Optimal Control and Nonlinear Taxation*

Todd Lensman

February 19, 2018

These notes describe basic results about continuous optimal control systems and their applications to static models of nonlinear taxation. The discussion of optimal control is based heavily on [3] and [4], and references to more detailed analyses in these texts are made where appropriate. The application of optimal control to a model of optimal nonlinear taxation follows the celebrated contribution of [5].

1 Control Systems and Control Problems

1.1 Control Systems

We consider control systems of the form

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)) \quad t \in [t_0, t_1], \\ x(t_0) = \bar{x}(t_0), \end{cases} \tag{1}$$

where $0 \leq t_0 < t_1 < \infty$ and

$$\begin{aligned} x &: [t_0, t_1] \to \mathbb{R}^n, \\ u &: [t_0, t_1] \to \mathbb{R}^m, \\ f &: [t_0, t_1] \times \mathbb{R}^{n+m} \to \mathbb{R}^n. \end{aligned}$$

Here $x(t)$ is the *state* vector of the system at time t, $u(t)$ is the *control* vector at time t, and f defines the evolution of the state vector as a function of the current state $x(t)$ and the current control $u(t)$. The function x is called the *trajectory* of the system and describes the state at any time t. The set of permissible *control functions* u is denoted U, and the set of permissible trajectories x is denoted X. Given a control function $u \in U$, a trajectory x solves the initial

*These notes benefited substantially from comments by Maxim Troshkin, who undertook the task of actually reading them and noting the many typos and unclear passages that appeared in previous drafts.

value problem (1) if x is absolutely continuous and satisfies the integral equation

$$x(t)-\bar{x}(t_0)-\int_{t_0}^{t} f(s,x(s),u(s))\,ds=0 \qquad t\in[t_0,t_1]. \tag{2}$$

In subsequent sections, we describe high-level regularity assumptions placed on f and U to ensure that the system (1) defines a trajectory x in an appropriate sense. This matter can be quite technical and belongs more to the theory of differential equations than to optimal control, so we will not consider it in much detail. Interested readers are referred to §3.3.1 of [3] for a short discussion and to chapters 3 and 10 of [2] for an extensive analysis.

1.2 Basic Control Problem

In what we will call the "Basic Control Problem," the goal is to minimize a functional that depends on the trajectory x and the control u, subject to constraints on the system dynamics, the terminal state $x(t_1)$, and the control vector $u(t)$. At the heart of the problem is a control system defined by the integral equation (2). We suppose that there exists a non-empty *target set* $S\subseteq\mathbb{R}^n$ such that the state $x(t)$ must be driven from $x(t_0)$ to S by time t_1. The corresponding constraint on the terminal state $x(t_1)$ is given by

$$x(t_1)\in S. \tag{3}$$

We also allow additional integral constraints on the system dynamics of the form

$$\int_{t_0}^{t_1} \alpha(t,x(t),u(t))\,dt=0, \tag{4}$$

where $\alpha:[t_0,t_1]\times\mathbb{R}^{n+m}\to\mathbb{R}^q$. The control vector $u(t)$ at any time $t\in[t_0,t_1]$ is also constrained to an arbitrary non-empty *control set* $\Omega\subseteq\mathbb{R}^m$:

$$u(t)\in\Omega \quad t\in[t_0,t_1]. \tag{5}$$

Finally, the objective of the control problem is given by the functional

$$\int_{t_0}^{t_1} l(t,x(t),u(t))\,dt+k(t_1,x(t_1)), \tag{6}$$

where $l : [t_0, t_1] \times \mathbb{R}^{n+m} \to \mathbb{R}$ is a *loss* function that describes the "running cost" of the system at each time t and $k : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}$ is a *terminal cost* function.[1]

The Basic Control Problem requires that we select a control function $u \in U$ and a trajectory $x \in X$ to minimize (6), subject to (2), (3), (4), and (5):

$$\inf_{u \in U, x \in X} \int_{t_0}^{t_1} l(t, x(t), u(t))\, dt + k(t_1, x(t_1))$$

subject to

$$x(t) - \bar{x}(t_0) - \int_{t_0}^{t} f(s, x(s), u(s))\, ds = 0 \quad t \in [t_0, t_1],$$

$$x(t_1) \in S,$$

$$\int_{t_0}^{t_1} \alpha(t, x(t), u(t))\, dt = 0,$$

$$u(t) \in \Omega \qquad t \in [t_0, t_1].$$

A few comments are in order concerning the generality of the Basic Control Problem:

1. It should be noted that we are considering only control problems in which the final time is fixed and finite. This is purely for convenience; in fact, the methods described below are readily adapted to problems in which the terminal time for the process is free or infinite. These adaptations will be discussed briefly below, but a more complete analysis can be found in chapter 4 of [3].

2. The integral constraint (4) is an equality constraint, and there are no pointwise constraints of the form

$$\beta(t, x(t), u(t)) = 0 \qquad t \in [t_0, t_1].$$

   Such pointwise constraints can introduce considerable complications to the theory, and they will not appear in the optimal taxation examples that we consider. Interested readers are referred to chapter 7 of [3] for an introduction. The equality in the integral constraint is again simply convenient, and inequality integral constraints can be incorporated in a manner similar to that described by the Kuhn-Tucker Theorem in finite-dimensional optimization.

3. In what follows we will assume that the target set S is a k-dimensional differentiable

[1] By an unfortunate coincidence, in the optimal control literature the loss function l is often called the "Lagrangian," not to be confused with the augmented objective function used in finite-dimensional optimization. The control-theoretic analog of the augmented objective function is the Hamiltonian function, defined in §3.

manifold for $k \in \{0,1,\ldots,n\}$ that can be described globally as the solution set of the equation $g(x)=0$. Here $g:\mathbb{R}^n \to \mathbb{R}^{n-k}$ is continuously differentiable with surjective derivative. Again, generalizations of the results below to other target sets are available.

