EE291E Lecture Notes 8. Optimal Control and Dynamic Games

Claire J. Tomlin

March 7, 2018

These notes represent an introduction to the theory of optimal control and dynamic games; they were written by S. S. Sastry [1].

There exist two main approaches to optimal control and dynamic games:

1. via the Calculus of Variations (making use of the Maximum Principle);
2. via Dynamic Programming (making use of the Principle of Optimality).

Both approaches involve converting an optimization over a function space to a pointwise optimization. The methods are based on the following simple observations:

1. For the calculus of variations, the optimal curve should be such that neighboring curves to not lead to smaller costs. Thus the 'derivative' of the cost function about the optimal curve should be zero: one takes small variations about the candidate optimal solution and attempts to make the change in the cost zero.
2. For dynamic programming, the optimal curve remains optimal at intermediate points in time.

In these notes, both approaches are discussed for optimal control; the methods are then extended to dynamic games.

1 Optimal Control based on the Calculus of Variations

There are numerous excellent books on optimal control. Commonly used books which we will draw from are Athans and Falb [2], Berkovitz [4], Bryson and Ho [5], Pontryagin et al [6], Young [7], Kirk [8], Lewis [9] and Fleming and Rishel[10]. The history of optimal control is quite well rooted in antiquity, with allusion being made to Dido, the first Queen of Carthage, who when asked to take as much land as could be covered by an ox-hide, cut the ox-hide into a tiny strip and proceeded to enclose the entire area of what came to be know as Carthage in a circle of the appropriate radius[1]. The calculus of variations is really the ancient precursor to optimal control. Iso perimetric problems of the kind that gave Dido her kingdom were treated in detail by Tonelli and later by Euler. Both Euler and Lagrange laid the foundations of mechanics in a variational setting culminating in the Euler Lagrange equations. Newton used variational methods to determine the shape of a body that minimizes drag, and Bernoulli formulated his brachistochrone problem in the seventeenth century, which attracted the attention of Newton and L'Hôpital. This intellectual heritage was revived and generalized by Bellman [3] in the context of dynamic programming and by Pontryagin and his school in the so-called Pontryagin principle for optimal control ([6]).

Consider a nonlinear possibly time varying dynamical system described by

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

with state $x(t) \in \mathbb{R}^n$ and the control input $u \in \mathbb{R}^{n_i}$. Consider the problem of minimizing the performance index

$$J = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} L(x(t), u(t), t)dt \tag{2}$$

where t_0 is the initial time, t_f the final time (free), $L(x, u, t)$ is the running cost, and $\phi(x(t_f), t_f)$ is the cost at the terminal time. The initial time t_0 is assumed to be fixed and t_f variable. Problems involving a cost only on the final and initial state are referred to as Mayer problems, those involving only the integral or running cost are called Lagrange problems and costs of the form of equation (2) are referred to as Bolza problems. We will also have a constraint on the final state given by

$$\psi(x(t_f), t_f) = 0 \tag{3}$$

where $\psi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^p$ is a smooth map. To derive necessary conditions for the optimum, we will perform the calculus of variations on the cost function of (2) subject to the constraints of equations (1), (3). To this end, define the modified cost function, using the Lagrange multipliers $\lambda \in \mathbb{R}^p, p(t) \in \mathbb{R}^n$,

$$\tilde{J} = \phi(x(t_f), t_f) + \lambda^T \psi(x(t_f), t_f) + \int_{t_0}^{t_f} \left[L(x, u, t) + p^T(f(x, u, t) - \dot{x})\right] dt \tag{4}$$

