

CLASSICAL OPTIMAL CONTROL IN CONTINUOUS TIME
WITH APPLICATIONS IN ECONOMICS

by

LINGFEI NI

B.A., California State University, Chico, 2012

B.S., California State University, Chico, 2012

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF ARTS

Department of Economics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Approved by:

Major Professor
Steven P. Cassou

Copyright

LINGFEI NI

2014

Abstract

This report shows the mathematics behind the solution to continuous time optimization problems. It shows how to specify the Hamiltonian function, how to use the Hamiltonian to obtain the optimal conditions for a typical economic optimal control problem and applies these techniques to several optimal control problems commonly encountered in macroeconomics. An appendix shows how to set up the optimal conditions for the case in which the state and co-state variables are both vectors. A second appendix shows how to approach the control situation for a system of optimal control problems where the co-state variable for the first sub-optimal control problem is the state variable for the second sub-optimal control problem.

Table of Contents

Acknowledgements.....	v
Dedication.....	vi
Chapter 1 - Introduction.....	1
Chapter 2 - Finite-Horizon OCP with a Single Pair of State and Co-state Variables.....	4
Chapter 3 - Infinite-Horizon Autonomous OCP with a Single Pair of State and Co-state Variables	9
Chapter 4 - Economic Applications of Autonomous OCPs with a Single Pair of State and Co- state Variables.....	11
Consumer Endowment Problem.....	11
Neoclassical Growth.....	14
Chapter 5 - Concluding Remarks.....	23
Reference	25
Appendix A - Finite-Horizon OCP with State and Co-state Variable Vectors.....	27
Appendix B - A System of OCPs with N Canonical Variables - A Proposal.....	29

Acknowledgements

I would like to take this opportunity to express my deepest gratitude to my major advisor, Dr. Steven P. Cassou, for his guidance and supports in the last two years; for his open mind on my non-conventional research objectives at the beginning stage. Without his toleration and scientist-like curiosity, this report would not even get started. He perfectly presented the American spirit to me. And his academic achievement always has been a major driving force for me during my studying and researching periods at Kansas State University.

I would also like to thank Dr. Yang-Ming Chang for his guidance, and constant encouragements throughout the years. And special thanks to Dr. William Blankenau, Dr. Anson Ho and Dr. Lloyd Thomas for their trusts and supports. And their helps on my various academic questions.

Thank you very much for Dr. Dong Li, Dr. Philip Gayle, Dr. Tracy Tuner, Dr. Michael Babcock and all the economics professors like them here at KSU for your hard work, so we the students, are able to obtain the universal truth we are searching for.

Dedication

I would like to dedicate this paper to god for my lucky survival, and my family for their consistent support on my careers, especially during the darkest moments of my life. And to my first economics professor Dr. Ruben Sargsyan who not only introduced me to economics but also spent significantly amount of time and energy guided me on various economics concepts and problems. To Dr. Michael Perelman and Dr. Frederica Shockley for their supports and helping me kept my life well balanced during my days in California. And to all my professors and friends.

Chapter 1 - Introduction

According to mathematics historian Cantor (1907), the earliest solved optimizing problem is in Euclid's "Elements" which considers, "of all the rectangles contained by the segments of a given straight line, the greatest is the square which is described on half the line" around 300 B.C. quoted from Byrne's classic study on Euclid (1847). It essentially gives the following solution: $\frac{a}{2} \in \arg \max_{x \in \mathbb{R}} \{x(a - x)\}$. After infinitesimal calculus (or later just calculus) was discovered by Newton and Leibniz in the 1670s, Johann Bernoulli from the well-known Bernoulli family of Basel, Switzerland, posed the famous brachistochrone problem to challenge his peers. The problem is to determine the path from point A to B in a vertical plane that minimizes the traveling time under the influence of gravity in between. The proofs were quickly submitted by Galilei (but not solved), Leibniz, de l'Hôpital, Newton, Johann Bernoulli himself and his brother Jakob Bernoulli. Among them, Newton's submission was anonymous, however, according to it was immediately recognized by Johann, "ex ungue leonem" ("one knows the lion by his claw") as translated by Bell (1937). Some of these proofs gave birth to a branch of calculus known as calculus of variations, which is the first sub-field in mathematics that focuses on dynamic optimization¹. Decades later, Johann Bernoulli's star pupil, Leonhard Euler revisited it as a general problem, $J \equiv \int_{t_0}^{t_s} f(x(t), u(t), t) dt$. His work discovered the now called Euler equation as a necessary condition to constrained solutions of dynamic optimization problems.

¹ It is useful to point out the difference between calculus of variations and optimal control theory. Both seek to find the maximum of an objective functional, but they differ in their approach in finding the maximum. For OCT, its maximum condition is Pontryagin's maximum principle - a global condition, while for calculus of variations, the maximum condition is to set the derivative of Hamiltonian function with respect to the control variable equal to zero. This difference is due to the different control set characteristics between the two methods. For calculus of variation, the control set is an open set which does not allow corner solutions for $u^*(t)$, as for OCT, the control set is a close set which allows corner solutions for $u^*(t)$. This means, when there is a corner solution, the global maximality condition (Pontryagin's maximum principle) has to be applied, instead of a local one (set the derivative of Hamiltonian function with respect to the control variable equals to zero). Also, in an optimal control program, the objective function contains a control (choice) variable, whereas in calculus of variation the control variable is replaced by the derivative of the state variable with respect to time.

Scholars in different fields have been using calculus of variations since then, including in the first long lasting example in economics in December 1928. This long lasting example was posited by the great logician, philosopher and economist Frank P. Ramsey in his third mathematic economics paper, “A Mathematical Theory of Saving,” where he applied calculus of variations to determine how much a country needs to save instead of consume in order to maximize its future utility according to Samuelson (1970).

The idea of optimal control theory (OCT) was first initiated during the cold war era. Both the U.S. and U.S.S.R. realized the importance of this type of mathematics for national defense after World War II. According to Pesch and Plail (2009), both countries started research on OCT for issues such as the minimum time interception problems for fighters. Not surprisingly, the two most important innovators in this field, Lev Semyonovich Pontryagin and Richard Bellmen, came from these two rival countries.

Richard E. Bellman worked at RAND on multi-stage decision problems in the early 1950s. His contributions to OCT are the discoveries of the principle of optimality, the Bellmen equation and the Hamilton-Jacobi-Bellmen (HJB) equation. The principle of optimality states that an optimal decision strategy in a multi-stage decision making process is made regardless of the status of the previous decisions. The rest of the decisions going forward have to be optimal choices. The Bellmen equation is a value function solution to a HJB equation, which describes the decision maker’s optimal payoff. And the HJB equation is a generalized partial differential equation (PDE) of the Hamilton-Jacobi equation from classical mechanics, which Bellmen replaced the velocity of the state variable with a control function. These findings were crucial in the development of modern OCT, and serves as the foundation of dynamic programming.

Lev S. Pontryagin, a leading topologist in the U.S.S.R., changed his research interest to OCT at the beginning of the 1950s while he was working on an automatic control problem posted by the military. He noticed the potential of A. A. Fel’dbaum’s work on the famous bang-bang solution and A. J. Lerner’s higher order generalization of the solution. According to Pesch and Plail (2009), Pontryagin’s first achievement is the discovery of the co-vector function which helps to solve a “Bushaw-Fel’dbaum type problem.” As a result, the admissible control set was first introduced. In 1956 a milestone in OCT known as Pontryagin’s maximum principle was presented in his work, "Towards a Theory of Optimal Process". The English version can be found in the book of his collective works published in 1987. We will briefly apply it later,

however, we will not focus on it because it is not easily tractable. This approach has also lost popularity in today's economic education, and it has been replaced by another local optimal condition from calculus of variations instead.

If Pontryagin's paper, "Towards a Theory of Optimal Process," is marked as the beginning of optimal control theory, it has been 58 years until this day. However, its applications in economics is still quite limited. Mostly it can be found in growth theory and sometimes in game theory and mechanism design. Mathematical economics textbooks by and large jump directly into how to use the Hamiltonian function instead of first thoroughly introducing where it comes from. This report offers a more detailed view on two classic types of single paired state and co-state variable OCPs in order to satisfy explore this issue with two economic applications to illustrate the usages.

The next chapter discusses the general method for solving a simple OCP, where there is only one state and one co-state variable. The third chapter investigates the autonomous infinite horizon case, which is the most commonly seen type of OCP in macroeconomics. Appendix A reviews the general solution for a finite-horizon OCP with state and co-state variable vectors and a further expansion can be found in Appendix B. This report considers contributions by the great mathematicians during the last century for almost all variations on OCPs. However, one interesting case is missing. In the Appendix B of this paper, I propose a new kind of OCP, and offer a new mathematical method for solving it. The new OCP is what I would call a system of OCPs, where the co-state variable from the first sub-OCP is the state variable for the second sub-OCP, and so on. The method is what I call the Nambu Hamiltonian method and is named after the founder of Nambu mechanics. However, its application in economics still needs to be discovered.

Chapter 2 - Finite-Horizon OCP with a Single Pair of State and Co-state Variables

There are three perspectives on how to interpret an optimal control problem. The first believes that a policy maker controls the control variable, within its control region, to affect the state variable through an ordinary differential equation (ODE) called the law of motion. This process eventually achieves the targeted value of the objective functional structured by the terminal value of the state variable. However, in reality, the targeted value is not always achievable. This brings up the second interpretation. The policy maker shifts their attention to the progress for optimizing an objective functional, while the objective functional is integrated as an overall payoff over a particular time period. In another words, the policy maker focuses on the input trajectory instead. The third one is a combination of the previous two views, where the objective functional has an integrated part and a terminal value part. The following finite-horizon OCP with one state and one control variables is one of the simplest OCPs in theory. The following content demonstrates the process of how a Hamiltonian function is set up; what a Hamiltonian function is; and how to obtain a set of optimal conditions for the second perspective. Considering the following OCP given by Sydsæter, Strøm, and Berck (2005):

Case 1

$$J(u(t)) \equiv \int_{t_0}^{t_s} f(x(t), u(t), t) dt \rightarrow \max_{u(t) \in \{U\}}$$

$$s.t.: \dot{x}(t) = g(x(t), u(t), t)$$

$$x(t_0) = x^0$$

$$x(t_s) \text{ free}$$

$$u(t) \in \{U\}$$

where, $J(u(t)) \equiv \int_{t_0}^{t_s} f(x(t), u(t), t) dt$ is the objective functional in its integrated form over time period $[t_0, t_s]$, $x(t)$ is a state variable from a state variable set $\{X\} \subseteq \mathbb{R}$, $\{U\}$ is an admissible control set (control region), admissible control function $u(t) \in \{U\} \subseteq \mathbb{R}$. Time set $\{T\} = \{t \mid t \in [t_0, t_s]\}$. $x(t_0) = x^0$; $g(\cdot)$ is C^∞ with respect to x and t . This is the simplest OCP, which asks us to find the optimal control path $u^*(t)$ to maximize the objective functional $f(\cdot)$.

