) + \langle p, \Psi(u, y) \rangle + \langle \lambda, h(u, y) \rangle, \\
\ell(u, \lambda) := & 
\ell(u, \lambda) := F(u) + \langle \lambda, G(u) \rangle.
\end{cases}
\end{align*}
\]
The costate equation at the point \( \bar{x} = (\bar{u}, \bar{y}) \), with \( \bar{y} \) the state associated with \( \bar{u} \), reads
\[
0 = D_yL(\bar{u}, \bar{y}, p, \lambda) = D_y\Phi(\bar{u}, \bar{y}) + D_y\Psi(\bar{u}, \bar{y})^\dagger p + D_yg(\bar{u}, \bar{y})^\dagger \lambda.
\]
The optimality conditions for the the original and reduced problems are resp. (assuming in the case of the original problem that the costate equation holds)
\[
\begin{align*}
\begin{cases}
(i) \quad 0 = D_uL(\bar{u}, \bar{y}, p, \lambda) = D_u\Phi(\bar{u}, \bar{y}) + D_u\Psi(\bar{u}, \bar{y})^\dagger p + D_ug(\bar{u}, \bar{y})^\dagger \lambda, \\
(ii) \quad 0 = D_u\ell(\bar{u}, \lambda) = F'(\bar{u}) + G'(\bar{u})^\dagger \lambda.
\end{cases}
\end{align*}
\]
Lemma 2.34. Let the costate equation (2.91) hold. Then $\lambda$ satisfies (2.92)(i) iff it satisfies (2.92)(ii). In other words, the Lagrange multipliers associated with the original and reduced formulation coincide.

Proof. It suffices to eliminate $p$ thanks to the costate equation (2.91) in (2.92)(i), and to express $F'(\bar{u})$ and $G'(\bar{u})$ (thanks to the IFT) in (2.92)(ii). □

Remark 2.35. In the same way we can check that the second order necessary optimality conditions for the original and reduced problem give the same information, since

$$D^2_{uv}(\bar{u},\lambda)(v,v) = D^2_{(u,y)p}L(\bar{u},\bar{y},p,\lambda)(v,v)^2$$

where $(v,z) \in U \times Y$ satisfy $D\Psi(\bar{u},\bar{y})(v,z) = 0$.

2.6. Calculus over $L^\infty$ spaces. We denote by $L^0(0,T)$ the set of measurable functions over $(0,T)$. With $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^m$ we associate the Nemitskii mapping $F : L^0(0,T)^n \to L^0(0,T)^m$ defined by

$$F(y)(t) := f(t,y(t)) \text{ for a.a. } t \in (0,T).$$

Lemma 2.36. If $f$ is of class $C^p$, then $F$ is of class $C^p : L^\infty(0,T)^n \to L^\infty(0,T)^m$, and satisfies

$$(D^jF(y)(z)^j)(t) = D^j_{y^j}f(t,y(t))(z(t))^j \text{ a.e.,}$$

for all $j \leq p$, and so, for all $y, z \in L^\infty(0,T)^n$ we have the following Taylor expansion:

$$F(y+z)(t) = f(t,y(t)) + \sum_{j=1}^{p-1} \frac{1}{j!} D^j_{y^j}f(t,y(t))(z(t))^j + r(t); \quad \|r\|_\infty = o(\|z\|_\infty^p).$$

Proof. Being continuous, $F$ is bounded over bounded sets. In addition, the composition of a measurable mapping by a continuous one is measurable. So, $F$ has image in $L^\infty(0,T)^m$.

In view of the measurable mapping by a continuous one is measurable. So, $F$ has image in $L^\infty(0,T)^m$. (2.32), the equality in (2.96) holds with

$$r(t) := a(t,y(t),z(t))(z(t))^p,$$

where the symmetric $p$ linear form $a$ is defined by

$$a(t,y,z) := \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} \left( D^p_{y^p}f(t,y+sz) - D^p_{y^p}f(t,y) \right) ds,$$

and therefore (the norm of the multilinear form is defined analogously to (2.2))

$$r(t) \leq \|a(t,y(t),z(t))\|\|z(t)\|^p.$$

Since continuous functions are uniformly continuous over compact sets, $a(t,y(t),z(t)) \to 0$ in $L^\infty(0,T)^m$ when $z \to 0$ in $L^\infty(0,T)^m$. So, (2.96) holds. Writing it with $p = 1$, we obtain that (2.95) holds for $p = 1$. Let it hold for some $j < p$. Then for $y, w, z \in L^\infty(0,T)^n$, we have a.e.: (2.100)

$$(D^jF(y+w)(z)^j)(t) = D^j_{y^j}f(t,y(t)+w(t))(z(t))^j$$

$$= (D^j_{y^j}F(y)(z)^j)(t) + \rho(t)$$

with $|\rho(t)| = o(|w(t)||z(t)|^j)$ uniformly in time. This proves (2.95) for $j + 1$, and so, the result follows by induction. □

Analyzing the above proof we may guess that smoothness w.r.t. time can be weakened, provided that the functions are measurable and that the property of uniformly small remainders in (2.96) holds.

Definition 2.37. Let $g_t$ be a family of mappings $\mathbb{R}^n \to \mathbb{R}^m$, defined for a.a. $t \in (0,T)$. We say that $g_t$ is a Carathéodory mapping if $t \mapsto g_t(x)$ is measurable for each $x \in X$, and $x \mapsto g_t(x)$ is continuous for a.a. $t$.

(ii) We say that $g$ has a uniform (in time) modulus of continuity over bounded sets if for any
$R > 0$, there exists a nondecreasing function $\omega_R : \mathbb{R}^+ \to \mathbb{R}^+ \cup \{+\infty\}$, such that $\omega_R(\varepsilon) \to 0$ when $\varepsilon \downarrow 0$, and

\[(2.101) \quad |g_t(x) - g_t(y)| \leq \omega_R(|x - y|), \quad \text{whenever } |x| < R \text{ and } |y| < R.\]

**Lemma 2.38.** Let $g_t$ be a Carathéodory function. Let $f : [0, T] \to \mathbb{R}^m$ be measurable. Then $t \mapsto g_t(f(t))$ is measurable.

**Proof.** (a) We start with the case when $f$ is a simple function, say $f(t) = \sum_k a_k \chi_{A_k}(t)$, where the sum is finite, the $A_k$ are measurable subsets of $[0, T]$, with negligible intersection, and $\chi_{A_k}$ is the characteristic function of $A_k$ (with value 1 over $A_k$ and 0 otherwise). Then

\[(2.102) \quad g_t(f(t)) = \sum_k g_t(a_k) \chi_{A_k}(t)\]

is measurable (being a sum of products of measurable mappings).

(b) Since any measurable function $f : [0, T] \to \mathbb{R}^m$ is the limit a.e. of simple functions say $f_j$, and since $g_t(\cdot)$ is a.e. continuous, we have that $g_t(f(t)) = \lim_j g_t(f_j(t))$ a.e., and the limit a.e. of measurable functions is measurable. 

**Definition 2.39.** We say that $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^m$ is uniformly quasi $C^p$, for $p \in \mathbb{N}$, if $f(t,x)$ is a measurable function of time for each $x$, is a $C^p$ function of $x$ for a.a. $t$, such that for any $0 \leq j \leq p$, the function $D^j f$ (partial derivative $j$ times w.r.t. $x$ of $f$):

1. has a uniform (over time) modulus of continuity w.r.t. $x$, continuous at 0, on bounded sets,
2. for fixed $x$, is an essentially bounded function of time.

**Remark 2.40.** A uniformly quasi $C^p$ function is a Carathéodory function. This is also true for $D^j f$ for $j \leq p$ (using the fact that derivatives are limits of quotients). From the inequality $|f(t,x)| \leq |f(t,0)| + |f(t,x) - f(t,0)|$ we easily deduce that a uniformly quasi $C^p$ function is bounded over bounded sets.

**Lemma 2.41.** The conclusion of lemma 2.36 still holds if we assume only that $f$ is uniformly quasi $C^p$.

**Proof.** Follows from the previously discussed arguments. 

### 3. Optimal control setting

#### 3.1. Weak derivatives.

**Definition 2.42.** Let $\mathcal{D}(0,T;\mathbb{R}^n)$ denote the set of $C^\infty$ function with compact support in $(0,T)$ and value in $\mathbb{R}^n$, and let $y \in L^1(0,T;\mathbb{R}^n)$. We say that $g \in L^1(0,T;\mathbb{R}^n)$ is the weak derivative of $y$ if it satisfies

\[(2.103) \quad \int_0^T y(t) \cdot \varphi(t)dt + \int_0^T g(t) \cdot \varphi(t)dt = 0, \quad \text{for all } \varphi \in \mathcal{D}(0,T;\mathbb{R}^n).\]

**Lemma 2.43.** The weak derivative is unique: given $y \in L^1(0,T;\mathbb{R}^n)$, there is at most one $g \in L^1(0,T;\mathbb{R}^n)$ satisfying (2.103).

**Proof.** (a) It is enough to consider the scalar case $n = 1$. Let $g'$ and $g''$ be weak derivatives of $y$. Then $g := g'' - g'$ satisfies

\[(2.104) \quad \int_0^T g(t) \cdot \varphi(t)dt = 0, \quad \text{for all } \varphi \in \mathcal{D}(0,T;\mathbb{R}).\]

A quick conclusion is obtained in the case when $g \in L^2(0,T)$; since $\mathcal{D}(0,T;\mathbb{R}^n)$ is known to be a dense subset of $L^2(0,T)$, there exists a sequence $\varphi_k \in \mathcal{D}(0,T;\mathbb{R}^n)$ converging to $g$ in $L^2(0,T)$, so that

\[(2.105) \quad \int_0^T g(t)^2dt = \lim_k \int_0^T g(t) \cdot \varphi_k(t)dt = 0\]
which implies that \( g = 0 \) so that the result holds.

(b) In the general case, observe that if \( g \neq 0 \), there exists \( \varepsilon > 0 \) and a measurable subset \( A \) of \( (0, T) \), of positive measure, such that (changing \( g \) into \(-g\) if necessary), \( g(t) > \varepsilon \) a.e. on \( A \). It is known that there exists a compact set \( K \subset A \) with positive measure, see [7, Ch. 1]. Obviously we can take \( K \) as a subset of \([\varepsilon_1, T - \varepsilon_1]\) for some \( \varepsilon_1 > 0 \). Set \( \Psi_k(t) := (1 - k \text{dist}_K(t))_+ \), where \( \text{dist}_K \) denotes the distance to the set \( K \). By the dominated convergence theorem, \( \int_0^T g(t)\Psi_k(t)dt \to \int_K g(t)dt > 0 \). So there exists \( k \) such that \( \int_0^T g(t)\Psi_k(t)dt > 0 \). Observe that \( \Psi_k \) is continuous and (taking \( k \) large enough) with compact support in \((0, T)\). So there exists a sequence \( \varphi_\ell \) in \( \mathcal{D}(0, T; \mathbb{R}^n) \) that uniformly converges to \( \Psi_k \). By the dominated convergence theorem,

\[
(2.106) \quad 0 < \int_0^T g(t)\Psi_k(t)dt = \lim_{\ell \to 0} \int_0^T g(t)\varphi_\ell(t)dt,
\]

but the r.h.s. is zero by the definition, which gives the desired contradiction. \( \square \)

Conversely, a given weak derivative is the one of a unique function up to a constant, as shows the following result: we give a proof in the \( L^2 \) setting.

**Lemma 2.44** (Du Bois-Reymond). Let \( y \in L^2(0, T, \mathbb{R}^n) \) be such that

\[
(2.107) \quad \int_0^T y(t) \cdot \dot{\varphi}(t) dt = 0, \quad \text{for all} \ \varphi \in \mathcal{D}(0, T; \mathbb{R}^n).
\]

Then \( y \) is constant.

**Proof.** We may assume that \( n = 1 \). If \( y \) satisfies (2.107), then so does the function \( y - (1/T) \int_0^t y(s)ds \). So, we may assume that \( \int_0^T y(t)dt = 0 \). Since \( \mathcal{D}(0, T) \) is a dense subset of \( L^2(0, T) \), there exist a sequence \( \psi_k \) in \( \mathcal{D}(0, T) \) converging to \( y \) in \( L^2(0, T) \), so that \( \alpha_k := \int_0^T \psi_k(t)dt \to 0 \). Let \( \eta \in \mathcal{D}(0, T) \) have a unit integral. Then \( \psi'_k(t) = \psi_k(t) - \alpha_k \eta(t) \) is another sequence in \( \mathcal{D}(0, T) \) converging to \( y \) in \( L^2(0, T) \), with zero integral. The primitives \( \varphi_k(t) := \int_0^t \psi'_k(s)ds \) also belong to \( \mathcal{D}(0, T) \), so that by hypothesis

\[
(2.108) \quad 0 = \lim_{k \to \infty} \int_0^T y(t)\dot{\varphi}_k(t)dt = \lim_{k \to \infty} \int_0^T y(t)\psi'_k(t)dt \to \int_0^T y^2(t)dt.
\]

The conclusion follows. \( \square \)

For \( s \in [1, \infty] \) we define the Sobolev space

\[
(2.109) \quad W^{1,s}(0, T, \mathbb{R}^n) := \{ y \in L^s(0, T, \mathbb{R}^n): \ \text{there exists} \ \tilde{y} \in L^s(0, T, \mathbb{R}^n) \}
\]

where here \( \tilde{y} \) denotes the weak derivative of \( y \), endowed with the norm

\[
(2.110) \quad \|y\|_{W^{1,s}(0, T, \mathbb{R}^n)} := \|y\|_{L^s(0, T, \mathbb{R}^n)} + \|	ilde{y}\|_{L^s(0, T, \mathbb{R}^n)}.
\]

One easily checks that \( W^{1,s}(0, T, \mathbb{R}^n) \) is a Banach space. The following is well-known, see Royden \[37, \text{Ch. 5}]:

**Lemma 2.45.** Let \( y \in L^1(0, T, \mathbb{R}^n) \) and \( s \in [1, \infty] \). Then \( y \) is the primitive of a function \( g \) in \( L^s(0, T, \mathbb{R}^n) \) iff \( y \in W^{1,s}(0, T, \mathbb{R}^n) \), and then \( g \) is the weak derivative of \( y \).

3.2. Controlled dynamical system and associated cost. Consider a controlled dynamical systems of the type

\[
(2.111) \quad \begin{cases} (i) \quad \dot{y}(t) = f(t, u(t), y(t)), & \text{for a.a.} \ \ t \in [0, T]; \\ (ii) \quad y(0) = y_0. \end{cases}
\]

The data are the horizon or final time \( T > 0 \), the initial condition \( y_0 \in \mathbb{R}^n \), and the dynamics

\[
(2.112) \quad f: \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n, \quad \text{uniformly quasi} \ C^r, \ r \geq 1, \ \text{Lipschitz w.r.t.} \ y.
\]

We call \( u(t) \) and \( y(t) \) the control and state at time \( t \). The control and state spaces are

\[
(2.113) \quad U := L^\infty(0, T; \mathbb{R}^m); \quad Y := W^{1,\infty}(0, T, \mathbb{R}^n).
\]
We may see (2.111)(i) as an equality in $L^\infty(0,T)^n$. By the Cauchy-Lipschitz theorem (adapted to the uniformly quasi $C^r$ setting, with the same proof based on a fixed-point theorem) for all $(u, y^0) \in U \times \mathbb{R}^n$, the state equation (2.111) has a unique solution in $Y$, denoted by $y[u, y^0]$ (or $y[u]$ if $y^0$ is fixed). In addition, if $f$ is uniformly Lipschitz in $u$, by Gronwall’s lemma, for some $C_f$ depending only on the Lipschitz constant of $f$:

\[(2.114)\]

\[\|y[u', (y')'] - y[u, y^0]\| \leq C_f (\|u' - u\|_1 + |(y')' - y^0|).\]

We denote by $z[v, z^0]$, or $z[v]$ if $z^0 = 0$, the unique solution of the linearized state equation

\[(2.115)\]

\[\begin{cases}
(i) & \dot{z}(t) = f'(t, u(t), y(t))(v(t), z(t)), \text{ for a.a. } t \in (0, T]; \\
(ii) & z(0) = z^0.
\end{cases}\]

Here, by $f'(t, u, y)$, we denote the partial derivative of $f$ w.r.t. $(u, y)$. The mapping $z[v, z^0]$ is well defined and continuous $U \times \mathbb{R}^n \rightarrow Y$. It has a continuous extension\(^2\) from $L^s(0, T, \mathbb{R}^m) \times \mathbb{R}^n$ into $W^{1,s}(0, T, \mathbb{R}^n)$, for any $s$ in $[1, \infty]$.

**Proposition 2.46.** The mapping $U \times \mathbb{R}^n \rightarrow Y$, $(u, y^0) \mapsto y[u, y^0]$, is of class $C^r$.

**Proof.** Let $F$ be the mapping $U \times Y \times \mathbb{R}^n$ to $L^\infty(0, T, \mathbb{R}^n)$, such that with $(u, y, y^0)$ associates the state equation (2.111), that is,

\[(2.116)\]

\[F(u, y, y^0) := \left( \begin{array}{c} \dot{y}(t) - f(t, u(t), y(t)), \\ y(0) - y^0 \end{array} \right), \quad t \in (0, T).\]

By lemma 2.41, $F$ is of class $C^r$. So the conclusion will follow from the implicit function theorem, provided that the partial derivative of $F$ w.r.t. the state is invertible. This holds if, for any $(g, e) \in L^\infty(0, T, \mathbb{R}^n) \times \mathbb{R}^n$, the following variant of the linearized state equation (2.115) has a unique solution in $Y$:

\[(2.117)\]

\[(i) \dot{z}(t) = f'(t, u(t), y(t))(v(t), z(t)) + g(t), \quad \text{for a.a. } t \in [0, T]; \quad (ii) z(0) = e,\]

which obviously is the case. The conclusion follows. ∎

We say that $(u, y) \in U \times Y$ is a trajectory if $y = y[u, y(0)]$. In the sequel we perform an analysis around the nominal trajectory $(\tilde{u}, \tilde{y})$. We denote

\[(2.118)\]

\[\tilde{f}(t) := f(t, \tilde{u}(t), \tilde{y}(t)), \quad \tilde{f}'(t) := (f'_u(t, \tilde{u}(t), \tilde{y}(t)), f'_y(t, \tilde{u}(t), \tilde{y}(t))),\]

the nominal dynamics and its derivative w.r.t. control and state, with a similar convention for derivatives at any order of $(\tilde{u}(t), \tilde{y}(t))$, the partial derivatives being denoted by e.g. $\tilde{f}_{u}(t)$. By proposition 2.46, $z[v, z^0]$ is the directional derivative of $y[u, y^0]$ at the point $(\tilde{u}, \tilde{y})$ in direction $(v, z^0) \in U \times \mathbb{R}^n$, i.e.,

\[(2.119)\]

\[z[v, z^0] = \lim_{s \to 0} \frac{1}{s} (y[\tilde{u} + sv, \tilde{y} + sz^0] - y[\tilde{u}, \tilde{y}^0]).\]

With the controlled system (2.111) we associated the cost function

\[(2.120)\]

\[J(u, y) := \int_0^T \ell(t, u(t), y(t))dt + \varphi(y(T)),\]

sum of an integral cost, with integrand $\ell$, and of a final cost $\varphi$; recalling the hypothesis on $f$, we assume that

\[(2.121)\]

\[f \text{ and } \ell \text{ are uniformly quasi } C^r, \varphi \text{ is } C^r.\]

Since a composition of $C^r$ mappings is $C^r$, $J : U \times Y \rightarrow \mathbb{R}$ is $C^r$, as well as the reduced cost

\[(2.122)\]

\[J_R(u, y^0) := J(u, y[u, y^0]).\]

---

2Let $X$ and $Y$ be Banach spaces and $E$ be a dense vector subspace of $X$. Let $A$ be a linear mapping $E \rightarrow Y$, such that for some $c > 0$, $\|Ax\| \leq c\|x\|$, for all $x \in E$. Then there exists a unique $A' \in L(X, Y)$ that extends $A$, i.e., such that $A'x = Ax$ for all $x \in E$, and $A'$ satisfies $\|A'\| \leq c$. We call $A'$ the continuous extension of $A$. 29
In addition, for \((v, z^0) \in U \times \mathbb{R}^n\), writing \(z = z[v, z^0]\), the expression of the derivative of the reduced cost is (adopting the notations \(\ell, \ell^\prime\) similar to \(\bar{f}, \bar{f}'\)):

\[
J^*_p(u, \bar{f}^0)(v, z^0) = \int_0^T \ell^\prime(t)(v(t), z(t))dt + \varphi'(\bar{y}(T))z(T).
\]

Our next step is to consider local (in time) control constraints of the type

\[
u(t) \in U_{ad}, \quad \text{for a.a. } t \in [0, T],
\]

where \(U_{ad}\) is a closed subset of \(\mathbb{R}^m\). Typical examples are the case of bound constraints, and the closed unit ball:

\[
\Pi_i^m[a_i, b_i]; \quad B := \left\{ u \in \mathbb{R}^m; \sum_{i=1}^m u_i^2 \leq 1 \right\}.
\]

### 3.3. Derivative of the cost function.

The reduced Lagrangian (defined in (2.64)) has in the present setting the following expression:

\[
J(u, y) + \int_0^T p(t) \cdot (f(t, u(t), y(t)) - \bar{y}(t)) dt + q \cdot (y(0) - y^0).
\]

Here we choose \(L^\infty(0, T; \mathbb{R}^n)\) as space for the state equation. So, a priori, we search for a costate in the dual space. But we assume more regularity, namely \(p \in \mathcal{Y}\), so that the above integral makes sense. Define the pre-Hamiltonian \(H : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}\) by

\[
H(t, u, y, p) := \ell(t, u, y) + p \cdot f(t, u, y).
\]

Integrating by parts, we can rewrite the reduction Lagrangian (2.126) as

\[
J(u, y) + \int_0^T (H(t, u(t), y(t), p(t)) + \dot{p}(t) \cdot y(t)) dt + \varphi(\bar{y}(T)) - p(T) \cdot \bar{y}(T) - (p(0) + q) \cdot y(0) - q \cdot y^0.
\]

Computing this amount at the point \((\bar{u}(t), \bar{y}(t), \bar{p}(t))\), and setting to zero its derivative w.r.t. \(y\) in an arbitrary direction \(z \in \mathcal{Y}\), we obtain the relation

\[
0 = \int_0^T \left( \nabla_y H(t, \bar{u}(t), \bar{y}(t), \bar{p}(t)) + \dot{\bar{p}}(t) \right) \cdot z(t) dt + \left( \nabla \varphi(\bar{y}(T)) - \bar{p}(T) \right) \cdot z(T) + (\bar{p}(0) + q) \cdot z(0).
\]

Taking \(z\) arbitrary with zero values at time 0 and \(T\) (which is a dense subset of \(L^2(0, T; \mathbb{R}^n)\)) we see that the above integrand has to be equal to zero, and then taking \(z(0)\) and \(z(T)\) arbitrary we see that \(q = -\bar{p}(0)\) and that the costate \(\bar{p} \in \mathcal{Y}\), is solution of the costate equation

\[
\begin{cases}
\nabla_y H(t, \bar{u}(t), \bar{y}(t), \bar{p}(t)) + \dot{\bar{p}}(t) = 0, & \text{for a.a. } t \in [0, T]; \\
\bar{p}(T) = \nabla \varphi(\bar{y}(T)).
\end{cases}
\]

Given the trajectory \((\bar{u}, \bar{y})\), this equation is backwards (the final condition is given) and has a unique solution in \(\mathcal{Y}\). We adopt the notation \(\bar{H}(t)\) in the spirit of (2.118), i.e., for instance:

\[
\bar{H}(t) := H(t, \bar{u}(t), \bar{y}(t), \bar{p}(t)), \quad \nabla_y \bar{H}(t) := \nabla_y H(t, \bar{u}(t), \bar{y}(t), \bar{p}(t)).
\]

Note that the r.h.s. of the costate dynamics is

\[
\nabla_y \bar{H}(t) = \nabla_y \bar{\ell}(t) + \bar{f}_y(t)^\dagger \bar{p}(t),
\]

Similarly, we have that

\[
\nabla_u \bar{H}(t) = \nabla_u \bar{\ell}(t) + \bar{f}_u(t)^\dagger \bar{p}(t).
\]

It follows from (2.132) that, for any \(z \in \mathcal{Y}\):

\[
\nabla_y \bar{H}(t) \cdot z(t) = \nabla_y \bar{\ell}(t) \cdot z(t) + \bar{p}(t)^\dagger \bar{f}_y(t)z(t).
\]
Lemma 2.47. The reduced cost $J_R$, defined in (2.122), has a derivative at $(\bar{u}, \bar{y}^0)$ characterized by

$$J'_R(\bar{u}, \bar{y}^0)(v, z^0) = \bar{p}(0) \cdot z^0 + \int_0^T \nabla_u H(t) \cdot v(t) dt.$$  \hfill (2.135)

Proof. This follows from the theory of reduction Lagrangian, see lemma 2.23, but we will give a direct argument. Using (2.115) and (2.130)-(2.134), we get:

$$\nabla \varphi(\bar{y}(T)) \cdot z(T) = \bar{p}(T) \cdot z(T)$$

$$= \bar{p}(0) \cdot z(0) + \int_0^T \frac{d}{dt} (\bar{p}(t) \cdot z(t)) dt$$

$$= \bar{p}(0) \cdot z(0) + \int_0^T (\dot{\bar{p}}(t) \cdot z(t) + \bar{p}(t) \cdot \dot{z}(t)) dt$$

$$= \bar{p}(0) \cdot z(0) + \int_0^T (\dot{\bar{p}}(t) \cdot f_u(t)v(t) - \nabla_y \bar{f}(t) \cdot z(t)) dt.$$  \hfill (∗)

The result follows using (2.123) and the expression of $\nabla_u H(t)$ given in (2.133). \hfill \Box

We say that $(\bar{u}, \bar{y}^0)$ is a local minimum point of $J_R$ in $\mathcal{U} \times \mathbb{R}^n$ if

$$J_R(\bar{u}, \bar{y}^0) \leq J_R(u, y^0), \text{ if } ||\bar{u} - u||_\infty + ||\bar{y}^0 - y^0|| \text{ is small enough.}$$  \hfill (2.136)

We define in the same way local minima of arbitrary function, paying attention to the fact that the definition depends on the norm. In the case of the space $\mathcal{U} \times \mathbb{R}^n$, local minimum points are also called \textit{weak minima}. Since the directional derivatives of a function (if they exist) are nonnegative at a local minimum point, we deduce from lemma 2.47 that:

Theorem 2.48. (i) If $\bar{u}$ is a local minimum point of $J_R(\cdot, y^0)$, then

$$H_u(t, \bar{u}(t), \dot{\bar{y}}(t), \bar{p}(t)) = 0, \text{ for a.a. } t \in [0, T].$$  \hfill (2.137)

(ii) If $\bar{y}^0$ is a local minimum point of $J_R(\bar{u}, \cdot)$, then $\bar{p}(0) = 0$.

Example: quadratic regulator. Consider the case with dynamics $f(u, y) := Ay + Bu$, cost integrand $\ell(u, y) := \frac{1}{2}(u^T R u + y^T Q y)$, final cost $\varphi(y) := \frac{1}{2}y^T Q_T y$, and fixed initial state, the matrices $R, Q, Q_T$ being symmetric. The costate equation is

$$\begin{cases} -\dot{\bar{p}}(t) = A^T \bar{p}(t) + Q \bar{y}(t), & \text{for a.a. } t \in [0, T], \\ \bar{p}(T) = Q_T \bar{y}(T), \end{cases}$$  \hfill (2.138)

and the local optimality condition (2.137) reads

$$R \bar{u}(t) + B^T \bar{p}(t) = 0, \text{ for a.a. } t \in [0, T].$$  \hfill (2.139)

If $R$ is invertible, eliminating the control with the previous equation, we see that $(\bar{y}, \bar{p})$ is solution of the two point boundary value problem

$$\begin{cases} \dot{y}(t) = Ay(t) - BR^{-1}B^T \bar{p}(t), & \text{for a.a. } t \in [0, T], \\ -\dot{p}(t) = A^T \bar{p}(t) + Qy(t), & \text{for a.a. } t \in [0, T], \\ y(0) = y^0; \quad p(T) = Q_T y(T). \end{cases}$$  \hfill (2.140)

3.4. Optimal conditions with control constraints. We assume here that the initial state is given, and that there are only control constraints of type (2.124). Eliminating the initial state as argument of the reduced cost, we can write the reduced problem in the form

$$\min_u J_R(u) \quad \text{s.t. } (2.124).$$  \hfill (2.141)

The set of \textit{admissible}, or \textit{feasible} controls is

$$\mathcal{U}_{ad} := \{u \in \mathcal{U}; \ u(t) \in U_{ad} \text{ for a.a. } t\}$$  \hfill (2.142)

We assume in this subsection that

$$U_{ad} \text{ is a closed convex set.}$$  \hfill (2.143)
Lemma 2.49. The set \( U_{ad} \) is a closed convex subset of \( U \).

Proof. The convexity of \( U_{ad} \) follows from the one of \( U_{ad} \). Now let \( u_k \rightarrow \bar{u} \) in \( U \), \( u_k \in U_{ad} \). From any subsequence we can extract another subsequence converging a.e., for which, by the dominated convergence theorem, \( \int_0^T \text{dist}(\dot{u}(t), U_{ad}) dt \) is the limit of \( \int_0^T \text{dist}(u_k(t), U_{ad}) dt \), equal to 0. The result follows. \( \square \)

Proposition 2.50. Let \( \bar{u} \) be a weak minimum (local minimum in \( U \)) of (2.141). Then
\[
\text{if (2.145) holds, the set } I := \{ t \in [0, T] ; H_u(t)(u^k - \bar{u}(t)) < 0 \}. \text{ If (2.145) holds, the set } I := \bigcup_k I_k \text{ has null measure, being a countable union of null measure sets. Over } [0, T] \setminus I, \text{ we have } H_u(t)(u^k - \bar{u}(t)) \geq 0 \text{ for all } k, \text{ and so, (2.144) holds.}
\]

Proof. (a) Let \( \{ u^k \} \) be a dense sequence in \( U_{ad} \). We first check that (2.144) is equivalent to
\[
H_u(t)(u^k - \bar{u}(t)) \geq 0, \text{ for all } k \in \mathbb{N}, \text{ for a.a. } t \in [0, T].
\]
Obviously (2.144) implies (2.145). Conversely, set \( I_k := \{ t \in [0, T] ; H_u(t)(u^k - \bar{u}(t)) < 0 \}. \) If (2.145) holds, the set \( I := \bigcup_k I_k \) has null measure, being a countable union of null measure sets. Over \( [0, T] \setminus I, \) we have \( H_u(t)(u^k - \bar{u}(t)) \geq 0 \) for all \( k, \) and so (2.144) holds.

(b) Let \( u \) be admissible, and \( s \in ]0, 1[. \) Then \( u_* := \bar{u} + s(u - \bar{u}) \) is admissible since it is a convex combination of \( u \) and \( \bar{u}, \) and hence, when \( s \) is small enough, \( J_R(u) \leq J_R(u_*) \). By lemma 2.47:
\[
0 \leq \lim_{s \downarrow 0} \frac{J_R(u_*) - J_R(u)}{s} = J'_R(\bar{u})(u - \bar{u}) = \int_0^T \bar{H}_u(t)(u(t) - \bar{u}(t)) dt.
\]
(c) If (2.145) does not hold, there exists \( E \subset [0, T] \) measurable with positive measure and \( k \in \mathbb{N} \) such that \( H_u(t)(u^k - \bar{u}(t)) < 0 \) a.e. on \( E. \) Define \( u \in U \) by \( u(t) = u^k \) if \( t \in E, \) and \( u(t) = \bar{u} \) otherwise. Then
\[
J'_R(\bar{u})(u - \bar{u}) = \int_E \bar{H}_u(t)(u^k - \bar{u}(t)) dt < 0,
\]
contradicting (2.146). The conclusion follows. \( \square \)

We obtain a characterization of optimality in the case of a convex problem.

Corollary 2.51. Let \( u \mapsto J_R(u) \) be convex. Then \( \bar{u} \) is solution of (2.141) iff (2.144).

Proof. By proposition 2.50, (2.144) is a necessary condition for optimality. Assume now that (2.144) holds. Let \( u \in U_{ad}, \) and set \( v := u - \bar{u}. \) Using the convexity of \( J_R, \) lemma 2.47, and (2.144), we obtain:
\[
J_R(u) - J_R(\bar{u}) \geq D J_R(\bar{u}) v = \int_0^T \bar{H}(t)v(t) dt \geq 0.
\]
The conclusion follows. \( \square \)

Remark 2.52. A sufficient condition for the convexity of \( J_R \) is that \( f \) is an affine function of \( (u, y), \) \( \ell \) is for a.a. \( t \) a convex function of \( (u, y), \) and \( \varphi \) is convex.

Remark 2.53. (i) If \( U_{ad} = \mathbb{R}^m, \) (2.144) boils down to (2.137).
(ii) The relation (2.144) is often written in the form
\[
- \nabla_u \bar{H}(t) \in N_{U_{ad}}(\bar{u}(t)) \text{ for a.a. } t \in [0, T].
\]
We can also say that \( \bar{u}(t), \) for a.a. \( t, \) attains the minimum of \( u \mapsto \bar{H}_u(t) u, \) for \( u \in U_{ad}. \)

We easily extend the previous results to the case when the initial state is not fixed, but subject to a constraint of the form
\[
y_0 \in K_0,
\]
where \( K_0 \) is a convex subset of \( \mathbb{R}^n. \) The optimal control problem can be stated as
\[
\min_{u,y} J_R(u, y) \text{ s.t. (2.124) and (2.150).}
\]
The proof of the next lemma is left as an exercise.