Defining the *Hamiltonian* $H(x, u, t)$ using what is referred to as a *Legendre transformation*

$$H(x, p, u, t) = L(x, u, t) + p^T f(x, u, t) \tag{5}$$

The variation of (4) is given by assuming independent variations in $\delta u(), \delta x(), \delta p(), \delta\lambda$, and δt_f:

$$\begin{aligned} \delta\tilde{J} = \;& (D_1\phi + D_1\psi^T\lambda)\delta x|_{t_f} + (D_2\phi + D_2\psi^T\lambda)\delta t|_{t_f} + \psi^T\delta\lambda \\ +\;& (H - p^T\dot{x})\delta t|_{t_f} \\ +\;& \textstyle\int_{t_0}^{t_f} \left[D_1 H\delta x + D_3 H\delta u - p^T\delta\dot{x} + (D_2 H^T - \dot{x})^T\delta p\right] dt \end{aligned} \tag{6}$$

[1]The optimal control problem here is to enclose the maximum area using a closed curve of given length.

The notation $D_i H$ stands for the derivative of H with respect to the i th argument. Thus, for example,

$$D_3 H(x,p,u,t) = \frac{\partial H}{\partial u} \quad D_1 H(x,p,u,t) = \frac{\partial H}{\partial x}$$

Integrating by parts for $\int p^T \delta \dot{x} dt$ yields

$$\begin{aligned} \delta \tilde{J} = & \ (D_1\phi + D_1\psi^T\lambda - p^T)\delta x(t_f) + (D_2\phi + D_2\psi^T\lambda + H)\delta t_f + \psi^T \delta\lambda \\ + & \int_{t_0}^{t_f} \left[(D_1 H + \dot{p}^T)\delta x + D_3 H \delta u + (D_2^T H - \dot{x})^T \delta p \right] dt \end{aligned} \tag{7}$$

An extremum of $\tilde{J}$ is achieved when $\delta \tilde{J} = 0$ for all independent variations $\delta\lambda, \delta x, \delta u, \delta p$. These conditions are recorded in the following

Table of necessary conditions for optimality:

Table 1

Description	Equation	Variation
Final State constraint	$\psi(x(t_f), t_f) = 0$	$\delta\lambda$
State Equation	$\dot{x} = \frac{\partial H}{\partial p}^T$	δp
Costate equation	$\dot{p} = -\frac{\partial H}{\partial x}^T$	δx
Input stationarity	$\frac{\partial H}{\partial u} = 0$	δu
Boundary conditions	$D_1\phi - p^T = -D_1\psi^T\lambda\vert_{t_f}$	$\delta x(t_f)$
	$H + D_2\phi = -D_2\psi^T\lambda\vert_{t_f}$	δt_f

The conditions of Table (1) and the boundary conditions $x(t_0) = x_0$ and the constraint on the final state $\psi(x(t_f), t_f) = 0$ constitute the necessary conditions for optimality. The end point constraint equation is referred to as the transversality condition:

$$\begin{aligned} D_1\phi - p^T = & \ -D_1\psi^T\lambda \\ H + D_2\phi = & \ -D_2\psi^T\lambda \end{aligned} \tag{8}$$

The optimality conditions may be written explicitly as

$$\begin{aligned} \dot{x} = & \ \frac{\partial H}{\partial p}^T(x, u^*, p) \\ \dot{p} = & \ -\frac{\partial H}{\partial x}^T(x, u^*, p) \end{aligned} \tag{9}$$

with the stationarity condition reading

$$\frac{\partial H}{\partial u}(x, u^*, p) = 0$$

and the endpoint constraint $\psi(x(t_f), t_f) = 0$. *The key point to the derivation of the necessary conditions of optimality is that the Legendre transformation of the Lagrangian to be minimized into a Hamiltonian converts a functional minimization problem into a static optimization problem on the function* $H(x, u, p, t)$.