4. The control set Ω is assumed time-invariant, but the methods below can be adapted to cases in which it is time-dependent, i.e., $\Omega(t)$.

1.3 Control Problem Simplifications

Before attempting to characterize necessary conditions for local optima, it will be useful to simplify the objective and constraint set of the Basic Control Problem through a few elementary observations. For example, the explicit dependence of the state evolution function f on t can be removed by appending the state vector $x(t)$ with an additional component $x_{n+1}(t)$, with time evolution

$$\begin{cases} \dot{x}_{n+1}(t) = 1 & t \in [t_0, t_1] \\ x_{n+1}(t_0) = t_0. \end{cases}$$

Hereafter we assume that this has been done in the original state vector $x(t)$. A similar trick can be applied to the integral constraint (4): Define the new state variables $x_i(t)$ for $i = n+1,\ldots,n+q$ that have time evolution

$$\begin{cases} \dot{x}_i(t) = \alpha_{i-q}(t, x(t), u(t)) & t \in [t_0, t_1] \\ x_i(t_0) = 0. \end{cases}$$

Lift the target set $S \subset \mathbb{R}^n$ to $\mathbb{R}^{n+q}$ by including the terminal constraints

$$x_i(t_1) = 0 \qquad i = n+1,\ldots,n+q.$$

Then the integral constraint (4) can be subsumed into the state evolution equation (2) and the terminal constraint (3), and we will assume that this has been done for the original state vector $x(t)$ and the original target set S. Regarding the objective, we will assume that the terminal cost function k is absolutely continuous, so we can write

$$\begin{aligned} k(x(t_1)) &= k(x(t_0)) + \int_{t_0}^{t_1} \frac{d}{dt} k(x(t))\, dt \\ &= k(x(t_0)) + \int_{t_0}^{t_1} k_x(x(t)) f(x(t), u(t))\, dt. \end{aligned}$$

The value $k(x(t_0))$ is fixed, so we can "remove" the terminal cost by replacing the objective (6) by

$$\int_{t_0}^{t_1} l(x(t), u(t)) + \frac{d}{dt} k(x(t))\, dt.$$

We will assume that this has been done with the original loss function l. Using the characterization of S as the solution set of $g(x) = 0$, we can without loss of generality restrict our attention to a simplified version of the Basic Control Problem:

$$\inf_{u \in U, x \in X} \int_{t_0}^{t_1} l(x(t), u(t))\, dt \tag{7}$$

subject to

$$x(t) - \bar{x}(t_0) - \int_{t_0}^{t} f(x(s), u(s))\, ds = 0 \quad t \in [t_0, t_1], \tag{8}$$

$$g(x(t_1)) = 0, \tag{9}$$

$$u(t) \in \Omega \quad t \in [t_0, t_1]. \tag{10}$$

Note that in subsequent sections, we will assume the existence of a solution to this problem. It should be emphasized that this assumption is highly nontrivial,[2] and in general establishing the existence of solutions to control problems can be a difficult task. Readers are referred to §4.5 of [3] and the references therein for a discussion.

2 Optimal Control - Variational Approach

This section provides an introduction to the variational approach to optimal control. Generally speaking, the variational approach entails placing sufficient regularity assumptions on the control system, objective, and constraints of the control problem to derive necessary conditions for local optima using variational calculus. A control problem is thus treated as an optimization problem over an infinite-dimensional Banach space, with no attention paid to the special structure of control problems (i.e., the presence of a dynamical system that describes the dependence of a state variable on a time-dependent control variable). The key advantage of this approach is the ease with which necessary conditions can be derived: The regularity assumptions ensure that all operators are differentiable in an appropriate sense, and the main tools for deriving necessary conditions are direct generalizations of those used in finite-dimensional optimization (i.e., the generalized Lagrange Multiplier and Kuhn-Tucker Theorems). However,

[2] Google "Perron's Paradox" to find an unfortunately persuasive argument for the importance of existence results in optimization.

these assumptions can be too restrictive for many control problems, and a more refined analysis is required to derive necessary conditions for optima. This analysis will be presented in the next section, where Pontryagin's celebrated Maximum Principle will be stated and proven for the Basic Control Problem.

2.1 Abstract-Operator Control Problem

In this section, we assume the following about the parameters of the Basic Control Problem:

1. $\Omega \subseteq \mathbb{R}^m$ is open, and X and U are open subsets of the sets of continuous functions from $[t_0, t_1]$ to $\mathbb{R}^n$ and Ω, respectively;
2. l is once continuously differentiable;
3. f is once continuously differentiable.

The first assumption allows us to vary the control function u and the trajectory x in an open set around a local optimum (x_*, u_*) while remaining in the sets of permissible control functions and trajectories. In addition, under this specification X and U are Banach spaces, so the generalized Lagrange Multiplier Theorem may be applied.

The second and third assumptions imply that $l(x(t), u(t))$ and $f(x(t), u(t))$ are integrable over $[t_0, t_1]$ for each $x \in X$ and $u \in U$. They also imply that the objective integral (7) and the evolution integral (8), viewed as operators on the set $X \times U$, are continuously Fréchet differentiable.[3] In particular, we can write (7) abstractly as a functional J on $X \times U$:

$$J(x,u) := \int_{t_0}^{t_1} l(x,u)\,dt.$$

Since X and U are open subsets of normed spaces, $X \times U$ is an open subset of the product norm space (with the standard norm). It is then possible to define the Fréchet derivative of an operator from $X \times U$ to another normed space, and a straightforward calculation shows that J has Fréchet differential

$$\delta J(x,u)[h,v] = \int_{t_0}^{t_1} l_x(x,u)h + l_u(x,u)v\,dt.$$

Here $(h, v) \in C([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^m)$ are permissible perturbations, i.e., $(x + \alpha h, u + \alpha v) \in$

[3] For a nice review of the calculus of variations, and specifically differential calculus on Banach spaces, see §7.1-7.3 of [4].