**Lemma 2.54.** If \((\bar{u}, \bar{y}^0)\) is a local solution of (2.151), then (2.144) holds, as well as the condition
\[
(2.152) \quad \bar{p}(0) \cdot (y^0 - \bar{y}^0) \geq 0, \quad \text{for all } y^0 \in K_0.
\]
Conversely, if \(J_R\) is convex and both (2.144) and (2.152) hold, then \((\bar{u}, \bar{y}^0)\) is solution of (2.151).

### 4. Examples

#### 4.1. Multiple integrators.

**Example:** double integrator, quadratic energy. Consider the following problem. The state equation is
\[
(2.153) \quad \dot{h} = v; \quad \dot{v} = u,
\]
with \(h(t)\) the distance, \(v(t)\) the velocity and \(u(t)\) the acceleration. The initial state is given, and the cost function is \(\frac{1}{2} \int_0^T u(t)^2 dt + \varphi(h(T), v(T))\). The pre-Hamiltonian is \(H = \frac{1}{2} u^2 + p_h v + p_v u\). The costate equation is
\[
(2.154) \quad \begin{cases} 
-\dot{p}_h = 0; \\
-\dot{p}_v = p_h; \\
p_h(T) = \nabla_h \varphi(h(T), v(T)); \\
p_v(T) = \nabla_v \varphi(h(T), v(T)).
\end{cases}
\]
So \(p_h\) is constant, and \(u = -p_v\) is an affine function of time. It follows that the optimal state is a cubic function of time.

As an illustration, consider the case when \(\varphi(h, y) := \frac{1}{2} h^2\): the criterion is to minimize distance to 0. Then
\[
(2.155) \quad p_h(t) = p_h(T) = h(T); \quad p_v(T) = 0 \quad \Rightarrow \quad u(t) = -p_v(t) = (t - T)h(T).
\]
Therefore
\[
(2.156) \quad v(t) = v_0 + (\frac{1}{2} t^2 - t T) h(T); \quad h(t) = h_0 + tv_0 + (\frac{1}{6} t^3 - \frac{1}{2} t^2 T) h(T)
\]
Setting \(t = T\) we can compute \(h(T) = (h_0 + T v_0)/(1 + T^2/3)\), which in turn allows to compute the optimal control.

**Exercice 2.55.** Extend the analysis to the case of \(n\) integrations.

#### 4.2. Control constraints.

**Example 2.56.** Bound constraints: \(U_{ad} = \Pi_{i=1}^m [a_i, b_i]\) with \(b_i > a_i\). We have for a.a. \(t\), and for \(i = 1 \text{ to } m: \quad H_{ui}(t) \geq 0\) if \(\bar{u}_i(t) = a_i, \quad \bar{H}_{ui}(t) \leq 0\) if \(\bar{u}_i(t) = b_i, \quad \bar{H}_{ui}(t) = 0\) if \(\bar{u}_i(t) \in ]a_i, b_i[\).

**Example 2.57.** Constraint on Euclidean norm: \(U_{ad} = \bar{B}\) (closed unit ball). We have for a.a. \(t\), if \(\bar{H}_u(t) \neq 0\), then \(\bar{u}(t) = -\bar{H}_u(t)/|\bar{H}_u(t)|\).
1. Pontryagin’s minimum principle

1.1. The easy Pontryagin’s minimum principle

1.1.1. Main result.

Consider again problem (2.141) (with fixed initial state and control constraints) with now $U_{ad}$ a nonempty closed set, possibly nonconvex. Note that initial state is fixed and that there is no constraint on the final state. We will weaken the hypotheses, in order to cover some cases when the dynamics and cost function are non differentiable functions of the control. We assume that

\[
\begin{align*}
 & f \text{ and } \ell \text{ are } C^1 \text{ functions of } y. \text{ The functions } f, f_y, \ell, \ell_y \text{ are } \\
 & \text{Carathéodory functions, uniformly (in time) locally Lipschitz w.r.t. } (u, y), \\
 & \text{and for given } (u, y), \text{ essentially bounded functions of time.} \\
 & \text{The final cost } \varphi \text{ is } C^1 \text{ with a locally Lipschitz gradient.}
\end{align*}
\]
Let \((\tilde{u}, \tilde{y})\) be a trajectory (it satisfies the state equation (2.111)). We recall that the pre-Hamiltonian is the function
\[
H(t, u, y, p) := \ell(t, u, y) + p \cdot f(t, u, y)
\]
and that the costate \(\bar{p}\) is the unique solution of the costate equation (2.130), reproduced below:
\[
\begin{align*}
(i) \quad & -\dot{\bar{p}}(t) = \nabla_y H(t, \bar{u}(t), \bar{y}(t), \bar{p}(t)), \quad \text{for a.a. } t \in [0, T], \\
(ii) \quad & \bar{p}(T) = \nabla \varphi(\bar{y}(T)).
\end{align*}
\]

**Definition 3.1.** We say that \((\tilde{u}, \tilde{y})\) is a Pontryagin extremal if the Hamiltonian inequality below holds:
\[
H(t, \tilde{u}(t), \tilde{y}(t), \bar{p}(t)) \leq H(t, u, \tilde{y}(t), \bar{p}(t)) \quad \text{for all } u \in \mathcal{U}_{\text{ad}}, \quad \text{for a.a. } t \in [0, T].
\]

Note that (3.4) can be written in the form
\[
H(t, \tilde{u}(t), \tilde{y}(t), \bar{p}(t)) = \inf_{u \in \mathcal{U}_{\text{ad}}} H(t, u, \tilde{y}(t), \bar{p}(t)) \quad \text{for a.a. } t \in [0, T].
\]

If the previous relation holds, we also say that \((\bar{u}, \bar{y}, \bar{p})\) or \((\tilde{y}, \tilde{p})\) if no ambiguity is possible is a Pontryagin biextremal.

One easily deduces from the Hamiltonian inequality the following:

**Lemma 3.2.** Assume that \(\mathcal{U}_{\text{ad}}\) is convex. Then a Pontryagin extremal satisfies the first order condition (2.144).

**Proof.** By (3.4), for a.a. \(t\), for any \(u \in \mathcal{U}_{\text{ad}}\), setting \(v := u - \tilde{u}(t)\) we have that:
\[
0 \leq \lim_{\sigma \downarrow 0} \frac{1}{\sigma} (H(t, \tilde{u}(t) + \sigma v, \tilde{y}(t), \bar{p}(t)) - H(t, \tilde{u}(t), \tilde{y}(t), \bar{p}(t)))
\]
\[
= H_u(t, \tilde{u}(t), \tilde{y}(t), \bar{p}(t))(u - \tilde{u}(t)).
\]
The conclusion follows. \(\square\)

Pontryagin’s minimum principle (PMP) is (in our framework) the following statement:

**Theorem 3.3.** Let \(\tilde{u}\) be solution of (2.141). Set \(\bar{y} := y[\tilde{u}]\). Then \((\tilde{u}, \bar{y})\) is a Pontryagin extremal.

We recall that the reduced cost was defined in (2.122) as
\[
J_R(u, y^0) := J(u, y[\bar{u}], y^0).
\]
The proof is based on the following lemma. Remember the definition of the control and state spaces \(\mathcal{U}\) and \(\mathcal{Y}\) in (2.113). In the sequel, for some \(M \geq ||\bar{u}||_\infty\), set
\[
\mathcal{U}^M := \{u \in \mathcal{U}; \; ||u||_\infty \leq M\}.
\]

**Lemma 3.4.** There exists \(c_M > 0\) such that, for all \(u \in \mathcal{U}^M\), we have that for some \(r(u) \in \mathbb{R}\):
\[
J_R(u) = J_R(\bar{u}) + \int_0^T (H(t, u(t), \bar{y}(t), \bar{p}(t)) - \bar{H}(t)) dt + r(u), \quad \text{with } |r(u)| \leq c_M ||u - \bar{u}||_1^2.
\]

**Proof.** Let \(y = y[\bar{u}]\). Integrating by parts and using the costate equation, we obtain
\[
J_R(u) = J_R(\bar{u}) + \int_0^T \bar{p}(t) \cdot (f(u(t), y(t)) - \dot{\bar{y}}(t)) dt = \Delta_1(y) + \Delta_2(u),
\]
with
\[
\begin{align*}
\Delta_1(y) & := \varphi(y(T)) - \varphi'(\bar{y}(T))y(T) + \bar{p}(0) \cdot \dot{\bar{y}}(0), \\
\Delta_2(u) & := \int_0^T (H(t, u(t), y(t), \bar{p}(t)) - \bar{H}(t)y(t)) dt.
\end{align*}
\]
Therefore, \(J_R(u) - J_R(\bar{u}) = \Delta_1(y) - \Delta_1(\bar{y}) + \Delta_2(u) - \Delta_2(\bar{u}).\) We have that \(u(t)\) and consequently \(y(t)\) remain in bounded subsets of \(\mathcal{U}\) and \(\mathcal{Y}\), over which the functions \(f, \ell, \varphi\) and their partial
derivatives w.r.t. the state are uniformly Lipschitz. So, setting \( \delta y(t) := y(t) - \bar{y}(t) \), that satisfies \( \delta y(0) = 0 \):

\[
\Delta_1(y) - \Delta_1(\bar{y}) = \varphi(y(T)) - \varphi(\bar{y}(T)) - \varphi'(\bar{y}(T))\delta y(T) = O(\|\delta y(T)\|)^2.
\]

In addition

\[
\Delta_2(u) - \Delta_2(\bar{u}) = \int_0^T (H(t, u(t), \bar{y}(t), \bar{\bar{p}}(t)) - H(t, u(t), \bar{y}(t), \bar{\bar{p}}(t)) - \bar{H}_{y}(t)\delta y(t))dt + \Delta_3,
\]

where

\[
\Delta_3 := \int_0^T (H(t, u(t), y(t), p(t)) - H(t, u(t), y(t), p(t)) - \bar{H}_{y}(t)\delta y(t))dt.
\]

Since

\[
H(t, u(t), y(t), p(t)) - H(t, u(t), y(t), p(t)) = \left( \int_0^1 H_{y}(t, u(t), y(t) + s\delta y(t), p(t))ds \right)\delta y(t),
\]

we have that

\[
|\Delta_3| = O \left( \|u - \bar{u}\|_1 + \|\delta y\|_\infty \right) \|\delta y\|_\infty).
\]

We conclude by noting that \( \|\delta y\|_\infty = O(\|u - \bar{u}\|_1) \).

**Proof of theorem 3.3.** (a) We first show that, if the trajectory \((u, y)\) is admissible, then

\[
H(t, \bar{u}(t), \bar{y}(t), \bar{p}(t)) \leq H(t, u(t), \bar{y}(t), \bar{p}(t)) \quad \text{a.e. on } [0, T].
\]

Otherwise, there would exist \( \varepsilon > 0 \) and a measurable subset \( E \) of \([0, T]\) with positive measure, such that

\[
H(t, u(t), \bar{y}(t), \bar{p}(t)) > H(t, \bar{u}(t), \bar{y}(t), \bar{p}(t)) - \varepsilon \quad \text{a.e. over } E.
\]

Let \( M := \max(\|u\|_\infty, \|\bar{u}\|_\infty) \), and \( E' \) be a measurable subset of \( E \). Define \( u' \in U \) by \( u'(t) = u(t) \) if \( t \in E' \), and \( u'(t) = \bar{u}(t) \) otherwise. Since \( \|u' - \bar{u}\|_\infty \leq 2M \), we have that \( \|u' - \bar{u}\|_1 \leq 2M \text{ meas}(E') \). By Lemma 3.4,

\[
J_R(u') \leq J_R(\bar{u}) - \varepsilon \text{ meas}(E') + 4cM^2 \text{ meas}(E')^2.
\]

This gives a contradiction by choosing \( E' \) with a sufficiently small, positive measure.

(b) Let \( u^k \) be a dense sequence in \( U_{ad} \), and \( u^k \) be the sequence in \( U \) defined by \( u^0 := \bar{u} \), and

\[
u^k(t) = \begin{cases} u^k & \text{if } H(t, u^k, \bar{y}(t), \bar{p}(t)) < H(t, u^{k-1}(t), \bar{y}(t), \bar{p}(t)), \\ u^{k-1}(t) & \text{otherwise.} \end{cases}
\]

The density property of \( u^k \) implies that \( H(t, u^k(t), \bar{y}(t), \bar{p}(t)) \to \inf_{u \in U_{ad}} H(t, u, \bar{y}(t), \bar{p}(t)) \) a.e. So, taking \( u = u^k \) in (3.14) and passing to the limit in this inequality we get the conclusion. \( \square \)

**Remark 3.5.** If \( U_{ad} = \mathbb{R}^m \), the Hamiltonian inequality (3.4) implies the following **Legendre-Clebsch** condition:

\[
\bar{H}_{uu}(t) \succeq 0 \quad \text{for a.a. } t \in [0, T].
\]

We next introduce a new notion of minimum.

**Definition 3.6.** A sequence \( u^k \) of \( U \) is said to converge to \( \bar{u} \) in **Pontryagin's sense** if there exists \( M > 0 \) such that \( \|u^k\|_\infty \leq M \), and that \( u^k \to \bar{u} \) a.e. We say that \( \bar{u} \) is a **Pontryagin minimum** if \( u^k \to \bar{u} \) in Pontryagin’s sense implies that \( J_R(\bar{u}) \leq J_R(u^k) \) for \( k \) large enough.

**Remark 3.7.** If \( u^k \to \bar{u} \) in Pontryagin’s sense, by the dominated convergence theorem, \( u^k \to \bar{u} \) in \( L^s(0, T)^m \) for all \( s \in [1, \infty[ \).

**Remark 3.8.** One easily checks that the proof of theorem 3.3 holds for a Pontryagin minimum. With any Pontryagin minimum is therefore associated a Pontryagin extremal.
1.1.2. A related nonlinear expansion of the state. We consider the nonlinear relation
\[
(3.19) \quad \dot{\xi}(t) = f_y(t, \xi(t)) + f(t, u(t), y(t)) - f(t, \bar{u}(t), \bar{y}(t)), \quad \text{for a.a. } t \in (0, T),
\]
with \( u \in \mathcal{U} \), and initial condition \( \xi(0) = 0 \), whose solution in \( \mathcal{Y} \) may viewed as a function of \( u \), denoted by \( \xi[u] \).

**Lemma 3.9.** For every \( M > 0 \), there exists \( c_M > 0 \) such that, if \( \|u\|_\infty \leq M \), then
\[
(3.20) \quad \|y[u] - \bar{y} - \xi\|_\infty \leq c_M \|u - \bar{u}\|_1.
\]

**Proof.** We may assume that the data are autonomous. Denote \( y[u] \) by \( y \), and let \( \eta := y - \bar{y} - \xi \). Then, skipping time arguments
\[
(3.21) \quad \dot{\eta} = f(u, y) - f(u, \bar{y}) - \bar{\xi}
\]
where
\[
(3.22) \quad A = A(t) := \int_0^1 (f_y(u, \bar{y} + \sigma(y - \bar{y})) - \bar{f}_y) d\sigma
\]
is a \( n \times n \) matrix, and since the state is uniformly bounded whenever \( \|u\|_\infty \leq M \), there exists \( c'_M > 0 \) not depending on \( u \) such that
\[
(3.23) \quad |A(t)| \leq c'_M(\|u(t) - \bar{u}(t)\| + \|y(t) - \bar{y}(t)\|).
\]
Since \( \|y - \bar{y}(t)\|_\infty = O(\|u - \bar{u}(t)\|_1) \), the conclusion follows with Gronwall’s lemma. \( \square \)

**Alternative proof of Lemma 3.4.** It is enough to consider the autonomous case, with only a final cost. Then, by (2.114) and (3.20):
\[
(3.24) \quad J_R(u) - J_R(\bar{u}) = \nabla \varphi(\bar{y}(T)) \cdot (y(T) - \bar{y}(T)) + O(\|y(T) - \bar{y}(T)\|^2),
\]
Also,
\[
(3.25) \quad \bar{p}(T) \cdot \xi(T) = \int_0^T \left( \bar{p}(T) \cdot \dot{\xi}(T) + \bar{p}(T) \cdot \dot{\xi}(T) \right) dt = \int_0^T (\bar{H}(t, u(t), \bar{y}(t), \bar{p}(t)) - \bar{H}(t)) dt.
\]
The result follows. \( \square \)

1.2. Pontryagin’s principle with two point constraints.

1.2.1. **Main result.** We now allow the initial and final state to vary, subject to some constraints that may couple them (as in the case of periodic trajectories). So, consider the cost function
\[
(3.26) \quad J^{IF}(u, y) := \int_0^T \ell(t, u(t), y(t)) dt + \varphi(y(0), y(T)),
\]
where ‘IF’ stands for initial-final, and the two point constraints
\[
(3.27) \quad \Phi(y(0), y(T)) \in K_\Phi.
\]
The hypotheses on the data are as follows:
\[
(3.28) \quad \left\{ \begin{array}{l}
\text{Hypothesis (3.1) holds and } \varphi, \Phi \text{ are } C^1 \text{ with locally Lipschitz derivatives.} \\
\text{The set } K_\Phi \text{ is a nonempty, closed convex subset of } \mathbb{R}^{n_\Phi}.
\end{array} \right.
\]
Consider the optimal control problem
\[
(3.29) \quad \left\{ \begin{array}{l}
\text{Min}_{u, y} J^{IF}(u, y); \quad \text{s.t. the state equation (2.111)(i),} \\
\text{control constraints (2.124), and (3.27).}
\end{array} \right.
\]
The pre-Hamiltonian, now expressed in non qualified form, is the function $H : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$H(\beta, t, u, y, p) := \beta \ell(t, u, y) + p \cdot f(t, u, y).$$

We call $\beta$ the cost multiplier. Denoting by $\Psi \in \mathbb{R}^{n*}$ the multiplier associated with the two point constraints. The end points Lagrangian is $L^{IF} : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n*}$ defined by

$$(3.31) \quad L^{IF}(\beta, y^0, y^T, \Psi) := \beta \varphi(y^0, y^T) + \Psi \cdot \Phi(y^0, y^T).$$

The Lagrangian of problem (3.29) (sum of weighted cost function and constraints) is

$$(3.32) \quad L(\beta, u, y, p, \Psi) := \beta J^{IF}(u, y) + \int_0^T p(t) \cdot (f(t, u(t), y(t)) - \dot{y}(t))dt + \Psi \cdot \Phi(y(0), y(T))$$

As usually we will obtain a costate equation by expressing the condition of stationarity (i.e., zero partial derivative) of the Lagrangian w.r.t. the state. After an integration by parts (valid since we assume that $p \in \mathcal{Y}$), this boils down to the fact that, for all $z \in \mathcal{Y}$:

$$D_y L z = \int_0^T \left( \nabla_y H(\beta, t, \bar{u}(t), \bar{y}(t), p(t)) + \dot{p}(t) \right) \cdot z(t) dt$$

$$+ (\nabla_y L^{IF}(\beta, \bar{y}(0), \bar{y}(T), \Psi) + \bar{p}(0)) \cdot z(0)$$

$$+ (\nabla_y L^{IF}(\beta, \bar{y}(0), \bar{y}(T), \Psi) - \bar{p}(T)) \cdot z(T) = 0.$$

By the same arguments than those following (2.128), the coefficients of the linearized state on each line must be zero. So, let us define the costate equation as

$$(3.34) \quad \begin{cases} (i) & -\dot{p}(t) = \nabla_y H(\beta, t, \bar{u}(t), \bar{y}(t), p(t)), \text{ for a.a. } t \in [0, T], \\ (ii) & -\dot{p}(0) = \nabla_y L^{IF}(\beta, \bar{y}(0), \bar{y}(T), \Psi), \\ (iii) & \bar{p}(T) = \nabla_y L^{IF}(\beta, \bar{y}(0), \bar{y}(T), \Psi). \end{cases}$$

**Remark 3.10.** For historical reasons, the above two last lines are called transversality conditions.

**Remark 3.11.** By the costate equation, if $z = z[v, z^0]$, the solution of the linearized state equation (2.115), then

$$\tilde{\beta} \varphi'(\tilde{y}(0), \tilde{y}(T))(z^0, z(T)) + \bar{\Psi} \cdot \Phi'(\tilde{y}(0), \tilde{y}(T))(z^0, z(T))$$

$$+ \bar{\beta} \int_0^T \ell'(\bar{u}(t), \bar{y}(t))(v(t), z(T))dt = \int_0^T \bar{H}_u(t)v(t)dt.$$

**Definition 3.12.** We call Pontryagin multiplier associated to the nominal trajectory $(\bar{u}, \bar{y})$, any triple $\lambda := (\tilde{\beta}, \bar{\Psi}, \bar{p})$ verifying (3.34), as well as

$$\tilde{\beta} \in \{0, 1\}, \quad \bar{\Psi} \in N_{K_{\pi}}(\Phi(\tilde{y}(0), \tilde{y}(T))),$$

the non nullity relation

$$(3.37) \quad \tilde{\beta} + |\bar{\Psi}| > 0,$$

and the Hamiltonian inequality that generalizes (3.5):

$$(3.38) \quad H(\tilde{\beta}, \bar{u}(t), \bar{y}(t), p(t)) = \inf_{u \in U_{\text{ad}}} H(\tilde{\beta}, u, \bar{y}(t), p(t)) \quad \text{for a.a. } t \in (0, T).$$

We say that $(\bar{u}, \bar{y})$ is a Pontryagin extremal if the set $\Lambda_P(\bar{u}, \bar{y})$ of associated Pontryagin multipliers is non empty.

Relation (3.37) prevents a multiplier to be zero. We say that an element of $\Lambda_P(\bar{u}, \bar{y})$ is a singular or abnormal multiplier if $\tilde{\beta} = 0$, and a normal (Pontryagin) multiplier if $\tilde{\beta} = 1$.

**Theorem 3.13.** A solution of (3.29) is a Pontryagin extremal.

**Proof.** The proof being technical, we postpone it to section 4.
Example 3.14. Let us apply theorem 3.13 to the problem of minimizing $\int_0^1 u(t)(1-u(t))dt$ s.t. $\dot{y}(t) = u(t) \in [0,1]$ for a.a. $t \in (0,1)$, $y(0) = 0$, and $y(0) = 1/2$. The pre-Hamiltonian is $H = \beta u(1-u) + pu$ so that $-\dot{p}(t) = 0$ a.e., whence $\dot{p}(t) = \Psi$. If $\beta = 0$ then $\Psi \neq 0$ and so, $u(t)$ is constant, either equal to 0 or 1, in contradiction with the initial-final state constraints. So, $\beta = 1$, $\Psi = 0$, and since $H$ is a strictly concave function of the control, we have that $\dot{u}(t) \in \{0,1\}$ for a.a. $t$. By the initial-final state constraint, $\text{meas} \{t; \dot{u}(t) = 0\} = \text{meas} \{t; \dot{u}(t) = 1\} = 1/2$. In this simple nonconvex example, one can check that any control satisfying these necessary conditions is optimal.

1.2.2. Specific structures. We next consider several examples that illustrate the above theorem.

Example 3.15. Assume that we have finitely many initial-final equality and inequality constraints, i.e. (with obvious interpretation if either $n_1$ or $n_2$ equals zero):

$$\begin{align*}
\Phi_i(y(0), y(T)) &= 0, \quad i = 1, \ldots, n_1, \\
\Phi_i(y(0), y(T)) &\leq 0, \quad i = n_1 + 1, \ldots, n_1 + n_2.
\end{align*}$$

This corresponds to the choice

$$K_\Phi := \{0\}^{n_1} \times \mathbb{R}^{n_2}.$$

In that case we easily check that $\Psi \in N_{K_\Phi}(\Phi(y(0), y(T)))$ if and only if the following sign and complementarity conditions for inequality constraints hold:

$$\Psi_i \geq 0, \quad \Psi_i \Phi_i(y(0), y(T)) = 0, \quad i = n_1 + 1, \ldots, n_1 + n_2.$$

Example 3.16. Problem (2.141), that has a fixed initial state and a free final state, is a particular case of the present setting where $\varphi$ depends only on the final state, $\Phi(y^0, y^T) = y^0$, and $K_\Phi$ is the singleton $\{y^0\}$. The two point conditions of the costate equation reduce then to

$$\dot{p}(T) = \beta \nabla \varphi(y(T)); \quad p(0) = -\bar{\Psi}.$$

The last relation expresses $\bar{\Psi}$ as a function of $\dot{p}(0)$. If $\bar{\beta}$ was equal to 0, we would have that $\dot{p}(T) = 0$, and then $p = 0$ by the costate equation, so that $\bar{\Psi} = 0$, contradicting (3.37). Therefore, there exists a regular Pontryagin multiplier, and no singular Pontryagin multiplier. We recover the conclusion of theorem 3.3.

Example 3.17. More generally, if the initial state is fixed, we may write $\varphi$ as function of the final state only, and assume the two point constraints of the form

$$\Phi(y^0, y^T) = (y^0, \Phi_F(y^T)); \quad K_\Phi = \{y^0\} \times K_F,$$

with $\Phi_F : \mathbb{R}^n \to \mathbb{R}^{n_F}$ and $K_F$ convex and closed subset of $\mathbb{R}^{n_F}$. Denoting by $(\Psi_0, \bar{\Psi}_F) \in \mathbb{R}^n \times \mathbb{R}^{n_F}$ the components of the multiplier $\Psi$ associated with the initial and final state constraint, we may replace the transversality conditions (3.34(ii)-(iii)) by

$$\begin{align*}
\dot{p}(T) &= \bar{\beta} \nabla \varphi(y(T)) + D\Phi_F(y(T))\bar{\Psi}_F; \\
\dot{\Psi}_F &\in N_{K_F}(\Phi_F(y(T))), \\
\bar{p}(0) &= \Psi_0.
\end{align*}$$

We see then that we can replace the non nullity condition (3.37) by

$$\bar{\beta} + |\bar{\Psi}_F| > 0.$$
Example 3.18. Another generalization of example 3.16 is when we assume only that the final state is free, and we have the initial constraint $\Phi(y(0)) \in K_\Phi$. If $\beta = 0$, then $\bar{p}(T) = 0$, and by the costate equation, $\dot{\bar{p}}(t) = 0$ for all $t \in [0, T]$. So, $0 = \bar{p}(0) = D\Phi(y(0))\hat{\Psi}$, with $0 \neq \hat{\Psi} \in N_{K_\Phi}(y(0))$. If $D\Phi(y(0))$ is surjective, then its transpose is injective and this gives a contradiction with (3.37). So, if $D\Phi(y(0))$ is surjective, there is no singular Pontryagin multiplier, and there exists a regular one, i.e., we must have $\beta = 1$.

1.2.3. Link with Lagrange multipliers; convex problems. When $U_{ad}$ is convex, we easily deduce from the Hamiltonian inequality (3.38) that

$$
\begin{align*}
\{ & \text{For a.a. } t \in [0, T], \text{ we have that} \\
& H_u(\beta, t, \bar{u}(t), \bar{y}(t), \bar{p}(t))(u - \bar{u}(t)) \geq 0, \text{ for all } u \in U_{ad}. 
\}
\end{align*}
$$

Definition 3.19. When $U_{ad}$ is convex, we define the set $\Lambda_L(\bar{u}, \bar{y})$ of Lagrange multipliers as the set of triple $\lambda := (\beta, \hat{\Psi}, \bar{p})$ satisfying (3.34)-(3.37) and (3.46). If $\beta = 0$ (resp. $\beta = 1$) we say that the Lagrange multiplier is singular (resp. normal).

Since (3.38) implies (3.46), the following holds:

$$
(3.47) \quad \text{If } U_{ad} \text{ is convex, then } \Lambda_P(\bar{u}, \bar{y}) \subset \Lambda_L(\bar{u}, \bar{y}).
$$

Lemma 3.20. Let $\lambda := (\beta, \hat{\Psi}, \bar{p})$ be a regular ($\beta = 1$) Lagrange multiplier, such that the associated Lagrangian defined in (3.32) is a convex function of $(u, y)$. Then $(\bar{u}, \bar{y})$ is solution of problem (3.29).

Proof. Let $(u, y)$ satisfy the constraints of problem (3.29). Skipping the argument $\beta$, we get:

$$
J^{IF}(u, y) - J^{IF}(\bar{u}, \bar{y}) \geq L(u, y, \bar{p}, \hat{\Psi}) - L(\bar{u}, \bar{y}, \bar{p}, \hat{\Psi})
$$

$$
\geq D_uL(\bar{u}, \bar{y}, \bar{p}, \hat{\Psi})(u - \bar{u}) = \int_0^T H(t)(u(t) - \bar{u}(t))dt \geq 0.
$$

The first inequality follows from the state constraint and the relation

$$
\hat{\Psi} \cdot (\Phi(y(0), y(T)) - \Phi(\bar{y}(0), \bar{y}(T))) \leq 0,
$$

which holds since $\hat{\Psi} \in N_{K_\Phi}(\Phi(y(0), y(T)))$ and $\Phi(y(0), y(T)) \in K_\Phi$. The second inequality follows from the convexity hypothesis for the Lagrangian, remembering that the partial derivative of the Lagrangian w.r.t. the state variable is zero (this is how is derived the costate equation), and then we use the definition of a Lagrange multiplier. The result follows.

Remark 3.21. For a sufficient condition for the convexity hypothesis on the Lagrangian we can make the same hypotheses as in remark 2.52. It remains then to check that

$$
(3.50) \quad (y^0, y^T) \mapsto \hat{\Psi} \cdot \Phi(y^0, y^T) \text{ is convex.}
$$

This always holds if $\Phi$ is an affine mapping. In the case of $n_1$ equality constraints and $n_2$ inequality constraints on the initial and final state, i.e. when $K_\Phi$ is of the form (3.39), since $\hat{\Psi}_i \geq 0$ whenever $i > n_1$ by (3.41), we see that (3.50) holds if the first $n_1$ components of $\Phi$ are affine, and the remaining ones are convex.

1.3. Variations of the pre-Hamiltonian.

1.3.1. Pontryagin extremal over $(0, T)$. Let $(\bar{u}, \bar{y})$ be a trajectory, and $\bar{p} \in \mathcal{Y}$ satisfy the first part (3.34)(i) of the costate equation. Set

$$
(3.51) \quad \bar{h}(t) := \inf_{u \in U_{ad}} H(\beta, t, u, \bar{y}(t), \bar{p}(t)).
$$

We say that $(\bar{u}, \bar{y}, \bar{p})$ is a Pontryagin extremal over $(0, T)$ (rather than $[0, T]$) if the Hamiltonian inequality (3.38) holds, that is, if

$$
(3.52) \quad \bar{h}(t) = \bar{H}(t), \quad \text{for a.a. } t.
$$
So, we forget the constraints on initial-final state as well as the transversality conditions for the costate. Set 
\begin{equation}
\dot{U}_{ad} := U_{ad} \cap B(0, \|\bar{u}\|_\infty); \quad \hat{h}(t) := \inf_{u \in U_{ad}} H(\beta, t, u, \bar{y}(t), \bar{p}(t)).
\end{equation}
Then \(\hat{U}_{ad}\) is compact, and so, since \(|\bar{u}(t)| \leq \|\bar{u}\|_\infty\) a.e.: 
\begin{equation}
\hat{h}(t) = \hat{h}(t) \quad \text{for a.a. } t \in [0, T].
\end{equation}

**Remark 3.22.** Note that \(y\) and \(\bar{p}\) are Lipschitz functions of time. It easily follows from (3.54) that if \(f(\cdot, u, y)\) has a continuity modulus (resp. Lipschitz constant) uniformly in time, over bounded sets, then \(\hat{h}\) is continuous (resp. Lipschitz).

1.3.2. **Constant pre-Hamiltonian for autonomous problem.** We say that the optimal control problem is autonomous if \(f\) and \(\ell\) do not depend on time, which then is also the case for the pre-Hamiltonian. In that case we say that these functions are autonomous, and we may skip the time from their arguments.

**Lemma 3.23.** Let \((\bar{u}, \bar{y}, \bar{p})\) be a Pontryagin extremal over \((0, T)\) of an autonomous problem. Then \(h(t)\) is, up to a null measure set, equal to some constant over \((0, T)\).

**Proof.** a) Consider first the easy case when \(\bar{u}(t)\) is differentiable over \((0, T)\), in the absence of control constraints. Then \(H_u(t) = 0\) for all \(t \in (0, T)\), and (using the state and costate equations) \(h(t)\) has a continuous derivative satisfying 
\begin{equation}
\dot{h}(t) = \bar{H}_y(t)\dot{y}(t) + \bar{H}_p(t)\dot{\bar{p}}(t) = -\dot{\bar{p}}(t) \cdot \dot{y}(t) + \dot{y}(t) \cdot \dot{\bar{p}}(t) = 0,
\end{equation}
from which the result follows.

b) We now now with the general case. By remark 3.22, \(\hat{h}\) is a Lipschitz function; it is can be shown that this is equivalent to say that \(\hat{h} \in W^{1,\infty}(0, T, \mathbb{R}^n)\). By lemma 2.45, \(\bar{h}\) is the primitive of its weak derivative, and it suffices to show that \(\bar{h}\) has zero derivative for a.a. \(t\). For a.a. \(t_0 \in (0, T)\), we have that \((\bar{h}, \bar{y}, \bar{p})\) are differentiable at \(t_0\), \(H(\beta, \cdot, \bar{y}(t_0), \bar{p}(t_0))\) attains its minimum over \(U_{ad}'\) at \(\bar{u}(t_0)\), and the state and costate equation hold (at time \(t_0\)). Therefore, 
\begin{equation}
\bar{h}(t_0) \leq \lim_{t \downarrow t_0} \frac{H(\beta, \bar{u}(t_0), \bar{y}(t_0), \bar{p}(t_0)) - H(\beta, \bar{u}(t_0), \bar{y}(t_0), \bar{p}(t_0))}{t - t_0} = H_y(\beta, \bar{u}(t_0), \bar{y}(t_0))\dot{y}(t_0) + H_p(\beta, \bar{u}(t_0), \bar{y}(t_0))\ddot{\bar{p}}(t_0)
\end{equation}
By taking \(t \uparrow t_0\) we would prove the opposite inequality in the same way. The result follows.

**Remark 3.24.** In the case of state constrained problems, we will see in lemma 5.7 another method of proof based on a time dilation approach.