The question of when these equations also constitute sufficient conditions for (local) optimality is an important one and needs to be ascertained by taking the second variation of $\tilde{J}$. This is an involved procedure but the input stationarity condition in Table (1) hints at the **sufficient condition for local minimality** of a given trajectory $x^*(\cdot), u^*(\cdot), p^*(\cdot)$ being a local minimum as being that the Hessian of the Hamiltonian,

$$D_2^2 H(x^*, u^*, p^*, t) \tag{10}$$

being positive definite along the optimal trajectory. A sufficient condition for this is to ask simply that the $n_i \times n_i$ Hessian matrix

$$D_2^2 H(x, u, p, t) \tag{11}$$

be positive definite. As far as conditions for **global minimality** are concerned, again the stationarity condition hints at a sufficient condition for global minimality being that

$$u^*(t) = \underset{\{\text{ min over u }\}}{\operatorname{argmin}} H(x^*(t), u, p^*(t), t) \tag{12}$$

Sufficient conditions for this are, for example, the convexity of the Hamiltonian $H(x, u, p, t)$ in u.

Finally, there are instances in which the Hamiltonian $H(x, u, p, t)$ is not a function of u at some values of x, p, t. These cases are referred to as *singular extremals* and need to be treated with care, since the value of u is left unspecified as far as the optimization is concerned.

1.1 Fixed Endpoint problems

In the instance that the final time t_f is fixed, the equations take on a simpler form, since there is no variation in δt_f. Then, the boundary condition of equation (8) becomes

$$p^T(t_f) = D_1\phi + D_1\psi^T\lambda|_{t_f} \tag{13}$$

Further, if there is no final state constraint the boundary condition simplifies even further to

$$p(t_f) = D_1\phi^T|_{t_f} \tag{14}$$

1.2 Time Invariant Systems

In the instance that $f(x, u, t)$ and the running cost $L(x, u, t)$ are not explicitly functions of time, there is no final state constraint and the final time t_f is fixed, the formulas of Table (1) can be rewritten as

State Equation	$\dot{x} = \frac{\partial H}{\partial p}^T = f(x, u^*)$
Costate Equation	$\dot{p} = -\frac{\partial H}{\partial x}^T = -D_1 f^T p + D_1 L^T$
Stationarity Condition	$0 = \frac{\partial H}{\partial u} = D_2 L^T + D_2 f^T p$
Transversality Conditions	$D_1\phi - p^T = -D_1\psi^T\lambda$
	$H(t_f) = 0$

In addition, it may be verified that

$$\frac{dH^*}{dt} = \frac{\partial H^*}{\partial x}(x,p)\dot{x} + \frac{\partial H^*}{\partial p}\dot{p} = 0 \tag{15}$$

thereby establishing that $H^*(t) \equiv 0$.

1.3 Connections with Classical Mechanics

Hamilton's principle of least action states (under certain conditions [2]) that a conservative system moves so as to minimize the time integral of its "action", defined to be the difference between the kinetic and potential energy. To make this more explicit we define $q \in \mathbb{R}^n$ to be the vector of generalized coordinates of the system and denote by $U(q)$ the potential energy of the system and $T(q, \dot{q})$ the kinetic energy of the system. Then Hamilton's principle of least action asks to solve an optimal control problem for the system

$$\dot{q} = u$$

with Lagrangian

$$L(q, u) = T(q, u) - U(q)$$

The equations (9) in this context have $H(q, u, p) = L(q, u) + p^T u$. $u^* = u^*(p, q)$ is chosen so as to minimize the Hamiltonian H. A necessary condition for stationarity is that $u^*(p, q)$ satisfies

$$0 = \frac{\partial H}{\partial u} = \frac{\partial L}{\partial u} + p^T \tag{16}$$

The form of the equations (9) in this context is that of the familiar *Hamilton Jacobi equations.* The costate p has the interpretation of momentum.

$$\begin{aligned} \dot{q} &= \frac{\partial H^*}{\partial p}(p,q) &&= u^*(p,q) \\ \dot{p} &= -\frac{\partial H^*}{\partial q}(p,q) \end{aligned} \tag{17}$$

Combining the second of these equations with (16) yields the familiar *Euler Lagrange equations*

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0 \tag{18}$$

[2] For example, there is no dissipation or no nonholonomic constraints. Holonomic or integrable constraints are dealt with by adding appropriate Lagrange multipliers. If nonholonomic constraints are dealt with in the same manner, we get equations of motion, dubbed vakonomic by Arnold [11] which do not correspond to experimentally observed motions. On the other hand, if there are only holonomic constraints, the equations of motion that we derive from Hamilton's principle of least action is equivalent to Newton's laws.