$X \times U$ for $\alpha \in \mathbb{R}$ with $|\alpha|$ sufficiently small. By the continuity of l_x and l_u, it is easily verified that the mapping $(x,u) \mapsto \delta J(x,u)$ is continuous, where the codomain of this mapping is the set of bounded, linear, real-valued functionals on $X \times U$, given the operator norm topology. We can similarly write (8) as

$$A(x,u) = 0,$$

where $A : X \times U \to X$ is an operator that maps x and u to the function of t on the left side of (8). This operator is also continuously Fréchet differentiable with differential

$$\delta A(x,u)[h,v] = t \mapsto \left\{ h(t) - \int_{t_0}^{t} f_x(x,u)h + f_u(x,u)v ds \right\}.$$

We can also identify g with an operator $g : X \to \mathbb{R}^{n-k}$, where the Fréchet differential is trivially given by

$$\delta g(x)h = g_x(x(t_1))h(t_1).$$

The Basic Control Problem can then be restated abstractly as

$$\begin{aligned} &\inf_{x \in X, u \in U} J(x,u) \\ &\text{subject to} \\ &A(x,u) = 0, \\ &g(x) = 0. \end{aligned} \tag{11}$$

2.2 Necessary Conditions

We have the following description of necessary first order conditions for an interior solution (x_*, u_*), with proof adapted from §9.5 of [4]:

Theorem 1. *Suppose (x_*, u_*) is a solution to the Basic Control Problem (11) and is a regular point of the constraint set. Then $\exists \lambda : [t_0, t_1] \to \mathbb{R}^n$ differentiable and $\exists \mu \in \mathbb{R}^{n-k}$ such that $\forall t \in [t_0, t_1]$,*

$$-\dot{\lambda}(t) = [f_x(x_*(t), u_*(t))]^T \lambda(t) + l_x(x_*(t), u_*(t))^T, \tag{12}$$

$$\lambda(t_1) = [g_x(x(t_1))]^T \mu, \tag{13}$$

$$0 = \lambda(t)^T f_u(x_*(t), u_*(t)) + l_u(x_*(t), u_*(t)). \tag{14}$$

Proof. Since $X \times U$ is open, the generalized Lagrange Multiplier Theorem (Theorem 1, §9.3, [4]) implies that $\exists \lambda : [t_0, t_1] \to \mathbb{R}^n$ of bounded variation and right-continuous and $\exists \mu \in \mathbb{R}^r$

such that

$$\delta J(x_*, u_*)[h, v] + \langle \delta A(x_*, u_*)[h, v], \lambda \rangle + \langle \delta g(x_*) h, \mu \rangle = 0$$
$$\forall (h, v) \in C([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^m).$$

Note that we use $\langle \cdot, \cdot \rangle$ to denote the pairing between a vector space and its dual, and recall that the dual of the set of continuous functions on an interval can be identified with the set of right-continuous functions of bounded variation via Lebesgue-Stieltjes integration. We can treat h and v separately to find that this condition is equivalent to

$$0 = \int_{t_0}^{t_1} l_x(x_*(t), u_*(t)) h(t) dt + \int_{t_0}^{t_1} \left[h(t) - \int_{t_0}^{t} f_x(x_*(s), u_*(s)) h(s) ds \right] d\lambda(t)^T + \mu^T g_x(x_*(t_1)) h(t_1), \quad (15)$$

$$0 = \int_{t_0}^{t_1} l_u(x_*(t), u_*(t)) v(t) dt - \int_{t_0}^{t_1} \int_{t_0}^{t} f_u(x_*(s), u_*(s)) v(s) ds d\lambda(t)^T, \quad (16)$$

$\forall (h, v) \in C([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^m)$.[4] We normalize λ such that $\lambda(t_1) = 0$ and integrate by parts in the first condition (15) to find

$$0 = \int_{t_0}^{t_1} l_x(x_*(t), u_*(t)) h(t) dt + \int_{t_0}^{t_1} h(t) d\lambda(t)^T + \int_{t_0}^{t_1} \lambda(t)^T f_x(x_*(t), u_*(t)) h(t) dt + \mu^T g_x(x_*(t_1)) h(t_1). \quad (17)$$

If $\lambda(t)$ had a jump in $[t_0, t_1)$, the Lebesgue-Stieltjes measure $d\lambda_i(t)^T$ would have a mass point in $[t_0, t_1)$ for some $i = 1, \ldots, n$. Standard Lebesgue measure is non-atomic, so we could construct an $h \in C([t_0, t_1], \mathbb{R}^n)$ to make the second term in (17) large relative to the others. (This argument suffices only on $[t_0, t_1)$ because the fourth term in (17) involves $h(t_1)$.) Hence λ must be continuous on $[t_0, t_1)$. Choose $h \in C([t_0, t_1, \mathbb{R}^n])$ such that $h(t_0) = h(t_1) = 0$ and

[4] The integral $\int_{t_0}^{t_1} a(t) d\lambda(t)^T$ is defined by

$$\int_{t_0}^{t_1} a(t) d\lambda(t)^T := \sum_{i=1}^{n} \int_{t_0}^{t_1} a_i(t) d\lambda_i(t).$$

h is differentiable. Then we can again integrate by parts:

$$
\begin{aligned}
0 &= \int_{t_0}^{t_1} l_x(x_*(t), u_*(t)) h(t)\, dt - \int_{t_0}^{t_1} \lambda(t)^T \dot{h}(t)\, dt \\
&\quad + \int_{t_0}^{t_1} \lambda(t)^T f_x(x_*(t), u_*(t)) h(t)\, dt \\
&= \int_{t_0}^{t_1} \left[l_x(x_*(t), u_*(t)) - \lambda(t)^T f_x(x_*(t), u_*(t)) \right] h(t) + \lambda(t)^T \dot{h}(t)\, dt.
\end{aligned}
$$

By the continuity of λ on $[t_0, t_1)$ and Lemma 3 in §7.5 of [4], we can conclude that λ is differentiable on $[t_0, t_1)$ and satisfies (12).