Meanwhile, it generates the optimal time path $x^*(t)$ to form an optimal pair (u^*, x^*) we are looking for.

Let's start by assume $V(x(t), t) \in \arg \sup_{u[t, t+dt]} \left\{ f(x(t), u(t), t) + \frac{\partial V(x, t)}{\partial x} g(x(t), u(t), t) \right\}$

² is unique on $\{[t_0, t_s]\} \times \{X\}$. The following is the reason why we can make such an assumption.

Proof 1.

\because For $V(\chi(t), \tau) \in \arg \sup_{u[t, t+dt]} \left\{ f(x(t), u(t), t) + \frac{\partial V(x, t)}{\partial x} g(x(t), u(t), t) \right\}$

when $t_0 = 0$,

\exists' initial data $(\chi(t), \tau) \in \{X\} \times \{(t_0, t_s)\}$; and $V(\chi(t), \tau) \in \{X\} \times \{[t_0, t_s]\}$

Also, $V(x(t), t)$ is continuous on $\{X\} \times \{(t_0, t_s)\}$

$\therefore \forall V(x(t), t) \in \arg \sup_{u[t, t+dt]} \left\{ f(x(t), u(t), t) + \frac{\partial V(x, t)}{\partial x} g(x(t), u(t), t) \right\}$ is unique on $\{[t_0, t_s]\} \times$

$\{X\}$.

■

Hence, we can assume there exists a unique optimal value function for a Hamilton-Jacobi-Bellman (HJB) PDE of this dynamical system. What should be pointed out is that, this solution assumes the unique existence of a particular value function (optimal function, the fastest value dropping function) that we choose from the other also unique functions which solve the problem. Again, there can be more than one solutions to a HJB equation, and all of them shell be unique in this case. This is important to the potential possibility of solving the problem in Appendix B.

Let value function $V: \{[t_0, t_s]\} \times \{X\} \rightarrow \mathbb{R}$ be a C^∞ function which satisfies a HJB equation. According to the principle of optimality, from t to $t + dt$, we have "Bellman equation":

$$V(x(t), t) = \sup_{u[t, t+dt]} \{f(x(t), u(t), t)dt + V(x(t + dt), t + dt)\}$$

This value function represents the optimal payoff of the system over $[t, t_s]$ period, when started at t with state variable level $x(t) = x$. Now loosely followed Dr. Ian Mitchell's class notes on Dynamic Programming and Approximate Dynamic Programming, and apply Taylor expansion on $(x(t + dt), t + dt)$:

² $f(x(t), u(t), t)$ is the instantaneous payoff at t ; $g(x(t), u(t), t)$ is the system velocity at t ; $\frac{\partial V(x, t)}{\partial x}$ is the co-state variable, it represents the per-unit value of the system velocity at t .

$$\begin{aligned}
V(x(t+dt), t+dt) &= V(x(t), t) + \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial x} dx + O(x, t) \\
&= V(x(t), t) + \frac{\partial V(x, t)}{\partial t} dt + \frac{\partial V(x, t)}{\partial x} \dot{x}(t) dt + O(x, t) \\
&\approx V(x(t), t) + \frac{\partial V(x, t)}{\partial t} dt + \frac{\partial V(x, t)}{\partial x} \dot{x}(t) dt
\end{aligned}$$

where, $O(x, t) := \{\text{Derivatives of 2nd order and above}\}$. Combine the above two value functions:

$$V(x(t), t) = \sup_{u[t, t+dt]} \left\{ f(x(t), u(t), t) dt + V(x(t), t) + \frac{\partial V(x, t)}{\partial t} dt + \frac{\partial V(x, t)}{\partial x} \dot{x}(t) dt \right\}$$

Cancel $V(x(t), t)$ from both sides:

$$0 = \sup_{u[t, t+dt]} \left\{ f(x(t), u(t), t) dt + \frac{\partial V(x, t)}{\partial t} dt + \frac{\partial V(x, t)}{\partial x} \dot{x}(t) dt \right\}$$

Divide dt on both sides, and we can get the HJB equation (01):

$$\begin{aligned}
0 &= \sup_{u[t, t+dt]} \left\{ f(x(t), u(t), t) + \frac{\partial V(x, t)}{\partial t} + \frac{\partial V(x, t)}{\partial x} \dot{x}(t) \right\} \\
0 &= \sup_{u[t, t+dt]} \left\{ f(x(t), u(t), t) + \frac{\partial V(x, t)}{\partial t} + \frac{\partial V(x, t)}{\partial x} g(x(t), u(t), t) \right\} \\
-\frac{\partial V(x, t)}{\partial t} &= \sup_{u[t, t+dt]} \left\{ f(x(t), u(t), t) + \frac{\partial V(x, t)}{\partial x} g(x(t), u(t), t) \right\} \quad (01)
\end{aligned}$$

where, $(\forall (x, t) \in \{X\} \times \{[t_0, t_s]\})$. The reason for the last step to take out $\frac{\partial V(x, t)}{\partial t}$ is that the supremum is only on $u(t)$ at t , and $V(x, t)$ is not a functional of $u(t)$.

From Eq. 01, we define the Hamiltonian function, $H: \{[t_0, t_s]\} \times \{X\} \times \{U\} \times \mathbb{R} \rightarrow \mathbb{R}$ such that:

$$(t, x, u, z) \mapsto H(t, x, u, z) = f(x(t), u(t), t) + z(t)g(x(t), u(t), t) \quad (02)$$

where, the co-state variable (also known as adjoint variable) $z(t) := \frac{\partial V(x, t)}{\partial x}$ (03)

Hence, we have: $\frac{\partial V(x, t)}{\partial t} = \inf_{u[t, t+dt]} H(t, x, u, z)$.

As we can tell from Eq. (01), Hamiltonian function indicates that there is a negative co-relationship between the changing of the value function with respect to time and the (value of) instantaneous payoff plus the value of the system velocity at time t .

Hence, we can obtain Hamiltonian equations of motion which consist two first order PDEs for this dynamic system (with a 2-dimensional xz phase plane):

$$\left\{ \begin{array}{l} \frac{\partial H(z(t), x^*(t), u^*(t), t)}{\partial z} = g(x(t), u(t), t) = \dot{x}(t) \quad (04) \\ \dot{z}(t) := \frac{\partial}{\partial t} \left(\frac{\partial V(x, t)}{\partial x} \right) = \frac{\partial}{\partial x} \left(\frac{\partial V(x, t)}{\partial t} \right) = - \frac{\partial H(z(t), x^*(t), u^*(t), t)}{\partial x} \quad (05) \end{array} \right.$$

Eq. (04) is the state function, and Eq. (05) is the co-state function or so called Euler equation.

According to Takeyama (2009) with Pontryagin's Maximum Principle (PMP), we can draw the following conclusions:

For, $u(t) \in \{U\}$,

$$\{ H(z(t), x^*(t), u^*(t), t) \geq H(z(t), x^*(t), u(t), t) \text{ for a maximum} \quad (06)$$

$$\{ H(z(t), x^*(t), u^*(t), t) \leq H(z(t), x^*(t), u(t), t) \text{ for a minimum} \quad (07)$$

Since this is a free end problem, hence, a transversality condition is needed. If we lack both a boundary condition and a transversality condition to leave either the initial point or the end point free, then, the original Euler equation as a second order nonlinear differential equation is unsolvable³. Setting Lagrangian for the objective functional following Woodward (2013):

$$\mathcal{L} = \int_{t_0}^{t_s} f(x(t), u(t), t) dt + z(t)[g(x(t), u(t), t) - \dot{x}(t)]$$

Since, the constraint has to hold at all time, hence:

$$\mathcal{L} = \int_{t_0}^{t_s} \{ f(x(t), u(t), t) + z(t)[g(x(t), u(t), t) - \dot{x}(t)] \} dt$$

∴ Assume $t \in \{[t_0, t_s]\}$ is at the optimum.

$$\therefore \frac{\partial \mathcal{L}}{\partial z} = 0$$

$$\therefore \mathcal{L}^* = \int_{t_0}^{t_s} f(x(t), u(t), t) dt = J^*$$

∴ We can write: $\mathcal{L}^* = J^* = \int_{t_0}^{t_s} [f(x(t), u(t), t) + z(t)g(x(t), u(t), t) - z(t)\dot{x}(t)] dt$

$$= \int_{t_0}^{t_s} [H(t, x, u, z) - z(t)\dot{x}(t)] dt$$

$$= \int_{t_0}^{t_s} H(t, x, u, z) dt - \int_{t_0}^{t_s} z(t)\dot{x}(t) dt$$

$$= \int_{t_0}^{t_s} H(t, x, u, z) dt - [z(t)x(t)]_{t_0}^{t_s} + \int_{t_0}^{t_s} x(t)\dot{z}(t) dt$$

³ Even though, in macroeconomics we usually get a first order PDE instead.