1.3.3. **Variation of the pre-Hamiltonian for non autonomous problem.** We may view the time as an additional state variable say \(\tau\) with initial condition zero and dynamics \(\dot{\tau} = 1\). In this new setting, \(f = f(\tau, u, y)\) and \(\ell = \ell(\tau, u, y)\) become autonomous functions. In view of (3.1), we need to assume that 
\begin{equation}
\begin{cases}
f \text{ and } \ell \text{ are } C^1 \text{ functions of } (t, y). \\
The functions } f, f_y, f_t, \ell, \ell_y, \ell_t \text{ are locally Lipschitz.} \\
\psi \text{ and } \Phi \text{ are } C^1 \text{ with locally Lipschitz derivatives.}
\end{cases}
\end{equation}
The new pre-Hamiltonian is \(H(\beta, \tau, u, y, p) + q\), the additional costate \(\bar{q}\) being solution of 
\begin{equation}
-\dot{\bar{q}}(t) = \bar{H}_t(t), \quad \text{for a.a. } t \in (0, T), \quad \bar{q}(T) = 0.
\end{equation}
Since \(\tau(0)\) is fixed, \(\bar{q}(0)\) is equal to the associated Lagrange multiplier (or its opposite) which gives no information. So, \(\bar{q}(t) = \int_0^T \bar{H}_t(s) ds\). In this autonomous reformulation, by lemma 3.23, the pre-Hamiltonian has constant value, equal to \(\bar{H}(t) + \int_0^T \bar{H}_t(s) ds\). Equivalently, the total and partial derivative of the pre-Hamiltonian along the trajectory are equal (and bounded), so that:


**Lemma 3.25.** Let (3.57) hold. Then we have that \( \bar{h}(t) \) is Lipschitz, and

\[
\frac{d}{dt} \bar{h}(t) = \bar{H}(t), \quad \text{a.e. on } (0, T).
\]

### 2. Extensions

**2.1. Decision variables.** Quite often some variables not depending on time have to be optimized as well as the control, think for instance of the mass of fuel to be taken by an airplane, or some design variables. So, consider, for \((u, y, \pi) \in U \times Y \times \mathbb{R}^{n_u}\), the cost function in this augmented setting is (we choose to express the initial-final cost as pre-Hamiltonian and)

\[
J(u, y, \pi) := \int_0^T \ell(t, u(t), y(t), \pi) dt + \varphi(y(0), y(T), \pi),
\]

and the state equation

\[
\dot{y}(t) = f(t, u(t), y(t), \pi), \quad \text{for a.a. } t \in [0, T].
\]

We also have initial-final state constraints

\[
\Phi(y(0), y(T), \pi) \in K_\Phi.
\]

Here \( \ell \) and \( f \) are mappings from \( \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n_u} \) into \( \mathbb{R} \) and \( \mathbb{R}^n \) resp., and \( \varphi : \mathbb{R}^n \times \mathbb{R}^{n_u} \to \mathbb{R} \).

We also have local (in time) control constraints of the type

\[
u(t) \in U_{ad}, \quad \text{for a.a. } t \in [0, T].
\]

The hypotheses on the data are as follows:

\[
\begin{aligned}
f \text{ and } \ell \text{ are } C^1 \text{ functions of } (y, \pi); \text{ the mappings } f, f_y, f_\pi, \ell, \ell_y, \ell_\pi \text{ are Carathéodory functions, uniformly locally Lipschitz w.r.t. } (u, y, \pi), \\
\varphi, \Phi \text{ are } C^1 \text{ with locally Lipschitz derivatives.} \\
\text{The set } K_\Phi \text{ is a nonempty, closed convex subset of } \mathbb{R}^{n_u}. \\
U_{ad} \text{ is a nonempty closed subset of } \mathbb{R}^m.
\end{aligned}
\]

The optimal control problem is

\[
\text{Min } J(u, y, \pi) \text{ s.t. } (3.61)-(3.63).
\]

We can reduce this problem to the general format (3.29) by interpreting \( \pi \) as an additional state variable denoted by \( \pi \), with zero dynamics. Then the augmented state is \( y^a := (y^\dagger, \pi)^\dagger \).

Identifying in the sequel \( y \) with the first \( n \) components of the augmented state, we can write the state equation for \( y^a \) as

\[
\dot{y}^a(t) = \begin{pmatrix}
f(t, u(t), y(t), \pi(t)) \\
0
\end{pmatrix} \quad \text{for a.a. } t \in [0, T].
\]

The cost function in this augmented setting is (we choose to express the initial-final cost as function of the final value of \( \pi \)):

\[
J^a(u, y^a, \pi) := \int_0^T \ell(t, u(t), y(t), \pi(t)) dt + \varphi(y(T), y(T), \pi(T)).
\]

The resulting optimal control problem is

\[
\text{Min } J^a(u, y^a, \pi) \text{ s.t. } (3.62)-(3.63) \text{ and } (3.66).
\]

We next discuss the optimality conditions of this problem. We have an augmented costate \( p^a := (p, q) \), where \( p(t) \in \mathbb{R}^n \) and \( q(t) \in \mathbb{R} \) are costates associated with \( y \) and \( \pi \), resp. The pre-Hamiltonian and end points Lagrangian in this setting are resp.

\[
\begin{aligned}
H(\beta, t, u, y, \pi, p, q) &:= \beta \ell(t, u, y, \pi) + p \cdot f(t, u, y, \pi), \\
L^IF(\beta, y^0, y^T, \pi, \Psi) &:= \beta \varphi(y^0, y^T, \pi) + \Psi \cdot \Phi(y^0, y^T, \pi).
\end{aligned}
\]

43
Since $\pi(t) = \bar{\pi}$ for all $t$, the costate equation (3.34) takes here the expression

$$
\begin{align*}
-\dot{p}(t) &= \nabla_y H(\beta, t, \bar{u}(t), \bar{y}(t), \bar{\pi}, \bar{p}(t)), & \text{for a.a. } t \in [0, T], \\
-\dot{q}(t) &= \nabla_\pi H(\beta, t, \bar{u}(t), \bar{y}(t), \bar{\pi}, \bar{p}(t)), & \text{for a.a. } t \in [0, T], \\
-\bar{p}(0) &= \nabla_y L^{IF}(\beta, \bar{y}(0), \bar{y}(T), \bar{\pi}, \bar{\Psi}), \\
\bar{p}(T) &= \nabla_\pi L^{IF}(\beta, \bar{y}(0), \bar{y}(T), \bar{\pi}, \bar{\Psi}), \\
\bar{q}(0) &= 0, \\
\bar{q}(T) &= \nabla_\pi L^{IF}(\beta, \bar{y}(0), \bar{y}(T), \bar{\pi}, \bar{\Psi}).
\end{align*}
$$

(3.70)

It follows that

$$
0 = \bar{q}(0) = -\int_0^T \dot{q}(t) dt + \bar{q}(T)
$$

(3.71)

$$
= \int_0^T \nabla_\pi H(\beta, t, \bar{u}(t), \bar{y}(t), \bar{\pi}, \bar{p}(t)) dt + \nabla_\pi L^{IF}(\beta, \bar{y}(0), \bar{y}(T), \bar{\pi}, \bar{\Psi}).
$$

This is nothing but the condition of stationarity of the Lagrangian of the optimal control problem w.r.t $\pi$. We also have the Hamiltonian inequality for the control variables. We have proved the following:

**Lemma 3.26.** For the optimal control problem (3.65) with decision variables, the PMP holds with the same expression as for the corresponding problem with decision variables set at their nominal value, with in addition the condition of stationarity of the Lagrangian w.r.t the decision variables.

**Exercise 3.27.** Show that we obtain the same result when in (3.67) we choose to express the initial-final cost function as $\varphi(y(0), y(T), \pi(0))$.

2.1.1. A design problem. Consider the following model of a ground vehicle with rectilinear trajectory:

$$
\dot{h} = v; \quad \dot{v} = (eu - D(\gamma, v))/m; \quad \dot{m} = -u,
$$

(3.72)

where $h$ is the distance to the initial position, $v$ is the velocity, $m$ is the (positive) mass, the control is the mass flow $u \in [0, U]$, with $U > 0$, $e > 0$ is the ejection speed, and $D(v, \gamma)$ is the drag force, function of the speed and of a design variable (think of a parametrized shape of the vehicle) $\gamma \in \mathbb{R}$. We assume that $D(v, \gamma)$ and $D_v(v, \gamma)$ have positive values (when $v(t) > 0$ which will be the case by hypothesis).

The initial state is given, and we have a final state constraint

$$
h(0) = 0; \quad v(0) = 0; \quad m(0) = m_0 > 0; \quad m(T) \geq m^T.
$$

(3.73)

We want to maximize the final distance, but have a cost $c(\gamma)$ for the design, and so we decide to minimize $rc(\gamma) - h(T)$ where $r > 0$ is a parameter.

We next express the PMP. The pre-Hamiltonian is

$$
H = phv + pv(eu - D(\gamma, v))/m - pmu.
$$

(3.74)

The costate equation is

$$
\begin{align*}
-\dot{p}_h &= H_h = 0; \\
-\dot{p}_v &= H_v = ph - pv D_v/m; \\
-\dot{p}_m &= H_m = -pv(eu - D)/m^2; \\
p_h(T) &= -\beta, \\
p_v(T) &= 0, \\
p_m(T) &= -\Psi \leq 0.
\end{align*}
$$

(3.75)

If $\beta = 0$ then $p_h(t) = 0$ for all $t$, so that $p_v/p_m$ is bounded along the trajectory. Since $p_v(T) = 0$, it follows that $p_v(t) = 0$ for all $t$, whence $-p_m = 0$, implying $p_m(t) = p_m(T) = -\Psi < 0$, so that

$$
H_u = ep_v/m - p_m
$$

is constant and positive, meaning that $u(t) = 0$ for all time, which is optimal only when $m^T = m_0$. In the sequel we exclude this case and therefore we may assume that $\beta = 1$. 

44
Obviously \( p_v(t) = -1 \) for all \( t \), and we have that
\[
(3.77) \quad \dot{p}_v = 1 + p_v D_v/m; \quad \dot{p}_v(T) = 1; \quad \dot{p}_m(T) = 0.
\]
Since \( \dot{p}_v(T) = 1 \) and \( p_v(T) = 0 \), \( p_v(t) < 0 \) for \( t \) close to \( T \). As \( \dot{p}_v > 0 \) when \( p_v \) is close to zero, necessarily
\[
(3.78) \quad \text{One has } p_v(t) < 0, \text{ for all } t \in [0,T).
\]
It follows that \( \dot{p}_m \) has the sign of \( D - e u \). Since \( p_v(T) = 0 \), \( \dot{p}_m(T) = 0 \), and so:
\[
(3.79) \quad H_u(T) = \Psi \geq 0; \quad \dot{H}_u(T) = e/m(T) > 0.
\]
**Case when \( \Psi = 0 \).** Then \( H_u(t) < 0 \) for \( t \) close to \( T \), and so, the trajectory has a final full thrust arc, say over \( (t_0,T) \). We assume that \( D \leq e U \) (i.e., the full thrust allows to increase the speed) over the trajectory. 

Over the final full thrust arc, \( p_m \) is positive since it decreases with final condition zero. Reminding that \( p_v(t) < 0 \) for all \( t \), we see that \( H_u(t_0) < -p_m(t_0) < 0 \) so that \( t_0 = 0 \): the optimal policy is the full thrust over \( (0,T) \).

**Case when \( \Psi > 0 \).** Then for \( t \) close to \( T \), \( H_u > 0 \) and so \( u = 0 \): the trajectory ends with a zero thrust arc (over which the mass equals \( m T \)) say \( (t_0,T) \). Over this arc \( p_m \) increases.

**Optimality w.r.t. \( \gamma \).** The condition of stationarity of the Lagrangian w.r.t. \( \gamma \) gives
\[
(3.80) \quad rc'(\gamma) - \int_0^T p_v(t) \gamma \dot{v}(t) dt = 0.
\]
We may assume that, for instance, \( D(\gamma,v) = \frac{1}{2} \gamma v^2 \), i.e., \( \gamma \) is the aerodynamic coefficient, and that \( c'(\gamma) < 0 \) (it is expensive to achieve a small aerodynamic coefficient). Then the above equation reads
\[
(3.81) \quad rc'(\gamma) = \frac{1}{2} \int_0^T \frac{p_v(t)}{m(t)} v(t)^2 dt.
\]
As expected, in view of (3.78), both sides have the same negative sign.

**2.2. Variable horizon.** In many problems, the final time is itself an optimization parameter. For instance, in the minimum time problem, we aim to reach a certain target in a minimum time. So, consider the case when the initial time is zero and the initial state is fixed, the final state being free, and assume that \( T \) is a decision variable. The original problem is
\[
\begin{align*}
\int & \ell(t, u(t), y(t)) dt + \varphi(y(T)) \\
\text{s.t.} & y(t) = f(t, u(t), y(t)) \text{ a.e. in } (0,T); \; y(0) = y^0, \\
& u(t) \in U_{ad} \text{ a.e.}, \quad \Phi_F(y(T)) \in K_F.
\end{align*}
\]
We introduce the *normalized time* \( \tau := t/T \), so that \( \tau \in [0,1] \). Then \( y^N(\tau) := y(\tau T) \) is the *time normalized state*, with a similar conventions for other variables such as the control. Then the normalized control and state satisfy the (normalized) state equation
\[
(3.83) \quad y^N(\tau) = T f(T \tau, u^N(\tau), y^N(\tau)), \quad \tau \in [0,1]; \; y^N(0) = y^0.
\]
On the other hand the cost function (3.26) can, in this setting, be expressed as
\[
(3.84) \quad J^N(u^N, y^N) := T \int_0^1 \ell(T \tau, u^N(\tau), y^N(\tau)) d\tau + \varphi(y^N(1)).
\]
The resulting *time normalized optimal control problem* is
\[
(3.85) \quad \int_{u^N} J^N(u^N, y^N) \quad \text{s.t. (3.83)}; \quad u^N(\tau) \in U_{ad} \text{ for a.a. } \tau, \quad \Phi_F(y^N(1)) \in K_F.
\]
The pre-Hamiltonian for the original and normalized problems are resp.
\[
\begin{align*}
H(\beta, t, u, y, p) & := \beta \ell(t,u,y) + p \cdot f(t,u,y); \\
H^N(\beta, \tau, T, w^N, y^N, p^N) & := TH(\beta, T \tau, w^N, y^N, p^N).
\end{align*}
\]
We see that the Lagrangian function of the time normalized optimal control problem is
\begin{equation}
\mathcal{L}^N := \int_0^1 \left( TH(\beta, T\tau, \mathbf{u}^N(\tau), \mathbf{y}^N(\tau), \mathbf{p}^N(\tau)) - \mathbf{p}^N(\tau) \cdot \dot{\mathbf{y}}^N(\tau) \right) d\tau + \beta \varphi(\mathbf{y}^N(1)) + \Psi_F \cdot \Phi_F(\mathbf{y}^N(1)).
\end{equation}

The horizon \( T \) appears then as a decision variable and we can apply the analysis of section 2.2. We know that we obtain the PMP for the same problem with a fixed horizon (theorem 3.13), with in addition the stationarity of the Lagrangian of the normalized problem w.r.t. the variable \( T \) (lemma 3.26). Set
\begin{equation}
\bar{\mathbf{p}}^N(1) = \beta^N \nabla \varphi(\mathbf{y}^N(1)) + D\Phi_F(\mathbf{y}^N(1))^{\dagger} \Psi^N_F; \quad \Psi^N_F \in K_F(\Phi_F(\mathbf{y}^N(1))).
\end{equation}

Set \( \bar{\mathbf{p}}(t) := \bar{\mathbf{p}}^N(t/T) \). It is easily checked that \( \bar{\mathbf{p}}(t) \) is solution of the costate equation for the original problem with multipliers \( (\bar{\beta}, \Phi_F) = (\beta^N, \Psi^N_F) \), that \( \Psi^N_F \in K_F(\Phi_F(\bar{\mathbf{y}}(T))) \), and that the Hamiltonian inequality for the normalized problem induces the Hamiltonian inequality for the original one. Indeed, one easily checks that this induces a bijective application between the Pontryagin extremals of the original problem with fixed horizon, and those of the normalized time problem (also with a fixed horizon) as follows:
\begin{equation}
(\bar{\beta}, \Phi_F) = (\beta^N, \Psi^N_F); \quad \bar{\mathbf{p}}(t) = \bar{\mathbf{p}}^N(t/T), \quad t \in (0, T).
\end{equation}

As a consequence of the PMP with a fixed horizon, we have seen in section 1.3.3 how to reformulate non autonomous problems (with smooth enough data) into autonomous ones, leading to relation (3.59), which in the setting of the normalized formulation gives
\begin{equation}
\tau \mapsto H^N(\tau) \text{ is Lipschitz, and } \frac{dH^N(\tau)}{d\tau} = \frac{\partial H(\tau)}{\partial \tau} \text{ a.e.}
\end{equation}

Since \( \bar{H}(\tau) := H^N(\tau)/T \), we deduce that
\begin{equation}
\frac{d\bar{H}(\tau)}{d\tau} = \frac{\partial H(\tau)}{\partial \tau} = T \dot{H}_t(\tau) \text{ a.e.}
\end{equation}

**Lemma 3.28.** Assume that the initial time is fixed and the final time is free, and that (3.57) holds. Then the PMP has the same expression as in the case of a fixed final time, with in addition (assuming that the final time does not enter in the final cost or constraints) the condition of zero final infimum of pre-Hamiltonian:
\begin{equation}
\bar{h}(T) = 0.
\end{equation}

**Proof.** We recall that \( \bar{h}^N(\tau) \) was defined in (3.89). Using lemma 3.26, (3.87) and (3.94), we get that:
\begin{equation}
0 = \frac{\partial \mathcal{L}^N}{\partial T} = \int_0^1 \left( \bar{H}(\tau) + T \dot{H}_t(\tau) \right) d\tau = \int_0^1 \left( \bar{h}^N(\tau) + T \frac{d\bar{h}^N(\tau)}{d\tau} \right) d\tau = \int_0^1 \left( \frac{d(\tau \bar{h}^N(\tau))}{d\tau} \right) d\tau = \bar{h}^N(1) = \bar{h}(T).
\end{equation}

The conclusion follows. \( \square \)
For an autonomous problem, we know by lemma 3.23 that the pre-Hamiltonian is constant for any Pontryagin extremal (on the problem with fixed horizon, and therefore also with a variable one), and therefore:

**Corollary 3.29.** If in addition the problem is autonomous then the pre-Hamiltonian has constant value 0 over $[0, T]$.

**Example 3.30.** Let the problem be autonomous, and let the horizon enter in the initial-final costs and constraint functions (think of a rendez-vous problem), whose expression are resp.

\[ \varphi(y(0), y(T), T); \quad \Phi(y(0), y(T), T). \]

Then the condition of stationarity of the Lagrangian for the normalized formulation w.r.t. $T$, denoting by $\tilde{h}$ the value (constant in time) of $H(t)$, is

\[ \tilde{h} + \tilde{\beta} DT\varphi(y(0), y(T), T) + \tilde{\Psi} \cdot DT\Phi(y(0), y(T), T) = 0. \]

**Example 3.31.** In the case of a simple integrator $\dot{y} = \alpha y$, with $y(t) \in \mathbb{R}^n$, consider the problem of minimizing the cost function $\int_0^T (1 + \frac{1}{2} |u(t)|^2) dt$, the final time being free. Here we minimize a compromise of the horizon and the quadratic energy, with constraints $y(0) = a$ and $y(T) = b$. The pre-Hamiltonian (in qualified form) is

\[ H = 1 + \frac{1}{2} |u|^2 + p \cdot (\alpha y + u). \]

The costate dynamics is $-p = \alpha p$, and so, $p(t) = e^{-\alpha t} p(0) = -u(t)$.

When $\alpha = 0$, the trajectories are straight lines from $a$ to $b$ with constant speed say $\dot{u} = (b - a)/T$, $p(t) = -\dot{u}$, $y(t) = a + t(b - a)/T$, and so since the time is free

\[ 0 = H = 1 - \frac{1}{2} \dot{u}^2 = 1 - \frac{1}{2} b - a)^2 / T^2. \]

This determines the optimal horizon $T = |b - a|/\sqrt{2}$. The optimal speed has always modulus $\sqrt{2}$.

### 2.3. Varying initial and final time.

Consider the case when both the initial time $t_0$ and the final time $T$ are free, these times entering in the initial-final state constraints

\[ \Phi(y(t_0), y(T), t_0, T) \in K_\Phi, \]

as well as in the cost function, of the form:

\[ \int_{t_0}^T \ell(t, u(t), y(t)) dt + \varphi(y(t_0), y(T), t_0, T). \]

The Lagrangian of the time normalized optimal control problem is

\[ \mathcal{L}^N = \int_0^1 ((T - t_0) H(\beta, t_0 + (T - t_0)\tau, u^N(\tau), y^N(\tau)) - p^N(\tau) \cdot y^N(\tau)) d\tau \]

\[ + \beta \varphi(y^N(t_0), y^N(1), t_0, T) + \Psi \cdot \Phi(y^N(t_0), y^N(1), t_0, T). \]

**Lemma 3.32.** Let (3.57) hold. If $(\bar{u}, \bar{y})$ is an optimal trajectory, there exist a Pontryagin multiplier $(\bar{\beta}, \bar{\Psi}, \bar{p})$ (for the problem with fixed initial and final time) such that

\[ \begin{cases} 
-\bar{h}(t_0) + \bar{\beta} DT_0 \varphi(y(t_0), y(T), T) + \bar{\Psi} \cdot DT_0 \Phi(y(t_0), y(T), t_0, T) = 0, \\
\bar{h}(T) + \bar{\beta} DT \varphi(y(t_0), y(T), T) + \bar{\Psi} \cdot DT \Phi(y(t_0), y(T), t_0, T) = 0.
\end{cases} \]

**Proof.** We only detail the derivation of the first relation. Redefine

\[ \bar{h}^N(\tau) := \inf_{u \in U_{ad}} H(\beta, (T - t_0)\tau, u, y^N(\tau), p^N(\tau)). \]

Again we apply lemma 3.26 to the time normalized optimal control problem where now $[t_0, T]$ is mapped into $[0, 1]$. So, the partial derivative of the Lagrangian (3.103) w.r.t. $t_0$ is equal to
0. Denoting by \( r \) the (easily computed) contributions of the initial-final cost and constraints, recalling the definition (3.88) of \( H(\tau) \), we get that:

\[
0 = \frac{\partial \mathcal{L}^N}{\partial t_0} = \int_0^1 \left( -\dot{H}(\tau) + (1 - \tau)(T - t_0)\dot{H}_t(\tau) \right) d\tau + r
\]

\[
(3.106)
\]

\[
\begin{align*}
\dot{H}(\tau) &= \dot{h}(\tau) + (1 - \tau)\dot{h}_t(\tau) d\tau \\
&= \int_0^1 \left( (1 - \tau)\dot{h}(\tau) + (1 - \tau)\frac{d}{d\tau} \dot{h}(\tau) \right) d\tau + r \\
&= \int_0^1 \frac{d}{d\tau} (1 - \tau)\dot{h}(\tau) d\tau + r = -\dot{h}(0) = -h(0) + r.
\end{align*}
\]

The first relation in (3.104) follows. \( \square \)

2.4. Interior point constraints.

2.4.1. A general setting. It may happen that some constraints hold on the state at finitely many times other that the initial and final ones. Typical examples are (i) interpolation problems, (ii) successive rendez-vous (e.g. of several asteroids with a satellite). We may formalize these constraints as follows. Consider a subdivision of \([0, T]\):

\[
(3.107) \quad 0 = t_0 < t_1 < \cdots < t_N = T,
\]

and the interior point constraints

\[
(3.108) \quad \Phi_i(\mathbf{y}(t_i)) \in K_i, \quad i = 0, \ldots, N,
\]

where \( K_i \) is a closed convex subset of \( \mathbb{R}^{n_{i*}} \) and \( \Phi_i : \mathbb{R}^n \to \mathbb{R}^{n_{i*}} \). Let us describe in general terms how to reduce these constraints to the setting (3.29). As in the case of variable horizon we can reparametrize each interval \([t_i, t_{i+1}]\) so that it starts at time 0 and ends at time 1, putting as decision variable the duration

\[
(3.109) \quad T_i := t_i - t_{i-1}; \quad i = 1, \ldots, N.
\]

The dynamics (after renormalizing the time) that may vary from one interval to another, over the \( i \)th interval, can be written, assuming that we have autonomous problems, as

\[
(3.110) \quad \dot{\mathbf{y}}_i(\tau) = T_i f_i(\mathbf{u}_i(\tau), \mathbf{y}_i(\tau)) \quad \text{for a.a. } \tau \text{ in } (0, 1), \quad i = 1, \ldots, N.
\]

The cost function is

\[
(3.111) \quad J(\mathbf{u}, \mathbf{y}) := \sum_{i=1}^N T_i \int_0^1 \ell^{i}(\mathbf{u}_i^{i}(\tau), \mathbf{y}_i^{i}(\tau)) d\tau + \varphi(\mathbf{y}_1^{1}(0), \mathbf{y}_1^{1}(1), \mathbf{y}_2^{2}(1), \ldots, \mathbf{y}_N^{N}(1)).
\]

We also have the continuity equations

\[
(3.112) \quad \mathbf{y}_i^{i}(1) - \mathbf{y}_i^{i+1}(0) = 0, \quad i = 1, \ldots, N - 1,
\]

whose Lagrange multipliers will be denoted by \( \eta_i \), and control constraints

\[
(3.113) \quad \mathbf{u}_i^{i}(\tau) \in U_{ad}^{i}, \quad \text{for a.a. } \tau \text{ in } (0, 1), \quad i = 1, \ldots, N.
\]

as well as possible interior constraints other than (3.112):

\[
(3.114) \quad \Phi(\mathbf{y}_1^{1}(0), \mathbf{y}_1^{1}(1), \mathbf{y}_2^{2}(1), \ldots, \mathbf{y}_N^{N}(1)) \in K,
\]

with \( K \) closed and convex. The Lagrangian of this problem is \( L := L^1 + L^2 \), with

\[
L^1 := \sum_{i=1}^N \int_0^1 \left( \beta T_i \ell^{i}(\mathbf{u}_i^{i}(\tau), \mathbf{y}_i^{i}(\tau)) + p_i^{i}(\tau) \cdot (T_i f_i(\mathbf{u}_i^{i}(\tau), \mathbf{y}_i^{i}(\tau)) - \dot{\mathbf{y}}_i^{i}(\tau)) \right) d\tau,
\]

\[
L^2 := \beta \varphi(\mathbf{y}_1^{1}(0), \mathbf{y}_1^{1}(1), \ldots, \mathbf{y}_N^{N}(1)) + \Psi \cdot \Phi(\mathbf{y}_1^{1}(0), \mathbf{y}_1^{1}(1), \ldots, \mathbf{y}_N^{N}(1)) + \Psi^0 \cdot \Phi^0(\mathbf{y}_1^{1}(0)) + \sum_{i=1}^N \Psi_i \cdot \Phi_i(\mathbf{y}_i^{i}(1)) + \sum_{i=1}^{N-1} \eta_i \cdot (\mathbf{y}_i^{i+1}(1) - \mathbf{y}_i^{i}(1)).
\]
By (an obvious extension of) lemma 3.23 it follows that

\[ H^i(\beta, u, y, p) := \beta \ell^i(u, y) + p \cdot f^i(y, u). \]

Integrating by parts, we have that

\[ L^1 = \sum_{i=1}^{N} \int_{0}^{1} (\dot{p}^i(\tau) \cdot y^i(\tau) + T_i H^i(\beta, u^i(\tau), y^i(\tau), p^i(\tau))) \, d\tau + p^i(0) \cdot y^i(0) - p^i(1) \cdot y^i(1). \]

Therefore we have the costate equation at the reference trajectory \((\dot{u}, y)\):

\[ -\dot{\bar{p}}^i(\tau) = T_i \nabla_y H^i(\tau) = T_i f^i_\tau(\bar{u}^i(\tau), y^i(\tau)), \quad i = 1, \ldots, N, \]

with transversality conditions, skipping arguments when possible:

\[ 0 = \nabla_{y^0} L = \dot{\bar{p}}^i(0) - \bar{\eta}^{j-1}, \quad i = 2, \ldots, N, \]

\[ 0 = \nabla_{y^1} L = \bar{\eta}^1 - \dot{p}^i(1) + \beta \nabla_{y^i} \varphi + (D_{y^i} \Phi)^\dagger \bar{\Psi} + (D \Phi^0)^\dagger \bar{\Psi}^0, \quad i = 1, \ldots, N - 1, \]

\[ 0 = \nabla_{y^N} L = -\dot{p}^{N}(1) + \beta \nabla_{y^N} \varphi + (D_{y^N} \Phi)^\dagger \bar{\Psi} + (D \Phi^N)^\dagger \bar{\Psi}^N. \]

So, we can eliminate \(\bar{\eta}\) and write the punctual conditions for the costate in the form

\[ \begin{align*}
\dot{p}^{N}(1) &= \beta \nabla_{y^N} \varphi + (D_{y^N} \Phi)^\dagger \bar{\Psi} + (D \Phi^N)^\dagger \bar{\Psi}^N, \\
\dot{p}^{i}(1) &= \dot{p}^{i+1}(0) + \beta \nabla_{y^i} \varphi + (D_{y^i} \Phi)^\dagger \bar{\Psi} + (D \Phi^i)^\dagger \bar{\Psi}^i, \quad i = 1, \ldots, N - 1, \\
-\dot{p}^{0}(0) &= \beta \nabla_{y^0} \varphi + (D_{y^0} \Phi)^\dagger \bar{\Psi} + (D \Phi^0)^\dagger \bar{\Psi}^0.
\end{align*} \]

The multiplier \(\bar{\Psi}\) and the \(\Psi^i\) satisfy

\[ \begin{align*}
\Psi &\in N_K(\Phi(y^1(0), y^1(1), \ldots, y^N(1))), \\
\Psi^0 &\in N_{K^0}(\Phi^0(y^1(0))), \\
\Psi^i &\in N_{K^i}(\Phi^i(y^i(1))), \quad i = 1, \ldots, N.
\end{align*} \]

The nontriviality condition reads

\[ \beta + |\Psi| + \sum_{i=0}^{N+1} |\Psi^i| > 0. \]

The Hamiltonian inequality can be decomposed between separated problems for each interval: we get that

\[ H^i(\beta, u^i(\tau), y^i(\tau), p^i(\tau)) = \min_{u^i \in U^{i}_{ad}} H^i(\tilde{\beta}, u, y^i(\tau), p^i(\tau)), \]

for a.a. \(\tau \in (0, 1), \quad i = 1, \ldots, N.\)

By (an obvious extension of) lemma 3.23 it follows that

\[ t \mapsto H^i(\beta, u^i(t), y^i(t), p^i(t)) \text{ is constant over } (0, 1). \]

**Exercise 3.33.** Write the PMP in the original formulation, i.e., with costate function of \(t \in [0, T]\).
2.4.2. Variable durations. In the study of interior point constraints, we assumed that the durations $T_i$ were given. But if some of them are free, say for $i \in I$, the others being fixed. Then for $i \in I$ we have the additional condition

$$0 = \frac{\partial L}{\partial T_i} = \int_0^1 \tilde{H}^i(\tau) d\tau.$$  

Equivalently, by (3.127):

$$0 = \frac{\partial L}{\partial T_j} = \gamma + \int_0^1 \tilde{H}^j(\tau) d\tau, \quad j = i, i + 1.$$  

Consider now the case when the final time of the $i$th interval is free. We can then see $T_i$ and $T_{i+1}$ as optimization variables, subject to the constraint $T_i + T_{i+1} = b$ ($b$ given). So we must add $\gamma(T_i + T_{i+1} - b)$ to the Lagrangian of the problem and we get that

$$0 = \frac{\partial L}{\partial T_j} = \gamma + \int_0^1 \tilde{H}^j(\tau) d\tau, \quad j = i, i + 1.$$  

Thanks again to (3.127) we obtain that

$$0 = \frac{\partial L}{\partial T_j} = \gamma + \int_0^1 \tilde{H}^j(\tau) d\tau, \quad j = i, i + 1.$$  

Exercice 3.34. Extend this result to the case of non autonomous problems. Hint: follow the proof of lemma 3.28.