At $t = t_1$, λ must have a jump of $-[g_x(x(t_1))]^T \mu$ so that (17) is satisfied $\forall h \in C([t_0, t_1], \mathbb{R}^n)$. (Consider $h \equiv 0$ on $[t_0, t_1 - \epsilon]$ with $h(t_1) \neq 0$, and take $\epsilon \downarrow 0$.) Without affecting the differentiability of λ on $[t_0, t_1)$, we can then define $\lambda(t_1)$ such that (13) is satisfied. This implies that λ is also continuous on $[t_0, t_1]$.

Finally, integrate (16) to find

$$
0 = \int_{t_0}^{t_1} l_u(x_*(t), u_*(t)) v(t)\, dt + \int_{t_0}^{t_1} \lambda(t)^T f_u(x_*(t), u_*(t)) v(t)\, dt.
$$

By Lemma 1 in §7.5 of [4], (i.e., the Reymond-Du Bois Lemma), this implies (14). ☐

The first order conditions given by Theorem 1 and the constraints of the problem (11) yield a system that can, in principle, be solved to find (x_*, u_*): equations (8), (12), and (14). However, two remarks are in order:

1. As with finite-dimensional optimization problems, the necessary conditions of Theorem 1 are not sufficient to characterize a local or global optimum without additional verifications. Sufficient second order conditions for a local minimum involve second variations and will not be addressed here, and conditions that ensure a local minimum is a global minimum mirror those from the finite-dimensional case (e.g., convex constraint set and quasi-convex objective).

2. The regularity assumption in Theorem 1 is non-trivial and has a nice interpretation in the context of control problems. In general, the regularity condition ensures that in a neighborhood of the local extremum (x_*, u_*), the constraint set is "well-behaved" in the sense that it can be reasonably approximated by linearizing the constraints around (x_*, u_*).

This is crucial, because the generalized Lagrange Multiplier Theorem follows from an application of the generalized Inverse Function Theorem (Theorem 1, §9.2, [4]). The intuition for the latter carries over from the finite-dimensional case: A continuously Fréchet differentiable mapping can be inverted locally by inverting its differential (modulo the kernel of the differential), and this requires that the differential be surjective. We should verify, then, that the differential of $(A, g) : X \times U \to X \times \mathbb{R}^n$ is onto when evaluated at a solution (x_*, u_*).

From the differential of g, we see that we must require $g_x(x_*(t_1))$ to have rank $n-k$, but this is ensured by our previous full rank assumption. We also assume that f is such that for any $e \in \mathbb{R}^n$, we can choose $v \in C([t_0, t_1, \mathbb{R}^m])$ such that

$$\begin{aligned} h(t) &= \int_{t_0}^{t} f_x(x_*, u_*) h + f_u(x_*, u_*) v ds, \qquad (18) \\ h(t_1) &= e. \end{aligned}$$

The right side of equation (18) is simply the original control system linearized about the solution (x_*, u_*). This *controllability condition* requires that the linearized system can be driven from 0 to any $e \in \mathbb{R}^n$ by time t_1 with an appropriate choice of control v. Under these assumptions, (x_*, u_*) is a regular point of the constraint set in problem (11) if $\forall (y, e) \in C([t_0, t_1, \mathbb{R}^n]) \times \mathbb{R}^n$, $\exists (h, v) \in C([t_0, t_1, \mathbb{R}^n]) \times C([t_0, t_1, \mathbb{R}^m])$ such that

$$y(t) = h(t) - \int_{t_0}^{t} f_x(x_*, u_*) h + f_u(x_*, u_*) v ds, \qquad (19)$$

$$h(t_1) = e. \qquad (20)$$

Note that the substitution $h(t_1)$ for $g_x(x_*(t_1))h(t_1)$ in (20) is without loss of generality since $g_x(x_*(t_1))$ has full rank. When $v \equiv 0$, the general existence theorem for ordinary differential equations implies a solution $\bar{h}$ to this system. By setting $w(t) = h(t) - \bar{h}(t)$, the system (19-20) is equivalent to

$$\begin{aligned} 0 &= w(t) - \int_{t_0}^{t} f_x(x_*, u_*) w + f_u(x_*, u_*) v ds, \\ w(t_1) &= 0. \end{aligned}$$

The controllability condition ensures that this system has a solution, so the constraint set is regular at (x_*, u_*).

3 Pontryagin's Maximum Principle

The variational approach used to prove Theorem 1 has several obvious shortcomings: Most significantly, use of the generalized Lagrange Multiplier Theorem requires that we restrict to control functions u that belong to a Banach space, and the most natural such space in an optimal control setting is the set of $\mathbb{R}^m$-valued continuous functions. However, it is straightforward to devise nontrivial examples of control problems that do not have solutions with continuous controls, but are easily solvable with piecewise-continuous controls. In addition, the variational approach used above requires that the control set Ω and the function spaces X and U be open in order to consider arbitrary variations from a local optimum (x_*, u_*). This method can be applied under some weaker restrictions on Ω, X, and U, but it requires a more subtle argument than that used above. Finally, the smoothness assumptions placed on l and f are quite strong, and it would be helpful to develop a theory that applies to a larger class of control problems. This section presents precisely such a theory, culminating in a proof of Pontryagin's Maximum Principle.