2 Optimal Control based on Dynamic Programming

To begin this discussion, we will embed the optimization problem which we are solving in a larger class of problems, more specifically we will consider the original cost function of equation (2) from an initial time $t \in [t_0, t_f]$ by considering the cost function on the interval $[t, t_f]$:

$$J(x(t), t) = \phi(x(t_f), t_f) + \int_t^{t_f} L(x(\tau), u(\tau), \tau) d\tau$$

Bellman's principle of optimality says that if we have found the optimal trajectory on the interval from $[t_0, t_f]$ by solving the optimal control problem on that interval, the resulting trajectory is also optimal on all subintervals of this interval of the form $[t, t_f]$ with $t > t_0$, provided that the initial condition at time t was obtained from running the system forward along the optimal trajectory from time t_0. The optimal value of $J(x(t), t)$ is referred to as the "cost-to go". To be able to state the following key theorem of optimal control we will need to define the "optimal Hamiltonian" to be

$$H^*(x, p, t) := H(x, u^*, p, t)$$

Theorem 1 The Hamilton Jacobi Bellman equation

Consider, the time varying optimal control problem of (2) with fixed endpoint t_f and time varying dynamics. If the optimal value function, i.e. $J^*(x(t_0), t_0)$ is a smooth function of x, t, then $J^*(x, t)$ satisfies the **Hamilton Jacobi Bellman** *partial differential equation*

$$\frac{\partial J^*}{\partial t}(x, t) = -H^*(x, \frac{\partial J^*}{\partial x}(x, t), t) \tag{19}$$

with boundary conditions given by $J^*(x, t_f) = \phi(x, t_f)$ for all $x \in \{x : \psi(x, t_f) = 0\}$.

Proof: The proof uses the principle of optimality. This principle says that if we have found the optimal trajectory on the interval from $[t, t_f]$ by solving the optimal control problem on that interval, the resulting trajectory is also optimal on all subintervals of this interval of the form $[t_1, t_f]$ with $t_1 > t$, provided that the initial condition at time t_1 was obtained from running the system forward along the optimal trajectory from time t. Thus, from using $t_1 = t + \Delta t$, it follows that

$$J^*(x, t) = \begin{array}{c} \min \\ u(\tau) \\ t \leq \tau \leq t + \Delta t \end{array} \left[\int_t^{t+\Delta t} L(x, u, \tau) d\tau + J^*(x + \Delta x, t + \Delta t) \right] \tag{20}$$

Taking infinitesimals and letting $\Delta t \to 0$ yields that

$$-\frac{\partial J^*}{\partial t} = \begin{array}{c} \min \\ u(t) \end{array} \left(L + (\frac{\partial J^*}{\partial x}) f \right) \tag{21}$$

with the boundary condition being that the terminal cost is

$$J^*(x, t_f) = \phi(x, t_f)$$

on the surface $\psi(x) = 0$. Using the definition of the Hamiltonian in equation (5), it follows from equation (21) that the Hamilton Jacobi equation of equation (19) holds.

$\square$

Remarks:

1. The preceding theorem was stated as a necessary condition for extremal solutions of the optimal control problem. As far as minimal and global solutions of the optimal control problem, the Hamilton Jacobi Bellman equations read as in equation (21). In this sense, the form of the Hamilton Jacobi Bellman equation in (21) is more general.

