Chapter 1

Introduction to Optimal Control

1.1 Historical Introduction to Optimal Control

- 1940's: Calculus of variation has become a standard math subject, associated with famous names such as Euler, Lagrange, Hilbert, and Jacobi.
- 1956: Potryagin and his collaborators develop Optimal Control, motivated by the US-Russian space race to reach the moon. This is a non-trivial development from calculus of variations.
- 1970's: Optimal control is applied to standard mathematical problems, including Riemannian and sub-Riemannian geometry.

1.2 Geometric Step

As in calculus of variation, we start with

$$\min \int_{t_0}^{t_1} L(x, \dot{x})dt$$

and introduce new notation for the control, $\dot{x} = u(t)$, and denote the cost, $\dot{x}^\circ = L(x, \dot{x}) = L(x, u)$. In addition, we further denote $\hat{x} = \begin{pmatrix} x \\ x^\circ \end{pmatrix}$ and rewrite the system as

$$\dot{\hat{x}} = F(\hat{x}, u)$$

$$\hat{x}(0) = \begin{pmatrix} x(0) \\ x^\circ(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$$

Then $u(\cdot) \to \hat{x}(t, \hat{x}(0), u) \in$ boundary of the Accessibiity set, A, at time $t = t_1$.

In general, for the system $\dot{x} = F(x, u)$, we need only ensure $t \to x(t, x_0, u)$ is well-defined. We assume F is smooth and by Caratheodory we can conclude we are well-defined as long as u is any measurable and bounded function.

We denote such a set of controls $u(t) \in U$, the *Control Domain.* Typically, further restrictions are placed on U based on the application under study.

1.3 The Weak Maximum Principle

Although stated as "weak", this version of the maximum principle covers all standard calculus of variation results we have discussed.

We define the following system:

$$\Sigma = \frac{dx}{dt} = F(x, u(t))$$

with $x \in \mathbb{R}^n$, a control $u(\cdot)$, and initial state $x(0) = x_0$. We assume F is smooth. We define $\mathcal{U} = \{u : [0, T(u)] \to U \subset \mathbb{R}^m | u(\cdot) \in L^\infty\}$, where we define the $\|u\| = \sup |u(t)|$, $t \in [0, T]$. Or we can use any other norm on $\mathbb{R}^m$.

For $u(\cdot) \in U$, we have $t \to x(t, x_0, u)$ for $t \in [0, T']$, with $T' < T$, and obtain a solution to the system $(u(\cdot), x(\cdot))$. Thus $u(\cdot) \to \dot{x} = F(x, u) \to x(t, x_0, u)$.

1.3.1 Input-Output Mapping

We define the *input-output mapping*, $u(\cdot) \in \mathcal{U} \to x(T, x_0, u) \in \mathbb{R}^n$. This mapping is smooth. To show this, we will use the *Frechet derivative*.

Choose a $u(\cdot)$, $x(\cdot)$, and $\dot{x} = F(x, u)$. We apply a variation, $v(\cdot)$:

$$u(\cdot) + v(\cdot) \to x(t) + \xi(t)$$

We now fix t and calculate the Taylor expansion of F.

$$F(x + \xi, u + v) = F(x, u) + \frac{dF}{dx}(x, u)\xi + \frac{dF}{du}(x, u)v + \dots$$

Then:

$$\dot{x} + \dot{\xi} = F(x + \xi, u + v) = F(x, u) + \frac{dF}{dx}(x, u)\xi + \frac{dF}{du}(x, u)v + \dots$$

We let $A(t) = \frac{dF}{dx}(x, u)$ and $B(t) = \frac{dF}{du}(x, u)$ to then conclude:

$$\dot{\xi} = A(t)\xi + B(t)v$$

Theorem 1. *The Frechet derivation of $E_u^{x_0,T}(v)$ is $\xi(T)$.*

The explicit formula for $\xi(T)$ is given by:

$$\xi(0) = 0$$
$$\xi(T) = M(T) \int_0^T M^{-1}(t)B(t)v(t)dt$$

where $M(t)$ is the matrix solution of $\dot{M} = AM$, with $M(0) = \text{Identity}$.