Before proceeding to the mathematics, a quick philosophical remark: There is a good reason why Pontryagin's Maximum Principle is typically called a "principle" rather than a "theorem." Optimal control problems come with any number of different constraints, objectives, state evolution functions, control sets, etc., and it is reasonable to expect that optimality conditions are highly sensitive to the problem specification. As a result, it would be surprising if there were a single theorem that could concisely describe the optimality conditions for every conceivable control problem. The Maximum Principle does not do this, and I know of no theorem that does. Rather, the Maximum Principle is best thought of as a methodology that describes how to derive optimality conditions for a particular problem. Indeed, common statements of the Maximum Principle, like that given below, only apply to special classes of control problems, and it may be that a particular problem of interest does not precisely meet the hypotheses of these statements. Optimality conditions must then be derived from scratch, and the proof of the Maximum Principle provides a guide on how to do this. With this in mind, I encourage the reader to at least read through the proof of the Maximum Principle given below to develop an intuition for the arguments. It would be particularly beneficial to try adapting the proof to, say, an optimal control problem in which the terminal time is also a choice variable, or to read through chapters 4 and 7 of [3] for proofs of Maximum Principles for other control problems.

In this section, we assume the following about the parameters of the Basic Control Problem:

1. $\Omega \subseteq \mathbb{R}^m$ is an arbitrary non-empty set;

2. U is the set of all piecewise-continuous functions from $[t_0, t_1]$ to Ω, and X is the set of all continuous functions from $[t_0, t_1]$ to $\mathbb{R}^n$;

3. l and f are continuous, and they are continuously differentiable in the state vector $x(t)$.

It should be clear that these assumptions are generically weaker than those used in §2, with the exception of the specification of U: It is both analytically convenient and practically attractive to permit piecewise-continuous controls, though the arguments below do not apply immediately when the control function u is restricted to be continuous.

Central to the statement and intuition behind the Maximum Principle is the *Hamiltonian function* $H : \mathbb{R}^n \times \Omega \times \mathbb{R}^n \times \mathbb{R}_- \to \mathbb{R}$, defined by

$$H(x, u, \lambda, \lambda_0) := \lambda_0 l(x, u) + \lambda \cdot f(x, u).$$

The name of this function is drawn from the Hamiltonian of classical mechanics, and the analogy will become clear below. Given the assumptions above, the version of Pontryagin's Maximum Principle for the Basic Control Problem is

Theorem 2 (Pontryagin's Maximum Principle). *Suppose (x_*, u_*) is a solution to the basic control problem (7). Then $\exists \lambda : [t_0, t_1] \to \mathbb{R}^n$ absolutely continuous and $\exists \lambda_0 \in \mathbb{R}_-$ with $(\lambda_0, \lambda) \neq (0, 0)$ $\forall t \in [t_0, t_1]$ such that*

1. *x_* and λ follow the evolution equations*

$$\begin{aligned} \dot{x}_* &= H_\lambda(x_*, u_*, \lambda, \lambda_0), \\ \dot{\lambda} &= -H_x(x_*, u_*, \lambda, \lambda_0); \end{aligned} \tag{21}$$

2. *x_* and λ satisfy the boundary conditions*

$$\begin{aligned} &x_*(t_0) = \bar{x}(t_0), \\ &x_*(t_1) \in S, \\ &\lambda(t_1) \perp T_{x_*(t_1)} S; \end{aligned} \tag{22}$$

3. *u_* maximizes the Hamiltonian pointwise:*

$$H(x_*(t), u_*(t), \lambda(t), \lambda_0) \geq H(x_*(t), u, \lambda(t), \lambda_0) \quad \forall t \in [t_0, t_1], \forall u \in \Omega. \tag{23}$$

Before presenting the proof, a few clarifying remarks: The evolution equations (21) are immediately recognizable as the canonical equations from the Hamiltonian formulation of classical mechanics, where the trajectory x_* plays the part of the "position function" and the *costate* λ plays the part of the "momentum" function. This is no accident – Hamilton's Principle from classical mechanics states that the trajectory of a physical system can be interpreted as the solution to a particular optimal control problem, and much of classical mechanics can be derived from this viewpoint.

In the statement of the boundary conditions (22), $T_{x_*(t_1)}S$ denotes the *tangent space* of the manifold S at $x_*(t_1)$. Readers familiar with differential geometry or basic differential topology will recognize that this constraint can be written

$$\lambda(t_1) \perp \ker g_x(t_1) \iff \lambda(t_1) \in \operatorname{span}\{\nabla g_i(t_1) : i = 1, \ldots, n-k\}.$$

Thus the condition $\lambda(t_1) \perp T_{x_*(t_1)}S$ is precisely equivalent to the existence of $\mu \in \mathbb{R}^{n-k}$ such that $\lambda(t_1) = [g_x(x_*(t_1))]^T \mu$, and this is recognizable as the boundary condition (13) from Theorem 1. Note also that the boundary conditions (22) place $n+k+(n-k) = 2n$ constraints on the values of (x_*, λ) at the boundary times t_0, t_1, so we expect that (21) and (22) uniquely specify x_* and λ, given λ_0 and the optimal control function u_*.

We will now embark on a proof of Theorem 2. The proof proceeds in several steps and is modeled after the proof given in §4.2 of [3].

4 Application: Mirrleesian Taxation

In this section, we consider an application of control methods to the classic model of optimal nonlinear taxation described in [5]. That paper represents a landmark achievement in the study of optimal income taxation: The Mirrlees model elegantly captures the *equity-efficiency* trade-off that is salient in questions of income taxation, and for this reason it serves as the basic framework for many theoretical and applied studies. Our purpose below is to describe the structure of the model and show how Pontryagin's Maximum Principle (Theorem 2) can be used to characterize the optimal income tax. This analysis largely follows [5], and we will also present an example from [1] in which it is possible to derive a formula for optimal marginal tax rates. It should be noted that the technical assumptions below differ slightly from those in [5], but this is purely for convenience and ease of explication.