$$\begin{aligned}
&= \int_{t_0}^{t_s} [H(t, x, u, z) + x(t)\dot{z}(t)]dt - [z(t_s)x(t_s) - z(t_0)x(t_0)] \\
&= \int_{t_0}^{t_s} [H(t, x, u, z) + x(t)\dot{z}(t)]dt + z(t_0)x(t_0) - z(t_s)x(t_s)
\end{aligned}$$

Since, the differentiable functional J is assumed at its optimal, hence, the first order necessary condition for the functional $\frac{\partial J^*}{\partial x(t_s)} \equiv 0$. So, $\frac{\partial J^*}{\partial x(t_s)} = -z(t_s) = 0$. Along with Eq. (04) (05) (06) and the boundary condition, we can structure the following necessary conditions to solve for the optimal pair of the optimal control problem:

For a maximum J , the following conditions have to be satisfied:

$$\left\{ \begin{array}{l} z(t_s) = 0 \\ x(t_0) = x^0 \\ \dot{x}(t) = \frac{\partial H(z(t), x^*(t), u^*(t), t)}{\partial z} \\ \dot{z}(t) = -\frac{\partial H(z(t), x^*(t), u^*(t), t)}{\partial x} \\ H(z(t), x^*(t), u^*(t), t) \geq H(z(t), x^*(t), u(t), t) \end{array} \right.$$

Instead of applying PMP, if H is differentiable with respect to u , more likely we will have the following simpler case, where, $\frac{\partial H^*}{\partial u^*} = 0$.

$$\left\{ \begin{array}{l} z(t_s) = 0 \\ x(t_0) = x^0 \\ \dot{x}(t) = \frac{\partial H(z(t), x^*(t), u^*(t), t)}{\partial z} \\ \dot{z}(t) = -\frac{\partial H(z(t), x^*(t), u^*(t), t)}{\partial x} \\ \frac{\partial H^*}{\partial u^*} = 0 \end{array} \right.$$

This change of optimal value condition makes the optimal solution changes from a global optimal of H (when chooses $u^*(t)$) to a local optimal. However, in this way, the "mathematical beauty" and the easiness to solve are fully presented. Also, if the boundary condition: $x(t_s) = x^{t_s}$ is added, the boundary equations becomes a set of Dirichlet boundary conditions, and the problem becomes a fixed end point problem, which does not need the transversality condition any more.

Chapter 3 - Infinite-Horizon Autonomous OCP with a Single Pair of State and Co-state Variables

In order to show the variety of different kinds of OCPs, the OCP in this section is almost the “opposite” of the previous case. It is an autonomous OCP (time is not a variable for function f , even though it appears in the discount term) with infinite horizon (time goes to infinity), which leads to a Current-Value Hamiltonian function.

Given the following autonomous infinite-horizon OCP mentioned by Cui in his book (2008):

Case 2

$$J(u(t)) \equiv \int_{t_0}^{\infty} e^{-\rho t} f(x(t), u(t)) dt \rightarrow \max_{u(t) \in \{U\}}$$

$$s.t.: \dot{x}(t) = g(x(t), u(t))$$

$$x(t_0) = x^0$$

$$u(t) \in \{U\}$$

where, $x(t)$ is a state variable from a state variable set $\{X\} \subseteq \mathbb{R}$; $\{U\}$ is a admissible control set; admissible control function $u(t) \in \{U\} \subseteq \mathbb{R}$. $\{T\} := \{t \mid t \in [t_0, \infty)\}$. $x(t_0) = x^0$; $g(\cdot)$ is C^∞ with respect to x and t . $e^{-\rho t}$ is the discount factor, where, $\rho \in (0,1)$ is the discount rate.

This case can be solved by applying the method from Case 1, however, if we assume its Hamiltonian is C^∞ with respect to x and u , then the discount factor may complicate the issue. Hence, usually when dealing with an OCP with discount factor, we can choose to set up a Current-Value Hamiltonian instead. By applying the same method in case 1, and taking the autonomous situation into consideration. We obtain the following Hamiltonian functions $H: \{X\} \times \{U\} \times \mathbb{R} \rightarrow \mathbb{R}$ such that:

$$(x, u, z) \mapsto H(x, u, z) = e^{-\rho t} f(x(t), u(t)) + z(t)g(x(t), u(t))$$

where, the co-state variable $z(t) := \frac{\partial V(x)}{\partial x}$

The Current-Value Hamiltonian function is defined as:

$$\tilde{H}(x, u, \mu) := e^{\rho t} H = f(x(t), u(t)) + \mu(t)g(x(t), u(t))$$

where, $\mu(t) = e^{\rho t} z(t)$. As we can tell "Current-Value Hamiltonian" by its name means no future discount factor is needed.

$$\because \mu(t) = e^{\rho t} z(t)$$

$$\because \dot{\mu}(t) = e^{\rho t} \dot{z}(t) + \rho e^{\rho t} z(t)$$

$$\because \text{From case 1: } \dot{z}(t) = -\frac{\partial H}{\partial x} = -e^{-\rho t} \frac{\partial \tilde{H}}{\partial x}$$

$$\begin{aligned} \therefore \dot{\mu}(t) &= -e^{\rho t} e^{-\rho t} \frac{\partial \tilde{H}}{\partial x} + \rho e^{\rho t} z(t) \\ &= -\frac{\partial \tilde{H}}{\partial x} + \rho \mu \end{aligned}$$

Hence, Hamiltonian equations of motion for this case:

$$\begin{cases} \dot{x}(t) = g(x(t), u(t)) = \frac{\partial \tilde{H}(\mu(t), x^*(t), u^*(t))}{\partial \mu} \\ \dot{\mu}(t) = \rho \mu - \frac{\partial \tilde{H}(\mu(t), x^*(t), u^*(t))}{\partial x} \end{cases}$$

Because this is an infinite horizon model, a transversality condition is needed to fill in as a terminal condition:

$$\lim_{t \rightarrow \infty} x^*(t) = \bar{x}$$

As time goes to infinity, we assume there exists a stable state for x^* which stays at \bar{x} . So \bar{x} equivalent to the terminal value in the finite-horizon case.

Therefore, we have the following optimal conditions:

$$\begin{cases} x(t_0) = x^0 \\ \lim_{t \rightarrow \infty} x^*(t) = \bar{x} \\ \dot{x}(t) = \frac{\partial \tilde{H}(\mu(t), x^*(t), u^*(t))}{\partial \mu} \\ \dot{\mu}(t) = \rho \mu - \frac{\partial \tilde{H}(\mu(t), x^*(t), u^*(t))}{\partial x} \\ \frac{\partial \tilde{H}(\mu(t), x^*(t), u^*(t))}{\partial u^*} = 0 \end{cases}$$

Chapter 4 - Economic Applications of Autonomous OCPs with a Single Pair of State and Co-state Variables

Bellman in his book Dynamic Programming first applied OCT to economic problems. Therefore, it is not an exaggeration to say that the building of OCT considered usage in economics from the beginning. The fifth chapter of Bellman's book (1957) discusses the problem of stocking a supply of an item for an uncertain future, where there are costs associated with the future undersupply and oversupply. And later, many problems in economic growth which were originally solved by calculus of variations were converted into OCPs. Most time in macroeconomic OCPs appears in autonomous situations. We focus our examples in this area. The first case is a consumer endowment problem, and the second one is a popular demonstration of OCP in growth theory.

Consumer Endowment Problem

Here we focus on solving a common consumer endowment problem. This is a problem with an explicit objective functional, in which we can actually solve for its optimal path. Considering the following case modified from Turkington's textbook homework question (2008).

A consumer receives his income through his wealth, $Y = rW$, where Y is income, r is the interest rate and W is wealth. The income is used in consumption or investment and investment increases his wealth according to $\dot{W} = Y - C = rW - C$ where \dot{W} is the rate of change in wealth and C is consumption. The initial endowment at time t_0 is given by $W(t_0) = W_0$. The consumer will live until $t = t_s$ and will use up all their wealth and no debts are allowed at time t_s . The consumer gains utility only through consumption and has utility function given by $U(C(t)) = \log C(t)$ for $C(t) > 0$. The consumer wishes to maximize life-time utility given by $\int_{t_0}^{t_s} e^{-\rho t} U(C(t)) dt$, where ρ is the consumer's time preference. The objective is to find the consumer's optimal consumption path.

First, we determine the objective functional, and constraints from the given scenario to set up the OCP:

$$U(C(t)) = \int_{t_0}^{t_s} e^{-\rho t} U(C(t)) dt = \int_{t_0}^{t_s} e^{-\rho t} \log C(t) dt \rightarrow \max_{C(t) \in \{C\}}$$

$$s.t.: \dot{W}(t) = rW(t) - C(t)$$

$$W(t_0) = W_0$$

$$W(t_s) = 0$$

$$C(t) \in \{C\}$$

where, $W(t)$ is the state variable, $C(t)$ is the control variable.

Assume, $\exists' V(W(t), t) \in \arg \max_{C[t, t+dt]} \left\{ e^{-\rho t} U(C(t)) + \frac{\partial V(W(t), t)}{\partial W} [rW(t) - C(t)] \right\}$ on $\{[t_0, t_s]\} \times \{W\}$. And $V(\cdot)$ is a C^∞ function which satisfies a HJB equation. Hence, we can yield a "Bellman equation" for the problem:

$$V(W(t), t) = \max_{C[t, t+dt]} \left\{ e^{-\rho t} U(C(t)) dt + V(W(t+dt), t+dt) \right\}$$

Then, apply Taylor expansion on $V(W(t+dt), t+dt)$:

$$\begin{aligned} V(W(t+dt), t+dt) &= V(W(t), t) + \frac{\partial V(W, t)}{\partial t} dt + \frac{\partial V(W, t)}{\partial W} dW + O(dt) \\ &\approx V(W(t), t) + \frac{\partial V(W, t)}{\partial t} dt + \frac{\partial V(W, t)}{\partial W} \dot{W}(t) dt \end{aligned}$$

Combine the above two value functions:

$$V(W(t), t) = \max_{C[t, t+dt]} \left\{ e^{-\rho t} U(C(t)) dt + V(W(t), t) + \frac{\partial V(W, t)}{\partial t} dt + \frac{\partial V(W, t)}{\partial W} \dot{W}(t) dt \right\}$$

After cancelling $V(x(t), t)$ from both sides, dividing by dt on both sides, plug in the law of motion to get the following HJB equation:

$$-\frac{\partial V(W, t)}{\partial t} = \max_{C[t, t+dt]} \left\{ e^{-\rho t} U(C(t)) + \frac{\partial V(W, t)}{\partial W} [rW(t) - C(t)] \right\}$$

where, $(\forall (W, t) \in \{W\} \times \{[t_0, t_s]\})$

From the above, we define the Hamiltonian function, $H: \{W\} \times \{C\} \times \mathbb{R} \rightarrow \mathbb{R}$ such that:

$$(W, C, z) \mapsto H(W, C, z) = e^{-\rho t} U(C(t)) + z(t)[rW(t) - C(t)] \quad (08)$$

where, the co-state variable $z(t) := \frac{\partial V(W, t)}{\partial W}$.