2. The **Eulerian conditions** of Table (1) are easily obtained from the Hamilton Jacobi Bellman equation by proving that $p^T(t) := \frac{\partial J^*}{\partial x}(x,t)$ satisfies the costate equations of that Table. Indeed, consider the equation (21). Since $u(t)$ is unconstrained, it follows that it should satisfy
$$\frac{\partial L}{\partial u}(x^*, u^*) + \frac{\partial f}{\partial u}^T p = 0 \tag{22}$$
Now differentiating the definition of $p(t)$ above with respect to t yields
$$\frac{dp^T}{dt} = \frac{\partial^2 J^*}{\partial t \partial x}(x^*, t) + \frac{\partial^2 J^*}{\partial x^2} f(x^*, u^*, t) \tag{23}$$
Differentiating the Hamilton Jacobi equation (21) with respect to x and using the relation (22) for a stationary solution yields
$$-\frac{\partial^2 J^*}{\partial t \partial x}(x^*, t) = \frac{\partial L}{\partial x} + \frac{\partial^2 J^*}{\partial x^2} f + p^T \frac{\partial f}{\partial x} \tag{24}$$
Using equation (24) in equation (23) yields
$$-\dot{p} = \frac{\partial f}{\partial x}^T p + \frac{\partial L}{\partial x}^T \tag{25}$$
establishing that p is indeed the co-state of Table 1. The boundary conditions on $p(t)$ follow from the boundary conditions on the Hamilton Jacobi Bellman equation.

2.1 Constrained Input Problems

In the instance that there are no constraints on the input, the extremal solutions of the optimal control problem are found by simply extremizing the Hamiltonian and deriving the stationarity condition. Thus, if the specification is that $u(t) \in U \subset \mathbb{R}^{n_i}$ then, the optimality condition is that
$$H(x^*, u^*, p^*, t) \leq H(x^*, u, p^*, t) \quad \forall u \in U \tag{26}$$
If the Hamiltonian is convex in u and U is a convex set, there are no specific problems with this condition. In fact, when there is a single input and the set U is a single closed interval, there are several interesting examples of Hamiltonians for which the optimal inputs switch between the endpoints of the interval, resulting in what is referred to as **bang bang control**. However, problems can arise when U is either not convex or compact. In these cases, a concept of a **relaxed control** taking values in the convex hull of U needs to be introduced. As far as an implementation of a control $u(t) \in \text{conv}U$, but not in U, a probabilistic scheme involving switching between values of U whose convex combination u is needs to be devised.

2.2 Free end time problems

In the instance that the final time t_f is free, the transversality conditions are that

$$\begin{aligned} p^T(t_f) = & \quad D_1\phi + D_1\psi^T\lambda \\ H(t_f) = & \quad -(D_2\phi + D_2\psi^T\lambda) \end{aligned} \tag{27}$$

2.2.1 Minimum time problems

A special class of minimum time problems of especial interest is minimum time problems, where t_f is to be minimized subject to the constraints. This is accounted for by setting the Lagrangian to be 1, and the terminal state cost $\phi \equiv 0$, so that the Hamiltonian is $H(x,u,p,t) = 1 + p^T f(x,u,t)$. Note that by differentiating $H(x,u,p,t)$ with respect to time, we get

$$\frac{dH^*}{dt} = D_1H^*\dot{x} + D_2H^*\dot{u} + D_3H^*\dot{p} + \frac{\partial H^*}{\partial t} \tag{28}$$

Continuing the calculation using the Hamilton Jacobi equation,

$$\frac{dH^*}{dt} = (\frac{\partial H^*}{\partial x} + \dot{p})f(x,u^*,t) + \frac{\partial H^*}{\partial t} = \frac{\partial H^*}{\partial t} \tag{29}$$

In particular, if H^* is not an explicit function of t, it follows that $H^*(x,u,p,t) \equiv H$. Thus, for minimum time problems for which $f(x,u,t)$ and $\psi(x,t)$ are not explicitly functions of t, it follows that $0 = H(t_f) \equiv H(t)$.