2.4.3. Variable dynamics structure. Think of the problem of the ascent trajectory of a space shuttle, with different stages that are expelled during the ascent. This fits in the present framework, but with a state $y^i$ whose dimension depends on the interval and, rather than continuity of the state, compatibility conditions on the form

$$\hat{\Phi}^i(y^i(1), y^{i+1}(0)) \in \tilde{K}^i, \quad i = 1, \ldots, N - 1,$$

with again $\tilde{K}^i$ closed convex subset of an Euclidean space and $\hat{\Phi}^i$ mapping between Euclidean spaces of appropriate dimension. So, we consider the same setting as before, with variable state dimension, and replacing the continuity equations (3.112) by (3.132). The multiplier $\hat{\Psi}^i$ associated with constraint (3.132) must satisfy

$$\hat{\Psi}^i \in N_{\tilde{K}^i}(\hat{\Phi}^i(y^i(1), y^{i+1}(0))), \quad i = 1, \ldots, N - 1.$$  

It appears that the costate dynamics has the same expression, the only changes are the end conditions for the costate, which instead of (3.119)-(3.122) now read, denoting by $\hat{L}$ the Lagrangian in the present setting (that we do not need to explicit):

$$0 = \nabla y^0 \hat{L} = \hat{p}^i(0) + \beta \nabla y^0 \hat{\varphi} + (Dy^0 \hat{\Phi})^\dagger \hat{\Psi} + (D \Phi^0)^\dagger \hat{\Psi}^0,$$

$$0 = \nabla y_0 \hat{L} = \hat{p}^i(0) + Dy^i \hat{\Phi}^i(y^i(1), y^{i+1}(0))^\dagger \hat{\Psi}^i, \quad i = 2, \ldots, N,$$

$$0 = \nabla y^i_1 \hat{L} = \hat{p}^i(1) + Dy^i \hat{\Phi}^i(y^i(1), y^{i+1}(0))^\dagger \hat{\Psi}^i + \beta \nabla y^i \hat{\varphi} + (Dy^i \hat{\Phi})^\dagger \hat{\Psi} + (D \Phi^i)^\dagger \hat{\Psi}^i, \quad i = 1, \ldots, N - 1,$$

$$0 = \nabla y^N_1 \hat{L} = \hat{p}^N(1) + \beta \nabla y^N \hat{\varphi} + (Dy^N \hat{\Phi})^\dagger \hat{\Psi} + (D \Phi^N)^\dagger \hat{\Psi}^N.$$

2.5. Minimal time problems, geodesics.
2.5.1. General results. Minimal time problems are the particular case of problems with fixed initial state, free final time, zero final cost, cost integrand \( \ell(u, y) = 1 \), and final constraint \( \Phi_F(\bar{y}(T)) \in K_F \). The pre-Hamiltonian is

\[
H(\beta, t, u, y, p) = \beta + p \cdot f(t, u, y).
\]

So, the Hamiltonian inequality (3.38) reads

\[
\dot{p}(t) \cdot f(t, \bar{u}(t), \bar{y}(t)) = \inf_{u \in U_{ad}} p(t) \cdot f(t, u, \bar{y}(t)) \quad \text{for a.a. } t \in (0, T).
\]

The costate dynamics and non triviality conditions are (see (3.44)-(3.45)):

\[
\begin{aligned}
-\dot{p}(t) &= \bar{f}_y(t)\cdot p(t) \quad \text{for a.a. } t \in [0, T]; \\
\bar{p}(T) &= D\Phi_F(\bar{y}(T))\cdot \bar{\Psi}_F; \quad \bar{\Psi}_F \in N_{K_F}(\bar{y}(T)).
\end{aligned}
\]

\[
\beta \in \{0, 1\}; \quad \beta + |\bar{\Psi}_F| > 0.
\]

So the PMP here reduces to (3.139)-(3.141), with (since the final time is free), by lemma 3.28, the condition that the final Hamiltonian has value 0.

\[
\bar{h}(T) = 0; \text{ if the problem is autonomous, then } \bar{h}(t) = 0, \text{ for all } t \in [0, T].
\]

One can prove the following key result:

**Lemma 3.35.** If \((\bar{u}, \bar{y})\) is solution of the minimal time problem, or more generally a Pontryagin extremal of this problem, then \(\bar{\Psi}_F \neq 0\), and if \(\Phi'_F(\bar{y}(T))\) is surjective, then \(\bar{p}(t) \neq 0\), for all \(t \in [0, T]\).

**Proof.** If on the contrary \(\bar{\Psi}_F = 0\), then by (3.140), \(\bar{p}(t) = 0\) for all \(t\). So the pre-Hamiltonian reduces to \(\beta\), but it has zero value at final time, by lemma 3.28, so that \(\beta = 0\), contradicting (3.141). So \(\bar{\Psi}_F \neq 0\). If \(\Phi'_F(\bar{y}(T))\) is surjective, then its transpose matrix is injective, and so \(\bar{p}(T) = \Phi'_F(\bar{y}(T))^\dagger \bar{\Psi}_F\) is not zero. Being the solution of an homogeneous linear ODE, the costate is nonzero at any time. The result follows. \(\square\)

**Remark 3.36.** Since the final pre-Hamiltonian has zero value, the l.h.s. of (3.139) is nonpositive, with value 0 if \(\beta = 0\). So (provided \(\bar{u}\) is continuous at time \(T\)), \(\bar{p}(T)\) and \(f(T, \bar{u}(T), \bar{y}(T))\) make an obtuse angle if \(\beta > 0\), and are orthogonal if \(\beta = 0\). The same relation holds for any \(t \in [0, T]\) if the problem is autonomous (since then the pre-Hamiltonian has a constant value).

**Remark 3.37.** We could formulate minimal time problems with a zero integral cost and final cost \(T\). The resulting optimality conditions are equivalent to those above. By pre-Hamiltonian of a minimal time problem one often understands the one of this second formulation, i.e., \(p \cdot f(u, y)\).

**Example 3.38 (A case when \(\beta = 0\)).** Let \(n = 2\), \(\Phi_F\) be the identity, \(K_F = \{x \in \mathbb{R}^2; \ x_2 \geq x_1^2\}\), \(y^0 = (-1, 0)^\dagger\), with dynamics \(\dot{y}_1(t) = u(t) \in [0, 1], \ \dot{y}_2(t) = 0\). Clearly the minimum time is 1 and the optimal solution is \(\bar{u}(t) = 1, \ \bar{y}_1(t) = t - 1, \ \bar{y}_2(t) = 0\). Since the normal cone to \(K_F\) at the final state \(\bar{y}(1) = 0\) is the half line generated by \((-1, -1)^\dagger\), we may assume that \(\bar{\Psi}_F = (-1, -1)^\dagger\); that \(\bar{p}(t) = 0\) implies \(\bar{p}(t) = (-1, -1)^\dagger\), over \([0, T]\). The Hamiltonian inequality is satisfied, as expected. Since \(\bar{p}(t) \cdot f(\bar{u}(t), \bar{y}(t)) = 0\), we have that \(\beta = 0\).

We next discuss some important classes of optimal time problems.

2.5.2. Linear dynamics. Let \(f(u, y) = Ay + Bu\) with \(A\) and \(B\) matrices of compatible dimensions. Then \(-\dot{p}(t) = A^\dagger \bar{p}(t)\), and so \(\bar{p}(t) = e^{(T-t)A^\dagger} \bar{p}(t)\) is an analytic function (i.e., has a power series expansion). Denote by \(\bar{p}(t)^{(k)}\) the \(k\)th derivative of \(\bar{p}(t)\). The control must minimize over \(U_{ad}\) the function \(\bar{p}(t)^\dagger Bu\). If the latter vanishes identically, then for \(k = 0\) to \(n - 1\), we have that

\[
0 = B^\dagger \bar{p}(t)^{(k)} = (-1)^k B^\dagger (A^\dagger)^k \bar{p}(t),
\]
whence \( \dot{p}(t) \) is orthogonal to \( A^{(k)}B \). It is said that the pair \((A, B)\) is \textit{controllable}, if (in the context of linear systems) the family \( A^{(k)}B, k = 0 \) to \( n - 1 \), has rank \( n \). If this condition holds then \( \dot{p} \) vanishes identically, and we know that this is impossible when \( \Phi_F'(\overline{y}(T)) \) is surjective.

**Example 3.39 (Bound constraints).** Assume that \( U_{ad} = [-1, 1]^n \) and denote by \( B_j \) the \( j \)th column of \( B \). Let \( j \in \{1, \ldots, m\} \). Then \( u_j(t) \) minimizes \( u_j \mapsto u_jB_j \cdot \dot{p}(t) \) for a.a. \( t \). If \( B_j \cdot \dot{p}(t) \) does not identically vanish, since a nonconstant analytic function has finitely many zeros over a bounded interval, we deduce that \( u_j(t) \) is bang-bang (equal to \( \pm 1 \) for a.a. \( t \in [0, T] \)), and more precisely, equal to \(-1\) when \( B_j \cdot \dot{p}(t) > 0 \), and to \( 1 \) when \( B_j \cdot \dot{p}(t) < 0 \), with finitely many commutation times.

In particular, if \( m = 1 \), the pair \((A, B)\) is controllable, and \( \Phi_F'(\overline{y}(T)) \) is surjective, the control is bang-bang with finitely many switching times.

**Remark 3.40.** It can be checked that, if \( A \) has only real eigenvalues, and \( B_j \cdot \dot{p}(t) \) does not identically vanish, then \( u_j(t) \) has at most \( n - 1 \) switches. This holds in particular for the standard \( n \)th order integrator (with here \( m = 1 \)):

\[
(3.144) \quad \dot{y}_1 = \dot{y}_2, \ldots, \dot{y}_{n-1} = y_n, \dot{y}_n = u.
\]

About this and the explicit resolution of two dimensional problems, see e.g. [14, Part III].

2.5.3. Geometric optics and Riemannian geometry.

**Example 3.41 (Geometric optics).** The nominal dynamics is \( \dot{\tilde{y}}(t) = F(\tilde{y}(t))\tilde{u}(t) \), where the positive valued \textit{light speed} \( F : \mathbb{R}^n \to \mathbb{R} \) is of class \( C^\infty \), and \( \tilde{u}(t) \in \tilde{B} \) (closed Euclidean unit ball) for a.a. \( t \). We assume \( \Phi_F'(\overline{y}(T)) \) to be surjective, so that the costate does not vanish. The pre-Hamiltonian \( \beta + F(\tilde{y}(t))\dot{p}(t) \cdot \tilde{u}(t) \) attains its minimum over \( \tilde{B} \) at \(-\dot{p}(t)/|\dot{p}(t)|\), and has value 0 at that point, so that \( \beta \) cannot have the value 0, and so, \( 0 = h(t) = 1 - F(\tilde{y})|\dot{p}| \). Eliminating \( \tilde{u}(t) = -\dot{p}(t)/|\dot{p}(t)| \), we see that the state-costate pair has the following dynamics, skipping the time argument:

\[
(3.145) \quad \dot{y} = -F(\tilde{y})\dot{p}/|\dot{p}|; \quad \ddot{p} = -\nabla_y \dot{H} = -(\dot{p} \cdot \tilde{u})\nabla F(\tilde{y}) = |\dot{p}| \nabla F(\tilde{y}).
\]

If over some open subset \( E \) of \( \mathbb{R}^n \) the function \( F \) is constant, then \( \dot{p} \) and consequently \( \tilde{u} \) are also constant, so that the extremal trajectories over \( E \) are lines. Since the final time is free, we have that

\[
(3.146) \quad 0 = \dot{H}(t) = 1 - F(\overline{y}(t))/|\dot{p}(t)|, \quad \text{for all } t \in [0, T].
\]

**Remark 3.42.** Given an horizon \( T \), the \textit{minimal energy transfer} for geometric optics consists in minimizing \( \frac{1}{2} \int_0^T \| \dot{u}(t) \|^2 dt \) with the above dynamics and final constraints. The pre-Hamiltonian is \( \frac{1}{2} \beta u^2 + F(y)\dot{p} \cdot u \). If \( \beta = 0 \) then \( \dot{p}(t) = 0 \) for all \( t \), otherwise the infimum of the pre-Hamiltonian has value \(-\infty \). Next conside the case when \( \beta = 1 \). The unique control minimizing the pre-Hamiltonian being \( \tilde{u}(t) = -\frac{1}{F(\tilde{y}(t))}\dot{p}(t) \), the state-costate pair has dynamics

\[
(3.147) \quad \dot{y} = -F(\tilde{y})^2 \dot{p}; \quad \ddot{p} = -\nabla_y \dot{H} = -(\dot{p} \cdot \tilde{u})\nabla F(\tilde{y}) = F(\tilde{y}) |\dot{p}|^2 \nabla F(\tilde{y}).
\]

This dynamics is, at each point, proportional to the one in (3.145). Therefore, \textit{the orbits of the state-costate pair in geometrical optics coincide with those of the minimal energy transfer} (by say the orbit of \( f(t) \) for \( t \in [0, T] \) we mean \( \{f(t), t \in [0, T]\} \)). Besides, in the minimal energy transfer problem, since the pre-Hamiltonian has a constant value equal to \(-\frac{1}{2}|\tilde{u}(t)|^2 \), \textit{the solution of an extremal of a minimal energy transfert problem has constant control norm}.

In the case of geometric optics, the medium is isotropic, that is, the speed at a given point does not depend on the direction of propagation. We next present an extension to the case of an anisotropic medium.

**Example 3.43 (Riemannian geometry).** Consider the minimal time problem with nominal dynamics \( \dot{y} = A(\tilde{y})\tilde{u}, \) the symmetric matrix \( A(y) \), of size \( n \), being invertible for all \( y \), and \( C^\infty \) function of \( y \), with the constraint \( \tilde{u}(t) \in \tilde{B} \) for a.a. \( t \), which means that \( \dot{y}^T A(\tilde{y})^{-2} \dot{y} \leq 1 \). We
still can reduce to the case of a regular multiplier and show that \( \mathbf{p} \) cannot vanish. The qualified pre-Hamiltonian (one easily checks that \( \beta > 0 \)) \( 1 + p y A(y)u \) attains, over a Pontryagin extremal, its minimum at \( -A(\mathbf{y})\mathbf{p}/|A(\mathbf{y})\mathbf{p}| \), and so,
\[
(3.148) \quad \dot{\mathbf{y}} = -A^2(\mathbf{y})\mathbf{p}/|A(\mathbf{y})\mathbf{p}|.
\]
In addition, for \( k = 1 \) to \( n \), setting \( A_k(y) := \partial A(y)/\partial y_k \), we have that
\[
(3.149) \quad \dot{\mathbf{p}}_k = -\nabla_{y_k} \dot{H} = -\mathbf{p}^A A_k(\mathbf{y})\mathbf{u} = \mathbf{p}^A A_k(\mathbf{y})\mathbf{p}/|A(\mathbf{y})\mathbf{p}|.
\]

If over some open subset \( E \) of \( \mathbb{R}^n \) the function \( A(y) \) is constant, then \( \mathbf{p} \) and consequently \( \mathbf{u} \) are also constant, so that the extremal trajectories over \( E \) are lines. The corresponding minimal energy transfe problem consists in minimizing \( \frac{1}{2} \int_0^T u(t)^2 dt \) over the previous dynamics and constraints. The pre-Hamiltonian \( \frac{1}{2} u^2 + p^A A(y)u \) attains its minimum at \( u = -A(y)p \), and so
\[
(3.150) \quad \dot{\mathbf{y}} = -A^2(\mathbf{y})\mathbf{p}; \quad \dot{\mathbf{p}}_k = -\nabla_{y_k} \dot{H} = -\mathbf{p}^A A_k(\mathbf{y})\mathbf{u} = \mathbf{p}^A A_k(\mathbf{y})\mathbf{p}/|A(\mathbf{y})\mathbf{p}|.
\]

Here again, the orbits of the two problems coincide. One calls geodesics the orbits of \( \mathbf{y} \) when the pair \( (\mathbf{y}, \mathbf{p}) \) satisfies the state-costate dynamics (forgetting the conditions at time 0 and \( T \)) for now \( t \in \mathbb{R} \). The pre-Hamiltonian having the constant value \( -\frac{1}{2} |\mathbf{u}(t)|^2 \), the control norm \( |\mathbf{u}(t)| = |A(\mathbf{y}(t))\mathbf{p}(t)| \) is also constant.

**Remark 3.44.** The usual way to introduce a variational problem related to these geodesics is to set a smooth mapping \( G(y) \), for \( y \in \mathbb{R}^n \), with value in the set of positive definite symmetric matrices and to define the length between \( y^0 \) and \( y^1 \) in \( \mathbb{R}^n \) as the infimum of
\[
I(y) := \int_a^b (\dot{y}(s))^1 G(y(s)) \dot{y}(s))^{1/2} ds
\]
over those absolutely continuous curves \( y(t) \) such that \( y(a) = y^0 \) and \( y(b) = y^1 \). It is easily shown that under a time change of the form \( s' = \varphi(s) \), with \( \varphi : \mathbb{R} \to \mathbb{R} \) of class \( C^1 \) with positive derivative, the cost is invariant. So it is possible to reduce the problem to the one of minimizing \( b - a \) under the constraint \( \dot{y}(s)^1 G(y(s)) \dot{y}(s))^{1/2} \leq 1 \). Taking \( A(y) := G(y)^{-1/2} \) we recover our setting.

### 3. Qualification conditions

We saw in several particular cases how to check that singular multipliers do not exist. See examples 3.16 and 3.18, and the above discussion of geometric optics and Riemannian geometry. We now give a general approach to this question, for problem (3.29), starting with the case of Lagrange multipliers. Let \( (\mathbf{u}, \mathbf{y}) \) be an admissible trajectory for problem (3.29).

**3.1. Lagrange multipliers.**

3.1.1. **General framework.** We assume here that \( U_{ad} \) is a nonempty, closed convex subset of \( \mathbb{R}^m \). Remember that \( z[v, z^0] \) denotes the solution of the linearized state equation (2.115), and that the Lagrange multipliers were introduced in definition 3.19. Consider the convex sets
\[
(3.152) \begin{align*}
C & := \{z[v, z^0]; \ z^0 \in \mathbb{R}^n; \ \mathbf{u}(t) + v(t) \in U_{ad} \ a.e.\}, \\
C_{0T} & := \{(z(0), z(T)); \ z \in C\}; \\
M & := D\Phi(\mathbf{y}(0), \mathbf{y}(T))C_{0T}; \\
M_\Phi & := M + \Phi(\mathbf{y}(0), \mathbf{y}(T)) - K_\Phi.
\end{align*}
\]

The main result concerning Lagrange multiplier is as follows:

**Proposition 3.45.** The set of singular Lagrange multipliers associated with \( (\mathbf{u}, \mathbf{y}) \) is empty iff the following qualification condition holds:
\[
(3.153) \quad M_\Phi \text{ is a neighbourhood of } 0.
\]

The proof needs some lemmas.

---

\(^1\) Indeed, set \( z(s') := y(s) \), and use \( ds' = \varphi'(s)ds \) as well as \( dz(s') = \varphi'(s)^{-1}dy(s) \).
**Lemma 3.46.** The positive polar cone (definition 2.9) of \(M_\Phi\) satisfies
\[ (M_\Phi)^+ = M^+ \cap N_{K_\Phi}(\Phi(\bar{y}(0), \bar{y}(T))). \]

**Proof.** Indeed, \(\Psi \in (M_\Phi)^+\) iff
\[ 0 \leq \Psi \cdot (m + \Phi(\bar{y}(0), \bar{y}(T)) - k), \quad \text{for all } m \in M \text{ and } k \in K. \]
Choosing \(m = 0\) (corresponding to \(v = 0\)) we deduce that \(\Psi \cdot (k - \Phi(\bar{y}(0), \bar{y}(T))) \leq 0\), so that \(\Psi \in N_{K_\Phi}(\Phi(\bar{y}(0), \bar{y}(T)))\). Choosing \(k = \Phi(\bar{y}(0), \bar{y}(T))\), we get \(\Psi \cdot m \geq 0\), so that \(\Psi \in M^+\). We just proved that \((M_\Phi)^+\) is included in the r.h.s. of (3.154). Conversely, if \(\Psi\) belongs to the r.h.s. of (3.154), then
\[ \Psi \cdot m \geq 0\] and \(\Psi \cdot (\Phi(\bar{y}(0), \bar{y}(T)) - k) \geq 0\), for any \(m \in M\) and \(k \in K\), so that (3.155) holds, i.e., \(\Psi \in (M_\Phi)^+\). This ends the proof.

**Lemma 3.47.** Let \(\Psi \in \mathbb{R}^{n_\Phi}\). Then there exists a singular Lagrange multiplier associated with \((u, \bar{y})\), of the form \((\beta = 0, \Psi, \bar{p})\), iff \(\Psi \neq 0\) and \(\Psi \in (M_\Phi)^+\).

**Proof.** Let \((\beta = 0, \Psi, \bar{p})\) be a singular Lagrange multiplier, so that \(\Psi \neq 0\), and let \(z = z[v, z^0]\). By remark 3.11,
\[ \Psi^1 D\Phi(\bar{y}(0), \bar{y}(T))(z(0), z(T)) = \int_0^T \bar{H}_a(t) v(t) dt. \]
The last integral is nonnegative in view of (3.46). So, any singular multiplier \(\Psi\) belongs to \(M^+\) and therefore (since \(\Psi \in N_{K_\Phi}(\Phi(\bar{y}(0), \bar{y}(T)))\)) also to \((M_\Phi)^+\) by lemma 3.46. Conversely, let \(\Psi \in (M_\Phi)^+\), \(\Psi \neq 0\). Let \(\vec{p}\) satisfy the costate dynamics, with final condition
\[ \vec{p}(T) = D_{\bar{y}} \Phi(\bar{y}(0), \bar{y}(T))^\dagger \Psi. \]
Set \(\eta := \bar{p}(0) + \nabla_{\bar{y}} \Phi(\bar{y}(0), \bar{y}(T))^\dagger \Psi. \)
Then
\[ 0 \leq \Psi^1 D\Phi(\bar{y}(0), \bar{y}(T))(z(0), z(T)) = [\bar{p} \cdot z]_0^T + \eta \cdot z(0) \]
\[ = \int_0^T (\bar{p}(t) \cdot z(t) + \bar{p}(t) \cdot \dot{z}(t)) dt + \eta \cdot z(0) \]
\[ = \int_0^T \bar{H}_a(t) v(t) dt + \eta \cdot z(0). \]
Since we may take \(z(0)\) arbitrarily if follows that \(\eta = 0\), i.e., \(\vec{p}\) is the costate associated with \(\Psi\). We deduce then easily that \(\bar{H}_a(t) v(t) \geq 0\) a.e., showing that \((\beta = 0, \Psi, \vec{p})\) is a singular Lagrange multiplier.

We can now prove the main result:

**Proof of prop. 3.45.** If \(M_\Phi\) is a neighbourhood of 0, then its polar cone reduces to 0. By the above lemma, the set of singular multipliers is empty. Conversely, if the convex set \(M_\Phi^+\) is not a neighbourhood of 0, by remark 2.14, we can separate \(M_\Phi\) from zero by some nonzero \(\Psi\), which means that \(\Psi \in (M_\Phi)^+\); we conclude with lemma 3.47.

**3.1.2. Specific structures.** We end the subsection by considering some particular cases.

**Exercise 3.48.** (i) If \(K_\Phi = \{0\}\) (initial-final constraints of equality type), show that the set of singular Lagrange multipliers is empty iff \(M\) is a neighbourhood of 0.
(ii) In addition there is no control constraints (i.e., \(U_{ad} = \mathbb{R}^{m}\)) show that the set of singular Lagrange multipliers is empty iff
\[ \text{The mapping } (v, z^0) \mapsto D\Phi(\bar{y}(0), \bar{y}(T))(z[v, z^0](0), z[v, z^0](T)) \text{ is surjective.} \]
In the last case above we can be slightly more precise:

**Lemma 3.49.** If (3.160) holds and \(U_{ad} = \mathbb{R}^{m}\), then there is no singular Lagrange multipliers, and there is a unique normal multiplier.
PROOF. In view of proposition 3.45, $\beta \neq 0$; so, we may take $\beta = 1$. By computations similar to those in (3.157), we find that for any $(v, z^0) \in U \times R^n$, setting $z = z[v, z^0]$:

\[
D\varphi(y(0), y(T))(z(0), z(T)) + \Psi^d D\Phi(y(0), y(T))(z(0), z(T)) = \int_0^T D_s H(t)v(t)dt \geq 0.
\]

Changing $(v, z^0)$ into its opposite we obtain the inequality in the reverse direction and so the integral is equal to zero. Had we two possible multipliers say $\Psi'$ and $\Psi''$, their difference $\Psi := \Psi'' - \Psi'$ would therefore satisfy

\[
\Psi^d D\Phi(y(0), y(T))(z(0), z(T)) = 0, \quad \text{for all } v \in U,
\]

which in view of (3.160) implies $\Psi = 0$. The conclusion follows.

We can also detail the case of inequality constraints, starting with the more general case when $K_\Phi$ has a nonempty interior.

**Lemma 3.50.** Let $K_\Phi$ have a nonempty interior. Then the set of singular Lagrange multipliers is empty iff there exists $(v, z^0) \in U \times R^n$ such that $\bar{u} + v$ satisfies the control constraints, and $z := z[v, z^0]$ satisfies

\[
\Phi(y(0), y(T)) + D\Phi(y(0), y(T))(z(0), z(T)) \in \text{int}(K_\Phi).
\]

**Proof.** By a translation argument we may assume that $\Phi(y(0), y(T)) = 0$. We can then write the above condition as $\text{int}(K_\Phi) \cap M \neq \emptyset$. So, in view of proposition 3.45, we must prove that $M_\Phi$ is a neighbourhood of 0 iff $\text{int}(K_\Phi) \cap M \neq \emptyset$.

If the intersection is non empty it is easily seen that $M_\Phi$ is a neighbourhood of 0. If the intersection is empty, we can separate these convex sets by some nonzero $\Psi$ and (since, $\text{int}(K_\Phi)$ being nonempty, its closure is equal to $K_\Phi$) we have that $\Psi \cdot x \geq 0$ for any $x \in M_\Phi$. But then $M_\Phi$ cannot be a neighbourhood of 0.

**Exercise 3.51.** When $K_\Phi = R_{n^*}$ (initial-final constraints of inequality type), show that the set of singular Lagrange multipliers is empty iff there exists $(v, z^0) \in U \times R^n$ such that $\bar{u} + v$ satisfies the control constraints, and $z := z[v, z^0]$ satisfies

\[
D\Phi_k(y(0), y(T))(z(0), z(T)) < 0 \quad \text{iff } \Phi_k(y(0), y(T)) = 0, \quad k = 1 \to n_\Phi.
\]

**3.2. Pontryagin multipliers.** We recall that $(\bar{u}, \bar{y})$ is an admissible trajectory for problem (3.29). We assume here that $U_{ad}$ is a nonempty, closed subset of $R^n$. Instead of the linearized state equation (2.115), we consider again the dynamics (3.19) for $\xi$, with $u \in U$, but in the case of a varying initial condition $\xi^0$. Its solution in $Y$ may viewed as a function of an initial value $\xi^0$ and $u$, denoted by $\xi[u, \xi^0]$. Consider the sets

\[
\begin{align*}
C' & := \{\xi[u, \xi^0]; \xi^0 \in R^n; \; u(t) \in U_{ad} \; \text{a.e.}\}, \\
C_{0T}' & := \{(\xi(0), \xi(T)); \xi \in C'\}; \\
M' & := D\Phi(\bar{y}(0), \bar{y}(T))C_{0T}'; \\
M'' & := \text{conv}(M'); \\
M_\Phi & := M'' + \Phi(\bar{y}(0), \bar{y}(T)) - K_\Phi.
\end{align*}
\]

By conv$(M')$ we mean the convex hull of $M'$, i.e., the smallest convex set containing $M'$, which is nothing but the set of convex combinations (linear combinations with nonnegative weights summing to 1) of elements of $M'$. Similarly to lemma 3.155, observing that $M'$ and $M''$ have the same polar sets, we can easily show that

\[
(M_\Phi)^+ = (M')^+ \cap N_{K_\Phi}(\Phi(\bar{y}(0), \bar{y}(T))).
\]

**Lemma 3.52.** Let $(\bar{u}, \bar{y})$ be an admissible trajectory for problem (3.29), and $\Psi \in R_{n^*}$. Then there exists a singular Pontryagin multiplier of the form $(\beta = 0, \Psi, \bar{p})$ iff $\Psi \neq 0$ and $\Psi \in (M_\Phi)^+$.
PROOF. Let \((\beta = 0, \Psi, \bar{p})\) be a singular Pontryagin multiplier, so that \(\Psi \neq 0\), and let \(\xi \in C\), \(\xi = \xi[u, \xi^0]\) with \(u(t) \in U_{ad}\) a.e. Then
\[
\Delta := \Psi^T D\Phi(y(0), y(T))(\xi(0), \xi(T)) = [\bar{p} \cdot \xi]^T
\]
(3.167)
\[
= \int_0^T (\dot{\bar{p}}(t) \cdot \xi(t) + \bar{p}(t) \cdot \dot{\xi}(t)) dt = \int_0^T (H(0, t, u(t), y(t), \bar{p}(t)) - \dot{H}(t)) dt.
\]
By the Hamiltonian inequality, the last integral is nonnegative, so that \(\Psi \in (M')^+\). Since \(\Psi \in N_{K\Phi}(\Phi(y(0), y(T)))\), it follows that \(\Psi \in (M\Phi)^+\). Conversely, let \(\Psi \in (M\Phi)^+, \Psi \neq 0\). Let \(\bar{p}\) satisfy the costate dynamics, with final condition
(3.168)
\[
\bar{p}(T) = D_{yT} \Phi(y(0), y(T))^T \Psi.
\]
Set \(\eta := \bar{p}(0) + \nabla_{y_T} \Phi(y(0), y(T))^T \Psi\). Then
\[
0 \leq \Psi^T D\Phi(y(0), y(T))(\xi(0), \xi(T)) = [\bar{p} \cdot \xi]^T + \eta \cdot \xi(0)
\]
(3.169)
\[
= \int_0^T (\dot{\bar{p}}(t) \cdot \xi(t) + \bar{p}(t) \cdot \dot{\xi}(t)) dt + \eta \cdot \xi(0)
\]
\[
= \int_0^T (H(0, t, u(t), y(t), \bar{p}(t)) - \dot{H}(t)) dt + \eta \cdot \xi(0).
\]
Since we may take \(\xi(0)\) arbitrarily if follows that \(\eta = 0\), i.e., \(\bar{p}\) is the costate associated with \(\Psi\). We deduce then easily that the Hamiltonian inequality holds, showing that \((\beta = 0, \Psi, \bar{p})\) is a singular Pontryagin multiplier.

We obtain then the main result for the characterization of the absence of Pontryagin multipliers:

PROPOSITION 3.53. The set of singular Pontryagin multipliers is empty iff the following qualification condition holds:
(3.170)
\[
M' \text{ is a neighbourhood of } 0.
\]

PROOF. Similar to the one of proposition 3.45.

LEMMA 3.54. If \(K\Phi = \{0\}\) (case of equality constraints), the set of singular Pontryagin multipliers is empty iff \(M''\) is a neighbourhood of \(0\).

PROOF. Obvious consequence of proposition 3.53.

LEMMA 3.55. Let \(K\Phi\) have a nonempty interior. Then the set of singular Pontryagin multipliers is empty if there exists \(u\) satisfying the control constraints and \(\xi^0 \in \mathbb{R}^n\) such that \(\xi := \xi[u, \xi^0]\) satisfies
(3.171)
\[
\Phi(y(0), y(T)) + D\Phi(y(0), y(T))(\xi(0), \xi(T)) \in \text{int}(K\Phi).
\]

PROOF. Similar to the proof of lemma 3.50.

EXAMPLE 3.56. Consider the problem of minimizing \(\frac{1}{2} \int_0^T u(t)^2 dt\), with state equation \(\dot{y}(t) = u(t)^3\), and with fixed initial and final time \(y(0) = y(T) = 0\) and no control constraints. The unique solution \((\bar{u}, \bar{y})\) is the null trajectory, and any costate has a constant value (as a function of time). There exists a singular Lagrange multiplier with \(\bar{p}(t) = 1\) for all \(t\), \(\Psi_F = 1\). But there exists no singular Pontryagin multiplier, due to the fact that a singular pre-Hamiltonian of the form \(\bar{p}(t)u^3\) cannot have a minimum at zero (since \(\bar{p}(t) = \Psi_F \neq 0\)).

4. Proof of Pontryagin’s principle

4.1. Ekeland’s principle. If \(f : X \to \mathbb{R}\) has a finite infimum, we say that \(\bar{x} \in X\) is an \(\varepsilon\)-optimal for the problem of minimizing \(f\) if \(f(\bar{x}) \leq \inf f + \varepsilon\). The following result demonstrates that, under weak assumptions, it is possible then to construct another \(\varepsilon\)-optimal solution, close to \(\bar{x}\), that becomes the minimizer of a slightly perturbed objective function.
THEOREM 3.57. (Ekeland’s variational principle) Let $(E, \rho)$ be a complete metric space and $f : E \to \mathbb{R} \cup \{+\infty\}$ a lower semicontinuous function. Suppose that $\inf_{e \in E} f(e)$ is finite and let, for a given $\varepsilon > 0$, $\hat{e} \in E$ be an $\varepsilon$-minimizer of $f$, i.e., $f(\hat{e}) \leq \inf_{e \in E} f(e) + \varepsilon$. Then for any $k > 0$, there exists a point $\hat{e} \in E$ such that $\rho(\hat{e}, \hat{e}) \leq k^{-1}$ and

\begin{align}
(3.172) & \quad f(\hat{e}) \leq f(\hat{e}) - \varepsilon k \rho(\hat{e}, \hat{e}), \\
(3.173) & \quad f(\hat{e}) - \varepsilon k \rho(\hat{e}, \hat{e}) < f(e), \quad \forall e \in E, \ e \neq \hat{e}.
\end{align}

PROOF. Consider the multifunction $M : E \to 2^E$ defined by

\[ M(e) := \{ e' : f(e') + k \varepsilon \rho(e, e') \leq f(e) \}. \]

It is not difficult to see that $M(\cdot)$ is reflexive, i.e., $e \in M(e)$, and transitive, i.e., $e' \in M(e)$ implies $M(e') \subset M(e)$. Consider the function $v : \text{dom } f \to \mathbb{R}$ defined by

\[ v(e) := \inf \{ f(e') : e' \in M(e) \}. \]

We have that $\inf_E f \leq v(e) \leq f(e)$, and

\[ \varepsilon k \rho(e, e') \leq f(e) - v(e), \quad \forall e' \in M(e). \]

Since $f(\hat{e}) - v(\hat{e}) \leq f(\hat{e}) - \inf_E f \leq \varepsilon$, it follows that

\[ k \rho(\hat{e}, e) \leq 1, \quad \forall e \in M(\hat{e}). \]

Consequently, the diameter (i.e., the supremum of distances between two points) of $M(\hat{e})$ is less than or equal to $2k^{-1}$.