The next step is to set the optimal conditions, which includes the initial and terminal conditions, the maximality condition and a set of Hamiltonian equations of motion:

$$\left\{ \begin{array}{l} W(t_0) = W_0 \quad (09) \\ W(t_s) = 0 \quad (10) \\ \dot{z}(t) = -\frac{\partial H}{\partial W} = -rz(t) \quad (11) \\ \dot{W}(t) = \frac{\partial H}{\partial z} = rW(t) - C(t) \quad (12) \\ \frac{\partial H^*}{\partial C^*} = 0 = C(t)^{-1}e^{-\rho t} - z(t) \quad (13) \end{array} \right.$$

From Eq. (13), we have

$$C(t) = e^{-\rho t} z(t)^{-1} \quad (14)$$

Plug Eq. (14) into Eq. (12):

$$\dot{W}(t) = rW(t) - e^{-\rho t} z(t)^{-1} \quad (15)$$

This equation along with the solution of $z(t)$ can help to solve the optimal path for $W(t)$ if we choose to. The key to the solution is to find the optimal path for $z(t)$, hence, solving equation (11):

$$\dot{z}(t) = -rz(t)$$

$$e^{rt} [\dot{z}(t) + rz(t)] = 0$$

$$e^{rt} \frac{dz(t)}{dt} + \frac{de^{rt}}{dt} z(t) = 0$$

$$\frac{d[e^{rt} z(t)]}{dt} = 0$$

$$\int \frac{d[e^{rt} z(t)]}{dt} dt = \int 0 dt$$

$$e^{rt} z(t) = c_1$$

$$z(t) = e^{-rt} c_1$$

Plugging this result into Eq. (14),

$$C(t) = e^{-\rho t} z(t)^{-1} = \frac{e^{(r-\rho)t}}{c_1}$$

$$\text{When, } t = 0, c_1 = C(0)^{-1}$$

Hence, the general solution for the optimal consumption path for the consumer is:

$$C(t)^* = e^{(r-\rho)t} C(0)$$

Given the optimal trajectory of $C(t)^*$, we shall be able to describe the initial consumption by the initial wealth endowment. Therefore, the lifetime discounted consumption is determined by the initial wealth endowment:

$$\begin{aligned}
W_0 &= \int_{t_0}^{t_s} e^{-\rho t} C(t)^* dt \\
&= \int_{t_0}^{t_s} e^{-\rho t} e^{(r-\rho)t} C(0) dt \\
&= \int_{t_0}^{t_s} e^{r-2\rho t} C(0) dt \\
&= \int_{t_0}^{t_s} e^{r-2\rho t} dt C(0)
\end{aligned}$$

$$C(0) = \frac{W_0}{\int_{t_0}^{t_s} e^{r-2\rho t} dt}$$

Therefore, we obtain the following optimal path with the initial wealth endowment:

$$C(t)^* = e^{(r-\rho)t} \frac{W_0}{\int_{t_0}^{t_s} e^{r-2\rho t} dt}$$

Neoclassical Growth

Considering a neoclassical growth model for a simple closed economy. Its production function is the following:

$$Y = F(L(t), K(t))$$

where, Y is the real output, L and K are labor and capital input respectively. $F: \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ is an at least C^2 homogeneous function. Assume “Inada condition” holds:

$$\left\{ \begin{array}{l}
F(0,0) = 0 \\
L(0) = L^0 \\
K(0) = K^0 \\
F(\cdot) \text{ is } C^\infty \\
F_L \in (0, \infty) \\
F_K \in (0, \infty) \\
F_{LL} \in (-\infty, 0) \\
F_{KK} \in (-\infty, 0) \\
\lim_{L \rightarrow 0} F'(L) \rightarrow \infty \\
\lim_{L \rightarrow \infty} F'(L) \rightarrow 0 \\
\lim_{K \rightarrow 0} F'(K) \rightarrow \infty \\
\lim_{K \rightarrow \infty} F'(K) \rightarrow 0
\end{array} \right.$$

Rewrite this strictly concave production function into its per effective labor form:

$$y = \phi(k) \quad (16)$$

where, per effective labor output $y = Y/L$; and per effective labor capital input $k = K/L$. Naturally, we have the following strict concavity properties for per effective labor output function:

$$\begin{cases} \phi'(k) \in (0, \infty) \\ \phi''(k) \in (-\infty, 0) \end{cases}$$

Assume the aggregate income Y (which equals to the aggregate output in this simple closed economy) is distributed between consumption and investment only, hence, we get the following:

$$Y - C - I = 0$$

$$Y - C - \delta K = I - \delta K$$

The changing to capital stock over time can be defined as the difference between inventory and the depreciated capital:

$$\frac{dK}{dt} = I - \delta K$$

where, depreciation rate $\delta \in [0,1]$. Also, given this is a closed economy, we have:

$$\frac{dK}{dt} = Y - C - \delta K$$

in per effective labor form:

$$\frac{1}{L} \frac{dK}{dt} = y - c - \delta k \quad (17)$$

Assume, the population and the labor force are identical. Define, the population growth rate n as:

$$n = \frac{dL}{dt} / L$$

$$\because K = kL$$

$$\therefore \frac{dK}{dt} = \frac{d(kL)}{dt} = k \frac{dL}{dt} + L \frac{dk}{dt} = knL + L \frac{dk}{dt} \quad (18)$$

Plug Eq. (16) (18) into (17) we have the complete law of motion for capital accumulation:

$$\frac{1}{L} \left(knL + L \frac{dk}{dt} \right) = \phi(k) - c - \delta k$$

$$\frac{dk}{dt} = \phi(k) - c - (\delta + n)k$$

Given a per capita social welfare function $U(c)$ with the following “Inada conditions”:

$$\begin{cases} U(0) = 0 \\ U(c) \text{ is } C^\infty \\ U'(c) \in (0, \infty) \\ U''(c) \in (-\infty, 0) \\ \lim_{c \rightarrow 0} U'(c) \rightarrow \infty \\ \lim_{c \rightarrow \infty} U'(c) \rightarrow 0 \end{cases}$$

Hence, the per capita social welfare function is concave.

According to Chiang and Wainwright (2005), if denotes ρ as the social discount rate and normalizes the initial population L_0 to one. We obtain the following objective functional to present value of a future utility flow for the society:

$$U(c(t)) = \int_{t_0}^{\infty} e^{-\rho t} U(c(t)) L_0 e^{nt} dt = \int_{t_0}^{\infty} e^{(n-\rho)t} U(c(t)) dt = \int_{t_0}^{\infty} e^{-rt} U(c(t)) dt$$

where, $r = \rho - n$.

Hence, from the setting above, like case 2, the optimal growth problem is:

$$U(c(t)) = \int_{t_0}^{\infty} e^{-rt} U(c(t)) dt \rightarrow \max_{c(t) \in \{C\}}$$

$$\text{s.t.: } \frac{dk}{dt} = \phi(k) - c - (\delta + n)k$$

$$k(0) = k^0$$

$$c(t) \in [0, \phi(k)]$$

where, k is the state variable, and c is the control variable from admissible control set $\{C\}$.

We obtain the following Hamiltonian function $H: \mathbb{R}_+^3 \rightarrow \mathbb{R}$ such that:

$$(k, c, z) \mapsto H(k, c, z) = e^{-rt} U(c(t)) + z(t) [\phi(k) - c - (\delta + n)k]$$

where, $z(t)$ is the co-state variable.

The Current-Value Hamiltonian function is defined as:

$$\tilde{H}(k, c, \mu) := e^{rt} H = U(c(t)) + \mu(t) [\phi(k) - c - (\delta + n)k]$$

where, $\mu(t) = e^{rt} z(t)$.

$$\because \mu(t) = e^{rt} z(t)$$

$$\because \dot{\mu}(t) = e^{rt} \dot{z}(t) + r e^{rt} z(t)$$

$$\because \text{From case 1: } \dot{z}(t) = -\frac{\partial H}{\partial k} = -e^{-rt} \frac{\partial \tilde{H}}{\partial k}$$

$$\because \dot{\mu}(t) = -e^{rt} e^{-rt} \frac{\partial \tilde{H}}{\partial k} + r e^{rt} z(t)$$

$$= -\frac{\partial \tilde{H}}{\partial k} + r\mu$$

Hence, Hamiltonian equations of motion for this case:

$$\left\{ \begin{array}{l} \dot{k}(t) = \phi(k) - c - (\delta + n)k = \frac{\partial \tilde{H}(\mu(t), k^*(t), c^*(t))}{\partial \mu} \\ \dot{\mu}(t) = r\mu - \frac{\partial \tilde{H}(\mu(t), k^*(t), c^*(t))}{\partial k} = r\mu - \mu \frac{\partial \phi}{\partial k} + \mu(\delta + n) = \mu \left[(r + \delta + n) - \frac{\partial \phi}{\partial k} \right] \end{array} \right.$$

For its transversality condition:

$$\lim_{t \rightarrow \infty} k^*(t) = \bar{k}$$

when time goes to infinity we assume exist a stable status that the optimal per effective labor capital input k^* stays at a constant level \bar{k} . This equivalent to the terminal value in the finite-horizon case.