4.1 Setup

The economy has a unit measure of agents who receive idiosyncratic *skills* $\theta \in \left[\underline{\theta}, \bar{\theta}\right] \subset \mathbb{R}_{++}$, where $0 < \underline{\theta} < \bar{\theta} < 0$. Skills are stochastic and distributed according to a continuous density function $f \in C\left(\left[\underline{\theta}, \bar{\theta}\right], \mathbb{R}_{+}\right)$. The agents exert *labor* $l \in [0,1)$ to earn *income* $y = \theta l$ in proportion to their skills, and they consume a non-negative quantity $c \in \mathbb{R}_{+}$ of a *consumption* good. Crucially, an agent's income y and consumption c are publicly observable, while his skill θ and labor choice l are privately known. Each agent's preferences over consumption and labor are described by a utility function $U(c,l)$, assumed twice continuously differentiable and weakly concave on $\mathbb{R}_{+} \times [0,1)$, with $U_c, -U_l > 0$ on the interior. It is also assumed that U satisfies the limit

$$\lim_{l \uparrow 1} U(c,l) = -\infty \quad c \in \mathbb{R}_{+}. \tag{24}$$

The economy has a government that seeks to provide a degree of redistribution while financing exogenous *expenditures* $E \in \mathbb{R}_{+}$. Since skills are unobservable, it does so by choosing an *income tax* $T : \mathbb{R}_{+} \to \mathbb{R}$ that describes the tax liability of an agent as a function of his income y. An income tax T induces the following choice problem for an agent with skill θ:

$$\begin{aligned} V^{\theta} := \max_{c \in \mathbb{R}_{+}, l \in [0,1)} & \; U(c,l) \\ \text{subject to} & \\ c \leq \theta l - T(\theta l)&. \end{aligned} \tag{25}$$

For technical convenience and to guarantee a solution to each agent's decision problem, we require that T be continuous. Let l^{θ} denote a solution function to the agent's problem (25).[5]

The government designs an income tax T to maximize a weighted expectation of the agents' utilities, subject to collecting revenues E. The weighting is given by a weakly increasing and concave function $G : \mathbb{R} \to \mathbb{R}$, which describes the redistribution motive of the government.

[5]Using (24), it is a straightforward exercise to show that the maximum in (25) is attained and so that V^{θ} is a well-defined, real-valued function. It does not necessarily hold that l^{θ} is a function rather than a correspondence, but it will be convenient to make this assumption. It can be checked that in many examples of interest, the agent's problem (25) has a unique solution.

The income tax T is then chosen to solve the problem

$$\sup_{T \in C(\mathbb{R}_+,\mathbb{R})} \mathbb{E}_f\left[G\left(V^\theta\right)\right] \tag{26}$$

subject to

$$\mathbb{E}_f\left[E - T\left(\theta l^\theta\right)\right] \leq 0, \tag{27}$$

$$U\left(\theta l^\theta - T\left(\theta l^\theta\right), l^\theta\right) \geq U\left(\theta l - T\left(\theta l\right), l\right) \quad \theta \in \left[\underline{\theta}, \bar{\theta}\right], l \in [0,1), \tag{28}$$

$$T(y) - y \leq 0 \quad y \in \left[0, \bar{\theta}\right]. \tag{29}$$

Note that the incentive constraints (28) are essentially a restatement of the definition of l^θ.

4.2 Government's Problem Simplification

Stated in the form (26), the government's problem is essentially intractable: The government must optimize over the Banach space $C(\mathbb{R}_+, \mathbb{R})$ subject to a double continuum of incentive constraints (28). Even with additional differentiability assumptions on the tax T and other primitive assumptions that ensure that the first order condition from the agent's problem (25) is both well-defined and sufficient to characterize l^θ, the resulting optimization problem is quite difficult to solve using variational calculus. A better approach can be found by more carefully characterizing the value function V^θ for the agent's problem (25), making use of the incentive constraints (28). The following lemma is crucial:

Lemma 3. *For any continuous $T : \mathbb{R}_+ \to \mathbb{R}$ that satisfies the constraints of (26), V^θ is Lipschitz continuous and satisfies*

$$V^\theta = V^{\underline{\theta}} + \int_{\underline{\theta}}^{\theta} -\frac{l^s}{s} U_l\left(sl^s - T\left(sl^s\right), l^s\right) ds. \tag{30}$$

Proof. The continuity of T implies that T attains a minimum over $\left[0, \bar{\theta}\right]$, and hence $\theta l - T(\theta l)$ is bounded above uniformly in $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$ and $l \in [0,1)$. Denote the upper bound by $\bar{c} \geq 0$. By the assumption (24), $U(\bar{c}, l) \downarrow -\infty$ as $l \uparrow 1$, so $\exists \bar{l} \in [0,1)$ such that

$$U\left(\bar{c}, \bar{l}\right) \leq V^{\underline{\theta}}.$$

Then for any $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$ and any $l \in \left(\bar{l}, 1\right)$,

$$U\left(\theta l - T(\theta l), l\right) \leq U\left(\bar{c}, \bar{l}\right) \leq V^{\underline{\theta}} \leq V^\theta.$$

This implies that we can restrict all agents' choice sets to $\left[0, \bar{l}\right]$:

$$V^{\theta} = \max_{l \in \left[0, \bar{l}\right]} U\left(\theta l - T\left(\theta l\right), l\right).$$