Consider a sequence $\{ e_n \}$ satisfying

\[ e_1 = \hat{e}, \quad e_{n+1} \in M(e_n), \quad \text{and} \quad f(e_{n+1}) \leq v(e_n) + \varepsilon 2^{-n}. \]

(By definition of $v(\cdot)$, such a sequence exists.) Since $M(e_{n+1}) \subset M(e_n)$ (by transitivity of $M(\cdot)$), we have $v(e_n) \leq v(e_{n+1})$, and since $v(e) \leq f(e)$, it follows that

\[ v(e_{n+1}) \leq f(e_{n+1}) \leq v(e_n) + \varepsilon 2^{-n} \leq v(e_{n+1}) + \varepsilon 2^{-n}, \]

implying $0 \leq f(e_{n+1}) - v(e_{n+1}) \leq \varepsilon 2^{-n}$. Combining this and (3.174), we get that $k \rho(e_{n+1}, e) \leq 2^{-n}$ for all $e \in M(e_{n+1})$, and hence the diameter of $M(e_{n+1})$ tends to zero as $n \to \infty$. Since in addition $e_{n+1} \in M(e_n)$ and $M(e_{n+1}) \subset M(e_n)$, it follows that $\{ e_n \}$ is a Cauchy sequence. By completeness of the space $E$ it follows that $\{ e_n \}$ converges to some $\hat{e} \in \text{dom}(f)$, and since the diameters of $M(e_n)$ tend to zero, $\bigcap_{n=1}^{\infty} M(e_n) = \{ \hat{e} \}$. Since $\{ e_n \}$ is included in the set $M(\hat{e})$ and $M(\hat{e})$ is closed, we have by (3.175) that $k \rho(\hat{e}, \hat{e}) \leq 1$, and by the definition of $M(\hat{e})$ that $f(\hat{e}) + \varepsilon k \rho(\hat{e}, \hat{e}) \leq f(\hat{e})$, so that (3.172) holds. In addition, since $M(\hat{e}) \subset M(e_n)$ for all $n$, and diam($M(e_n)$) $\to 0$, we have $M(\hat{e}) = \{ \hat{e} \}$, which implies (3.173).

Note that condition (3.172) of the above theorem implies that $\hat{e}$ is an $\varepsilon$-minimizer of $f$ over $E$, and that condition (3.173) means that $\hat{e}$ is the unique minimizer of the “perturbed” function $f(\cdot) + \varepsilon k \rho(\cdot, \hat{e})$ over $E$. In particular, by taking $k = \varepsilon^{-1/2}$ we obtain that for any $\varepsilon$-minimizer $\tilde{e}$ of $f$ there exists another $\varepsilon$-minimizer $\hat{e}$ such that $\rho(\tilde{e}, \hat{e}) \leq \varepsilon^{1/2}$ and $\hat{e}$ is the minimizer of the function $f(\cdot) + \varepsilon^{1/2} \rho(\cdot, \hat{e})$.

4.2. Ekeland’s metric and the penalized problem. Consider the set

\[ U_{ad} := L^\infty(0, T; U_{ad}) \]

endowed with Ekeland’s metric

\[ \rho_E(u, v) = \text{meas}\{ t ; u(t) \neq v(t) \}. \]

(3.177)

It is not difficult to check that $(U_{ad}, \rho_E)$ is a complete metric space. Since the initial state is not fixed, we need to consider the product space $U_{ad} \times \mathbb{R}^n$, endowed with the augmented Ekeland metric

\[ \rho_A[(u, y^0), (v, z^0)] := |z^0 - y^0| + \rho_E(u, v). \]

(3.178)
Again, it is easily checked that \((\mathcal{U}_{ad} \times \mathbb{R}^n, \rho_A)\) is a complete metric space.

We next introduce the \textit{composite cost function}, defined for \(\varepsilon > 0\):
\[
J_R^\varepsilon(u, y^0) := (\|J_R(u, y^0) - \bar{J} + \varepsilon^2\|^2 + d_{K_\Phi}(\Phi(y^0, y|u, y^0(T)))^2)^{1/2}.
\]

**Remark 3.58.** We will need the following results:
(i) The Euclidean distance to \(K_\Phi\), denoted by \(d_{K_\Phi}(\cdot)\), has a continuously differentiable square, and the expression of its derivative is, denoting by \(P_{K_\Phi}\) the orthogonal projection over \(K_\Phi\), see e.g. [9, Thm 4.77]:
\[
\nabla \left( \frac{1}{2} d_{K_\Phi}^2(w) \right) = w - P_{K_\Phi}(w), \quad w \in \mathbb{R}^n.
\]
(ii) The function \(\mathbb{R} \to \mathbb{R}, x \mapsto (x_+)^2\) is of class \(C^1\) with derivative \(2x_+\).

### 4.3. Proof of theorem 3.13.

a) We first deal with the case when \(U_{ad}\) is bounded. Setting \(\bar{J} := J^{1F}(\bar{u}, \bar{y})\), apply Ekeland’s principle to the perturbed problem
\[
\begin{equation}
\min_{u}\ J_R^\varepsilon(u, y^0); \quad u \in \mathcal{U}_{ad}; \quad y^0 \in \mathbb{R}^n. \tag{P_\varepsilon}
\end{equation}
\]
Since \(J_R^\varepsilon\) is a nonnegative function, and \(J_R^\varepsilon(\bar{u}, \bar{y}^0) = \varepsilon^2\), we have that \((\bar{u}, \bar{y}^0)\) is a \(\varepsilon^2\)-minimum of \((P_\varepsilon)\). Since \(U_{ad}\) is bounded, it is easily checked that the function \((u, y^0) \mapsto J_R^\varepsilon(u, y^0)\) is continuous for the augmented metric. By theorem 3.57, there exists \((u^\varepsilon, y^\varepsilon) \in \mathcal{U} \times \mathcal{Y}\) such that
\[
\begin{equation}
|y^\varepsilon(0) - \bar{y}^0| + \rho_E(u^\varepsilon, u) \leq \varepsilon,
\end{equation}
\]
\[
\begin{equation}
J_R^\varepsilon(u^\varepsilon, y^\varepsilon(0)) \leq J_R^\varepsilon(\bar{u}, \bar{y}^0) + \varepsilon(|y^0 - y^\varepsilon(0)|) + \rho_E(u, u^\varepsilon),
\end{equation}
\]
for all \((u, y^0) \in \mathcal{U}_{ad} \times \mathbb{R}^n\).

Necessarily \(J_R^\varepsilon(u^\varepsilon, y^\varepsilon(0)) > 0\), as was previously discussed. Define
\[
\beta_\varepsilon := \frac{(J_R(u^\varepsilon, y^\varepsilon(0)) - \bar{J} + \varepsilon^2)_+}{J_R(u^\varepsilon, y^\varepsilon(0))}; \quad \Psi_\varepsilon := \frac{\Phi(y^\varepsilon(0), y^\varepsilon(T)) - P_{K_\Phi}(\Phi(y^\varepsilon(0), y^\varepsilon(T)))}{J_R(u^\varepsilon, y^\varepsilon(0))};
\]
Then \(\beta_\varepsilon \in [0, 1]\), and \(|\Psi_\varepsilon| = \frac{J_R(u^\varepsilon, y^\varepsilon(0))}{J_R(u^\varepsilon, y^\varepsilon(0))} d_{K_\Phi}(\Phi(y^\varepsilon(0), y^\varepsilon(T)))\). Using the definition of \(J_R^\varepsilon(u^\varepsilon, y^\varepsilon(0))\),
\[
\beta_\varepsilon^2 + |\Psi_\varepsilon|^2 = 1.
\]
Using for \(x > 0\) and \(x' > 0\) the first-order expansion
\[
\sqrt{x} = \sqrt{x + \frac{1}{2}(x' - x)/\sqrt{x' + o(|x' - x|)},
\]
we get with remark 3.58, denoting by \(y\) the state associated with \((u, y^0)\):
\[
J_R^\varepsilon(u, y^0) = J_R^\varepsilon(u^\varepsilon, y^\varepsilon(0)) + \beta_\varepsilon (J_R(u, y^0) - J_R(u^\varepsilon, y^\varepsilon(0)))
+ (|\Psi_\varepsilon|, \Phi(y^\varepsilon(0), y^\varepsilon(T))(y_0 - y^\varepsilon(0), y(T) - y^\varepsilon(T)))
+ o(\rho_E(u - u^\varepsilon, 1) + |y^0 - y^\varepsilon(0)|).
\]
Let \(p^\varepsilon \in W^{1, \infty}(0, T, \mathbb{R}^n)\) be the unique solution of
\[
\begin{equation}
\begin{cases}
-\dot{p}^\varepsilon(t) = \nabla_y H(\beta_\varepsilon, t, u^\varepsilon(t), y^\varepsilon(t), p^\varepsilon(t)), \text{ a.e. } t \in (0, T); \\
p^\varepsilon(T) = \nabla_y T^{1F}(\beta_\varepsilon, y^\varepsilon(0), y^\varepsilon(T), \Psi_\varepsilon).
\end{cases}
\end{equation}
\]
**Lemma 3.59.** The following expansion holds:
\[
J_R^\varepsilon(u, y^0) = J_R^\varepsilon(u^\varepsilon, y^\varepsilon(0)) + \int_0^T \left( H(\beta_\varepsilon, t, u(t), y^\varepsilon(t), p^\varepsilon(t)) - H(\beta_\varepsilon, t, u^\varepsilon(t), y^\varepsilon(t), p^\varepsilon(t)) \right) dt
+ (\dot{p}^\varepsilon + \nabla_y p^{1F}(\beta_\varepsilon, y^\varepsilon(0), y^\varepsilon(T), \Psi_\varepsilon))(y^0 - y^\varepsilon(0)) + o(\rho_E(u - u^\varepsilon, 1) + |y^0 - y^\varepsilon(0)|).
\]
**Proof.** Combine (3.185) with arguments essentially similar to those in the proof of lemma 3.4, taking into account the fact that the initial state may vary and that it enters into the initial-final cost. \(\square\)
b) Setting \( u = u^\varepsilon \) in (3.182), obtain
\[
J^\varepsilon_R(u^\varepsilon, y^\varepsilon(0)) \leq J^\varepsilon_R(u^\varepsilon, y^0) + \varepsilon|y^0 - y^\varepsilon(0)|, \quad \text{for all } y^0 \in \mathbb{R}^n.
\]
Substituting in (3.188) the expression of \( J^\varepsilon_R(u^\varepsilon, y^0) \) given in (3.187) (note that, since \( u = u^\varepsilon \), the integral term vanishes) we get
\[
0 \leq (p^\varepsilon_0 + \nabla y^0 L^{IF}(\beta^\varepsilon, y^\varepsilon(0), y^\varepsilon(T), \Psi^\varepsilon))(y^0 - y^\varepsilon(0)) + \varepsilon|y^0 - y^\varepsilon(0)| + o(|y^0 - y^\varepsilon(0)|).
\]
Taking \( y^0 = y^\varepsilon(0) - \rho(p^\varepsilon_0 + \nabla y^0 L^{IF}(\beta^\varepsilon, y^\varepsilon(0), y^\varepsilon(T), \Psi^\varepsilon)) \), with \( \rho \downarrow 0 \), deduce that
\[
|p^\varepsilon_0 + \nabla y^0 L^{IF}(\beta^\varepsilon, y^\varepsilon(0), y^\varepsilon(T), \Psi^\varepsilon)| \leq \varepsilon.
\]

Taking \( y^0 = y^\varepsilon(0) \) in (3.182), obtain
\[
J^\varepsilon_R(u^\varepsilon, y^\varepsilon(0)) \leq J^\varepsilon_R(u, y^\varepsilon(0)) + \varepsilon \rho_E(u, u^\varepsilon), \quad \text{for all } u \in U_{ad}(u).
\]
Substituting in (3.191) the expression of \( J^\varepsilon_R(u, y^\varepsilon(0)) \) provided by (3.187), get
\[
0 \leq \int_0^T (H(\beta^\varepsilon, t, u(t), y^\varepsilon(t), p^\varepsilon(t)) - H(\beta^\varepsilon, t, u^\varepsilon(t), y^\varepsilon(t), p^\varepsilon(t))) \, dt
\ + \varepsilon \rho_E(u, u^\varepsilon) + o(||u - u^\varepsilon||_1).
\]
Equivalently, denoting \( I_{\varepsilon} := \{ t \in (0, T); u(t) \neq u^\varepsilon(t) \} \), we have that
\[
0 \leq \int_{I_{\varepsilon}} (H(\beta^\varepsilon, t, u(t), y^\varepsilon(t), p^\varepsilon(t)) - H(\beta^\varepsilon, t, u^\varepsilon(t), y^\varepsilon(t), p^\varepsilon(t)) + \varepsilon) \, dt + o(\text{meas}(I_{\varepsilon})�).
\]
By techniques similar to those presented after (3.15), we obtain that
\[
H(\beta^\varepsilon, t, u^\varepsilon(t), y^\varepsilon(t), p^\varepsilon(t)) \leq H(\beta^\varepsilon, t, v, y^\varepsilon(t), p^\varepsilon(t)) + \varepsilon,
\quad \text{for all } v \in U_{ad}, \text{ for a.a. } t \in (0, T).
\]

d) We now pass to the limit in (3.183)-(3.194). Since \( \rho_E(u, u^\varepsilon) \to 0 \) and \( U_{ad} \) is bounded, we have that \( y^\varepsilon \to \bar{y} \) uniformly. By (3.183), \( (\beta^\varepsilon, \Psi^\varepsilon) \) has a unit norm limit point \( (\bar{\beta}, \bar{\Psi}) \), so that the Hamiltonian inequality holds. Also, passing to the limit in the relation (consequence of the definition of projection)
\[
\Psi^\varepsilon \cdot (k - p_{K_{\Psi}}(\Phi(y^\varepsilon(0), y^\varepsilon(T)))) \leq 0, \quad \text{for all } k \in K_{\Psi},
\]
and since \( \Phi(\bar{y}^0, \bar{y}(T)) \in K_{\Psi} \), we get that \( \bar{\Psi} \in N_{K_{\Psi}}(\Phi(\bar{y}^0, \bar{y}(T))) \). There is no difficulty in passing to the limit in (3.190), and in (3.186) (the costate \( p^\varepsilon \) remains bounded since \( \beta^\varepsilon \) and \( \Psi^\varepsilon \) are bounded). Therefore the costate equation is satisfied. Finally, let \( \{\varepsilon_k\} \) be a sequence associated with the limit-point \( (\bar{\beta}, \bar{\Psi}) \). For each \( k \), (3.194) is satisfied on a set \( I_{\varepsilon_k} \) of full measure in \( [0, T] \). We may then pass to the limit in (3.194) on the set \( \bigcap_k I_{\varepsilon_k} \), that is also of full measure. The conclusion follows.

e) We end the proof by dealing with the case when \( U_{ad} \) is not bounded. Set
\[
U_R := U_{ad} \cap B_R,
\]
where \( B_R \) is the ball of radius \( R \) and center 0 in \( \mathbb{R}^m \), and \( U_R := L^\infty(0, T, U_R) \). If \( R \geq ||\bar{u}||_\infty \), then \( \bar{u} \) is solution of the problem
\[
\text{Min } J_R(u); \quad u \in U_R.
\]
By point d), there exists \( (\beta_R, \Psi^R) \in \mathbb{R}_+ \times N_{K_{\Psi}}(\Phi(\bar{y}^0, \bar{y}(T))) \), with \( (\beta, \Psi) \neq 0 \), and \( p^R \in \mathcal{Y} \), such that
\[
\begin{align*}
-\dot{p}^R(t) &= \nabla_y H(\beta_R, t, u(t), y(t), p^R(t)), \quad \text{for a.a. } t, \\
u(t) &\in \text{argmin}_{u \in U_R} H(\beta_R, t, u, y(t), p^R(t)) \quad \text{for a.a. } t,
\end{align*}
\]
\[
\begin{align*}
-\dot{p}^R_0 &= \nabla_y L^{IF}(\beta_R, \bar{y}^0, \bar{y}(T), \Psi^R); \\
p^R(T) &= \nabla_y L^{IF}(\beta_R, \bar{y}^0, \bar{y}(T), \Psi^R); \\
\Psi^R &\in N_{K_{\Psi}}(\Phi(\bar{y}^0, \bar{y}(T))).
\end{align*}
\]

Multiplying \((\beta_R, \Psi^R, p^R)\) by \(|(\beta_R, \Psi^R)|^{-1}\), we may assume that \(|(\beta_R, \Psi^R)| = 1\). It easily follows that \(p^R\) is bounded in \(\mathcal{Y}\). We conclude by passing to the limit in an extracted sequence for which \((\beta_R, \Psi^R)\) converges, in (3.197)-(3.201), using again the fact that a countable intersection of sets of full measure is of full measure.
1. Shooting algorithm

1.1. Unconstrained problems. We consider for the sake of simplicity an autonomous unconstrained optimal control problem: the state equation and cost function are resp.

\[
\dot{y}(t) = f(u(t), y(t)), \quad \text{for a.a. } t \in (0, T); \quad y(0) = y^0,
\]

\[
J(u, y) := \int_0^T \ell(u(t), y(t)) dt + \varphi(y(T)).
\]

Here \(y^0\) is given and the functions \(f, \ell, \varphi\) are of class \(C^2\) with locally Lipschitz second derivatives, and \(f\) is Lipschitz. As before the control and state spaces are

\[
\mathcal{U} := L^\infty(0, T, \mathbb{R}^m); \quad \mathcal{Y} := W^{1,\infty}(0, T, \mathbb{R}^n).
\]

So, the control to state mapping \(y[u]\) is well-defined \(\mathcal{U} \rightarrow \mathcal{Y}\) and we denote the reduced cost by

\[
J_R(u) := J(u, y[u]).
\]

The optimal control problem is

\[
\min_{u \in \mathcal{U}} J_R(u).
\]

We assume that

\[
\text{The control variable } \bar{u} \in \mathcal{U} \text{ is a local solution, and a continuous function of time.}
\]
We denote by \( \bar{y} := y[\bar{u}] \) the associated state. Then, by theorem 2.48:

\[
H_u(u(t), \bar{y}(t), \bar{p}(t)) = 0, \quad \text{for all } t \in [0, T], 
\]

(4.7) 

(We write this relation for all \( t \) and not a.e. since \( \bar{u} \) is continuous.) We assume that the following strong Legendre-Clebsch condition holds, where \( I_d \) is the identity matrix: there exists \( \alpha > 0 \) such that

\[
\bar{H}_{uu}(t) \succeq \alpha I_d, \quad \text{a.e. over } [0, T].
\]

(4.8) 

Remember that the corresponding condition with \( \alpha = 0 \) is a necessary condition for Pontryagin minima, see remark 3.5.

Applying the implicit function theorem 2.19 to (4.7), we deduce thanks to (4.8) that, for \( \tau \in [0, T] \) and \((u, y, p)\) close enough of \((\bar{u}(\tau), \bar{y}(\tau), \bar{p}(\tau))\), there exists \( \Upsilon : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{m} \) of class \( C^1 \) such that the relation \( H_u(u, y, p) = 0 \) is equivalent to \( u = \Upsilon(y, p) \). Putting if necessary the time as an additional state, we can “stick” together the functions \( \Upsilon \) for different times and we obtain in this way that for a certain \( \varepsilon > 0 \), and all \( t \in [0, T] \):

\[
\text{If } |y - \bar{y}(t)| + |p - \bar{p}(t)| < \varepsilon, \text{ then } H_u(u, y, p) = 0 \iff u = \Upsilon(y, p).
\]

(4.9) 

Set

\[
\Upsilon(t) := \Upsilon(\bar{y}(t), \bar{p}(t)), \quad t \in [0, T].
\]

Remark 4.1. When in the sequel we make reference to (4.8), one should understand \( \bar{H}_{uu}(t) \) as the second derivative w.r.t. the control variable of the pre-Hamiltonian specific to each problem.

Remark 4.2. Derivating in time the relation \( \nabla_u H(\bar{u}(t), \bar{y}(t), \bar{p}(t)) = 0 \), we get then

\[
\bar{H}_{uu}(t)\dot{\bar{y}}(t) + \bar{H}_{uy}(t)\dot{\bar{y}}(t) - f(u(t))\nabla_y H(t) = 0,
\]

which allows to express \( \dot{\bar{u}}(t) = \dot{\Upsilon}(t) \) as a function of \((\bar{u}(t), \bar{y}(t), \bar{p}(t))\).

Eliminating the control in the state and costate equations we see that the pair \((\bar{y}, \bar{p})\) is solution of the autonomous differential equation

\[
\begin{cases}
\dot{\bar{y}}(t) = f(\Upsilon(t), y(t)), \\
-\dot{\bar{p}}(t) = \nabla_y H(\Upsilon(t), y(t), p(t)),
\end{cases}
\]

(4.11) 

for all \( t \in [0, T] \), with initial-final conditions

\[
y(0) = y^0; \quad p(T) = \nabla \varphi(y(T)).
\]

(4.12) 

Remark 4.3. In the particular case of a linear quadratic problem, i.e. when the state equation is linear and the cost function is quadratic, we recover relations (2.140) which are themself linear.

Let us introduce the shooting parameter \( p^0 \in \mathbb{R}^n \), and denote (if it exists) by \((y[p^0], p[p^0])\) the solution of (4.11) with the initial condition \((y^0, p^0)\). Consider the shooting function \( \mathbb{R}^n \to \mathbb{R}^n \) defined by

\[
\Upsilon(p^0) := \nabla \varphi(y[p^0](T)) - p[p^0](T).
\]

(4.13) 

Then the following obviously holds:

Lemma 4.4. The two point boundary value problem (TPBVP) (4.11)-(4.12) has a solution such that \( p(0) = p^0 \) iff \( p^0 \) is a zero of the shooting function.

So, we reduce the search for a solution of the optimality conditions of the optimal control problem, to the resolution of the finite dimensional shooting equation

\[
\Upsilon(p^0) = 0.
\]

(4.14) 

By our hypothesis of \( C^2 \) data with Lipschitz second derivatives, we see that the shooting function is \( C^1 \) with Lipschitz derivatives. The shooting equation (4.14) is usually solved by a variant of
Newton’s method, see section 2.2 of chapter 2. The latter needs the evaluation of the Jacobian of \( T \). Let us give its expression when \( p^0 = \bar{p}(0) \) and so \((y[p^0], p[p^0]) = (\bar{y}, \bar{p})\). By the chain rule, we have that 

\[
D\mathcal{T}(p^0)q^0 = D^2\varphi(\bar{y}(T))z(T) - q(T),
\]

where \((z(T), q(T))\) is solution of the linearization of (4.11) with initial condition

\[
(z(0), q(0)) = (0, q^0).
\]

We can express this linearization as follows. Relation (4.11) is locally equivalent to

\[
\begin{align*}
\dot{y}(t) &= f(u(t), y(t)), \\
-p(t) &= \nabla_y H(u(t), y(t), p(t)), \quad \text{for all } t \in [0, T],
\end{align*}
\]

Indeed, (4.11) was obtained by eliminating the control variable in the above system, and the local equivalence of (4.11) and (4.17) follow from the IFT. It can be easily deduced from the IFT that both systems evaluated at a point where \( u(t) = \bar{y}(y(t), p(t)) \), have equivalent linearizations.

In particular, \((z, q)\) satisfies the linearization of (4.11) at \((\bar{y}, \bar{p})\) iff there exists \( v \in \mathcal{U} \) such that \((v, z, q)\) satisfies the linearization of (4.17) at \((\bar{u}, \bar{y}, \bar{p})\) and the latter is

\[
\begin{align*}
\dot{z}(t) &= \frac{\partial f}{\partial v}(t)(v(t), z(t)), \\
-\dot{q}(t) &= \frac{\partial f}{\partial \mathcal{U}}(t)q(t) + H_{yu}(t)v(t) + H_{yy}(t)z(t), \quad \text{for all } t \in [0, T],
\end{align*}
\]

We have proved the following:

**Lemma 4.5.** The expression of the Jacobian of the shooting function is given by (4.15), where \((v, z, q)\) satisfy (4.16) and (4.18).

### 1.2. Problems with initial-final state constraints

We restrict the analysis to the case of the minimization of the function (already defined in (3.26), but in the autonomous case)

\[
J^{IF}(u, y) := \int_0^T \ell(u(t), y(t))dt + \varphi(y(0), y(T)),
\]

with initial-final equality state constraints

\[
\Phi(y(0), y(T)) = 0.
\]

Now the initial state \( y^0 \) is free. As before, we assume that the data are \( C^2 \) with Lipschitz second derivatives, and denote by \( y[u, y^0] \) the state associated with \((u, y^0) \in \mathcal{U} \times \mathbb{R}^n\), and by \( J_R \) the reduced cost defined by

\[
J_R(u, y^0) = J^{IF}(u, y[u, y^0]).
\]

So, the optimal control problem is

\[
\begin{align*}
\min_{u \in \mathcal{U}, y^0 \in \mathbb{R}^n} \quad J_R(u, y^0) \quad \text{s.t. (4.20).}
\end{align*}
\]

Let \((\bar{u}, \bar{y}) \in \mathcal{U} \times \mathcal{Y}\) be a locally optimal trajectory whose initial state is denoted by \( y^0 \). We assume that it satisfies the following qualification condition, taken from exercise 3.48:

\[
\begin{align*}
\text{(4.23)} \quad \text{The mapping } \mathcal{U} \times \mathbb{R}^n \to \mathcal{Y}, (v, z^0) \mapsto \mathcal{D}\Phi(\bar{y}(0), \bar{y}(T))(z[v, z^0](0), \bar{z}[v, z^0](T)) \text{ is onto.}
\end{align*}
\]

By lemma 3.49, the regular Lagrange multiplier (with \( \beta = 1 \)) is unique. In this setting, the shooting function has for parameters \((y^0, p^0, \Psi)\), the initial values of the state and costate, as well as the multiplier associated with the constraints. This function expresses the initial-final conditions for the costate as well as the constraints:

\[
T^{IF}(y^0, p^0, \Psi) := \left( \begin{array}{c}
p^0 + \nabla_y L^{IF}(1, y^0, y(T), \Psi) \\
-p(T) + \nabla_y L^{IF}(1, y^0, y(T), \Psi) \\
\Phi(y^0, y(T))
\end{array} \right),
\]

\[
\text{(4.24)}
\]
where \((y, p)\) is the result of the integration of (4.11) with the initial condition \((y^0, p^0)\), and \(L^{IF}\) is defined in (3.31). Note that we have the same number of equations and unknown. here again, the shooting function is \(C^1\) with Lipschitz derivatives.

We recall that \((\hat{y}^0, p(0), \hat{\Psi})\) is a regular zero of \(T^{IF}\) if the Jacobian of \(T^{IF}\) at this point is invertible. This is known to ensure the local quadratic convergence of Newton’s method (and, by the IFT, the stability w.r.t. to perturbations of the solution of the shooting function). We will see in section 2.4 that this property is closely related to the second-order optimality conditions, and allows to ensure the convergence of the discretized problem.

### 1.3. Time discretization.

We discretize the state equation (2.111) by the Euler method with constant time step, assuming the control to be constant over each time step:

\[
y^{k+1} = y^k + h f(u^k, y^k), \quad k = 0, \ldots, N - 1; \quad y^0 = y^0,
\]

with \(h = T/N, N\) positive integer, \(u^k \in \mathbb{R}^m\) and \(y^k \in \mathbb{R}^n\). The space for the discrete controls variables is \(U^N := \mathbb{R}^{Nm}\), and the one for discrete states is \(Y^N := \mathbb{R}^{(N+1)n}\). With each \(u \in U^N\) is associated a unique state \(y[u, N] \in Y^N\) solution of (4.25). The corresponding discretization of the cost function (2.120) is

\[
J^N(u, y) := h \sum_{k=0}^{N-1} \ell(u^k, y^k) + \varphi(y^N).
\]

So, the unconstrained discrete optimal control problem (with \(y^0\) fixed) is

\[
\text{Min } J^N(u, y); \quad \text{subject to } (4.25).
\]

Writing the Lagrangian in the form

\[
\varphi(y^N) + \sum_{k=0}^{N-1} (h \ell(u^k, y^k) + p^{k+1} \cdot (y^k + h f(u^k, y^k) - y^{k+1})) + q \cdot (y^0 - y^0),
\]

we obtain the discrete costate equation by setting to zero the partial derivative of the Lagrangian function w.r.t. the state, i.e.

\[
\begin{align*}
p^k &= p^{k+1} + h (\nabla_y \ell(u^k, y^k) + f_y(u^k, y^k) \cdot p^{k+1}), \quad k = 1, \ldots, N - 1; \\
p^N &= \nabla \varphi(y^N), \\
-q &= p^1 + h (\nabla_y \ell(u^0, y^0) + f_y(u^0, y^0) \cdot p^1).
\end{align*}
\]

The reduced cost is \(J^N_R(u) := J^N(u, y[u, N])\). By the same kind of arguments as for the case of continuous time, and denoting \(y = y[u, N]\), we can check that

\[
D J^N_R(u) v = \sum_{k=0}^{N-1} H_u(u^k, y^k, p^{k+1}) v^k.
\]

A necessary condition of local optimality is therefore (compare to (4.7)):

\[
H_u(u^k, y^k, p^{k+1}) = 0, \quad k = 0, \ldots, N - 1.
\]

Set \(t_k = kh, k = 0\) to \(N\), and let the strong Legendre-Clebsch condition (4.8) be satisfied. If

\[
|\dot{u}(t_k) - u^k| + |\dot{y}(t_k) - y^k|
\]

is for all \(k\) small enough, we can (adding if necessary the time as a state variable) eliminate the control as function of the state and costate:

\[
u^k = \Upsilon(y^k, p^{k+1}),
\]

which allows, abbreviating \(\Upsilon(y^k, p^{k+1})\) into \(\Upsilon^k\), to reduce the discrete optimality problem to the discrete implicit dynamical system in \((y, p)\), for \(k = 0, \ldots, N - 1:\)

\[
\begin{align*}
y^{k+1} &= y^k + h f(\Upsilon^k, y^k), \\
p^k &= p^{k+1} + h (\nabla_y \ell(\Upsilon^k, y^k) + f_y(\Upsilon^k, y^k) \cdot p^{k+1})
\end{align*}
\]

64
with the end conditions
\begin{equation}
(4.34) \quad y^0 = y^0; \quad p^N = \nabla \varphi(y^N).
\end{equation}
Taking \( p^0 = p^0 \) as parameter and denoting by \( y[p^0], p[p^0] \) the solution of (4.33) with initial condition \((y^0, p^0)\), we obtain the discrete shooting function
\begin{equation}
(4.35) \quad T^N(p^0) := \nabla \varphi(y[p^0]) - p^N[p^0].
\end{equation}
When \( h \) is small enough, under the hypothesis (4.8), \( T^N \) is defined and uniformly converges (as well as its derivative) over a neighbourhood of \( \bar{p}(0) \).

In the sequel we assume that
\begin{equation}
(4.36) \quad \bar{p}(0) \text{ is a regular zero of the shooting function } T.
\end{equation}
We recall that this hypothesis means that \( T \) has an invertible Jacobian at the point \( \bar{p}(0) \). We can then show the following:

Lemma 4.6. The shooting function \( T^N \) has in the neighbourhood of \( \bar{p}(0) \) a unique solution, with which is associated a solution \((u_h, y_h, p_h)\) of the discrete optimality conditions, and the order of convergence is 1 in the sense that
\begin{equation}
(4.37) \quad \max_{0 \leq k \leq N-1} \left( |\tilde{u}(t_k) - u_h^k| + |\tilde{y}(t_k) - y_h^k| + |\bar{p}(t_k) - p_h^{k+1}| \right) + |y(T) - y^N| = O(h).
\end{equation}

Proof. Let \( V \) be a closed ball of center \( \bar{p}(0) \) over which the shooting function \( T^N \) is well-defined and converges uniformly, as well as its Jacobian \( T^N \). Given \( \varepsilon > 0 \) we have the existence of \( h_\varepsilon > 0 \) such that, for \( h < h_\varepsilon \), we have
\begin{equation}
(4.38) \quad |T^N(p) - T(p)| + |D T^N(p) - DT(p)| \leq \varepsilon, \quad \text{for all } p \in V.
\end{equation}
In fact, reducing \( V \) and \( h_\varepsilon \) if necessary and setting \( A := DT(\bar{p}(0)) \) we may assume that
\begin{equation}
(4.39) \quad |T^N(p)| + |D T^N(p) - A| \leq \varepsilon, \quad \text{for all } p \in V.
\end{equation}
Since we discretized by Euler’s method we have that for some \( c_0 > 0 \)
\begin{equation}
(4.40) \quad |T^N(\bar{p}(0))| \leq c_0 h.
\end{equation}
Let \( p^j \), \( j \in \mathbb{N} \) the sequence such that
\begin{equation}
(4.41) \quad p^0 := \bar{p}(0); \quad T^N(p^j) + A(p^{j+1} - p^j) = 0, \quad j \in \mathbb{N},
\end{equation}
so that
\begin{equation}
(4.42) \quad |p^{j+1} - p^j| = |A^{-1} T^N(p^j)| \leq \|A^{-1}\| |T^N(p^j)|.
\end{equation}
On the other hand, by (4.41):
\begin{equation}
(4.43) \quad T^N(p^{j+1}) = T^N(p^j) + \int_0^1 D T^N(p^j + \theta(p^{j+1} - p^j))(p^{j+1} - p^j)d\theta.
\end{equation}
Let \( j_0 \) be such that \( p^{j+1} \in V \) for \( j \leq j_0 \). Then for \( j \leq j_0 \), we have that \( p^j + \theta(p^{j+1} - p^j) \in V \). With (4.39) and (4.43) we get that
\begin{equation}
(4.44) \quad |T^N(p^{j+1})| \leq \max_{\theta \in [0,1]} |T^N(p^j + \theta(p^{j+1} - p^j)) - A||p^{j+1} - p^j| \leq \varepsilon|p^{j+1} - p^j|.
\end{equation}
Choosing \( \varepsilon \) such that \( \varepsilon\|A^{-1}\| \leq 1/2 \), we deduce with (4.42) and the previous display that
\begin{equation}
(4.45) \quad |p^{j+2} - p^{j+1}| \leq \varepsilon\|A^{-1}\||p^{j+1} - p^j| \leq \frac{1}{2}\|p^{j+1} - p^j| \leq \frac{1}{2^{j+1}}|p^1 - p^0|, \quad \text{for all } j \leq j_0.
\end{equation}
Since
\begin{equation}
(4.46) \quad |p^1 - p^0| \leq \|A^{-1}\||T^N(p^0)| = O(h)
\end{equation}
we may take $j_0 \geq 1$ for $h$ small enough. By (4.45),

\[ |p^{j_0+2} - p^0| \leq |p^{j_0+2} - p^1| + |p^1 - p^0| \leq 2|p^1 - p^0| = O(h), \]

so that all the sequence belongs to $V$ and $p^j$ linearly converges to $p_h$ such that

\[ |p_h - p^0| \leq 2|p^1 - p^0| \leq 2\|A^{-1}\|T^N(p^0)|. \]

By (1.3), $T^N(p_h) = \lim_j T^N(p^j) = 0$. Relation (4.37) follows. Finally if the shooting function has two zeros $p$ and $p'$ then setting $q = p' - p:

\[ 0 = T^N(p') - T^N(p) = A'q, \]

where $A' := \int_0^1 DT^N(p + \theta q)d\theta$.