Therefore, we have the following optimal conditions:

$$\left\{ \begin{array}{l} k(t_0) = k^0 \\ \lim_{t \rightarrow \infty} k^*(t) = \bar{k} \quad (19) \\ \dot{k}(t) = \phi(k) - c - (\delta + n)k = \frac{\partial \tilde{H}(\mu(t), k^*(t), c^*(t))}{\partial \mu} \quad (20) \\ \dot{\mu}(t) = r\mu - \frac{\partial \tilde{H}(\mu(t), k^*(t), c^*(t))}{\partial k} = \mu \left[(r + \delta + n) - \frac{\partial \phi}{\partial k} \right] \quad (21) \\ \frac{\partial \tilde{H}(\mu(t), k^*(t), c^*(t))}{\partial c^*} = 0 = \frac{dU}{dc} - \mu \quad (22) \end{array} \right.$$

Up until this point, the optimal conditions have been set up for the illustration purpose for the previous chapter. In order to gaining more insights of this model, we need to structure a two dimensional phase plane for it. First, we need to determine which two variables we would like to choose to build it. We have three options here, state variable k , co-state variable μ and control variable c . Traditionally, for a growth model we chose k and c . However, for the discussions in Appendix B, although there is no phase diagram is drawn (impossible to draw an $n \times N$ dimensional phase hyper-space anyway) please notice that the discussion was on the state and co-state variables space instead. In order to eliminate μ , we replace it through the maximality condition from Eq. (22):

$$\frac{dU}{dc(t)} = \mu(t) \quad (23)$$

Differentiate it with respect to t , yields:

$$\dot{\mu} = \frac{d^2U}{dc^2} \dot{c} \quad (24)$$

Plug Eq. (23) and (24) into Eq. (21):

$$\begin{aligned} \frac{d^2U}{dc^2} \dot{c} &= \frac{dU}{dc} \left[(r + \delta + n) - \frac{\partial \phi}{\partial k} \right] \\ \dot{c}(t) &= \frac{\frac{dU}{dc}}{\frac{d^2U}{dc^2}} \left[(r + \delta + n) - \frac{\partial \phi}{\partial k} \right] \quad (25) \end{aligned}$$

where, $1/\left(\frac{dU}{dc}\right)$ is the marginal utility elasticity. If we divided the equation by c :

$$\frac{\dot{c}(t)}{c} = -\frac{\frac{dU}{dc}}{\frac{d^2U}{dc^2} c} \left[\left(\frac{\partial \phi}{\partial k} - \delta - n \right) - r \right]$$

This indicates that the consumption growth rate has a positive co-relationship with the difference between the net marginal output and the time preference rate. Hence, if we look at this equation:

$$\dot{c}(t)c = -\frac{\frac{dU}{dc} c}{\frac{d^2U}{dc^2}} \left[\left(\frac{\partial \phi}{\partial k} - \delta - n \right) - r \right]$$

For the same time preference, the higher net marginal output the more worthy to wait for the higher future consumption. Hence, consumption growth rate $\frac{\dot{c}(t)}{c} \in (0, \infty)$.

To continue the construction of the phase diagram. Gather Eq. (25) along with Eq. (20) the Keynes-Ramsey rule, we have a system of autonomous differential equations:

$$\begin{cases} \dot{k}(t) = \phi(k) - c - (\delta + n)k & (20) \\ \dot{c}(t) = \frac{\frac{dU}{dc}}{\frac{d^2U}{dc^2}} \left[(r + \delta + n) - \frac{\partial \phi}{\partial k} \right] & (25) \end{cases}$$

Since we have chosen the axis as k and c . The next step is to write out the isoclines (zero derivative curve).

When $\dot{k}(t) = 0$ and $\dot{c}(t) = 0$ (Per capita social welfare function $U(c)$ follows: $U'(c) \in (0, \infty)$):

$$\begin{cases} c = \phi(k) - (\delta + n)k & (26) \\ \frac{\partial \phi}{\partial k} = r + \delta + n & (27) \end{cases}$$

Since we assumed concavity properties for per effective labor output function, hence, the $\dot{k}(t) = 0$ curve (Eq. (26)) is strictly concave. For the same reason, since the slope on the per effective labor output curve ($\phi(k)$) is unique for each value of k , hence, there is only one unique k can satisfy Eq. (27). And it is clear that $\frac{\partial \phi}{\partial k}$ is not a function of c , hence, Eq. (27) is represented as a vertical line in the kc phase plane (let k be the horizontal axis). In terms of the relative position between $\dot{k}(t) = 0$ and $\dot{c}(t) = 0$ curves, if we take the derivative of c with respect to k for Eq. (26), and from its first order condition, we have the golden rule:

$$\phi'(k_{GR}) = \delta + n \quad (19)$$

Compare it with Eq. (27):

$$\phi'(k_{GR}) < \phi'(k)$$

Since we have the following assumptions:

$$\left\{ \begin{array}{l} \phi(0,0) = 0 \\ \phi'(k) \in (0, \infty) \\ \phi''(k) \in (-\infty, 0) \\ \lim_{k \rightarrow 0} \phi'(k) \rightarrow \infty \\ \lim_{k \rightarrow \infty} \phi'(k) \rightarrow 0 \end{array} \right.$$

Hence, $k_{GR} > k$. i.e. $\dot{c}(t) = 0$ curve is to the left of the golden rule capital level.

Since the isoclines intersect and divide the plane into four quadrants. The next step is to determine the general motion laws of the trajectories in these quadrants. Mathematically, the two sides of an isocline have the opposite directions. In another words, in our case, in terms of isocline $\dot{k}(t) = 0$, on one side of it, $\dot{k}(t)$ is positive; on the other side it is negative.

When $\dot{k}(t) = 0$, from Eq. (26) $c = \phi(k) - (\delta + n)k$. Let c^{**} denote an arbitrary capital level given a specific k . And c^* be the corresponding consumption level for the given k from the $\dot{k}(t) = 0$ curve.

When $c^{**} < c^*$, then, for the next period: $k^{**} > k^*$

$$\therefore \phi(k^{**}) > \phi(k^*) = c^* + (\delta + n)k^*$$

$$\text{From Eq. (20), } \dot{k}(t) = \phi(k^{**}) - \phi(k^*) = \phi(k^{**}) - c^* - (\delta + n)k^* > 0$$

Hence, below the $\dot{k}(t) = 0$ curve: $\dot{k}(t) > 0$.

When $c^{**} < c^*$, $\dot{k}(t) = \phi(k^{**}) - c^* - (\delta + n)k^* < 0$. Therefore, above the $\dot{k}(t) = 0$ curve: $\dot{k}(t) < 0$. In another words, if consumption level is less than the capital growth over a unit

period, then, capital level will increase. In this case the direction in the area below the $\dot{k}(t) = 0$ curve points to the right, and therefore, to the left for the area above the curve.

Applying Romer's (2001) approach on how to determine the motion directions on both sides of the linear isoclines :

When $\dot{c}(t) = 0$, from Eq. (27):

Let k^{**} denote an arbitrary capital level.

When $k^{**} < k^*$, then, $\frac{\partial \phi}{\partial k^{**}} > \frac{\partial \phi}{\partial k^*} = r + \delta + n$, therefore, from Eq. (25), to the left of $\dot{c}(t) = 0$ curve: $\dot{c}(t) > 0$.

When $k^{**} > k^*$, then, $\frac{\partial \phi}{\partial k^{**}} < \frac{\partial \phi}{\partial k^*} = r + \delta + n$, therefore, from Eq. (25) to the right of $\dot{c}(t) = 0$ curve: $\dot{c}(t) < 0$.

Summarizing all the results we can obtain the following phase diagram⁴.

Figure 1.

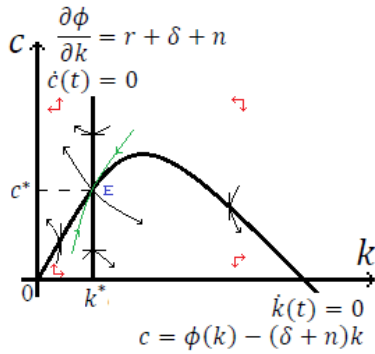
⁴ The more rigorous approach is through the sign of the characteristic roots of the differential equations. Since Eq. (20) and (25) are non-linear, we have to linearize them first to look for their characteristic roots. Assume, the equilibrium point is (k^*, c^*) . Apply Taylor expansion around equilibrium's neighborhood on Eq. (20) and (25), then write both expansions into the matrix form, reorganized them to get:

$$\begin{bmatrix} \dot{k} \\ \dot{c} \end{bmatrix} = \begin{bmatrix} \phi'(k^*) - \delta - n & -1 \\ -\frac{U_{c^*}}{U_{c^*}^2} \phi''(k^*) & [(r + \delta + n) - \phi'(k^*)] - \left\{ U_{c^*} (U_{c^*}^2)^{-2} U_{c^*}^3 [(r + \delta + n) - \phi'(k^*)] \right\} \end{bmatrix} \begin{bmatrix} k \\ c \end{bmatrix} + \begin{bmatrix} [\phi(k^*) - c^* - (\delta + n)k^*] - [\phi'(k^*) - \delta - n]k^* + c^* \\ \left[\frac{U_{c^*}}{U_{c^*}^2} [(r + \delta + n) - \phi'(k^*)] + \left[\frac{U_c}{U_c^2} \phi''(k^*) \right] k^* + \left\{ [(r + \delta + n) - \phi'(k^*)] - \left\{ U_c (U_c^2)^{-2} U_c^3 [(r + \delta + n) - \phi'(k^*)] \right\} \right\} c^* \right] \end{bmatrix}$$

. Then find the trace and the determinant of the parameter matrix. Plug them into the following to get the characteristic roots:

$$\begin{cases} R_k = \frac{\text{tr}(M) + \sqrt{[\text{tr}(M)]^2 - 4 \det(M)}}{2} \\ R_c = \frac{\text{tr}(M) - \sqrt{[\text{tr}(M)]^2 - 4 \det(M)}}{2} \end{cases} . \text{ These will yield real number solutions with opposite signs under certain}$$

conditions.