3 Two person zero sum dynamical games

The theory of games also has a long and distinguished, if not as long a history. Borel encountered saddle points in a matrix in 1915, but it really took von Neumann to prove his fundamental theorem about the existence of mixed strategies for achieving a saddle solution in games in 1936. In a classic book, von Neumann and Morgenstern ([12]) laid out the connections between static games and economic behavior. Nash, von Stackelberg and others extended the work to N person non-cooperative games. This work has continued in many important new directions in economics. Differential or dynamical games showed up in the work of Isaacs in 1969 and rapidly found fertile ground in the control community where it has progressed. There are several nice books on the subject of dynamical games, our treatment is drawn heavily from Basar and Olsder ([13]).

Consider a so-called dynamical, two person zero sum, perfect state information game modeled by

$$\dot{x} = f(x,u,d,t) \qquad x(t_0) = x_0 \tag{30}$$

on a fixed duration $[t_0, t_f]$. Here $x \in \mathbb{R}^n$ models the state of the system and $u \in \mathbb{R}^{n_1}, d \in \mathbb{R}^{n_2}$ represent the actions of the two players. The game is said to be *zero sum* if one player seeks to minimize and the other to maximize the same cost function taken to be of the form

$$J = \phi_(x(t_f), t_f) + \int_{t_0}^{t_f} L(x, u, d, t)dt \tag{31}$$

We will assume that player 1 (u) is trying to minimize J and player 2 (d) is trying to maximize J. For simplicity we have omitted the final state constraint and also assumed the end time t_f to be fixed. These two assumptions are made for simplicity but we will discuss the t_f free case when we study pursuit evasion games. The game is said to have perfect information if both players have access to the full state $x(t)$. The solution of two person zero sum games proceeds very much along the lines of the optimal control problem by setting up the Hamiltonian

$$H(x, u, d, p, t) = L(x, u, d, t) + p^T f(x, u, d, t) \tag{32}$$

Rather than simply minimizing $H(x, u, d, p, t)$ the game is said to have a *saddle point solution* if the following analog of the saddle point condition for two person zero sum static games holds:

$$\min_{\mathrm{u}} \max_{\mathrm{d}} H(x, u, d, p, t) = \max_{\mathrm{d}} \min_{\mathrm{u}} H(x, u, d, p, t) \tag{33}$$

If the minmax is equal to the maxmin, the resulting optimal Hamiltonian is denoted $H^*(x, p, t)$ and the optimal inputs u^*, d^* are determined to be respectively,

$$u^*(t) = \operatorname*{argmin}_{\mathrm{u}} \left(\max_{\mathrm{d}} H(x, u, d, p, t) \right) \tag{34}$$

and

$$d^*(t) = \operatorname*{argmax}_{\mathrm{d}} \left(\min_{\mathrm{u}} H(x, u, d, p, t) \right) \tag{35}$$

The equations for the state and costate and the transversality conditions are given as before by

$$\begin{aligned} \dot{x} &= \frac{\partial H^*}{\partial p}^T(x, p) \\ \dot{p} &= -\frac{\partial H^*}{\partial x}^T(x, p) \end{aligned} \tag{36}$$

with boundary conditions $x(t_0) = x_0$ and $p^T(t_f) = D_1\phi(x_{t_f})$, and the equation is the familiar *Hamilton Jacobi equation.* As before, one can introduce the optimal cost to go $J^*(x(t), t)$ and we have the following analog of Theorem (1):

Theorem 2 The Hamilton Jacobi Isaacs equation

Consider, the two person zero sum differential game problem of (31) with fixed endpoint t_f. If the optimal value function, i.e. $J^*(x(t_0), t_0)$ is a smooth function of x, t, then $J^*(x, t)$ satisfies the **Hamilton Jacobi Isaacs** *partial differential equation*

$$\frac{\partial J^*}{\partial t}(x, t) = -H^*(x, \frac{\partial J^*}{\partial x}(x, t), t) \tag{37}$$

with boundary conditions given by $J^*(x, t_f) = \phi(x)$ for all x.