1.3.2 Singular Control

Definition 1. *Control $u(\cdot) \in U$ is called* regular *if at $u(\cdot)$, E' is of full rank. $u(\cdot)$ is called* singular *if E' is not of full rank (ie. not regular).*

The computation of the singular control is given via the weak maximum principle.

Denote $F =$ image of E' (a finite dimensional vector space). Assume $u(\cdot)$ is singular, therefore $\dim F < n$. Take $\bar{P}$ orthogonal to F, $\bar{P} \neq 0$. Then

$$\bar{P} \cdot F = \bar{P}M(T) \int_0^t M^{-1}(t)B(t)v(t)dt = 0$$

for each $v(\cdot) \in L^\infty$. This occurs if and only if $\bar{P}M(T)M^{-1}(t)B(t) = 0$ for a.e. $t \in [0, T]$.

We set $p(t) = \bar{P}M(T)M^{-1}(t)$, $z(t) = (x(t), p(t))$, and denote the *pseudo-Hamiltonian*:

$$H(z(t), u(t)) = p(t) \cdot F(x(t), u(t))$$

Then we make the following claim: $u(\cdot)$ is singular if there exists $p(t)$ such that:

$$\dot{x} = \frac{\partial H}{\partial p} = F(x, u)$$

$$\dot{p} = -\frac{\partial H}{\partial x}$$
$$\frac{\partial H}{\partial u} = 0$$

It is this third condition that allows us to compute the u.

Theorem 2. *Let $\dot{x} = F(x,u)$, $x(0) = x_0$. U open set in $\mathbb{R}^m$. Assume $x(T, x_0, u) \in$ Boundary of accessibility set $A(x_0, T) = \bigcup_{u \in U} x(T, x_0, u)$. This it must satisfy:*

$$\dot{x} = \frac{\partial H}{\partial p}$$
$$\dot{p} = -\frac{\partial H}{\partial x}$$
$$\frac{\partial H}{\partial u} = 0$$

for almost every t.

Proof. If $x(T, x_0, u) \in \partial A(x_0, T)$, then $\dim E_u'^{x_0,T} < n$, and therefore $u(\cdot)$ is singular. If E' were of full rank, then the mapping would give us an open set instead. □

1.4 Application to Computation of Geodesics in Riemannian geometry

Let

$$\dot{x} = \sum_{i=1}^{n} u_i(t) F_i(x(t))$$

with $\dot{x} \in \mathbb{R}^n$, $\{F_i\}$ a set of n vector fields. We assume the F_i are orthonormal. Then we can define the length of $x(\cdot)$:

$$l(x) = \int_0^T \left(\sum u_i^2(t) \right)^{1/2} dt$$

Remark: $\dot{x} = (F_1, \ldots, F_n) \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = Fu$, and therefore $u = F^{-1}\dot{x}$. Thus $\sum u_i^2$ has a quadratic form with respect to $\dot{x}$.

By the Moupetuis Principle, we know that the minimum energy $\left(\sum u_i^2(t)\right)$ solution must be the same as the minimum length solution.

We now write the extended system:

$$\dot{x} = \sum u_i F_i$$
$$\dot{x}^\circ = \sum u_i^2 \quad \text{the cost}$$

Then for the extended system, $u(\cdot)$ is singular. We denote $\hat{z} = (\hat{x}, \hat{p}) = \left(\begin{pmatrix} x \\ x^\circ \end{pmatrix}, \begin{pmatrix} p \\ p^\circ \end{pmatrix} \right)$ and define the Hamiltonian.

$$H(\hat{z}, u) = p \sum u_i F_i + p^\circ \sum u_i^2$$

p° is constant, and by homogeneity, $p^\circ = -\frac{1}{2}$.

We use $\frac{\partial H}{\partial u} = 0$ to get $u_i = pF_i = H_i$. We plug this in to obtain:

$$H(z) = \frac{1}{2}\sum_{i=1}^{n} H_i^2$$

$$\dot{x} = \frac{\partial H}{\partial p}$$

$$\dot{p} = -\frac{\partial H}{\partial x}$$