In our case, like most endogenous growth models, the intersection is a mathematically unstable saddle point equilibrium, but in economics we consider it to be a stable equilibrium as long as the policy is on the stable branch before it reaches the equilibrium. It is obviously that both from the transversality condition and the law of motions of the quadrants we just obtained, only one saddle path is remained and the rest trajectories are ruled out. Since the function $\phi(k)$ is not explicitly given, hence, we do not know what the exact separatrices (optimal path in this case) are.

After we draw the phase diagram, there is an interesting discussion can be made on the different dynamic behavior of the different initial consumption level. We can tell from figure 2, for instance, there are totally five different types of trajectories for an initial $k(0)$ left to the $\dot{c}(t) = 0$ curve ($k(0) \neq 0$). Denote, the corresponding consumption level for the $k(0)$ on the left optimal branch as c_{opt} ; and the one for $k(0)$ on the $\dot{k}(t) = 0$ curve as c_k .

If we assume the corresponding consumption level for the $k(0)$ is $c(0) \in (0, c_{opt})$, then, its trajectory will first rise until it reaches the $\dot{c}(t) = 0$ curve, then c starts to drop to zero. The economic meaning behind this is that when the initial consumption level below the optimal consumption level for a given capital endowment, the economy will first start to grow its consumption level backed up by its capital accumulation, and the consumption will grow for a while, but not strong enough, due to the low consumption level it started with. The economy will not be able to reach the equilibrium when its capital level hits k^* , and right after the economy missed this opportunity, with a low consumption level, the economy starts to sink. Agents truly starts to save rather than consume, and eventually, to an extreme, no one consumes, and that is when the economy dies with all the capital accumulated in its “life time”.

If we assume the corresponding consumption level for the $k(0)$ is c_{opt} , then, its trajectory will lead it to the equilibrium. This means that if an economy is already on the optimal economic

growth path, given the law of motion on the capital accumulation, for each stage there will be a unique level of capital be reached by a unique level of consumption from the “last period”, and a unique future consumption will be determined as well. Until the economy reaches the equilibrium level of capital and consumption level.

The third case is when the initial corresponding consumption level for the $k(0)$ is $c(0) \in (c_{opt}, c_k)$, then, its trajectory will first rise rapidly until it reaches the $\dot{k}(t) = 0$ curve, then k starts to drop to zero while the c increases. This means the economy consumes too much today than saving for the future consumption. And due to the failure on accumulating capital, it fails to reach the equilibrium (although it is already predetermined since the initial value for k and c are not on the saddle path), and starts to consuming on its previously little accumulated capital and then initial (endowed) capital and eventually “dies” with nothing left to consume. This scenario explains many warfare in the naïve era. In the ancient world, given a relatively much lower overall connection (such as international laws, economic treaties and mutual respects among different cultures and regions, etc.) among civilizations. If a feudal country’s moral is governed by hedonism, with a low productivity capacity, the most efficient way to maintain or increase the utility for the ruling class is to pillage the other civilizations. Because if the civilization keep partying but choose not to pillage, it will eventually consume up all of its capitals, given the unpredictable nature of technological innovation, low population growth rate (due to high infant mortality rate and short life expectancy in the ancient world in general) and relatively much lower inter-civilization trading volume in the ancient time (due to the low transportation technology and the limited understanding on trading advantages) (i.e. with a low ancient productivity level most production outputs have to fulfill the domestic demands first). At this stage, sadly, due to the low capital remains there is not too much left to either fight with or fight for. Interestingly enough though, the relatively less capital remained can be a huge fortune for the non-ruling classes. Here is when we see another kind of war, revolts.

If we assume the corresponding consumption level for the $k(0)$ is $c(0) = c_k$, then, once time starts to lapse, similar to the second stage of the last case, the economy feeds its consumption with the initial capital imminently.

The last case is that the corresponding consumption level for the $k(0)$ is $c(0) \in (c_k, \infty)$, where $\dot{c}(t) > 0$ and $\dot{k}(t) < 0$. Thus, while c start to increase, k starts to decrease.

Chapter 5 - Concluding Remarks

The paper has been describes how the Hamiltonian function is derived for an OCP, how the commonly used OCPs are set up, and what their applications are in economics. We learned that for a given OCP, theoretically, we need to assume a value function as a unique solution to a HJB equation first. Using a Taylor expansion of the continuous time version of the Bellman equation one can find the value function for the "next time period". After combining this with "today's value function" we can get the actual HJB equation, which can be redefined as a Hamiltonian function. During this process, the co-state variable is also created. Given the state and co-state variables we can structure a set of Hamiltonian equations of motion, which is part of optimality conditions. The rest conditions, in most cases the initial condition of the state variable, are given. We can choose the easier-to-solve local maximality condition ($\frac{\partial H^*}{\partial \text{control variable}^*} = 0$) or the global maximality condition (PMP). In terms of the terminal condition, for finite horizon OCP, there are three cases. The first one is the one we mentioned in Case 1: a free end terminal condition, where a transversality condition is needed. The second one is when $x(t_s) = x^1$, which appears in the first example. And there is no conditions on $z(t_s)$. The last case which we did not discuss in the paper, but worth mentioning is when $x(t_s) \geq x^1$, it needs $z(t_s) \geq 0$ ($z(t_s) = 0$, if $x(t_s)^* > x^1$) (Sydsæter, Strøm, and Berck 113). For the infinite horizon OCP, in most cases in economics, we set $\lim_{t \rightarrow \infty} x(t)^* = \bar{x}$.

We also need to be aware of that the optimal control theory used in economics is not as strict as the one in mathematics. First, the admissible control set is not restricted. In mathematics, the admissible control set $\{U\}$ is restricted for OCT, i.e. it's a closed set. By assume $\{U\}$ is an open set in economics leads to the second problem. Second, as I mentioned in the first foot note, there is a difference for the maximality condition between calculus of variations ($\frac{\partial H^*}{\partial \text{control variable}^*} = 0$) and OCT (PMP). While applying OCT in economics, we often use $\frac{\partial H^*}{\partial \text{control variable}^*} = 0$ instead. As a matter of fact, it is correct to do so, only if the optimal control u^* is an internal solution inside of its admissible control set $\{U\}$ (and the Humiliation function is continuously differentiable with respect to the control variable). However, if u^* is a corner

solution, which the strict OCT allows (since its $\{U\}$ is a closed set). Then the maximality condition has to be PMP.

Reference

- [1] Byrne, O. 1847. *The First Six Books of the Elements of Euclid*. Pickering, London.
- [2] Bell, E. T. 1937. *Men of Mathematics*. Simon and Schuster Publications, New York.
- [3] Bellman, R. 1957. *Dynamic Programming*. Princeton University Press, New Jersey.
- [4] Cantor, M. B. 1907. *Lectures via History of Mathematics*. B. G. Teubner, Leipzig.
- [5] Chiang, A., and K. Wainwright. 2005. *Fundamental Methods of Mathematical Economics, Forth Edition*. McGraw-Hill, New York.
- [6] Cui, D. 2008. *Methods of Dynamic Analysis in Advanced Macroeconomics*. China Financial & Economic Publishing House, Beijing.
- [7] da Dilver, A.C. 2001. *Hamiltonian Vector Fields*. Springer, New York.
- [8] Deriglazov, Alexei. 2010. *Classical Mechanics*. Springer, New York.
- [9] Mitchell, I. 2008. *Dynamic Programming and Approximate Dynamic Programming*. University of British Columbia. Kelowna, BC. Retrieved from <http://www.cs.ubc.ca/~mitchell/Class/CS532M.2007W2/Talks/hjbSham.pdf>
- [10] Nambu, Y. 1973. Generalized Hamiltonian Dynamics. *Physical Review D* 7.8, 2405-2412. Retrieved from <http://journals.aps.org/prd/pdf/10.1103/PhysRevD.7.2405>
- [11] Pesch, H. J., and M. Plail. 2009. The Maximum Principle of Optimal Control: A History of ingenious ideas and missed opportunities. *Control and Cybernetics* vol. 38 No. 4A, 973-995. Retrieved from <http://www.pages.drexel.edu/~hgk22/courses/MEM733/Pesch%20and%20Plail%202009.pdf>
- [12] Pontryagin, L. S. 1986. *L.S. Pontryagin: Mathematical Theory of Optimal Process*. Gordon and Breach Science Publishers, New York.
- [13] Romer, D. 2001. *Advanced Macroeconomics, 2nd Edition*. McGraw-Hill Higher Education, New York.
- [14] Samuelson, P. 1970. What Makes a Beautiful Problem in Science? *Journal of Political Economy* Vol. 78, No. 6, 1372-1377. Retrieved from <http://www.jstor.org/stable/pdfplus/1830631.pdf?acceptTC=true&jpdConfirm=true>
- [15] Sydsæter, K., A. Strom & P. Berck. 2005. *Economists' Mathematical Manual, 4th Edition*. Springer, New York.

- [16] Takayama, A. 2009. *Mathematical Economics, 2nd Edition*. China Renmin University Press, Beijing.
- [17] Turkington, D. A. 2008. *Mathematical Tools for Economics*. Beijing: Higher Education Press.
- [18] Woodward, R. 2013. *Dynamic Optimization in Agricultural & Applied Economics*. Texas A&M University. College Station, Texas. Retrieved from <http://agecon2.tamu.edu/people/faculty/woodward-richard/637/notes/04.pdf>

Appendix A - Finite-Horizon OCP with State and Co-state Variable Vectors

In order to discuss Appendix B, a rough idea on the following theoretical OCP is worth mentioning, in which case all decision maker's choice variables are collected in a vector valued control, say $u_k(t)$. And the same for the state variable as well.