For any $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$, let $\delta_{\theta} > 0$ such that $\frac{\theta+\delta_{\theta}}{\theta-\delta_{\theta}}\bar{l} < 1$. Then $\forall \theta', \theta'' \in B_{\delta_{\theta}}\left(\theta\right)$, we have

$$\begin{aligned}
\left|V^{\theta'} - V^{\theta''}\right| &\leq \sup_{y \in \left[0, (\theta+\delta)\bar{l}\right]} \left|U\left(y - T\left(y\right), \frac{y}{\theta'}\right) - U\left(y - T\left(y\right), \frac{y}{\theta''}\right)\right| \\
&= \sup_{y \in \left[0, (\theta+\delta)\bar{l}\right]} \left|\int_{\theta''}^{\theta'} -\frac{y}{s^2} U_l\left(y - T\left(y\right), \frac{y}{s}\right) ds\right| \\
&\leq \left|\theta' - \theta''\right| K_{\theta},
\end{aligned}$$

where

$$K_{\theta} := \max_{y \in \left[0, (\theta+\delta)\bar{l}\right], s \in \left[\underline{\theta}, \bar{\theta}\right]} \frac{y}{s^2} \left|U_l\left(y - T\left(y\right), \frac{y}{s}\right)\right|$$

is finite because $0 < \underline{\theta}$, U_l is continuous on $\mathbb{R}_+ \times [0,1)$, and the choice set is compact. This implies that V^{θ} is locally Lipschitz continuous on $\left[\underline{\theta}, \bar{\theta}\right]$. But $\left[\underline{\theta}, \bar{\theta}\right]$ is compact, so V^{θ} is globally Lipschitz on $\left[\underline{\theta}, \bar{\theta}\right]$.

Since V^{θ} is Lipschitz on $\left[\underline{\theta}, \bar{\theta}\right]$, it must be differentiable a.e. on $\left[\underline{\theta}, \bar{\theta}\right]$ and satisfy

$$V^{\theta} = V^{\underline{\theta}} + \int_{\underline{\theta}}^{\theta} \dot{V}^s ds.$$

In particular, the right- and left-derivatives $\frac{\partial}{\partial \theta^+} V^{\theta}$ and $\frac{\partial}{\partial \theta^-} V^{\theta}$ must exist and be equal (to $\dot{V}^{\theta}$) a.e. The incentive constraints (28) imply the inequality

$$\begin{aligned}
\frac{\partial}{\partial \theta^+} V^{\theta} &= \lim_{h \downarrow 0} \frac{1}{h}\left(V^{\theta+h} - V^{\theta}\right) \\
&\leq \lim_{h \downarrow 0} \frac{1}{h}\left(U\left(\theta l^{\theta} - T\left(\theta l^{\theta}\right), \frac{\theta}{\theta+h} l^{\theta}\right) - U\left(\theta l^{\theta} - T\left(\theta l^{\theta}\right), l^{\theta}\right)\right) \\
&= -\frac{l^{\theta}}{\theta} U_l\left(\theta l^{\theta} - T\left(\theta l^{\theta}\right), l^{\theta}\right).
\end{aligned}$$

A similar calculation yields the inequality

$$\frac{\partial}{\partial \theta^-} V^{\theta} \geq -\frac{l^{\theta}}{\theta} U_l\left(\theta l^{\theta} - T\left(\theta l^{\theta}\right), l^{\theta}\right).$$

We can then conclude $\dot{V}^{\theta} = -\frac{l^{\theta}}{\theta} U_l\left(\theta l^{\theta} - T\left(\theta l^{\theta}\right), l^{\theta}\right)$, so (30) holds. □

Intuitively, this lemma holds because the incentive constraints (28) include *local incentive-compatibility constraints* that require that an agent with type θ weakly prefer his own labor choice l^θ to the labor choice of an agent with type $\theta + h$, where $|h| \downarrow 0$. In the limit, these local constraints determine completely how the value function V^θ changes with θ. This is a profound property of the incentive structure of the model, and it is even more remarkable when we recognize that the derivative $\dot{V}^\theta$ can be expressed

$$\dot{V}^\theta = -\frac{l^\theta}{\theta} U_l\left(\theta l^\theta - T\left(\theta l^\theta\right), l^\theta\right) = \frac{\partial}{\partial \theta}\left\{U\left(y - T(y), \frac{y}{\theta}\right)\right\}\Bigg|_{y=\theta l^\theta}.$$

This last equality shows that when the agent's choice variable is taken to be income y, the derivative of the value function V^θ is precisely the partial derivative of the utility function U with respect to θ, evaluated at the optimal choice. In this sense, the incentive constraints (28) require that the value function V^θ change in lockstep with utility U when the skill θ is perturbed marginally.

It is particularly helpful to interpret Lemma 3 through the lens of optimal control: Given a tax policy T, V^θ can be considered a sort of "state" variable whose trajectory is driven by the "control" l^θ via the evolution equation (30). However, we can do even better than this and remove the tax T entirely from (30). To see this, note simply that since $U(\cdot, l)$ is strictly increasing, it must be invertible, so we can define the function $c(V, l)$ implicitly by

$$V = U(c(V, l), l). \tag{31}$$

In fact, if (V, l) is interior to the set $U(\mathbb{R}_+ \times [0,1)) \times [0,1)$, then the Implicit Function Theorem guarantees that c is actually C^1. The evolution equation (30) can then be written

$$V^\theta = V^{\underline{\theta}} + \int_{\underline{\theta}}^{\theta} -\frac{l^s}{s} U_l(c(V^s, l^s), l^s)\, ds. \tag{32}$$

We can similarly express the resource constraint (27) and the non-negativity constraint (29) in terms of V^θ and l^θ using the function c. This suggests that it may be possible to formulate the government's problem in terms of the choice functions V^θ and l^θ and then apply techniques from optimal control to derive optimality conditions. This method for deriving optimality conditions is known as the "Mirrlees Trick," and it is precisely what we plan to do.