Remarks

1. We have dealt with saddle solutions for unconstrained input signals u, d thus far in the development. If the inputs are constrained to lie in sets U, D respectively the saddle solutions can be guaranteed to exist if

$$\min_{u \in U} \max_{d \in D} H(x, u, d, p, t) = \max_{d \in D} \min_{u \in U} H(x, u, d, p, t) \tag{38}$$

   Again, if the input sets are not convex, relaxed controls may be needed to achieve the min-max.

2. The sort of remarks that were made about free endpoint optimal control problems can also be made of games.

3. In our problem formulation for games, we did not include explicit terminal state constraints of the form $\psi(x(t_f), t_f) = 0$. These can be easily included, and we will study this situation in greater detail under the heading of pursuit evasion games.

4. The key point in the theory of dynamical games is that the Legendre transformation of the Lagrangian cost function into the Hamiltonian function converts the solution of the "dynamic" game into a "static" game, where one needs to find a saddle point of the Hamiltonian function $H(x, u, d, p, t)$. This is very much in the spirit of the calculus of variations and optimal control.

4 N person dynamical games

When there are N persons playing a game, many new and interesting new possibilities arise. There is a scenario in which the N agents are non-cooperative, and another in which they cooperate in teams. Of course, if the information available to each of them is different, this makes the solution even more interesting. In this section, we will assume that each of the agents has access to the full state of the system. In this section, we will only discuss non-cooperative solution concepts: first the Nash solution concept and then briefly the Stackelberg solution concept. Cooperative games with total cooperation are simply optimal control problems. If there is cooperation among teams, this can be viewed as a noncooperative game between the teams. When however there are side payments between teams the scope of the problem increases quite considerably.

4.1 Non-cooperative Nash solutions

When there are N players each able to influence the process by controls $u_i \in \mathbb{R}^{n_i}, i = 1, \ldots, N$, modeled by

$$\dot{x} = f(x, u_1, \ldots, u_N, t) \tag{39}$$

and each cost functional (to be minimized) is of the form

$$J_i(u_1(\cdot),\ldots,u_N(\cdot)) = \phi_i(x(t_f),t_f) + \int_{t_0}^{t_f} L_i(x,u_1,\ldots,u_N,t)dt \tag{40}$$

different solution concepts need to be invoked. The simplest non-cooperative solution strategy is a so-called *non-cooperative Nash equilibrium.* A set of controls $u_i^*, i = 1,\ldots,N$ is said to be a Nash strategy if for each player modifying that strategy, and assuming that the others play their Nash strategies, results in an increase in his payoff, that is for $i = 1,\ldots,N$

$$J_i(u_1^*,\ldots,u_i,\ldots,u_N^*) \geq J_i(u_1^*,\ldots,u_i^*,\ldots,u_N^*) \qquad \forall u_i(\cdot) \tag{41}$$

It is important to note that Nash equilibria may not be unique. It is also easy to see that for 2 person zero sum games, a Nash equilibrium is a saddle solution.

As in the previous section on saddle solutions, we can write Hamilton Jacobi equations for Nash equilibria by defining Hamiltonians $H_i(x,u_1,\ldots,u_N,p,t)$ according to

$$H_i(x,u_1,\ldots,u_N,p,t) = L_i(x,u_1,\ldots,u_N) + p^T f(x,u_1,\ldots,u_N,t) \tag{42}$$

The conditions for a Nash equilibrium of equation (41) are there exist $u_i^*(x,p,t)$ such that

$$H_i(x,u_1^*,\ldots,u_i,\ldots,u_N^*,p,t) \geq H_i(x,u_1^*,\ldots,u_i^*,\ldots,u_N^*,p,t) \tag{43}$$

Then, we have N sets of Hamilton Jacobi equations for the N players satisfying the Hamilton Jacobi equations with $H_i^* = H_i^*(x,u_1^*,\ldots,u_N^*,p_i,t)$. Note that we have changed the costate variables to p_i to account for different Hamiltonians and boundary conditions.

$$\begin{aligned} \dot{x} &= \tfrac{\partial H_i^*}{\partial p_i}^T \\ \dot{p}_i &= -\tfrac{\partial H_i^*}{\partial x}^T \end{aligned} \tag{44}$$

with transversality conditions $p_i^T(t_f) = -D_1\phi_i(x(t_f),t_f)$.