Given the following Finite-Horizon OCP with one state and one control variables by Sydsæter, Strøm, and Berck (2005):

Case 3

$$J(u_k(t)) \equiv \int_{t_0}^{t_s} f(x_i(t), u_k(t), t) dt \rightarrow \max_{u(t) \in \{U\}}$$

$$s.t.: \dot{x}_i(t) = g_i(x_i(t), u_k(t), t) \quad (i = 1, 2, \dots, n; k = 1, 2, \dots, m)$$

$$x_i(t_0) = x_i^0$$

$$x_i(t_s) \text{ free}$$

$$u(t) \in \{U\}$$

where, $x_i(t)$ is a set of the state variables (i.e.: $x_i(t) := \{X\} := \{x \mid x = x_1(t), x_2(t), \dots, x_n(t)\} \subseteq \mathbb{R}^n$); $u_k(t)$ is the admissible control set (i.e.: $u_k(t) := \{U\} := \{u \mid u = u_1(t), u_2(t), \dots, u_m(t)\} \subseteq \mathbb{R}^m$)⁵; $u(t) \in \{U\} = \mathbb{R}^m$ is a admissible control function. $\{T\} := \{t \mid t \in [t_0, t_s]\}$. $x_i(t_0) = x_i^0$, i.e.: $\forall x_i(t) \in \{X\} \times \{T\}, \exists (x_0, t_0). \forall g_i(\cdot)$ are C^∞ with respect to x_i and t .

Similar to case 1, we can define Hamiltonian function $H: \{[t_0, t_s]\} \times \{X\} \times \{U\} \times \mathbb{R}^n \rightarrow \mathbb{R}$ as: $H(t, x_i, u_k, z_i) := f(x_i(t), u_k(t), t) + \langle z_i, g_i(x_i(t), u_k(t), t) \rangle$

We now have a multiple state variable case, so $z_i g_i(x_i(t), u_k(t), t)$ is written in inner product form. The optimal conditions for this case are:

⁵ Note that the amount of control variables do not necessary have to equal to the amount of state variables. If reality requires, in most cases, a control variable can be divided into many sub-controls, even they are divided evenly.

$$\left\{ \begin{array}{l} z_i(t_s) = 0 \\ x_i(t_0) = x_i^0 \\ \dot{x}_i(t) = \frac{\partial H(z_i(t), x_i^*(t), u_k^*(t), t)}{\partial z_i} \\ \dot{z}_i(t) = -\frac{\partial H(z_i(t), x_i^*(t), u_k^*(t), t)}{\partial x_i} \\ H(z_i(t), x_i^*(t), u_k^*(t), t) \geq H(z_i(t), x_i^*(t), u_k(t), t) \end{array} \right. \quad (i = 1, 2, \dots, n'; k = 1, 2, \dots, m)$$

In this case, the Hamiltonian system is structured within a $2n$ -phase hyperspace.

Appendix B - A System of OCPs with N Canonical Variables - A Proposal

Almost all different variations on the optimal control problem has been considered. However, one interesting case is missing. In this appendix, I propose a new kind of optimal control problem (OCP), and offer a new-to-mathecon method to solve it. The new OCP is what I would call it a system of OCP, where, the co-state variable from the first sub-OCP is the state variable for the second sub-OCP, and so on. And the method is Nambu Hamiltonian.

Prior to further discussion, several concepts need to be clarified first. We starting by redefine the following state and co-state variable sets from Case 3 into a canonical variable set:

Form Case 3, we obtained

$$\{x_i(t)\} := \{\text{State Variable}\} \quad (i = 1, 2, \dots, n)$$

$$\{z_i(t)\} := \{\text{Co - state Variable}\} \quad (i = 1, 2, \dots, n)$$

A canonical variable in our case is defined as: $\xi_{ij}(t) \in \{\xi(t)\} := \{x_{i1}(t)\} \cup$

$$\{z_{i2}(t)\} \cup_{a=3}^N C_{ia} \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, N)$$

where, C_{ia} is the a th column vector of the rest of the canonical variables, where $a \in [3, N]$. In another words, the entire canonical variable set is a $n \times N$ matrix (i.e.: N canonical variable sets, which each one contains n variables.) For the first and second column they are the state column vector and the co-state column vector in Case 3 respectively. But for the canonical variables in $j > 2$ column vectors (i.e. C_{ia}), we do not have a terminology in mathematic economics for them yet.

From the definitions above, the canonical variables from Case 1 which constructed a two dimensional phase plane (or a three dimensional general phase space) are now denoted as: $\xi_{11}(t) := x(t); \xi_{12}(t) := z(t)$.

Given Hamiltonian equations (canonical equations) of motion (04) and (05), assume, $\exists F(\xi_{11}(t), \xi_{12}(t), t)$, then, we can obtain the following:

$$\begin{aligned} \dot{F} &\equiv \frac{dF}{dt} = \frac{\partial F}{\partial \xi_{11}} \frac{\partial \xi_{11}}{\partial t} + \frac{\partial F}{\partial \xi_{12}} \frac{\partial \xi_{12}}{\partial t} + \frac{\partial F}{\partial t} \\ &= \frac{\partial F}{\partial \xi_{11}} \dot{\xi}_{11}(t) + \frac{\partial F}{\partial \xi_{12}} \dot{\xi}_{12}(t) + \frac{\partial F}{\partial t} \end{aligned}$$

$$\begin{aligned}
&= \frac{\partial F}{\partial \xi_{11}} \frac{\partial H_1}{\partial \xi_{12}} - \frac{\partial F}{\partial \xi_{12}} \frac{\partial H_1}{\partial \xi_{11}} + \frac{\partial F}{\partial t} \\
&= \frac{\partial(F, H_1)}{\partial(\xi_{11}, \xi_{12})} + \frac{\partial F}{\partial t} \quad (28)
\end{aligned}$$

where, the Jacobian determinant of F and H_1 with respect to ξ_{11} and ξ_{12} is:

$$\frac{\partial(F, H_1)}{\partial(\xi_{11}, \xi_{12})} = \begin{vmatrix} \frac{\partial F}{\partial \xi_{11}} & \frac{\partial H_1}{\partial \xi_{11}} \\ \frac{\partial F}{\partial \xi_{12}} & \frac{\partial H_1}{\partial \xi_{12}} \end{vmatrix} \quad (29)$$

And notice that the canonical equations of motion are essentially the same as Eq. (04) and (05), since we are discussing the two canonical variable case from Case 1:

$$\begin{cases} \dot{\xi}_{11}(t) = \frac{\partial H_1}{\partial \xi_{12}} \\ \dot{\xi}_{12}(t) = -\frac{\partial H_1}{\partial \xi_{11}} \end{cases} \quad (30)$$

If we have more than two canonical variables (other than state and co-state variables, for example N canonical variables), we shall have the following case:

$$\exists \xi_{ij}(t) \quad i = 1, 2, \dots, n; j = 1, 2, \dots, N \quad (31)$$

According to Nambu mechanics (1973) and given Eq. (29), the Hamiltonian equations of motions in a $n \times N$ dimensional phase hyperspace are:

$$\begin{cases} \dot{\xi}_{i1}(t) = \frac{\partial(H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i2}, \xi_{i3}, \dots, \xi_{iN})} = \frac{\partial(\xi_{i1}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} \\ \dot{\xi}_{i2}(t) = \frac{\partial(H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i3}, \xi_{i4}, \dots, \xi_{iN})} = \frac{\partial(\xi_{i2}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} \\ \vdots \\ \dot{\xi}_{iN}(t) = \frac{\partial(H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN-1})} = \frac{\partial(\xi_{iN}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} \end{cases} \quad (32)$$

where, there are $N-1$ amount of Hamiltonian functions. The reason it can be done this way is that (32) just like the two canonical variable case, it also fits Liouville theorem. According to Nambu⁶ (1973), Liouville theorem is actually the only rule a Nambu Hamiltonian need to obey in a phase space. And it does, here is the proof. Since Liouville theorem states that despite the

⁶ Yoichiro Nambu in his paper Generalized Hamiltonian Dynamics states that the only guiding principle of the existence of Generalized Hamiltonian is if it following Liouville theorem.

shape change of a region in a phase space over time, its "volume"⁷ maintains the same, hence, what we need to do is to prove the volume of this Nambu Hamiltonian over time is zero (i. e.: $\frac{\partial v}{\partial t} \equiv 0$).

Proof 2

Given an infinitesimal volume moving for an infinitesimal time. Assume, \exists point ξ_{ij} in a $n \times N$ dimensional phase hyperspace.

Its volume is defined as:

$$v := \prod_{i=1}^n \prod_{j=1}^N d\xi_{ij}$$

At time dt , we can define the canonical variables with new status form Eq. system (32):

$$\begin{cases} \tilde{\xi}_{i1} \equiv \xi_{i1} + \dot{\xi}_{i1} dt = \xi_{i1} + \frac{\partial(\xi_{i1}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} dt \\ \vdots \\ \tilde{\xi}_{iN} \equiv \xi_{iN} + \dot{\xi}_{iN} dt = \xi_{iN} + \frac{\partial(\xi_{iN}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} dt \end{cases}$$

Hence, the new volume is:

$$\tilde{v} = \prod_{i=1}^n \prod_{j=1}^N d\tilde{\xi}_{ij} = \prod_{i=1}^n \prod_{j=1}^N \xi_{ij} + \dot{\xi}_{ij} dt$$

$\because H_1(\xi_{ij}, t), H_2(\xi_{ij}, t), \dots, H_{N-1}(\xi_{ij}, t)$ are Hamiltonian functions, $\forall j = 1, 2, \dots, N$.

$\therefore \forall H(\cdot), \exists$ Hamiltonian velocity fields, according to da Silva (2001).

\because According to Deriglazov (2010) Hamiltonian velocity field is divergenceless. i.e.:

$$\dot{\xi}_{ij} = \frac{\partial(\xi_{ij}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} = 0 \quad \forall j = 1, 2, \dots, N$$

$\therefore \tilde{\xi}_{ij} \equiv \xi_{ij} \quad \forall j = 1, 2, \dots, N$

$\therefore v \equiv \tilde{v}$

$\therefore \frac{\partial v}{\partial t} \equiv 0$

■

⁷ "Volumes" actually is phase space density, it can be defined as $\rho(\xi_1, \xi_2, \dots, \xi_N, t)$, it tells us the probability of finding a Hamiltonian system near the neighborhood of $(\xi_1, \xi_2, \dots, \xi_N)$ at time t .