Taking $\varepsilon$ small enough we obtain that $A'$ is so close to $A$ that it must be invertible and therefore $q = 0$. 

**Remark 4.7.** With higher order discretization scheme, the above proof also holds. It follows from (4.48) that the error order for the state, costate and control is of the error order of the numerical integration scheme. In the case of Runge-Kutta schemes, it can be shown that the discretization of the optimal control problems provides a discretization of the state-costate dynamics that may have a smaller error order. See Hager [26], Bonnans and Laurent-Varin [12].

## 2. Second-order optimality conditions

In this section we connect the well-posedness of the shooting algorithm to the second-order optimality conditions for the optimal control problem.

### 2.1. Expansion of the reduced cost.**

We have seen in chapter 2 that the reduced cost $J_R(\cdot) : U \to \mathbb{R}$ defined in (4.4) is of class $C^1$, but since the data are $C^2$, we deduce by similar arguments that it is of class $C^2$. Denote by $Q_0(\nu) := J''_R(\bar{u})(v,v)$ the quadratic form associated with its Hessian (second derivative). The second order Taylor expansion of $J_R$ can be written as:

\[ J_R(\bar{u} + v) = J_R(\bar{u}) + J'_R(\bar{u})v + \frac{1}{2}Q_0(v) + R_2(v), \]

where the remainder of the Taylor expansion satisfies, by (2.34)-(2.35):

\[ R_2(v) = \int_0^1 (1 - \sigma)(J''_R(\bar{u} + \sigma v) - J''_R(\bar{u}))(v,v)d\sigma = o(\|v\|_2^3). \]

A necessary condition for local optimality in $U$ is therefore

\[ J'_R(\bar{u}) = 0 \text{ and } Q_0(v) \geq 0, \text{ for all } v \in U. \]

Consider next the quadratic growth condition:

\[ J''_R(\bar{u}) = 0 \text{ and } Q_0(v) \geq 0, \text{ for all } v \in U. \]

For some $\alpha > 0$: if $\|v\|_\infty$ is small enough, then $J_R(\bar{u} + v) \geq J_R(\bar{u}) + \frac{1}{2}\alpha\|v\|_2^2$, and the second order sufficient condition:

\[ Q_0(v) \geq 2\alpha\|v\|_2^2 \text{ for all } v \in U. \]

**Theorem 4.8.** Let $\bar{u} \in U$ be a critical point of $J_R$, i.e., $J'_R(\bar{u}) = 0$. Then conditions (4.53) and (4.54) are equivalent.

**Proof.** If the quadratic growth condition (4.53) holds, then $\bar{u}$ is a local solution of the problem of minimizing $J_{R,\alpha}(u) := J_R(u) - \frac{1}{2}\alpha\|u - \bar{u}\|_2^2$. This is the reduced cost for the optimal control problem with the same dynamics and final cost, and perturbed integrand of the integral cost: $\ell_{\alpha}(t,u,y) := \ell(u,y) - \frac{1}{2}\alpha|u - \bar{u}(t)|^2$, and $J''_{R,\alpha}(\bar{u})(v,v) = J''_R(\bar{u})(v,v) - \alpha\|v\|_2^2$. So the necessary condition $J''_{R,\alpha}(\bar{u})(v,v) \geq 0$ for any $\nu$ holds, which amounts to the second order sufficient condition (4.54).
Conversely, let the second order sufficient condition (4.54) hold. In lemma 4.10 below, we check that
\begin{equation}
|R_2(v)| = O(\|\nu\|_\infty \|v\|_2^2).
\end{equation}
So, when $\|\nu\|_\infty$ is small enough, $|R_2(v)| \leq \frac{1}{2} \alpha' \|v\|_2^2$. In view of the Taylor expansion (4.50), we deduce that
\begin{equation}
J_R(\bar{u} + v) = J_R(\bar{u}) + \frac{1}{2} \alpha' \|v\|_2^2,
\end{equation}
so that the quadratic growth condition (4.53) holds for any $\alpha < \alpha'$. 

Remember that we will complete the above proof by establishing (4.55) in lemma 4.10.

2.2. More on the Hessian of the reduced cost. The Lagrangian associated with the unconstrained problem (4.5) is, see (3.32):
\begin{equation}
J(\bar{u}, \bar{y}) + \int_0^T \mathbf{p}(t) \cdot (f(\bar{u}(t), \bar{y}(t)) - \dot{\bar{y}}(t))dt = \int_0^T H(\bar{u}(t), \bar{y}(t), \bar{p}(t))dt + \varphi(\bar{y}(T)) - \int_0^T \mathbf{p}(t) \cdot \dot{\bar{y}}(t)dt.
\end{equation}
Denote by $D^2\bar{H}(t)$ the second derivative of the pre-Hamiltonian w.r.t. to $(\bar{u}, \bar{y})$, at the point $(\bar{u}(t), \bar{y}(t), \bar{p}(t))$. The second derivative of the Lagrangian in the direction $(\nu, z) \in \mathcal{U} \times \mathcal{Y}$ is
\begin{equation}
\Omega_0(\nu, z) := \int_0^T D^2\bar{H}(t)(\nu(t), z(t))^2dt + \varphi''(\bar{y}(T))(z(T))^2.
\end{equation}
We can check that, denoting by $z[\nu]$ the solution of (2.115) with $z^0 = 0$:

**Lemma 4.9.** We have $Q_0(\nu) = \Omega_0(\nu, z[\nu]),$ for all $\nu \in \mathcal{U}$.

**Proof.** We have shown in lemma 2.24 that the Hessian of a reduced cost in a given direction is equal to the Hessian of the 'reduction' Lagrangian, provided the state and costate are those associated with the control. The result follows. 

**Lemma 4.10.** Relation (4.55) holds.

**Proof.** It is enough to get the result in the case of a zero final cost. Let $\nu \in \mathcal{U}$ and $z = z[\nu]$. Set, for $\sigma \in [0, 1],$
\begin{equation}
H_\sigma(t) := H(t, \bar{u}(t) + \sigma \nu(t), \bar{y}(t) + \sigma \bar{z}(t), \bar{p}(t)); \quad \delta H_\sigma(t) := H_\sigma(t) - H_0(t),
\end{equation}
and denote by $\delta H''_\sigma(t)$ the Hessian of $\delta H_\sigma(t)$ w.r.t. $(u, y)$. The mapping $u \mapsto y[u]$ being Lipschitz $\mathcal{U} \to \mathcal{Y}$, we have that for some $c > 0$ not depending on $t$:
\begin{equation}
\|\delta H''_\sigma(t)\| \leq c\|\nu\|.
\end{equation}
By lemma 4.9 and the similar expression of the Hessian of Lagrangian at the point $\bar{u} + \sigma \nu$:
\begin{equation}
|J''_{R_\bar{u}}((\bar{u} + \sigma \nu) - J''_{R_\bar{u}}(\bar{u}))(\nu, \nu)| \leq \int_0^T |\delta H''_\sigma(t)||\nu(t)||^2dt \leq c\|\nu\|\|\nu\|_2^2.
\end{equation}
We conclude then with (4.51). 

**Theorem 4.11.** Let second order sufficient condition (4.54) hold. Then $\bar{p}(0)$ is a regular zero of the shooting function $\mathcal{T}$.

**Proof.** We prove a more general statement in theorem 4.14 (since the qualification condition automatically holds when the initial state is given and there is no terminal constraints). 

**Remark 4.12.** Conversely it can be proved that, if (4.54) does not hold, $\bar{p}(0)$ is not a regular zero of the shooting function $\mathcal{T}$.
2.3. Case of initial-final equality constraints. We come back to problem (4.22), assuming the qualification condition (4.23) to hold so that we skip the argument \( \beta = 1 \) and set \( \lambda := (\bar{p}, \bar{w}). \) The Lagrangian (to be taken with \( \beta = 1 \)) is stated in (3.32). Its Hessian at the point \((\bar{u}, \bar{y}, \bar{p}, \bar{w})\), in the direction \((v, z) \in U \times Y\) is

\[
(4.61) \quad \Omega_0[\bar{\lambda}](v, z) := \int_0^T \dot{D}^2 H(t)(v(t), z(t))^2 dt + \dot{D}^2 L_{IF}^\beta(\bar{y}(0), \bar{y}(T), \bar{w})(z(0), z(T))^2,
\]

where by \( \dot{D}^2 L_{IF}^\beta \) we understand the second derivatives of \( L_{IF}^\beta \) w.r.t. \((y^0, y_T)\). Set

\[
Q[\bar{\lambda}](v, z^0) := \Omega[\bar{\lambda}](v, z[v, z^0]).
\]

Since this quadratic form is continuous in the norm of \( U_2 \), we can extend it to \((v, z^0) \in U_2 \times \mathbb{R}^n\).

The second-order conditions involve the critical cone whose elements, called critical directions, are defined as follows. Set \( z = z[v, z^0] \)

\[
(4.62) \quad C_2(\bar{u}, \bar{y}) := \left\{ (v, z^0) \in U_2 \times \mathbb{R}^n; \varphi'(\bar{y}(0), \bar{y}(T))(z(0), z(T)) = 0 \right\}.
\]

Note that for equality constraints the critical cone is actually a closed subspace of \( U \times Y \).

\textbf{Theorem 4.13.} If \((\bar{u}, \bar{y})\) is a qualified local solution of (4.22), with unique associated Lagrange multiplier \( \lambda \), for any critical direction \((v, z^0)\), we have that \( Q[\bar{\lambda}](v, z^0) \geq 0 \).

\textbf{Proof.} This follows from proposition 2.32. \( \square \)

2.4. Links with the shooting algorithm. We keep the framework of section 1.2 (the initial-final constraints are equalities) with \((\bar{u}, \bar{y})\) local solution of the problem. The shooting function \( T_{IF} \) is defined in (4.24). We have discussed in section 1.2 the interest of checking if the Jacobian of this function is invertible at the solution. This boils down to check if the kernel of the Jacobian is reduced to \( \{0\} \). We know that the qualification condition (4.23) ensures the uniqueness of the multiplier denoted by \( \lambda \) (with the normalisation \( \beta = 1 \)), and that, by theorem 4.8, (4.53) is equivalent then for some \( \alpha > 0 \) to

\[
(4.63) \quad Q[\bar{\lambda}](v, z^0) \geq \alpha(||v||^2 + |z^0|^2), \text{ for all } (v, z^0) \in C_2(\bar{u}, \bar{y}).
\]

In particular, the quadratic cost \( Q[\bar{\lambda}](v, z^0) - \alpha(||v||^2 + |z^0|^2) \) attains its minimum at \((v, z^0) = 0\). Applying the PMP to this problem we deduce that the strong Legendre-Clebsch condition (4.8) holds.

\textbf{Theorem 4.14.} We assume satisfied the qualification condition (4.23) and the second order condition (4.63). Then the Jacobian of \( T_{IF} \) at \((y^0, p^0, \bar{w})\) is invertible.

\textbf{Proof.} An element \((z^0, q^0, \delta \Psi)\) of \( \text{Ker } DT_{IF}^\beta(y^0, p^0, \bar{w})\) is, by (4.24), characterized by

\[
(4.64) \quad \begin{cases}
q^0 + L_{IF}^\beta(y^0)(z^0) + L_{IF}^\beta(y_T)(\bar{y}(T))(z(0), z(T)) + \Phi_{y^0}(y^0, \bar{y}(T))\delta \Psi = 0, \\
-q^0 + L_{IF}^\beta(y^0)(z^0) + L_{IF}^\beta(y_T)(z(0), z(T)) + \Phi_{y^0}(y^0, \bar{y}(T))\delta \Psi = 0,
\end{cases}
\]

with \((v, z, q)\) solution of (4.18). We recognize the optimality conditions of the problem

\[
(4.65) \quad \min_{v, z^0} Q[\bar{\lambda}](v, z[v, z^0]); \quad D\Phi(y^0, \bar{y}(T))(z^0, z[v, z^0]) = 0.
\]

By the second-order necessary condition (theorem 4.13), this problem has a convex cost over the feasible domain (which is a vector space), and so \((v, z^0)\) minimizes \( Q[\bar{\lambda}]\) over the critical cone.

If (4.63) holds then \( Q[\bar{\lambda}]\) is strictly convex over the critical cone, therefore its optimality condition is satisfied only at zero, i.e., \((v, z^0) = 0\), and therefore \( z[v, z^0] = 0\). Then

\[
(4.66) \quad -q = \hat{f}^\dagger q.
\]
Now for \( a(t) \in L^\infty(0,T;\mathbb{R}^n) \), let \( z = z[v,z^0,a] \) be the solution of
(4.67) \[
\dot{z}(t) = D\dot{f}(v(t),z(t)) + a(t).
\]
We obtain that
(4.68) \[
D\Phi(z^0,z(T)) = [q \cdot z^0_T] = \int_0^T (q(t) \cdot z(t) + a(t)) \, dt.
\]
For any \( a(t) \in L^\infty(0,T;\mathbb{R}^n) \), by the qualification condition (4.23), there exists \( (v,z^0) \in \mathcal{U} \times \mathbb{R}^n \) such that \( z = z[v,z^0,a] \) satisfies \( D\Phi(z^0,z(T)) \). It follows that \( \int_0^T q(t) \cdot a(t) \, dt = 0 \) for any such \( a \), proving that \( q = 0 \) and therefore \( D\Phi \) is onto, so that \( D\Phi^\dagger \) is injective, implying that \( \Psi = 0 \) as well, whence the injectivity of the linearized shooting function, as was to be shown.

\[ \square \]

### 3. Control constraints

We briefly show how to express the shooting function in the case of a scalar control \( (m=1) \) subject to the bound constraint
(4.69) \[
u(t) \geq 0, \quad t \in (0,T),
\]
the optimal control being continuous, and the constraint being active on an interval of the form \((\tau,T)\) with \( \tau \in (0,T) \). The optimality system reads
(4.70) \[
\dot{y}(t) = f(u(t),y(t)) - p(t) = \nabla_y H(u(t),y(t),p(t)), \quad t \in (0,\tau),
\]
(4.71) \[
H_u(u,y,p) = 0, \quad t \in (0,\tau), \quad u(t) = 0, \quad t \in (\tau,T),
\]
with initial-final conditions
(4.72) \[
y(0) = y^0; \quad p(T) = \nabla \varphi(y(T)),
\]
and the junction condition
(4.73) \[
H_u(0,y(\tau),p(\tau)) = 0.
\]
As in the study of variable horizon problems in section 2.2, we may rewrite an equivalent system with functions of the normalized time \( s \in [0,1] \). Indeed, set \( u_N(s) := u(s \tau) \) and
(4.74) \[
y_N^1(s) := y(s \tau); \quad y_N^2(s) := y(\tau + s(T - \tau)),
\]
(4.75) \[
p_N^1(s) := p(s \tau); \quad p_N^2(s) := y(\tau + s(T - \tau)).
\]
We easily see that the optimality system (4.70)-(4.73) is equivalent to, for \( s \in [0,1] \):
(4.76) \[
\dot{y}_N^1(s) = \tau f(u_N(s),y_N^1(s)),
\]
(4.77) \[
\dot{y}_N^2(s) = (T - \tau) f(0,y_N^2(s)),
\]
(4.78) \[
- p_N^1(s) = \tau \nabla_y H(u_N(s),y_N^1(s),p_N^1(s)),
\]
(4.79) \[
- p_N^2(s) = (T - \tau) \nabla_y H(0,y_N^2(s),p_N^2(s)),
\]
(4.80) \[
H_u(u_N(s),y_N^1(s),p_N^1(s)) = 0,
\]
(4.81) \[
H_u(0,y_N^2(s),p_N^2(s)) = 0,
\]
with initial-final conditions (matching those in (4.72), with in addition what corresponds to the continuity of state and costate at time \( \tau \)):
(4.82) \[
y_N^1(0) = y^0; \quad p_N^2(1) = \nabla \varphi(y_N^2(1)),
\]
(4.83) \[
y_N^1(1) = y_N^2(0); \quad p_N^1(1) = p_N^2(0),
\]
69
and the junction condition
\[(4.84)\quad H_u(u_N(s), y_N^1(s), p_N^1(s)) = 0.\]

We integrate the 'normalized' system (4.76)-(4.81) with initial condition \((y^0, y^{02})\) for the state and \((p^{01}, p^{02})\) for the costate, the control being expressed as a function of the state and costate. The shooting equation reads
\[(4.85)\quad T(y^{02}, p^{01}, p^{02}, \tau) = \begin{pmatrix}
y_N^1(1) - y^{02} 
p_N^1(1) - p^{02} 
p_N^2(1) - \nabla \varphi(y_N^2(1)) 
H_u(0, y_N^1(1), p_N^1(1))
\end{pmatrix}.
\]

4. Notes

Classical references on shooting methods are [20, 34, 41]. For problems with singular arcs, see [31]; there are typically more shooting equations than unknowns, see [2]. For the extension to state constrained problems, see [32] and [10].
1. Pontryagin’s principle

Consider now problems with additional inequality constraints on the state at each time. We will need an adapted version of Pontryagin’s principle in order to express the optimality conditions of such problems.

1.1. Setting. We recall the following definitions: the cost function

\[ J^{IF}(u, y) := \int_0^T \ell(t, u(t), y(t))dt + \varphi(y(0), y(T)), \]

the state equation

\[ \dot{y}(t) = f(t, u(t), y(t)), \quad \text{for a.a. } t \in [0, T]; \quad y(0) = y^0, \]

the control constraint, where \( U_{ad} \) is a closed subset of \( \mathbb{R}^m \),

\[ u(t) \in U_{ad}, \quad \text{for a.a. } t \in [0, T], \]

and the two point constraints

\[ \Phi(y(0), y(T)) \in K_\Phi. \]
In this chapter we add state constraints of the form
\[(5.5) \quad g_j(t, y(t)) \leq 0, \quad j = 1, \ldots, n_g, \quad \text{for all } t \in [0, T].\]

We consider the following optimal control problem:
\[(5.6) \quad \min_{(u, y) \in U \times Y} J^{IF}(u, y) \quad \text{s.t. } (5.2)-(5.5),\]
where we recall that
\[(5.7) \quad U := L^{\infty}(0, T, \mathbb{R}^m); \quad Y := W^{1, \infty}(0, T, \mathbb{R}^n).\]

About the data we assume (here and for extensions to problems with optimization parameters) that
\[(5.8) \quad \text{We take the same hypotheses on data as in chapter 3 with in addition } \mu \text{ of class } C^1.\]

1.2. Duality in spaces of continuous functions. Pontryagin’s principle involves multipliers associated with constraints. In the case of the state constraint we must first choose the function space in which this constraint is expressed, and see what are the properties of the dual space, where the multipliers live. It appears that a convenient choice is
\[(5.9) \quad X := C([0, T])^{n_g}, \text{ the space of continuous functions over } [0, T] \text{ with value in } \mathbb{R}^{n_g} \text{ since it is a Banach space whose dual has a nice structure.}\]

1.2.1. Bounded variation functions. We recall that the total variation of a function \(\mu\) over \([0, T]\) is
\[(5.10) \quad \var(\mu) := \sup \left\{ \sum_{i=0}^{n_{\mu}} |\mu(t_{i+1}) - \mu(t_i)|; \quad \text{where } (t_i) \text{ is a subdivision of } [0, T] \right\},\]
(subdivisions were defined in (3.107)). We say that \(\mu\) has bounded variation if \(\var(\mu) < \infty\), and denote by \(BV(0, T)\) the set of functions with bounded variation. The Stieltjes integral of \(h \in C([0, T])\) associated with the bounded variation function \(\mu\) is
\[(5.11) \quad \int_0^T h(t) d\mu(t) := \lim \sum_{j=0}^{n_{\mu}} h(\tau_j)(\mu(t_{i+1}) - \mu(t_i)),\]
where \(\tau_j \in [t_i, t_{i+1}]\), the limit being taken over subdivisions of vanishing maximal increment. We endow the space \(C([0, T])\) of continuous functions over \([0, T]\) with the supremum norm
\[(5.12) \quad \|h\|_{\infty} := \max\{|h(t)|; \quad t \in [0, T]\}.\]

We obviously have that
\[(5.13) \quad \left| \int_0^T h(t) d\mu(t) \right| \leq \var(\mu)\|h\|_{\infty}.
\]

Therefore, any Stieltjes integral defines a linear continuous operator over \(C([0, T])\). The converse statement, due to Riesz, holds, see e.g. Malliavin [29]:

**Lemma 5.1.** Any continuous linear form on \(C([0, T])\) is of the type \(h \mapsto \int_0^T h(t) d\mu(t)\), where \(\mu\) is a bounded variation function.

In the sequel we may assume that \(\mu\) belongs to the normed space
\[(5.14) \quad BV_0(0, T)^{n_g} := \{\bar{\mu} \in BV(0, T)^{n_g}; \quad \bar{\mu}(T) = 0\},\]
endowed with the norm
\[(5.15) \quad \|\bar{\mu}\|_{BV_0} := \var(\mu).\]
1.2.2. Costate equation. We recall that, for optimal control problems in the above format, but without state constraints, the Lagrangian was defined in (3.32) as

\[ L(\beta, u, y, p, \Psi) := \beta J^F(u, y) + \int_0^T p(t) \cdot (f(t, u(t), y(t)) - \dot{y}(t)) \, dt + \Psi(t(0), y(T)) \]

\[ = \int_0^T (H(\beta, t, u(t), y(t), p(t)) - p(t) \cdot \dot{y}(t)) \, dt + L^F(\beta, y(0), y(T), \Psi). \]

Here \( \beta \in \{0, 1\} \), \( u \in \mathcal{U} \), \( y \in \mathcal{Y} \), \( p \) belongs to the costate space \( \mathcal{P} := BV(0, T)^n \) (we explain why below) \( \Psi \in \mathbb{R}^n \). The Lagrangian say \( L^g \) of the state constrained problem (5.6) has a similar expression, with the additional contribution of the state constraint:

\[ L^g(\beta, u, y, p, \Psi, \mu) := L(\beta, u, y, p, \Psi) + \Delta^g(y, \mu), \]

where \( \mu \in BV_0(0, T)^n \) and \( \Delta^g : \mathcal{Y} \times BV_0(0, T)^n \to \mathbb{R} \), is defined by

\[ \Delta^g(y, \mu) := \sum_{j=1}^n \int_0^T g_j(t, y(t)) \, d\mu_j(t). \]

Again, the costate equation is obtained by expressing the stationarity of the Lagrangian w.r.t. the state, i.e. the relation

\[ D_y L^g(\beta, u, y, p, \Psi, \mu)z = 0, \quad \text{for all } z \in \mathcal{Y}. \]

As for problem (3.29), we obtain the same terms as in (3.33), except for the above linearization of \( \Delta^g(y, \mu) \), and for the integration by parts in the term involving \( \dot{z} \).

**Lemma 5.2.** Let \( p \in \mathcal{P} \) and \( z \in \mathcal{Y} \). Then the following integration by parts formula is valid:

\[ \int_0^T p(t) \cdot \dot{z}(t) \, dt = p(T) \cdot z(T) - p(0) \cdot z(0) - \sum_{i=1}^n \int_0^T z_i(t) \, dp_i(t). \]

**Proof.** We may assume that \( n = 1 \). The sum similar to (5.10), with \( t_0 = 0, t_{N+1} = T, \tau_i := t_i, i = 0 \) to \( N \), reads

\[ \sum_{i=0}^N z(t_i)(p(t_{i+1}) - p(t_i)) = \sum_{i=1}^{N+1} z(t_{i-1})p(t_i) - \sum_{i=0}^N z(t_i)p(t_i) = - \sum_{i=1}^N p(t_i)(z(t_i) - z(t_{i-1})) + z(n)p(T) - z(0)p(0) \]

\[ = - \sum_{i=1}^N p(t_i) \int_{t_{i-1}}^{t_i} \dot{z}(t) \, dt + z(n)p(T) - z(0)p(0) \]

\[ = - \int_{t_0}^{t_N} p(\tau(t)) \dot{z}(t) \, dt + z(n)p(T) - z(0)p(0), \]

where \( \tau(t) := \sum_{i=1}^N t_i \mathbf{1}_{t_i \in (t_{i-1}, t_i)}(t) \). Denote by \( h_N := \max\{t_{i+1} - t_i; i = 0, \ldots, N\} \) the supremum of the stepsizes of the subdivision. When \( N \uparrow \infty \) (by this we understand that also \( h_N \to 0 \)), then \( \tau(t) \to t \), for all \( t \in [0, T] \). Having in mind that a bounded variation function is a.e. continuous, it follows that \( p(\tau(t)) \to p(t) \) a.e. By the dominated convergence theorem, the above integral converges to the l.h.s. of (5.19). So, since \( z \) is continuous and therefore \( z(n) \to z(T) \),

\[ \int_0^T z(t) \, dp(t) = \lim_{N} \sum_{i=0}^N z(t_i) (p(t_{i+1}) - p(t_i)) = - \int_0^T p(t) \cdot \dot{z}(t) \, dt + p(T) \cdot z(T) - p(0) \cdot z(0) \]

as was to be proved. \( \square \)
Let $\beta \in \{0,1\}$, $\bar{u} \in U$, $\bar{y} \in Y$, $\bar{p} \in P$, $\bar{\Psi} \in \mathbb{R}^{n_x}$, $\bar{\mu} \in BV_0(0,T)^n$; Thanks to (5.19), one can easily express the directional derivative of the Lagrangian w.r.t. the state in a direction $z \in Y$, and deduce that the costate equation is, the pre-Hamiltonian $H(\cdot)$ and end points Lagrangian $L^IF(\cdot)$ being defined in (3.30)-(3.31):

\begin{equation}
\begin{cases}
-\dot{p}(t) = \nabla_y H(\bar{\beta}, t, \bar{u}(t), \bar{y}(t), \bar{p}(t))dt + \sum_{j=1}^{n_g} \nabla_y g_j(t, \bar{y}(t))d\bar{\mu}_j(t), \quad t \in [0, T], \\
-\bar{p}(0) = \nabla_y L^IF(\bar{\beta}, \bar{y}(0), \bar{\Psi}), \\
\bar{p}(T) = \nabla_y L^IF(\bar{\beta}, \bar{y}(T), \bar{\Psi}).
\end{cases}
\end{equation}

5.21

\begin{itemize}
\item REMARK 5.3. Let $c(t) := \nabla_y H(\bar{\beta}, t, \bar{u}(t), \bar{y}(t), \bar{p}(t))$. The first equality in (5.21), being an equality between measures, means that, for any $\varphi \in C([0,T])$:

\begin{equation}
\sum_{i=1}^{n} \int_0^T \varphi_i(t)d\bar{p}_i(t) + \int_0^T \varphi(t) \cdot e(t)dt + \sum_{j=1}^{n_g} \int_0^T \varphi(t) \cdot \nabla_y g_j(t, \bar{y}(t))d\bar{\mu}_j(t) = 0.
\end{equation}

5.22

We call Pontryagin multiplier, associated with the nominal trajectory $(\bar{u}, \bar{y})$, any $\lambda = (\bar{\beta}, \bar{\Psi}, \bar{\mu}, \bar{p})$ satisfying (5.21), the relations

\begin{equation}
\bar{\Psi} \in N_{K^d}(\Phi(\bar{y}(0), \bar{y}(T))), \quad \bar{\beta} \in \{0,1\},
\end{equation}

5.23

the Hamiltonian inequality similar to (3.38), i.e.,

\begin{equation}
H(\bar{\beta}, \bar{u}(t), \bar{y}(t), \bar{p}(t)) = \inf_{u \in U_{ad}} H(\bar{\beta}, u, \bar{y}(t), \bar{p}(t)) \quad \text{for a.a. } t \in (0, T),
\end{equation}

5.24

and the complementarity relations

\begin{equation}
d\bar{\mu} \geq 0; \quad \int_0^T g_j(t, \bar{y}(t))d\bar{\mu}_j(t) = 0, \quad j = 1, \ldots, n_g, \quad \bar{\mu}(T) = 0,
\end{equation}

5.25

and of non triviality

\begin{equation}
\bar{\beta} + |\bar{\Psi}| + \|\bar{\mu}\|_{BV_0} > 0.
\end{equation}

5.26

\begin{itemize}
\item REMARK 5.4. By $d\bar{\mu} \geq 0$ we mean that $\int_0^T \varphi(t)d\bar{\mu} \geq 0$ whenever $\varphi \in X_g$ has nonnegative values. This is equivalent to say that the function $\bar{\mu}$ is nondecreasing.

\item THEOREM 5.5. Any $(\bar{u}, \bar{y})$ solution of (2.141), is a Pontryagin extremal.

\item PROOF. We postpone to section 3 the proof of this difficult result.

\end{itemize}

1.2.3. Link with Lagrange multipliers; convex problems. We have a natural extension of the analysis of section 1.2.3 in chapter 3 to the state constrained case. Assuming $U_{ad}$ to be convex, we say that $\lambda := (\bar{\beta}, \bar{\Psi}, \bar{p}, \bar{\mu})$ is a Lagrange multiplier if it satisfies the same relations as for Pontryagin multipliers, but with (3.46) instead of the Hamiltonian inequality. Then obviously $\Lambda_P(\bar{u}, \bar{y}) \subset \Lambda_L(\bar{u}, \bar{y})$.

\begin{itemize}
\item LEMMA 5.6. Let $\bar{\lambda} := (\bar{\beta}, \bar{\Psi}, \bar{p}, \bar{\mu})$ be a regular ($\beta = 1$) Lagrange multiplier, such that the associated Lagrangian defined in (3.32) is a convex function of $(\bar{u}, \bar{y})$. Then $(\bar{u}, \bar{y})$ is solution of problem (3.29).

\item PROOF. Adapt the arguments in the proof of lemma 3.20.

\end{itemize}

1.3. Variation of the pre-Hamiltonian.
1.3.1. Constant pre-Hamiltonian for autonomous systems. We say that the optimal control problem is autonomous if the dynamics and running state and control constraints do not depend on time. For autonomous state-constrained problems, we still can prove that the pre-Hamiltonian is constant. The proof is based on a time dilation approach.

**Lemma 5.7.** If \((\bar{u}, \bar{y})\) is solution of the state constrained optimal control problem, the data being autonomous, then it has an associated Pontryagin multiplier such that the pre-Hamiltonian is constant.

**Proof.** We may assume that there is no integral cost. Consider the change of time \(t = \int_0^t \mathbf{w}(s) ds\), with \(\mathbf{w}(s) \in \left[ \frac{1}{2}, 2 \right]\) a.e. and \(\tau \in [0, T]\), for \(t \in (0, T)\). We have that \(t = \theta(\tau)\), with \(\dot{\theta}(\tau) = \mathbf{w}(\tau)\). Set
\[
\tilde{u}(\tau) := u(\theta(\tau)); \quad \tilde{y}(\tau) := y(\theta(\tau)).
\]

We easily check that for \(\tau \in [0, T]\):
\[
\begin{align*}
\frac{d}{d\tau} \tilde{y}(\tau) &= \mathbf{w}(\tau)f(\tilde{u}(\tau), \tilde{y}(\tau)), \\
\frac{d}{d\tau} \theta(\tau) &= \mathbf{w}(\tau).
\end{align*}
\]

So, we may consider an optimal control problem with state \((\tilde{y}, \theta)\), control variable \((\tilde{u}, \mathbf{w})\), whose expression is
\[
\begin{align*}
\min & \varphi(\tilde{y}(0), \tilde{y}(T)); \quad \text{subject to} \ (5.28), \ \Phi(\tilde{y}(0), \tilde{y}(T)) \in K_\delta \text{ and} \\
\tilde{u}(\tau) & \in U_{ad} \text{ a.e.; } g_j(\tau, \tilde{y}(\tau)) \leq 0, \ \tau \in [0, T]; \ \theta(T) = T.
\end{align*}
\]

By (5.27) any feasible trajectory of (5.29) is associated with a feasible trajectory of the original problem with same cost (and conversely). It follows that a solution of (5.29) is obtained when \(\mathbf{w}(\tau) = 1\) for a.a. \(\tau\), so that \(\theta(\tau) = \tau\), and \((\tilde{u}, \tilde{y}) = (\bar{u}, \bar{y})\). The costate of (5.29) can be denoted as \((\bar{p}, \bar{q})\), where \(\bar{p}, \bar{q}\) are associated resp. with \(\tilde{y}\) and \(\theta\); having zero derivative, \(\bar{q}\) is constant (equal to the multiplier associated with the terminal constraint \(\theta(T) = T\)). The pre-Hamiltonian and its value for the considered trajectory are resp.
\[
w(p \cdot f(u, y) + q); \quad \mathbf{w}(t)(\bar{H}(t) + \bar{q}(t)).
\]

By the Hamiltonian inequality, for a.a. \(t \in (0, T)\), its minimum over \([1/2, 2]\) is attained at \(\mathbf{w}(t) = 1\), which implies that \(\bar{H}(t) = -\bar{q}(t)\) is constant. In addition, since \(\mathbf{w}(t) = 1\) and \(\tau = t\) a.e., the costates \(\bar{p}\) and \(\bar{\mu}\) coincide through the considered change of time. In this way, any Pontryagin multiplier for (5.29) corresponds in an obvious way to a Pontryagin multiplier for the original one, with constant Hamiltonian. The conclusion follows. \(\square\)

1.3.2. Non autonomous systems. We can rewrite the optimal control problem (5.6) as an autonomous one, viewing the time as an additional state variable \(\tau\), with dynamics \(\dot{\tau} = 1\) and initial value \(\tau(0) = 0\).