4.2 Noncooperative Stackelberg solutions

Stackelberg or hierarchical equilibria refer to noncooperative solutions, where one or more of the players act as leaders. We will illustrate the solution concept for a two person game where player 1 is the leader and player 2 the follower. Once player 1 announces a strategy $u_1^o(\cdot)$, if player 2 is rational he choose his strategy so as to minimize his cost J_2 with the dynamics

$$\dot{x} = f(x,u_1^o,u_2,t) \tag{45}$$

and

$$J_2(u_2) = \phi_2(x(t_f),t_f) + \int_{t_0}^{t_f} L_2(x,u_1^o(t),u_2,t)dt$$

Thus, $u_2^*(u_1^o)$ is chosen to minimize $H_2(x,u_1^o,u_2,p_2,t)$, where p_2 satisfies the differential equation

$$\dot{p}_2 = -\frac{\partial H_2}{\partial x}^T(x,u_1^o,u_2(u_1^o),p,t) \quad p_2(t_f) = D_1^T\phi_2(x(t_f),t_f)$$

In turn the leader chooses his strategy to be that u_1^* which minimizes J_1 subject to the assumption that player 2 will rationally play $u_2^*(u_1^*)$. Thus, the system of equations that he has to solve to minimize J_1 subject to

$$\begin{aligned} \dot{x} &= f(x, u_1, u_2, t) & x(t_0) &= x_0 \\ \dot{p}_2 &= -\frac{\partial H_2}{\partial x}^T(x, u_1^o, u_2(u_1^o), p, t) & p_2(t_f) &= D_1^T \phi_2(x(t_f), t_f) \\ 0 &= D_3 H_2(x, u_1, u_2, p_2, t) \end{aligned} \tag{46}$$

The last equation in (46) is the stationarity condition for minimizing H_2. The optimization problem of the system in (46) is not a standard optimal control in $\mathbb{R}^{2n}$ because there is an equality to be satisfied. Thus, Lagrange multipliers (co-states) taking values in $\mathbb{R}^{2n+n_2}$ for $t \in [t_0, t_f]$ are needed. We will omit the details in these notes.

References

[1] S. S. Sastry, *Lectures in Optimal Control and Dynamic Games*, Notes for the course EECS290A, Advanced Topics in Control Theory, University of California, Berkeley, 1996.

[2] M. Athans and P. Falb, *Optimal Control*, Mc Graw Hill, 1966.

[3] R. Bellman, *Dynamic Programming*, Princeton University Press, 1957.

[4] L. D. Berkovitz, *Optimal Control Theory*, Springer-Verlag, 1974.

[5] A. E. Bryson and Y-C. Ho, *Applied Optimal Control*, Blaisdell Publishing Company, Waltham, 1969.

[6] L. Pontryagin, V. Boltyanskii, R. Gamkrelidze, and E. Mischenko, *The Mathematical Theory of Optimal Processes*, Wiley-New York, 1962.

[7] L. C. Young, *Optimal Control Theory*, Cambridge University Press, 2nd edition, 1980.

[8] D. Kirk, *Optimal Control Theory: An Introduction*, Prentice Hall, 1970.

[9] F. Lewis, *Optimal Control*, Wiley-New York, 1986.

[10] W. Fleming and R. Rishel, *Deterministic and Stochastic Optimal Control*, Springer Verlag, 1975.

[11] V. Arnold, *Mathematical Methods of Classical Mechanics*, Springer Verlag, 2nd edition, 1989.

[12] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 1947.

[13] T. Basar and G. Olsder, *Dynamic Noncooperative Games*, Academic Press, 2nd Edition, 1995.