Since, the Nambu Hamiltonian satisfied the Liouville theorem, therefore, Eq. system (32) exists. At this point, stage has been set. The core interest of this appendix is to solve the following N canonical variables OCP system:

Case 4

$$J(u_{k1}(t)) \equiv \int_{t_0}^{t_s} f_1(\xi_{i1}(t), u_{k1}(t), t) dt \rightarrow \max_{u(t) \in \{U\}}$$

$$s.t.: \dot{\xi}_{i1}(t) = g_{i1}(\xi_{i1}(t), u_{k1}(t), t) \quad (i = 1, 2, \dots, n; k = 1, 2, \dots, m)$$

$$\xi_{i1}(t_0) = \xi_{i1}^0$$

$$\xi_{i1}(t_s) = 0$$

$$u_{k1}(t) \in \{U_{k1}\}$$

$$J(u_{k2}(t)) \equiv \int_{t_0}^{t_s} f_2(\xi_{i2}(t), u_{k2}(t), t) dt \rightarrow \max_{u(t) \in \{U\}}$$

$$s.t.: \dot{\xi}_{i2}(t) = g_{i2}(\xi_{i2}(t), u_{k2}(t), t) \quad (i = 1, 2, \dots, n; k = 1, 2, \dots, m)$$

$$\xi_{i2}(t_0) = \xi_{i2}^0$$

$$\xi_{i2}(t_s) = 0$$

$$u_{k2}(t) \in \{U_{k2}\}$$

⋮

$$J(u_{kN-1}(t)) \equiv \int_{t_0}^{t_s} f_{N-1}(\xi_{iN-1}(t), u_{kN-1}(t), t) dt \rightarrow \max_{u(t) \in \{U\}}$$

$$s.t.: \dot{\xi}_{iN-1}(t) = g_{iN-1}(\xi_{iN-1}(t), u_{kN-1}(t), t) \quad (i = 1, 2, \dots, n; k = 1, 2, \dots, m)$$

$$\xi_{iN-1}(t_0) = \xi_{iN-1}^0$$

$$\xi_{iN-1}(t_s) = 0$$

$$u_{kN-1}(t) \in \{U_{kN-1}\}$$

where, $\xi_{ij}(t)$ is a canonical variable from a canonical variable matrix $\Xi_{i,j}$ (i.e.: $\xi_{ij}(t) \in \Xi_{i,j} :=$

$\begin{bmatrix} \xi_{11} & \cdots & \xi_{1N} \\ \vdots & \ddots & \vdots \\ \xi_{n1} & \cdots & \xi_{nN} \end{bmatrix}$). This also infers that in the first sub-OCP the "co-state variable" of state

variable ξ_{i1} is ξ_{i2} , and in the second sub-OCP the "co-state variable" of state variable ξ_{i2} is ξ_{i3} .

All the way to the N - I th sub-OCP's "co-state variable" of its state variable ξ_{iN-1} is ξ_{iN} . $u_{kj}(t)$ is

a control variable from a control variable matrix $U_{k,j}$ (i.e.: $u_{kj}(t) \in U_{k,j} :=$

$$\begin{bmatrix} u_{11} & \cdots & u_{1N-1} \\ \vdots & \ddots & \vdots \\ u_{m1} & \cdots & u_{mN-1} \end{bmatrix}; \{T\} := \{t \mid t \in [t_0, t_s]\}; \xi_{ij}(t_0) = \xi_{ij}^0, \text{ i.e.: } \forall \xi_{ij}(t) \in \Xi \times T, \exists (\xi_0, t_0).$$

$\forall g_{ij}(\cdot)$ are C^∞ with respect to ξ_{ij} and t .

For this OCP system, its Nambu Hamiltonians are:

$$\left\{ \begin{array}{l} H_1(t, \xi_{i1}, u_{k1}, \xi_{i2}) := f_1(\xi_{i1}, u_{k1}, t) + \langle \xi_{i2}, g_{i1}(\xi_{i1}, u_{k1}, t) \rangle \\ H_2(t, \xi_{i2}, u_{k2}, \xi_{i3}) := f_2(\xi_{i2}, u_{k2}, t) + \langle \xi_{i3}, g_{i2}(\xi_{i2}, u_{k2}, t) \rangle \\ \vdots \\ H_{N-1}(t, \xi_{iN-1}, u_{kN-1}, \xi_{iN}) := f_{N-1}(\xi_{iN-1}, u_{kN-1}, t) + \langle \xi_{iN}, g_{iN-1}(\xi_{iN-1}, u_{kN-1}, t) \rangle \end{array} \right.$$

And similar to Case 3, the necessary condition for the Nambu Hamiltonian functions is the following:

$$\left\{ \begin{array}{l} \xi_{ij}(t_0) = \xi_{ij}^0 \\ \xi_{ij}(t_s) = 0 \\ \dot{\xi}_{i1}(t) = \frac{\partial(H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i2}, \xi_{i3}, \dots, \xi_{iN})} = \frac{\partial(\xi_{i1}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} \\ \dot{\xi}_{i2}(t) = \frac{\partial(H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i3}, \xi_{i4}, \dots, \xi_{iN})} = \frac{\partial(\xi_{i2}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} \\ \vdots \\ \dot{\xi}_{iN}(t) = \frac{\partial(H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN-1})} = \frac{\partial(\xi_{iN}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} \\ H_1(\xi_{i2}(t), \xi_{i1}^*(t), u_{k1}^*(t), t) \geq H_1(\xi_{i2}(t), \xi_{i1}^*(t), u_{k1}(t), t) \\ H_2(\xi_{i3}(t), \xi_{i2}^*(t), u_{k2}^*(t), t) \geq H_2(\xi_{i3}(t), \xi_{i2}^*(t), u_{k2}(t), t) \\ \vdots \\ H_{N-1}(\xi_{iN}(t), \xi_{iN-1}^*(t), u_{kN-1}^*(t), t) \geq H_{N-1}(\xi_{iN}(t), \xi_{iN-1}^*(t), u_{kN-1}(t), t) \end{array} \right. \quad \left(\begin{array}{l} i = 1, 2, \dots, n \\ \forall, j = 1, 2, \dots, N \\ k = 1, 2, \dots, m \end{array} \right)$$

or, if for all Hamiltonian functions are differentiable with respect to their control variables. We have the following:

$$\left\{ \begin{array}{l} \xi_{ij}(t_0) = \xi_{ij}^0 \\ \xi_{ij}(t_s) = 0 \\ \dot{\xi}_{i1}(t) = \frac{\partial(H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i2}, \xi_{i3}, \dots, \xi_{iN})} = \frac{\partial(\xi_{i1}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} \\ \dot{\xi}_{i2}(t) = \frac{\partial(H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i3}, \xi_{i4}, \dots, \xi_{iN})} = \frac{\partial(\xi_{i2}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} \\ \vdots \\ \dot{\xi}_{iN}(t) = \frac{\partial(H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN-1})} = \frac{\partial(\xi_{iN}, H_1, H_2, \dots, H_{N-1})}{\partial(\xi_{i1}, \xi_{i2}, \dots, \xi_{iN})} \quad \left(\begin{array}{l} i = 1, 2, \dots, n \\ \forall j = 1, 2, \dots, N \\ k = 1, 2, \dots, m \end{array} \right) \\ \frac{\partial H_1^*}{\partial u_{k1}^*} = 0 \\ \frac{\partial H_2^*}{\partial u_{k2}^*} = 0 \\ \vdots \\ \frac{\partial H_{N-1}^*}{\partial u_{kN-1}^*} = 0 \end{array} \right.$$

The reason we have the two sets of necessary conditions above is that the legitimacy of Case 3 has already been established for decades, and Case 4 is just a logical expansion of Case 3. This means that for each one of the Nambu Hamiltonian functions, PMP has to be satisfied respectively. And along with the law of motions we have established and proved, we obtained the first set of necessary condition. Similarly for the second set of conditions as well.

One may doubt that Case 4 can be solved by solving each of the sub-OCP with traditional method like in Case 3. The problem is that unless those independent OCPs are carefully defined, the relationship between the state variable column vector in a sub-OCP (which is the co-state variable column vector in the previous sub-OCP) and the new co-state variable column vector in the same sub-OCP is ignored. For instance, the relationship between vector ξ_{i2} and ξ_{i3} .

Even though this method is proved to be mathematically correct, its application in economics is yet to be discovered. This is due to the structure of the OCP, where the co-state variables of the first sub-OCP are $\frac{\partial V_1}{\partial \xi_{i1}}$ ($i = 1, 2, \dots, n$), but the co-state variables of the second sub-OCP are $\frac{\partial V_2}{\partial \xi_{i2}} = \frac{\partial V_2}{\partial \left(\frac{\partial V_1}{\partial \xi_{i1}} \right)}$, and even worse, for the third co-state variables $\frac{\partial V_3}{\partial \xi_{i3}} = \frac{\partial V_3}{\partial \left(\frac{\partial V_2}{\partial \left(\frac{\partial V_1}{\partial \xi_{i1}} \right)} \right)}$

extend all the way until the N-1th co-state variables $\frac{\partial V_{N-1}}{\partial \xi_{iN-1}} = \frac{\partial V_{N-1}}{\partial \left(\frac{\partial V_{N-2}}{\partial \left(\frac{\partial V_{N-3}}{\partial (\cdot)} \right)} \right)}$, which is not infinitely

worse but “finitely worse”. I believe such free boundary problem is indeed quite interesting but tough to solve. Hence, finding an economic example to fit the structure can be rare. However, it is still possible if only for a two sub-OCP. And in my wildest dream, if an economy by chance is able to be described by a finite amount of OCPs and fits the requirements, then, this method can be a powerful tool to set up the optimal conditions for them.