The corresponding pre-Hamiltonian is \(H^a := H + q\), and the additional costate \(\bar{q}\) satisfies
\[
- d\bar{q}(t) = \bar{H}(t) dt + \sum_{j=1}^{n_0} D_j g_j(t, \bar{y}(t)) d\bar{\mu}_j(t), \quad \text{for } t \in [0, T]; \quad \bar{q}(T) = 0.
\]

Set
\[
\bar{h}(t) := \inf_{u \in U_{ad}} H(\beta, t, u, \bar{y}(t), \bar{p}(t)).
\]

Since \(\bar{H}(t) = \bar{H}(t) + \bar{q}(t)\) is constant by lemma 5.7, this means that \(\bar{h}(t)\) (equal to \(\bar{H}(t)\) a.e.) has bounded variation and satisfies \(d\bar{h}(t) = -d\bar{q}(t)\). We obtain:
Lemma 5.8. We have that
\begin{equation}
\frac{d\tilde{h}(t)}{dt} = -\tilde{q}(t) = \tilde{H}_t(t)dt + \sum_{j=1}^{n_q} D_tg_j(t, \tilde{y}(t))d\tilde{\mu}_j(t), \text{ for } t \in [0, T].
\end{equation}

In particular:
\begin{equation}
\frac{d}{dt} \tilde{h}(t) = \tilde{H}_t(t), \text{ for a.a. } t \in (0, T), \text{ if the state constraints are autonomous.}
\end{equation}

1.4. Decision variables, variable horizon.

1.4.1. Decision variables. As in chapter 3 we can extend Pontryagin’s principle to the case when decision variables (not depending on time) are involved, by viewing them as additional state variables, with zero dynamics.

Lemma 5.9. Consider the state constrained optimal control problem similar to (5.6), but with decision variables \( \pi \), entering as additional arguments in functions \( f, \ell, \Phi, g \). Assume that the data satisfy (3.64) and, in addition:
\begin{equation}
\begin{cases}
g \text{ is a } C^1 \text{ function, the mappings } g, g_y, g_\pi \text{ being} \\
\text{uniformly locally Lipschitz w.r.t. } (y, \pi).
\end{cases}
\end{equation}

Then the PMP holds with the same expression as for the corresponding problem with decision variables set at their nominal value, with in addition the condition of stationarity of the Lagrangian w.r.t. the decision variables, i.e. (compare to (3.71)):
\begin{equation}
\begin{cases}
\int_0^T \nabla_y H(\beta, t, \bar{u}(t), \bar{y}(t), \bar{p}(t))dt + \nabla_\pi L^{IF}(\beta, \bar{y}(0), \bar{y}(T), \bar{\pi}, \bar{\Psi}) \\
\quad + \int_0^T \nabla_\pi g(t, \bar{y}(t), \bar{\pi}))d\bar{\mu}(t) = 0.
\end{cases}
\end{equation}

Proof. The proof is similar to the one of the corresponding result for problems without state constraints, i.e., lemma 2.26.

1.4.2. Variable horizon. We can then deal with variable horizon problems, again as in chapter 3, by formulating a ‘normalized time’ problem with a normalized time \( \tau \in [0, 1] \), where the horizon \( T \) enters in the dynamics and integral cost. We obtain the following extension of lemma 3.28:

Lemma 5.10. If the initial time is fixed and the final time is free, the final cost not depending on the final time, the PMP has the same expression as in the case of a fixed final time, with in addition (assuming that the final time does not enter in the constraints) the condition
\begin{equation}
\tilde{h}(T) = 0.
\end{equation}

Proof. Let \( \mathcal{L}^N \) be defined as in (3.87). The Lagrangian function of the normalized time problem is
\begin{equation}
\tilde{\mathcal{L}}^N := \mathcal{L}^N + \sum_{j=1}^{n_q} \int_0^1 g_j(T\tau, y^N(\tau))d\mu_j^N(\tau).
\end{equation}

As in (3.88), set \( \tilde{H}(\tau) := H(\beta, T\tau, u^N(\tau), y^N(\tau), p^N(\tau)) \), which is a.e. equal to
\begin{equation}
\tilde{h}^N(\tau) := \inf_{u \in U_{ad}} H(\beta, T\tau, u, y^N(\tau), p^N(\tau)).
\end{equation}

The optimality condition is, using (5.33):
\begin{equation}
0 = \frac{\partial \tilde{\mathcal{L}}^N}{\partial T} = \int_0^1 \left( \tilde{H}(\tau) + T\tau \tilde{H}_t(\tau) \right) d\tau + T \int_0^1 \sum_{j=1}^{n_q} \tau D_tg_j(T\tau, \bar{y}(\tau))d\tilde{\mu}_j^N(\tau)
\end{equation}
\begin{equation}
= \int_0^1 (\tilde{h}^N(\tau)d\tau + \tau \tilde{h}^N(\tau)) = \int_0^1 d(\tau \tilde{h}^N(\tau)) = \tilde{h}^N(1).
\end{equation}
The conclusion follows. \hfill \square

\textbf{Remark 5.11.} The costate equation for the normalized time problem is
\begin{equation}
-d\bar{p}^N(\tau) = T\nabla_y \hat{H}(\tau)d\tau + \sum_{j=1}^{n_p} \nabla_y g_j(T\tau, \bar{y}(\tau))d\bar{\mu}^N_j(\tau).
\end{equation}

Since $\bar{p}^N(\tau) = \bar{p}(T\tau)$, $d\bar{p}^N(\tau) = Td\bar{p}(T\tau)$ and therefore $d\bar{\mu}^N(\tau) = Td\bar{\mu}(T\tau)$. Since both sides vanish when $\tau = 1$ if follows that
\begin{equation}
\bar{\mu}^N(\tau) = \bar{\mu}(T\tau), \quad \tau \in [0, 1].
\end{equation}

\textbf{Remark 5.12.} (i) Although $\bar{h}(t)$ may be discontinuous, its value at time $T$ makes sense since this is a bounded variation function.
(ii) In the more general case of varying initial and final time and of initial-final cost and constraints depending on the initial and final time, we have a natural extension of lemma 3.32 for which (3.104) still holds.

\section{2. Junction conditions}

In this section we assume that the functions involved in the expression of dynamics and constraints are of class $C^\infty$.

\subsection{2.1. Representation of control constraints.}

Consider the case when
\begin{equation}
U_{ad} = \{u \in \mathbb{R}^m; \ c_i(u) \leq 0, \ i = 1, \ldots, n_{cu}\},
\end{equation}
where the functions $c_i(u)$ are $C^2$ with Lipschitz second derivatives. Set
\begin{equation}
h(t, u) := H(\beta, t, u, \bar{y}(t), \bar{p}(t)).
\end{equation}
By the Hamiltonian inequality, for a.a. $t$, $\bar{u}(t)$ is solution of the problem
\begin{equation}
\min_{u \in \mathbb{R}^m} h(t, u); \ c_i(u) \leq 0, \ i = 1, \ldots, n_{cu}.
\end{equation}
Denote the set of active control constraints at time $t$ by
\begin{equation}
I_0(t) := \{1 \leq i \leq n_{cu}; \ c_i(\bar{u}(t)) = 0\}.
\end{equation}
We assume in the sequel that the following \textit{qualification condition for the control constraints} holds: \{$\nabla_u c_i((\bar{u}(t)), \ i \in I_0(t))$ is uniformly linearly independent, i.e.,
\begin{equation}
\text{There exists } \alpha > 0 \text{ such that for all } t: \ \alpha |\lambda| \leq \left| \sum_{i \in I_0(t)} \lambda_i \nabla_u c_i(\bar{u}(t)) \right| \quad \text{if } \lambda_i = 0, \ i \notin I_0(t).
\end{equation}
From the Lagrange multiplier theory (see \cite[Section 5.2]{15}) we deduce that there exists a unique Lagrange multiplier $\bar{\lambda}(t) \in \mathbb{R}^{n_{cu}}$ such that
\begin{equation}
\nabla_u h(t, \bar{u}(t)) + \sum_{i \in I_0(t)} \bar{\lambda}_i(t) \nabla_u c_i(\bar{u}(t)) = 0; \quad \bar{\lambda}(t) \geq 0; \quad \bar{\lambda}(t) \cdot c(\bar{u}(t)) = 0.
\end{equation}
It can easily be checked that $\bar{\lambda}$ is measurable and belongs to $L^\infty(0, T; \mathbb{R}^{n_{cu}})$. Defining the \textit{augmented Hamiltonian} as
\begin{equation}
H^a(\beta, t, u, y, p, \lambda) := H(\beta, t, u, y, p) + \lambda \cdot c(u)
\end{equation}
we observe that (5.48) can be rewritten in the form
\begin{equation}
\begin{cases}
\nabla_u H^a(\beta, t, \bar{u}(t), \bar{y}(t), \bar{p}(t), \bar{\lambda}(t)) = 0; \\
\bar{\lambda}(t) \cdot c(\bar{u}(t)) = 0, \quad \text{for a.a. } t.
\end{cases}
\end{equation}

\subsection{2.2. Constraint order and junction conditions.}
2.2.1. Order of a state constraint. Let the total derivative of \( g_j \) (along the dynamics \( f \)) be defined by
\[
(5.51) \quad g_j^{(1)}(t, u, y) := D_yg_j(t, y)f(t, u, y) + D_tg_j(t, y).
\]
Note that
\[
(5.52) \quad D_u g_j^{(1)}(t, u, y) = D_yg_j(t, y)f_u(t, u, y).
\]
Along the trajectory \((u, y)\), \( g_j^{(1)}(t, u(t), y(t)) \) is the time derivative of \( g_j(t, y(t)) \). If \( g_j^{(1)}(u, y) \) depends on \( u \), i.e., if \( D_u g_j^{(1)}(t, u, y) \) does not identically vanishes, we say that \( g_j \) is a first order state constraint. Otherwise, writing \( g_j^{(1)} \) as function of \( t \) and \( y \) only, we may compute the total derivative of the latter, i.e., skipping some arguments:
\[
(5.53) \quad g_j^{(2)}(t, u, y) := D_yg_j^{(1)}(t, y)f(t, u, y) + D_tg_j^{(1)}(t, y) = D^2_yg_j(t, y)(f, f) + D_yg_j(t, y)f_yf + D^2 yg_j(t, y)f + D_tg_j^{(1)}(t, y).
\]
If \( g_j^{(2)}(t, u, y) \) depend on \( u \), we say that \( g_j \) is a second order state constraint, etc. We can interpret the control constraints as zero order state constraints.

In the sequel we denote by \( g_j \) the order, assumed to be well-defined, of \( g_j \), by \( I_r \subset \{1, \ldots, n_g\} \), the set of constraints of order \( r \), and \( I_r(t) \) those active at time \( t \).

2.2.2. Costate jumps. We denote the jumps of function of time having left and right limits at \( \tau \in [0, T] \) (this is the case for functions with bounded variation) by a bracket, e.g.
\[
(5.54) \quad [\bar{p}(\tau)] := \bar{p}(\tau^+) - \bar{p}(\tau^-),
\]
where we set \( \bar{p}(0-) := \bar{p}(0) \) and \( \bar{p}(T^+) := \bar{p}(T) \). We also set \( \bar{\nu}_j(\tau) := [\bar{\mu}_j(\tau)] \). In view of \((5.21))\), we have, for all \( \tau \in [0, T] \):
\[
(5.55) \quad -[\bar{p}(\tau)] = \sum_{j=1}^{n_g} \bar{\nu}_j(\tau)\nabla_y g_j(\tau, \bar{y}(\tau)).
\]

2.2.3. Continuity of the multipliers.

**Lemma 5.13.** Let \( \bar{\ell}_u \) and \( \bar{f}_u \) be continuous at time \( \tau \in (0, T) \) (this holds in particular if \( \bar{u} \) is continuous at time \( \tau \)). Then
\[
(5.56) \quad [\nabla_u \bar{H}(\tau)] = -\sum_{j \in I_1(\tau)} \bar{\nu}_j(\tau)\nabla_y g_j^{(1)}(\tau).
\]

**Proof.** In view of \((5.55))\), we have that
\[
(5.57) \quad [\nabla_u \bar{H}(\tau)] = f_u(\tau)^T[p(\tau)] = -\sum_{j=1}^{n_g} \nu_j(\tau)f_u(\tau)^T
abla_y g_j(\tau, \bar{y}(\tau)).
\]
Note that
\[
(5.58) \quad f_u(\tau)^T\nabla_y g_j(\tau, \bar{y}(\tau)) = (D_y g_j(\tau, \bar{y}(\tau))\bar{f}_u(\tau))^T = (D_u g_j^{(1)}(\tau))^T = \nabla_u g_j^{(1)}(\tau).
\]
if \( j \notin I_1 \), by the definition of the order of a state constraint, this is equal to zero. On the other hand, if \( g_j(\tau, \bar{y}(\tau)) < 0 \) then \( \bar{\nu}_j(\tau) = 0 \). So, in \((5.57))\) we may reduce the sum over \( j \) to \( I_1(\tau) \) and the conclusion follows using again \((5.58))\). \( \square \)

**Proposition 5.14.** (i) Assume that
\[
(5.59) \quad \bar{\ell}_u, \bar{f}_u \text{ and } \nabla c_j(\bar{u}(\tau)), \ j \in I_0(\tau) \text{ are continuous at time } \tau \in (0, T).
\]
Then
\[
(5.60) \quad \sum_{j \in I_0(\tau)} [\bar{\lambda}_j(\tau)]\nabla c_j(\bar{u}(\tau)) - \sum_{j \in I_1(\tau)} \bar{\nu}_j(\tau)\nabla_u g_j^{(1)}(\tau) = 0.
\]
(ii) Assume in addition that

\begin{equation}
\{\nabla c_j(\bar{u}(\tau)), \ j \in I_0(\tau)\} \cup \{\nabla u \bar{g}_j^{(1)}(t, \bar{u}(\tau), \bar{y}(\tau)), \ j \in I_1(\tau)\} \text{ is linearly independent.}
\end{equation}

Then \(\bar{\lambda}\) and \(\{\bar{\mu}_j, j \in I_1\}\) are continuous at time \(\tau\).

**Proof.** (i) If \(c_j(\bar{u}(\tau)) < 0\) for some \(1 \leq j \leq n_{cu}\), then \(\bar{\lambda}_j(\tau^\pm) = 0 = [\bar{\lambda}_j(\tau)]\). Combining with (5.50) we get that

\begin{equation}
0 = [\nabla u \bar{H}^n(\tau)] = [\nabla u \bar{H}(\tau)] + \sum_{j \in I_0(\tau)} [\bar{\lambda}_j(\tau)] \nabla c_j(\bar{u}(\tau)).
\end{equation}

We conclude using Lemma 5.13.

(ii) Immediate consequence of point (i), having in mind that if a state constraint is not active, the corresponding component is constant near time \(\tau\).

**Remark 5.15.** An obvious sufficient condition for (5.59) is that the control is continuous at time \(\tau\). Another case is when \(f, \ell, \) and \(c\) are affine functions of the control. This is also the case when \(u = (u', u'') \in \mathbb{R}^m \times \mathbb{R}^{m''}\) with \(u'\) continuous at time \(\tau\) and for some functions \(t^0, \ell', f^0, f^1, c^0\), and constant \(c \in \mathbb{R}^{m''}:

\begin{align}
\ell(t, u, y) &= \ell^0(t, u', y) + \sum_{i=1}^{m'} u''_i \ell^i(t, y), \\
f(t, u, y) &= f^0(t, u', y) + \sum_{i=1}^{m''} u''_i f^i(t, y), \\
c(u) &= c^0(u') + \sum_{i=1}^{m''} u''_i c^i.
\end{align}

2.2.4. Continuity of the control. Let \((u, y)\) be a Pontryagin extremal, \((\beta, p, \mu)\) be an associated Pontryagin multiplier, and let \(\tau \in (0, T)\). For \(\tau \in (0, T)\), set

\begin{equation}
p^- := \bar{p}(\tau^-); \quad p^+ := \bar{p}(\tau^+); \quad p^\sigma := (1 - \sigma)p^- + \sigma p^+, \quad \sigma \in [0, 1].
\end{equation}

We may define in the same way \(u^\pm, u^\sigma, \bar{\lambda}^\pm, \bar{\lambda}^\sigma\), assuming that \(\bar{u}\) and \(\bar{\lambda}\) have right and left limits at time \(\tau\).

**Definition 5.16.** We say that \(H^a(\bar{\beta}, t, \cdot, \bar{y}(\tau), p^\sigma, \lambda^\sigma)\) is uniformly strongly convex in average between \(\tau^-\) and \(\tau^+\), in the direction \([\bar{u}(\tau)]\), if there exists \(\alpha > 0\) such that

\begin{equation}
\alpha ||\bar{u}(\tau)||^2 \leq \int_0^1 D_{uu} H^a(\bar{\beta}, t, u^\tau, \bar{y}(\tau), p^\sigma, \lambda^\sigma) d\sigma \quad ([\bar{u}(\tau)], [\bar{u}(\tau)]), \quad \text{for all } \sigma \in [0, 1].
\end{equation}

**Lemma 5.17.** Let \(H^a(\bar{\beta}, \tau, \cdot, \bar{y}(\tau), p^\sigma, \lambda^\sigma)\) be uniformly strongly convex in average between \(\tau^-\) and \(\tau^+\), in the direction \([\bar{u}(\tau)]\). Then \(\bar{u}\) is continuous at time \(\tau\).

**Proof.** Use \(F(1) - F(0) = \int_0^1 F'(\sigma) d\sigma\), with

\begin{equation}
F(\sigma) := \nabla_{uu} H^a(\bar{\beta}, \tau, u^\sigma, \bar{y}(\tau), p^\sigma, \lambda^\sigma).
\end{equation}

We get using (5.50):

\begin{equation}
0 = [\nabla u H^a(\tau)]
= \int_0^1 \left( H_{uuu}^a(\beta, \tau, u, \bar{y}(\tau), p, \lambda)[\bar{u}(\tau)] + c_u(u^\sigma)[\lambda(\tau)] + f_u(t, u^\sigma, \bar{y}(\tau)) [\bar{p}(\tau)] \right) d\sigma.
\end{equation}

With (5.65) and (5.55), and observing that \(g_{i,y} f_u = g_{j,u}^{(1)}(\tau) = 0\) if \(q_i > 1\), we obtain that

\begin{equation}
\alpha ||\bar{u}(\tau)||^2 \leq \int_0^1 H_{uu}^a(\beta, \tau, u^\sigma, \bar{y}(\tau), p^\sigma, \lambda^\sigma) d\sigma \quad ([\bar{u}(\tau)], [\bar{u}(\tau)])
\end{equation}

\begin{equation}
= \int_0^1 \left( - \sum_{i \in I_0(\tau)} [\lambda_i(\tau)] \nabla c_i(u^\sigma) + \sum_{i \in I_1(\tau)} \bar{\mu}_i \int_0^1 \nabla u g^{(1)}_i(\tau, u^\sigma, \bar{y}(\tau)) \right) [\bar{u}(\tau)] d\sigma.
\end{equation}
Next observe that
\begin{equation}
(5.69) \quad \int_0^1 \nabla c_i(u^\sigma)[\bar{u}(\tau)]d\sigma = [\bar{c}_i(\tau)]; \quad \int_0^1 \nabla_a \bar{g}_i^{(1)}(\tau, u^\sigma, \bar{y}(\tau))[\bar{u}]d\sigma = [\bar{g}_i^{(1)}(\tau)].
\end{equation}
So, we have proved that
\begin{equation}
(5.70) \quad \alpha ||\bar{u}(\tau)||^2 \leq -\sum_{i \in l_0} [\tilde{\lambda}_i(\tau)] [\bar{c}_i(\tau)] + \sum_{i \in l_1(\tau)} \bar{\nu}_i [\bar{g}_i^{(1)}(\tau)].
\end{equation}
If \( i \in I_0 \), by the complementarity conditions we have that
\begin{equation}
(5.71) \quad [\tilde{\lambda}_i(\tau)] [\bar{c}_i(\tau)] = -\tilde{\lambda}_i(\tau^+) [\bar{c}_i(\tau_-)] - \tilde{\lambda}_i(\tau^-) [\bar{c}_i(\tau_+)] \geq 0.
\end{equation}
Also, for \( i \in I_1(\tau) \), \( t \mapsto g_i(t, \bar{y}(t)) \) attains its maximum at time \( \tau \), so that \( \bar{g}_i^{(1)}(\tau) \leq 0 \). Since \( \bar{\nu}_i \geq 0 \), we have proved that the r.h.s. of (5.70) is nonpositive, implying \( [\bar{u}(\tau)] = 0 \). The conclusion follows.

REMARK 5.18. Sometimes on can deduce from the Hamiltonian inequality that \( \bar{u} \) is a Lipschitz function of \((t, \bar{y}(t), \bar{p}(t))\), and then it has left and right limits at time \( \tau \). Then so has \( \tilde{\lambda} \) if in addition the qualification condition (5.47) holds, in view of (5.50).

2.2.5. Jumps of the derivative of the control. We next show the link between a jump of the derivative of the control and those of \( \mu \) or of its derivative. Below \( \tau \in (0, T) \) is such that for some small \( \varepsilon > 0, \bar{y}, \bar{u}, \bar{p}, \bar{\lambda} \) and \( \bar{\mu} \) are continuously differentiable over \((\tau - \varepsilon, \tau) \cup (\tau, \tau + \varepsilon)\).

**Lemma 5.19.** Let the control be continuous at time \( \tau \in (0, T) \), and the linear independence hypothesis (5.61) hold. Then
(i) \( \bar{H}_{uu}(t) \) and \( \bar{H}_{uu}^a(t) \) are continuous, and
(ii) if the jumps below are well defined, then, for all \( \tau \in ]0, T[ \):
\begin{equation}
(5.72) \quad \bar{H}_{uu}^a(\tau)[\bar{u}(\tau)] = -\sum_{j \in I_0(\tau)} [\tilde{\lambda}_i(\tau)] [\nabla c_i(\bar{u}(\tau))] + \sum_{j \in I_1(\tau)} [\tilde{\mu}_j(\tau)] [\nabla_a \bar{g}_j^{(1)}(\tau)] - \sum_{j \in I_2(\tau)} \bar{\nu}_j(\tau) [\nabla_a \bar{g}_j^{(2)}(\tau)].
\end{equation}

**Proof.** (i) Note first that, for \( 1 \leq i, k \leq m \):
\begin{equation}
-\bar{H}_{uiuk}(\tau) = -[\bar{p}(\tau)] \cdot \bar{f}_{uiuk}(\tau) = \sum_{j=1}^{n_a} \bar{\nu}_j(\tau) \bar{g}_j^{(2)}(\tau) \bar{f}_{uiuk}(\tau) = \sum_{j=1}^{n_a} \bar{\nu}_j(\tau) D^2_{uiuk} \bar{g}_j^{(1)}(\tau).
\end{equation}
If \( j \in I_1(\tau) \), then \( \bar{\nu}_j(\tau) = 0 \) by proposition 5.14(ii). Otherwise, \( g_j^{(1)}(t, u, y) \) does not depend on \( u \) and so \( D^2_{uiuk} \bar{g}_j^{(1)}(\tau) = 0 \). Therefore, \( \bar{H}_{uu}(t) \) is continuous. Since \( \tilde{\lambda} \) is continuous at time \( \tau \) by proposition 5.14(ii), so is \( \bar{H}_{uu}^a(t) \).

(ii) It is enough to give the proof for an autonomous problem, since for a non autonomous one, we may add the time as a state variable with derivative 1, which leaves invariant the expression of \( \bar{H}_u \). Let \( i \in \{1, \ldots, m\} \). Skipping time arguments, we get for any time close to \( \tau \) but different from it:
\begin{equation}
(5.73) \quad \frac{d}{dt} \bar{H}_{ui} = \frac{d}{dt} (\bar{f}_{ui} + \bar{p} \cdot \bar{f}_{ui}) = \bar{e}_{uiy} \bar{f} + \bar{H}_{uiu} \bar{u} + \bar{p}^T \bar{f}_{uiy} \bar{f} + \bar{p} \cdot \bar{f}_{ui}.
\end{equation}
Set \( \eta := \bar{e}_{uiy} \bar{f} - \nabla_v \bar{e} \cdot \bar{f}_{ui} \). Since \( \bar{g}_j^{(1)} \bar{f}_{ui} = 0 \) if \( j \notin I_1(\tau) \), using the costate equation, we get
\begin{equation}
(5.74) \quad \frac{d}{dt} \bar{H}_{ui} = \eta(t) + \bar{H}_{uiu} \bar{u} + \bar{p} \cdot (\bar{f}_{uiy} \bar{f} - \bar{f}_y \bar{f}_{ui}) - \sum_{j \in I_1} \bar{\mu}_j \bar{g}_j^{(1)} \bar{f}_{ui}.
\end{equation}
On the other hand
\begin{equation}
(5.75) \quad \frac{d}{dt} (\tilde{\lambda} \cdot c_i(\bar{u})) = \tilde{\lambda} \cdot c_i(\bar{u}) + \sum_{j \in I_0(t)} \tilde{\lambda}_j \frac{\partial^2 c_j}{\partial u_i \partial u} \bar{u}.
\end{equation}
Adding both relations and using (5.50), we obtain that
\begin{equation}
0 = \frac{d}{dt} H_{ui} = \eta(t) + H_{ui}^a \dot{u} + \ddot{\lambda} = c_u(u) + \dot{p} \cdot (f_{uy} \bar{f} - \bar{f}_y f_{ui}) - \sum_{j \in I_1} \dot{\mu}_j \bar{g}_j \bar{f}_{ui}.
\end{equation}
Consequently, using (i) and the continuity of \( \bar{\eta} \):
\begin{equation}
H_{ui}^a [\dot{u}] = -[\ddot{\lambda}] = c_u(u) + \sum_{j \in I_1} [\dot{\mu}_j] \bar{g}_j \bar{f}_{ui} - [\dot{p}] \cdot (f_{uy} \bar{f} - \bar{f}_y f_{ui}).
\end{equation}
If \( j \notin I_1(\tau) \), we have that \( \bar{g}_j \bar{f}_{ui} = 0 \), and therefore
\begin{equation}
0 = (D_y (\bar{g}_j \bar{f}_{ui})) \bar{f} = \bar{g}_j'' (\bar{f}, \bar{f}_u) + \bar{g}_j' \bar{f}_{uy} \bar{f} + \bar{g}_j' \bar{f}_y f_{ui},
\end{equation}
and so, with (5.53):
\begin{equation}
\bar{g}_j'^{(2)} (\bar{f}_{ui}) = D_{ui} (\bar{g}_j'' (\bar{f}, \bar{f})),
\end{equation}
\begin{equation}
= 2 \bar{g}_j'' (\bar{f}, \bar{f}_u) + \bar{g}_j' \bar{f}_{uy} \bar{f} + \bar{g}_j' \bar{f}_y f_{ui},
\end{equation}
Therefore,
\begin{equation}
- [\dot{p}] \cdot (f_{uy} \bar{f} - \bar{f}_y f_{ui}) = \sum_{j=1}^{n_u} \bar{p}_j \bar{g}_j'' (f_{uy} \bar{f} - \bar{f}_y f_{ui}) = - \sum_{j=1}^{n_u} \bar{p}_j \bar{g}_j'^{(2)} (\bar{f}_{ui}).
\end{equation}
We conclude with (5.77), noting that \( \bar{p}_j = 0 \) when \( q_j = 1 \), and \( \bar{g}_j'^{(2)} (\bar{f}_{ui}) = 0 \) if \( q_j > 2 \).

\section*{2.3. Examples}

\begin{example}
Let the control be continuous at time \( \tau \), (5.61) hold, \( H_{uu} (\tau) \) be invertible, no control constraints, the state constraint being scalar and of first order. Then
\begin{equation}
\begin{cases}
H_{uu} (\tau) [\dot{\hat{u}} (\tau)] = [\dot{\hat{\mu}} (\tau)] \nabla_u g^{(1)} (\tau); \\
\text{If } \bar{g}_u^{(1)} (\tau) \neq 0, \text{ then } \hat{\hat{u}} (\tau) \text{ is discontinuous iff } \bar{\hat{\mu}} \text{ is.}
\end{cases}
\end{equation}
With the same hypotheses, but a second (instead of first) order state constraint, we have that
\begin{equation}
\begin{cases}
H_{uu} (\tau) [\dot{\hat{u}} (\tau)] = - \bar{\hat{\nu}} (\tau) \nabla_u g^{(2)} (\tau); \\
\text{If } \bar{g}_u^{(2)} (\tau) \neq 0, \text{ then } \hat{\hat{u}} (\tau) \text{ is discontinuous iff } \bar{\hat{\nu}} \text{ is.}
\end{cases}
\end{equation}
With the same hypotheses, but a state constraint of order greater than two, \( \hat{\hat{u}} (\tau) \) is continuous. This illustrates the fact that the behavior of the solution depends on an essential way on the order of the state constraint.
\end{example}

\begin{example}
For \( T = 2 \), minimize \( \int_0^T y(t) dt + y(T) \) s.t. \( \dot{y}(t) = u(t) \) a.e. over \( ( -1, 1 ) \), initial condition \( y(0) = 1 \), and state constraint \( y(t) \geq 0 \). The solution is obviously \( u(t) = -1 \) and \( y(t) = 1 - t \) for \( t \in ( 0, 1 ) \), \( u(t) = y(t) = 0 \) over \( ( 1, 2 ) \). The pre-Hamiltonian being \( H = \beta y + pu \), the costate equation is
\begin{equation}
- \bar{d} \bar{p} = \beta dt - d\bar{\mu}; \quad \bar{p}(T) = \beta.
\end{equation}
Over the open constrained arc \( C = ( 1, 2 ) \), since \( \bar{u} = 0 \), by the Hamiltonian inequality, \( \bar{p} = 0 \) so that \( \bar{\mu} = \beta \). In addition \( \bar{p} \) and therefore \( \bar{\mu} \) has a final jump of value \( \beta \).
\end{example}

We claim that \( \beta = 0 \) is impossible. Indeed, we would have \( \bar{p} = \bar{\mu} \) in view of (5.83), so that \( \bar{p}(1_+) = \bar{\mu}(1_+) = 0 \). Since \( \bar{\mu} \) cannot be zero and is constant over \( [0, 1) \) it must have a positive jump at \( t = 1 \), so that it is negative as well a \( \bar{p} \) over \( [0, 1) \). But then the control cannot have negative values over \( (0, 1) \) in view of the Hamiltonian inequality.

So, necessarily \( \beta = 1 \), \( \bar{p} \) and \( \bar{\mu} \) have a jump of 1 at the final time, and over \( (1, 2) \), \( \bar{\mu}(t) = t-3 \).

A nonzero jump at time 1 would imply negative values of \( \bar{p}(t) \) for \( t < 1 \) close to 1, contradicting the Hamiltonian inequality. So, over \( (0, 1) \), \( \bar{p}(t) = -1 \) so that \( \bar{\mu}(t) = 1 - t \) is positive, and the Hamiltonian inequality implies as expected that \( \bar{u}(t) = -1 \) a.e.
Example 5.22. Same state equation, cost function and control constraint, but now $T = 3$ and the state constraint is
\begin{equation}
\dot{x}(t) \geq -\frac{1}{2}(t - 1)^2.
\end{equation}
The optimal control $\bar{u}(t)$ has value -1 over $(0, 1)$, 1 - $t$ over $(1, 2)$, and -1 over $(2, 3)$, so it is discontinuous at time 1, and continuous at time 2. The state constraint is active over $[1, 2]$. Consider only the case $\beta = 1$. Over $(2, 3]$, since the state constraint is not active and the final cost is $y(T)$, $\bar{p}(t) = 1$ and $\bar{\mu}$ is zero. Over $(1, 2)$, $\bar{p}(t) = 0$ since $\bar{u}$ is out of bounds. So, at time 2, there is a jump of $\bar{p}$ and $\bar{\mu}$ of size 1. The analysis over $(0, 1)$ is similar to the one of the previous example. Observe that, at time 2, there is a jump despite the fact that the control is continuous.

Example 5.23. Consider a problem of equilibrium of an elastic string with length $T$, fixed at endpoints, with obstacle. Specifically, we minimize the sum of elastic and gravity energies
\begin{equation}
E(y) := \int_0^T \left(\frac{1}{2}y^2(t) + y(t)\right)dt,
\end{equation}
where $y(t)$ represents the vertical deformation, with constraints
\begin{equation}
y(0) = y(T) = 1; \quad y(t) \geq 0 \text{ for all } t \in [0, T].
\end{equation}
We can reformulate this problem as an optimal control one, with state equation, integrand of integral cost, state constraint and initial-final constraints given by
\begin{equation}
\left\{ \begin{array}{l}
\dot{y}(t) = u(t), \quad \ell(u, y) := \frac{1}{2}u^2 + y; \quad g(y) := -y; \\
\Phi(y^0, y^T) := (y^0 - 1, y^T - 1); \quad K\Phi := \{0\}_{\mathbb{R}^2}.
\end{array} \right.
\end{equation}
For large enough $T$, the optimal state is, for some $t_0 \in (0, \frac{1}{2}T)$, equal to zero over $[t_0, t_1]$ with $t_1 = T - t_0$, and positive outside. The pre-Hamiltonian and costate dynamics are
\begin{equation}
\left\{ \begin{array}{l}
H(u, y, p) = \frac{1}{2}u^2 + y + pu, \\
-d\bar{p}(t) = dt - d\bar{\mu}(t).
\end{array} \right.
\end{equation}
Since $\bar{\mu}(T) = 0$, it follows that
\begin{equation}
\bar{p}(t) = T - t + \bar{\mu}(t) + \bar{p}(T), \quad t \in [0, T].
\end{equation}
By the Hamiltonian inequality, $\bar{u}(t) = -\bar{p}(t)$. Over $[t_0, t_1]$, we have therefore
\begin{equation}
0 = \dot{\bar{y}}(t) = \bar{u}(t) = -\bar{p}(t),
\end{equation}
which implies $\dot{\bar{\mu}} = 1$. Over each unconstrained arc, we have that $\bar{p}(t) = c - t$ for some $c \in \mathbb{R}$, and hence, $\bar{u}(t) = t - c$. By lemma 5.17, both the control and state constraint multiplier are continuous at times $t_0$ and $t_1$, and so, the control is equal to $t - t_0$ over $[0, t_0]$ and $t - t_1$ over $[t_1, T]$. By proposition 5.14, $\bar{\mu}$ is continuous. The relation (5.72) gives $[\bar{u}(\tau)] = -[\bar{\mu}(\tau)]$. While $\bar{u}$ has a jump of -1 at $t_0$ and 1 at $t_1$, since $\bar{\mu} = 0$ out of $[t_0, t_1]$, the jump of $\bar{\mu}$ is 1 at $t_0$ and -1 at $t_1$, so that (5.72) holds, as expected.

Example 5.24. Consider a problem of equilibrium of an elastic beam with length $T$, fixed at endpoints and with obstacle. We minimize the sum of elastic and gravity energies: $E(x) := \int_0^T \left(\frac{1}{2}x^2(t) + x(t)\right)dt$, where $x(t)$ represents the vertical deformation, subject to $x(0) = x(T) = 1$ and $x(t) \geq 0$, for all $t$. The equivalent optimal control problem, in which $y_1(t) = x(t)$, has state equation $\dot{y}_1(t) = y_2(t)$, $\dot{y}_2(t) = u(t)$, an integral cost function with integrand $\ell(u, y) = \frac{1}{2}u^2 + y_1$, the state constraint $g(y) = y_1$, and the initial-final constraints $y_1(0) = y_1(T) = 1$. For large enough $T$ the optimal state is, for some $t_0 \in (0, \frac{1}{2}T)$, equal to zero over $[t_0, t_1]$ with $t_1 = T - t_0$, and positive outside. The pre-Hamiltonian is $H(u, y, p) = \frac{1}{2}u^2 + y_1 + p_1 y_2 + p_2 u$, and the costate has dynamics
\begin{equation}
\left\{ \begin{array}{l}
-d\bar{p}_1(t) = dt - d\bar{\mu}(t), \\
-d\bar{p}_2(t) = \bar{p}_1(t).
\end{array} \right.
\end{equation}
Since \( \dot{\mu}(T) = 0 \), it follows that \( \dot{p}_1(t) = T - t + \dot{\mu}(t) + \dot{p}_1(T) \). By the Hamiltonian inequality, \( \ddot{u}(t) = -\dot{p}_2(t) \). Over \((t_0, t_1)\), we have that
\[
0 = \dot{y}_2(t) = \ddot{u}(t) = -\dot{p}_2(t) = -\dot{p}_2(t) = \dot{p}_1(t),
\]
implies \( \ddot{\mu} = 1 \). Over each unconstrained arc, we have that \( \dot{p}_1(t) = c_1 - t \) for some \( c_1 \), and hence, \( \dot{p}_2(t) = \frac{1}{2}t^2 - c_1t - c_2 \) for some \( c_2 \), so that \( \ddot{u}(t) = c_1t + c_2 - \frac{1}{2}t^2 \). By lemma 5.17, the control is continuous at times \( t_0 \) and \( t_1 \), and is therefore of the form \( -c_1(t - t_0) - \frac{1}{2}(t - t_0)^2 \) over \([0, t_0]\) and \( c_1'(t - t_1) - \frac{1}{2}(t - t_1)^2 \) over \([t_1, T]\). The curvature of the beam (i.e. the optimal control) being nonnegative and continuous with value zero on \([t_0, t_1]\), we have that \( |\dot{u}(t_0)| = c_1' \geq 0 \) and \( |\dot{u}(t_1)| = c_1'' \geq 0 \). By (5.72), the state constraint being of second order, since \( g_{j,i}^{(2)}(t) = -1 \), we have that \( |\ddot{u}(t_i)| = \dot{v}(t_i) \geq 0, i = 1, 2 \), which is compatible with the nonnegativity of the jumps of \( \ddot{u} \), established just before.

3. Proof of Pontryagin’s principle

This section is devoted to the proof of theorem 5.5. We need some preparatory lemmas.

3.1. Some results in convex analysis. In what follows \( X, Y \) are Banach spaces. Given \( f : X \to \mathbb{R} \) we define the subdifferential (in the sense of convex analysis) of \( f \) at \( x \in X \) by
\[
\partial f(x) := \{ x^* \in X^* ; f(x') \geq f(x) + \langle x^*, x' - x \rangle , \ \text{for all} \ x' \in X \}.
\]

**Lemma 5.25.** Let \( A \in L(X, Y) \) and \( F(x) := f(x) + g(Ax) \) with \( f : X \to \mathbb{R} \) and \( g : Y \to \mathbb{R} \), both convex and continuous. Then
\[
\partial F(x) = \partial f(x) + A^T \partial g(Ax), \ \text{for all} \ x \in X.
\]

**Proof.** See e.g. [22, Ch. 1, prop. 5.6 and 5.7].

We say that \( f : X \to \mathbb{R} \) is Gâteaux differentiable at \( x \in X \) iff there exists \( x^* \in X^* \) (called the Gâteaux derivative of \( f \) at \( x \)) such that
\[
\lim_{t \downarrow 0} (f(x + ty) - f(x))/t = \langle x^*, y \rangle, \ \text{for all} \ y \in X.
\]
We say that \( f \) is Hadamard differentiable at \( x \in X \) if
\[
\lim_{t \downarrow 0} (f(x + ty(t)) - f(x))/t = \langle x^*, y \rangle, \ \text{for all} \ y \in X, \ \text{whenever} \ y(t) \to y \ \text{when} \ t \downarrow 0.
\]
We may write in short G, or H differentiable. Clearly, the Hadamard differentiability implies the Gâteaux differentiability; it is easily shown that the converse holds if \( f \) is Lipschitz near \( x \).

**Lemma 5.26.** Let \( f : X \to \mathbb{R} \) be convex and continuous. Then \( f \) is H-differentiable at \( x \) with G-derivative \( x^* \in X^* \) iff \( \partial f(x) = \{ x^* \} \).

**Proof.** By [22, Ch. 1, prop. 5.3], a convex function, continuous at \( x \), whose subdifferential is a singleton, is G-differentiable at \( x \). On the other hand, by [22, Cor. 2.4], being continuous, \( f \) is Lipschitz near \( x \). So, as observed above, \( f \) is also H-differentiable at \( x \).

We say that a sequence \( x_k^* \) in \( X^* \) weakly converges to \( x^* \in X^* \) if
\[
\langle x_k^*, x \rangle \to \langle x^*, x \rangle, \ \text{for all} \ x \in X.
\]

We recall that a Banach space \( X \) is said to be separable if it contains a dense sequence. By \( X^* \) we denote its topological dual, i.e., the set of continuous linear forms over \( X \).

**Lemma 5.27.** Let \( X \) be a separable Banach space. Then any bounded sequence on \( X^* \) has a weakly convergent subsequence.
Proof. (a) Let \( x_k, k \in \mathbb{N} \), be a dense sequence in \( X \), and \( x_j^* \) be a bounded sequence in \( X^* \). We claim that \( x_j^* \) weakly converges to \( x^* \) in \( X^* \) iff
\[
\langle x_j^*, x_k \rangle \to \langle x^*, x_k \rangle, \quad \text{for all } k \in \mathbb{N}.
\]
The condition is obviously necessary. Conversely, if it holds, then given \( x \in X \) and \( \varepsilon > 0 \), there exists \( n \in \mathbb{N} \) such that \( \| x - x_n \| < \varepsilon \), and then
\[
\lim \sup_j |\langle x_j^* - x^*, x \rangle| \leq \lim \sup_j |\langle x_j^* - x^*, x - x_n \rangle| + \lim \sup_j |\langle x_j^* - x^*, x_n \rangle| = \lim \sup_j |\langle x_j^* - x^*, x - x_n \rangle| \\
\leq \varepsilon \left( \sup_j \| x_j^* \| + \| x^* \| \right)
\]
can be made arbitrarily small. Our claim follows.

(b) Let now \( y_i^* \) be a bounded sequence in \( X^* \). By a standard diagonal argument we have that some subsequence denoted by \( x_j^* \) satisfies (5.97). The conclusion follows.

Remark 5.28. See another type of proof in [18, Cor. 3.30].

Lemma 5.29. Let \( X \) be a Banach space, \( K \) a convex subset of \( X \) with nonempty interior, \( x_k \in K \) converge to \( \bar{x} \), \( x_k^* \) weakly converge to \( x^* \) in \( X^* \), be such that
\[
\langle x_k^*, x - x_k \rangle \leq 0 \quad \text{for all } x \in K.
\]
If \( \lim \inf_k \| x_k^* \| > 0 \), then \( x^* \neq 0 \).

Proof. We may assume that \( B(0, \varepsilon) \subset \text{int}(K) \), for some \( \varepsilon > 0 \). By (5.99):
\[
\varepsilon \| x_k^* \| = \sup_{x \in B(0, \varepsilon)} \langle x_k^*, \bar{x} \rangle \leq \sup_{x \in K} \langle x_k^*, \bar{x} \rangle \leq \langle x_k^*, x_k \rangle,
\]
so that
\[
0 < \lim_k \langle x_k^*, x_k \rangle = \lim_k \langle x_k^*, \bar{x} \rangle = \langle x^*, \bar{x} \rangle,
\]
proving that \( \langle x_k^*, \bar{x} \rangle > 0 \). The result follows.

3.2. Renormalization and distance to a convex set. We need the renormalization lemma 5.30 below. It is a particular case of Day [19, Ch. VII, Sec. 4, Thm 1, p. 160 of third ed.]. For other results on renormalization see Diestel [21, Ch. 4].

Lemma 5.30. Let \( X \) be a separable Banach space. Then it has equivalent norm, say \( N \), such that the corresponding dual norm is strictly convex.

Proof. Let \( x_k, k \in \mathbb{N} \), be a dense sequence in \( X \). Set \( y_k := x_k/\| x_k \| \) (we may assume the \( x_k \) to be nonzero). Consider the following function on \( X^* \):
\[
N^*(x^*) := \| x^* \| + N_1^*(x^*), \quad N_1^*(x^*) := \left( \sum_{k \in \mathbb{N}} 2^{-k} \langle x^*, y_k \rangle^2 \right)^{1/2}.
\]
Since \( \langle x^*, y_k \rangle \leq \| x^* \| \), we have that \( N_1^*(x^*) \leq \sqrt{2} \| x^* \| \), and we easily check that \( N^*(\cdot) \) is a norm, \( N^* \) is equivalent to the dual norm, since
\[
\| x^* \| \leq N^*(x^*) \leq \gamma \| x^* \|, \quad \gamma := 1 + \sqrt{2},
\]
The sum of two convex functions, one of them being is strictly convex, is strictly convex. So, the strict convexity of \( N^* \) will follow from the strict convexity of \( N_1^* \). We can write \( N_1^*(x^*) = N_2^*(Ax^*) \) where \( N_2^*(\cdot) \) is the strictly convex norm of \( \ell^2 \) (the Hilbert space of square summable sequences) and \( A \in \mathcal{L}(X^*, \ell^2) \), defined by \( \langle Ax^*, y_k \rangle := 2^{-k} \langle x^*, y_k \rangle \) for all \( k \in \mathbb{N} \), is obviously continuous and injective. It is easily checked that the composition of an injective linear mapping by a strictly convex function is strictly convex. So, \( N^* \) is a strictly convex norm over \( X^* \). Let \( B_N^* \) denote its unit ball:
\[
B_N^* = \{ x^* \in X^*; \| x^* \| + N_1^*(x^*) \leq 1 \}.
\]
Denoting by $B$ the unit ball of the primal space $X$, we have that
\begin{equation}
B_N^* = \cap_{x \in B} B_N^*(x), \quad \text{where } B_N^*(x) := \{x^* \in X^*; \langle x^*, x \rangle + N^*_1(x^*) \leq 1\}.
\end{equation}
The continuous convex function $N^*_1(x^*)$ being positively homogeneous, it can be expressed at the support function of its subdifferential at zero (see e.g. [15, equation (2.240)]), that is,
\begin{equation}
N^*_1(x^*) = \sup\{\langle x^{**}, x^* \rangle; x^{**} \in \partial N^*_1(0)\}.
\end{equation}
Here $x^{**}$ is an element of the bidual space $X^{**}$. Since $A \in L(X^*, \ell_2)$, the subdifferential calculus rule in lemma 5.25 applies. It follows that
\begin{equation}
\partial N^*_1(0) = A^T \bar{B}_2,
\end{equation}
where $\bar{B}_2$ is the closed unit ball of $\ell_2$. For $z \in \ell_2$ we have that
\begin{equation}
A^T z = \sum_{k \in \mathbb{N}} 2^{-k} z_k y_k \in X.
\end{equation}
So, setting $K := B + A^T \bar{B}_2$, we have that $K \subset X$ and
\begin{equation}
B_N^* = \{x^* \in X^*; \langle x^*, x \rangle \leq 1, \text{ for all } x \in K\}
\end{equation}
is an intersection of closed half spaces corresponding to elements of $X$. The associated primal norm
\begin{equation}
N(x) := \inf\{\lambda > 0; \lambda^{-1} x \in K\}
\end{equation}
is equivalent to the original one since, by (5.103), $K$ contains $\bar{B}$ and is contained in $\gamma \bar{B}$. The closed unit ball of this new norm is the closure of $K$. Therefore, the associated dual norm is nothing but $N^*$. The conclusion follows.

**Remark 5.31.** For a characterization of equivalent dual norms that are the dual of an equivalent primal norm, see Penot [35, Lemma 3.94, p. 251].

**Lemma 5.32.** Let $X$ be a separable Banach space endowed with the above norm $N$, and let $K$ be a closed convex subset of $X$. Then the distance function $d_K$ is H-differentiable over $X \setminus K$.

**Proof.** Obviously, $d_K$ is convex and nonexpansive. Let $x \in X \setminus K$, $x^* \in \partial d_K(x)$, and $y \in K$. Then
\begin{equation}
0 = d_K(y) \geq d_K(x) + \langle x^*, y - x \rangle.
\end{equation}
Let $y_k \in K$ be such that $\|y_k - x\| \to d_K(x)$. Then
\begin{equation}
d_K(x) \leq -\langle x^*, y_k - x \rangle \leq \|x^*\| \|y_k - x\| \to \|x^*\| d_K(x),
\end{equation}
proving that $\|x^*\| = 1$. Therefore every element of the subdifferential has a unit norm. Since the dual norm is strictly convex it follows that $\partial d_K(x)$ is a singleton. The conclusion follows with lemma 5.26.

**3.3. Diffuse perturbations.** We next consider the diffuse perturbation theory; our results are a variant of Li and Yong [28, p. 143 and 175].

**Lemma 5.33.** Let $w \in L^1(0,T)^m$, $\varepsilon > 0$ and $\rho \in (0,1)$. Then there exists a measurable subset $E$ of $(0,T)$, of measure $T\rho$, such that
\begin{equation}
\left| \int_0^T (1 - \frac{1}{\rho} \mathbf{1}_E) w(t) dt \right| \leq \varepsilon.
\end{equation}

**Proof.** The set of simple functions being dense in $L^1(0,T)^m$, for any $\varepsilon > 0$, there exists $\beta_k \in \mathbb{R}^m$ and disjoint measurable subsets $A_k$ of $(0,T)$, $k = 1$ to $n_\varepsilon$ such that
\begin{equation}
\|w - w_\varepsilon\|_1 \leq \rho\varepsilon, \quad \text{where } w_\varepsilon := \sum_{k \leq n_\varepsilon} \beta_k \mathbf{1}_{A_k}.
\end{equation}
For any \( k \leq n \) and \( \rho \in (0,1) \), there exists a measurable subset \( A'_k \) of \( A_k \) such that \( |A'_k| = \rho |A_k| \) (where by \( | \cdot | \) we denote the Lebesgue measure). Then \( E_k := \cup_{k \leq n} A'_k \) satisfies

\[
\left| \int_0^T (1 - \frac{1}{\rho} 1_{E_k}) w(t) dt \right| \leq \left| \int_0^T (1 - \frac{1}{\rho} 1_{E_k}) w_z dt \right| + \| 1 - \frac{1}{\rho} 1_{E_k} \|_\infty \| w - w_z \|_1 \leq \varepsilon,
\]

since the above integral on the r.h.s. is equal to 0, and \( \| 1 - \frac{1}{\rho} 1_{E_k} \|_\infty \leq 1/\rho \). The result follows.

We next need a variant of the previous lemma where we consider now integrals from 0 to \( t \), for all \( t \in [0,T] \).

**Lemma 5.34.** Let \( w \in L^\infty(0,T)^m \), and \( \rho_k \downarrow 0 \) with values in \((0,1)\). Then there exists a sequence \( E_k \) of subsets of \((0,T)\) such that \( |E_k| = T \rho_k \), and

\[
\left| \int_0^t w(s) ds - \frac{1}{\rho_k} \int_0^t 1_{E_k} w(s) ds \right| \leq \rho_k, \quad \text{for all } t \in (0,T).
\]

**Proof.** For each \( k \) in \( \mathbb{N} \), let \( N(k) \in \mathbb{N} \), \( N(k) \neq 0 \). Set \( G(j) := [0,jT/N(k)] \), for \( j = 1 \) to \( N(k) \), and (skipping the time argument)

\[
\varphi^k := (1_{G(1)}w, \ldots, 1_{G(N(k))}w).
\]

Let \( \varepsilon_k \downarrow 0 \). By lemma 5.33, there exists a measurable subset \( E_k \) of \((0,T)\) such that \( |E_k| = T \rho_k \), and

\[
\left| \int_0^T (1 - \frac{1}{\rho_k} 1_{E_k}) 1_{G(j)} w(t) dt \right| \leq \varepsilon_k, \quad \text{for all } 1 \leq j \leq N(k).
\]

Equivalently

\[
\left| \int_0^{jT/N(k)} w(s) ds - \frac{1}{\rho_k} \int_0^{jT/N(k)} 1_{E_k} w(s) ds \right| \leq \varepsilon_k, \quad \text{for all } 1 \leq j \leq N(k).
\]

Let \( t \in (0,T) \) and \( j \in \mathbb{N} \), with \( Tj/N(k) \leq t < T(j+1)/N(k) \). Then

\[
\Delta := \left| \int_0^t w(s) ds - \frac{1}{\rho_k} \int_0^t 1_{E_k} w(s) ds \right|
\leq \left| \int_0^{jT/N(k)} w(s) ds - \frac{1}{\rho_k} \int_0^{jT/N(k)} 1_{E_k} w(s) ds \right|
+ \left| \int_{jT/N(k)}^t w(s) ds - \frac{1}{\rho_k} \int_{jT/N(k)}^t 1_{E_k} w(s) ds \right|
\leq \varepsilon_k + \frac{T}{N(k)} \left( 1 + \frac{1}{\rho_k} \right) \| w \|_\infty.
\]

By a proper choice of \( \varepsilon_k \) and \( N(k) \), the conclusion follows. \( \square \)

**3.4. Proof of theorem 5.5.** It suffices to discuss the case when the problem is autonomous, the initial state is fixed, and the final state is free (since the essential difficulties remain in that case). Set \( K := C([0,T])^m \). In the spirit of the proof of theorem 3.13 we introduce, for \( \varepsilon > 0 \), the ‘penalized’ cost function (compare to (3.179)):

\[
J^*_p(u) := (|J_R(u) - \overline{J} + \varepsilon|^2 + d_R^2((g(y[u]))^{1/2}.
\]

The distance function \( d_K \) is computed after renormalization of the separable space \( C([0,T])^m \), so that it is Hadamard differentiable out of \( K \). For \( R \geq \| \overline{u} \| \), set

\[
U_R := L^\infty(0,T, U_{ad} \cap B(0,R)).
\]
Since $J_R^e$ is continuous over this set endowed with Ekeland’s metric (3.177), and $u$ is an $\varepsilon^2$ minimizer of $J_R^e$ over $U_R$, by Ekeland’s principle (theorem 3.57), there exists $u^* \in U_R$ such that

\begin{equation}
(5.123) \quad \left\{ \begin{array}{l}
p_E(u^*, u) \leq \varepsilon; \\
J_R^e(u^*) \leq J_R^e(u) + \varepsilon p_E(u, u^*), \quad \text{for all } u \in U.
\end{array} \right.
\end{equation}

We have that $J_R^e(u^*) > 0$ (otherwise this would contradict the fact that $\bar{u}$ is a solution of the problem). Denote by $y^*$ the associated state. Set

\begin{equation}
(5.124) \quad \left\{ \begin{array}{l}
\beta^e := \frac{(J(u^*) - J(\bar{u}) + \varepsilon^2)_+}{J_R^e(u^*)}; \\
\eta^e := \frac{d_K(g(y^*))(Dd_K(g(y^*)))}{J_R^e(u^*)} \quad \text{if } g(y^*) \notin K, \quad 0 \text{ sinon.}
\end{array} \right.
\end{equation}

Since $\|Dd_K(g(y^*))\| = 1$, we have that $\beta^2 + \|\eta^e\|^2 = 1$. In addition, since $d_K(\cdot)$ is a convex function, we have, for all $w \in C([0, T])$:

\begin{equation}
(5.125) \quad \langle \eta^e, w - g(y^*) \rangle \leq d_K(w) - d_K(g(y^*)) \leq 0.
\end{equation}

Expanding the square root function as in (3.184) we obtain that

\begin{equation}
(5.126) \quad J_R^e(u) = J_R^e(u^*) + \beta^e (J(u) - J(u^*)) + \frac{1}{2} \frac{d_K^2(g(y)) - d_K^2(g(y^*))}{J_R^e(u^*)} + o(\|u - u^*\|_1).
\end{equation}

We next analyze how to get an expansion of the term involving the distance $d_K$. Set $f_y^e[t] := f_y(u^*(t), y^*(t))$, and let $\xi^e[u]$ be defined as in (3.19), but with $u^*$ instead of $\bar{u}$, i.e.,

\begin{equation}
(5.127) \quad \dot{\xi}^e(t) = f_y^e[t] \xi^e(t) + f(u(t), y^*(t)) - f(u^*(t), y^*(t)), \quad \text{for a.a. } t \in (0, T),
\end{equation}

with $u \in U$, and initial condition $\xi(0) = 0$. This equation has a unique solution in $\Upsilon$, denoted by $\xi^e[u]$. Next, fix $u \in U$. Let $E_k, \rho_k$ be as in lemma 5.34, where

\begin{equation}
(5.128) \quad w(t) := f(u(t), y^*(t)) - f(u^*(t), y^*(t)), \quad t \in (0, T).
\end{equation}

Define $u_k$ by

\begin{equation}
(5.129) \quad u_k(t) = \left\{ \begin{array}{ll}
\begin{align*}
\quad u(t) & \quad \text{if } t \in E_k, \\
u^* & \quad \text{otherwise},
\end{align*}
\end{array} \right.
\end{equation}

with associated state $y_k$. Set $\xi_k := \xi^e[u_k]$, and $\xi := \xi^e[u]$.

**Lemma 5.35.** Given $u$, $E_k$ and $u_k$, $y_k$ as above, we have that

\begin{equation}
(5.130) \quad \|y_k - y^* - \rho_k \xi\|_\infty = o(\rho_k).
\end{equation}

**Proof.** (i) By lemma 3.9, $\|y_k - y^* - \xi_k\|_\infty = O(\|u_k - u^*\|_1) = O(\rho_k^2)$. So, it suffices to prove that $\|\xi_k - \rho_k \xi\|_\infty = o(\rho_k)$, or equivalently:

\begin{equation}
(5.131) \quad \|\xi - \xi_k/\rho_k\|_\infty = o(1).
\end{equation}

(ii) We have that $\dot{\xi}(t) = f_y^e[t] \xi(t) + w(t)$, and $\xi(0) = 0$. The solution of this linear ODE is of the form

\begin{equation}
(5.132) \quad \xi(t) = \int_0^t A(t, s) w(s) ds,
\end{equation}

where the $n \times n$ matrix $A(t, s)$ is a Lipschitz function of $(t, s)$, and $A(t, t)$ is equal to the identity. Let $W(t) := \int_0^t w(s) ds$. Integrating by parts in (5.132) (which is valid for a product of Lipschitz functions) we find that

\begin{equation}
(5.133) \quad \xi(t) = [A(t, s) W(s)]_0^t - \int_0^t \frac{\partial A(t, s)}{\partial s} W(s) ds = W(t) - \int_0^t \frac{\partial A(t, s)}{\partial s} W(s) ds.
\end{equation}

Next, set, for $t \in (0, T)$:

\begin{equation}
(5.134) \quad w_k(t) := f(u_k(t), y^*(t)) - f(u^*(t), y^*(t)); \quad W_k(t) := \int_0^t w_k(s) ds.
\end{equation}
By the above arguments we have that $\xi_k(t) = W_k(t) - \int_0^t \frac{\partial A(t,s)}{\partial s} W_k(s)ds$, and therefore

$$\xi(t) - \xi_k(t)/\rho_k = W(t) - W_k(t)/\rho_k - \int_0^t \frac{\partial A(t,s)}{\partial s} (W(s) - W_k(s)/\rho_k)ds.$$  

(5.135)

By lemma 5.34, $W_k/\rho_k \to W$ uniformly. Therefore, (5.131) follows from (5.135). \hfill \Box

In the sequel, $u$, $E_k$ and $u_k$ are defined as before the above lemma. Let $\eta^\varepsilon \in BV(0,T)^n$ be such that $d\eta^\varepsilon = \eta^\varepsilon$ and $\eta^\varepsilon(T) = 0$. Let $p^\varepsilon \in BV([0,T])^n$ be solution of (note that, since $\ell = 0$, the pre-Hamiltonian does not depend on $\beta$):

$$-dp^\varepsilon(t) = \nabla_y H(u^\varepsilon(t), y^\varepsilon(t), p^\varepsilon(t))dt + \sum_{i=1}^r \nabla g_i(y^\varepsilon(t))d\eta_i^\varepsilon(t), \text{ for a.a. } t \in [0,T],$$

$$p^\varepsilon(T) = \beta \varepsilon \nabla \phi(y^\varepsilon(T)).$$  

(5.136)

COROLLARY 5.36. We have the following expansion:

$$J_R^\varepsilon(u_k) = J_R^\varepsilon(\tilde{u}) + \rho_k \int_0^T (H(t, u(t), y^\varepsilon(t), p^\varepsilon(t)) - H(t, u^\varepsilon(t), y^\varepsilon(t), p^\varepsilon(t))) dt + o(\rho_k).$$  

(5.137)

**Proof.** i) By lemma 3.4, since

$$\|u_k - u^\varepsilon\|^2 = O(\rho_k),$$  

(5.138)

we have that

$$J(u_k) = J(u^\varepsilon) + \int_{E_k} (H(t, u(t), y^\varepsilon(t), \tilde{p}^\varepsilon(t)) - H(t, u^\varepsilon(t), y^\varepsilon(t), \tilde{p}^\varepsilon(t))) dt + o(\rho_k),$$  

(5.139)

where $\tilde{p}^\varepsilon$ is solution of

$$\begin{cases}
-\tilde{dp}^\varepsilon(t) = \nabla_y H(u^\varepsilon(t), y^\varepsilon(t), \tilde{p}^\varepsilon(t)), \text{ for a.a. } t \in (0,T), \\
\tilde{p}^\varepsilon(T) = \nabla \phi(y^\varepsilon(T)).
\end{cases}$$  

(5.140)

ii) Since $\frac{1}{2}d_k^2$ is H-differentiable with derivative $\eta^\varepsilon$ (see (5.124)), and a composition of H-differentiable functions is H-differentiable, using lemma 5.35, we get that

$$\int_{E_k} (H(t, u(t), y^\varepsilon(t), \tilde{p}^\varepsilon(t)) - H(t, u^\varepsilon(t), y^\varepsilon(t), \tilde{p}^\varepsilon(t))) dt + o(\rho_k).$$  

(5.141)

Extending lemma 3.4 and using again (5.138), we can check that

$$\langle \eta^\varepsilon, y_k - y^\varepsilon \rangle = \int_{E_k} (H(t, u(t), y^\varepsilon(t), \tilde{p}^\varepsilon(t)) - H(t, u^\varepsilon(t), y^\varepsilon(t), \tilde{p}^\varepsilon(t))) dt + o(\rho_k),$$  

(5.142)

where $\tilde{p}^\varepsilon$ is solution of

$$\begin{cases}
-\tilde{dp}^\varepsilon(t) = \nabla_y H(u^\varepsilon(t), y^\varepsilon(t), \tilde{p}^\varepsilon(t))dt + J_R(\varepsilon, u^\varepsilon, \tilde{p}^\varepsilon) \sum_{i=1}^r \nabla g_i(y^\varepsilon(t))d\eta_i^\varepsilon(t), \text{ for a.a. } t \in (0,T), \\
\tilde{p}^\varepsilon(T) = 0.
\end{cases}$$  

(5.143)

(iii) Observe that $p^\varepsilon$ solution of (5.136) satisfies

$$p^\varepsilon = J_R(\varepsilon, u^\varepsilon)\tilde{p}^\varepsilon + \beta \varepsilon \tilde{p}^\varepsilon.$$  

(5.144)

Since a composition of H-differentiable mappings is H-differentiable, combining the previous steps, we get the conclusion. \hfill \Box

Then, by arguments similar to those in the proof of theorem 3.13 (taking advantage of the integration by parts formula (5.19)) we deduce that

$$H(\beta \varepsilon, u^\varepsilon(t), y^\varepsilon(t), p^\varepsilon(t)) \leq H(\beta \varepsilon, u, y^\varepsilon(t), p^\varepsilon(t)) + \varepsilon,$$

for all $u \in U_R$, and a.a. $t \in (0,T)$.  

(5.145)

Let us take $R = R_\varepsilon = \max(\|\tilde{u}\|, 1/\sqrt{\varepsilon})$. Since $\rho_E(\varepsilon, \bar{u}) \leq \varepsilon$, we have that, for small enough $\varepsilon$, $\|u^\varepsilon - \bar{u}\| \leq 2\varepsilon R_\varepsilon = 2\sqrt{\varepsilon}$ proving the uniform convergence of $y^\varepsilon$ towards $\bar{y}$. Next since
\[ \beta^2 + \| \eta^\varepsilon \|^2 = 1, \] in view of lemma 5.27 and (5.125) we can extract a subsequence \( \varepsilon_k \downarrow 0 \), such that (denoting by \( \rightharpoonup^* \) the *weak convergence)

\[
(5.146) \quad \begin{cases} 
\beta_{\varepsilon_k} \rightharpoonup^* \bar{\beta} \in [0, 1]; \\
\eta_{\varepsilon_k} \rightharpoonup^* \bar{\eta}; \\
\langle \bar{\eta}, y - g(\bar{y}) \rangle \leq 0, \quad \text{for all } y \in K.
\end{cases}
\]

We deduce from the costate equation that \( \mathbf{p}_{\varepsilon_k} \) converges in the weak * topology of \( M([0, T], \mathbb{R}^n) \) to \( \mathbf{p} \) solution of the costate equation (5.21). This implies \( \mathbf{p}_{\varepsilon_k} \rightharpoonup^* \mathbf{p} \) in \( L^1(0, T)^n \) and (up to the extraction of a subsequence) also a.e., see [1, Thm. 3.23]. Passing to the limit in (5.145), we obtain the Hamiltonian inequality. Passing to the limit in the last relation of (5.146) we obtain (5.25). If \( \bar{\beta} > 0 \), dividing \( (\bar{\eta}, \mathbf{p}) \) by \( \bar{\beta} > 0 \), we obtain the conclusion with \( \bar{\beta} = 1 \). Finally, if \( \bar{\beta} = 0 \), then \( \| \eta_k \| \to 1 \), and then by lemma 5.29, the non nullity relation (5.26) follows.
Bibliography


augmented metric, 57
autonomous, 42
batch process, 12
biextremal
  Pontryagin, 36
boundary, 16
bounded variation, 72
cell population, 11
closure, 16
composite cost function, 58
condition
  sufficient, 66
cone, 17
  normal, 17
  polar, 18
  positive, 17
constraint
  interior point, 48
continuity
  modulus, 20
  quasi uniform, 27
  uniform, 20
continuous extension, 29
control, 28
controlled dynamical systems, 28
convergence
  Pontryagin, 37
  weak, 83
cost
  final, 29
  integral, 29
  reduced, 29
costate, 30
costate equation, 73
derivative, 18
  partial, 19
  weak, 27
derivatives
  directional, 18
diameter, 57
differentiable
  Fréchet, 18
directional derivatives, 31
extremal
  Pontryagin, 36
fish harvesting, 10
Fuller, 10
Gâteaux differentiable, 83
geodesics, 53
gradient, 19
Hadamard differentiable, 83
Hessian, 23
Implicit function theorem, 20
inequality
  Hamiltonian, 36
integral
  Stieltjes, 72
  integrand, 29
integration by parts, 73
interior, 16
invertible linear operator, 20
Lagrange multipliers, 41
Lagrangian, 23
  reduction, 23
linear
  form, 16
Lokta-Volterra, 12
matrix
  Jacobian, 19
Mayer form, 10
minimum
  local, 17
  Pontryagin, 37
  weak, 31
Minkowski sum, 16
multiplier
  abnormal, 39
  cost, 39
  normal, 39
  singular, 39
Nemitskii mapping, 26
Newton’s method, 21
norm
  dual, 16
  normal
    cone, 17
    direction, 17
    normalized time, 45
orthogonal space, 17
pharmacodynamics, 11
pharmacokinetics, 11
polar cone, 18
quadratic growth, 66
reduced
cost, 22
problem, 22
reflexive
multifunction, 57
segment, 17
separation, 18
shooting, 61
discrete, 65
equation, 62
function, 62
parameter, 62
space
Banach, 15
Hilbert, 16
state, 28
constraints, 72
equation, 29
stationary point, 23
steady-state, 11
subdifferential, 83
Taylor expansion, 19
theorem
Fréchet-Riesz, 16
time dilation, 75
TPBVP, 62
trajectory, 29
transitive
multifunction, 57
transpose operator, 16
transversality conditions, 39
two point boundary value problem, 31, 62