However, there is yet a key problem with this plan: Lemma 3 shows that the evolution equation (32) follows from the incentive constraints (28), but it does not prove that any direct mechanism defined by $\left(V^{\theta}, l^{\theta}\right)$ which satisfies (32) is necessarily incentive-compatible.[6] In general, the evolution equation (32) is *not* sufficient to guarantee that the incentive-compatibility holds. Intuitively, this is because only local incentive constraints are necessary to derive the evolution equation, and it is not always true that local incentive-compatibility implies global incentive-compatibility. To salvage the analysis, we can place a relatively weak assumption on the utility function U to ensure that (32) and an additional condition are equivalent to incentive-compatibility.

In particular, we will assume that U is such that the expression

$$-l\frac{U_l(c,l)}{U_c(c,l)} \tag{33}$$

is strictly increasing in l for $c \in \mathbb{R}_+$. This is equivalent to assuming that the marginal rate of substitution

$$-\frac{1}{\theta}\frac{U_l\left(c,\frac{y}{\theta}\right)}{U_c\left(c,\frac{y}{\theta}\right)}$$

is strictly decreasing in θ for $c \in \mathbb{R}_+$. This is known as the *Spence-Mirrlees condition*, a *single-crossing condition*, or the *agent monotonicity condition*, and variants of it are used frequently in contract theory and mechanism design to deal precisely with the issue that we are facing here. It implies that higher-skilled agents find it strictly less costly to produce income than lower-skilled agents. This condition is relatively weak, and it is satisfied by most common utility specifications. We then have the following characterization of incentive-compatibility:

Lemma 4. *Under the Spence-Mirrlees condition, $\left(V^{\theta}, l^{\theta}\right)$ is incentive-compatible if and only if (32) is satisfied and $y^{\theta} = \theta l^{\theta}$ is strictly increasing in θ when $y^{\theta} > 0$.*

Proof. For simplicity, we will prove the result in the case that the direct mechanism $\left(V^{\theta}, l^{\theta}\right)$ is twice continuously differentiable in θ. For the forward direction, Lemma 3 immediately implies that (32) is satisfied. To see that $\dot{y}^{\theta} > 0$ when $y^{\theta} > 0$, note that local incentive-compatibility

[6] Under the direct mechanism induced by $\left(V^{\theta}, l^{\theta}\right)$, an agent reports a type $\hat{\theta}$, and in return he must produce effective labor $\hat{\theta} l^{\hat{\theta}}$ and receive consumption $c^{\hat{\theta}} := c\left(V^{\hat{\theta}}, l^{\hat{\theta}}\right)$. This mechanism is *incentive-compatible* if

$$V^{\theta} \geq U\left(c^{\hat{\theta}}, \frac{\hat{\theta}}{\theta} l^{\hat{\theta}}\right) \quad \theta, \hat{\theta} \in \left[\underline{\theta}, \bar{\theta}\right].$$

implies the first order condition

$$
\begin{aligned}
0 &= \dot{c}^{\hat{\theta}} U_c\left(c^{\hat{\theta}}, \frac{y^{\hat{\theta}}}{\theta}\right) + \frac{\dot{y}^{\hat{\theta}}}{\theta} U_l\left(c^{\hat{\theta}}, \frac{y^{\hat{\theta}}}{\theta}\right)\Bigg|_{\hat{\theta}=\theta} \\
&= \frac{U_c\left(c^{\hat{\theta}}, \frac{y^{\hat{\theta}}}{\theta}\right)}{y^{\hat{\theta}}} \left(y^{\hat{\theta}} \dot{c}^{\hat{\theta}} + \dot{y}^{\hat{\theta}} \frac{y^{\hat{\theta}}}{\theta} \frac{U_l\left(c^{\hat{\theta}}, \frac{y^{\hat{\theta}}}{\theta}\right)}{U_c\left(c^{\hat{\theta}}, \frac{y^{\hat{\theta}}}{\theta}\right)} \right)\Bigg|_{\hat{\theta}=\theta}
\end{aligned}
$$

The necessary second order condition requires that the right side be non-increasing in $\hat{\theta}$ at $\hat{\theta} = \theta$. The factor outside the parentheses is strictly positive since $y^{\theta} > 0$, so the right side is non-increasing in $\hat{\theta}$ at $\hat{\theta} = \theta$ if and only if

$$
0 \geq \frac{\partial}{\partial \hat{\theta}} \left(y^{\hat{\theta}} \dot{c}^{\hat{\theta}} + \dot{y}^{\hat{\theta}} \frac{y^{\hat{\theta}}}{\theta} \frac{U_l\left(c^{\hat{\theta}}, \frac{y^{\hat{\theta}}}{\theta}\right)}{U_c\left(c^{\hat{\theta}}, \frac{y^{\hat{\theta}}}{\theta}\right)} \right)\Bigg|_{\hat{\theta}=\theta}
$$

□

for the evolution equation (32) to imply incentive-compatibility is for effective labor $y^{\theta} = \theta l^{\theta}$ to be strictly increasing in θ, unless $y^{\theta} = 0$.

$$
\left(1 - T'\right) U_c^{\theta} + \frac{1}{\theta} U_l^{\theta} = 0 \iff \frac{U_c^{\theta}}{y^{\theta}} \left(y^{\theta} \left(1 - T'\right) + \frac{y^{\theta}}{\theta} \frac{U_l^{\theta}}{U_c^{\theta}} \right) = 0.
$$

Note that here we have taken the agent's choice variable to be y for convenience. The sufficient second order condition requires that the left side be strictly decreasing at y^{θ}. Since the left side is precisely zero at y^{θ}, we can drop the positive factor $\frac{U_c^{\theta}}{y^{\theta}}$. The sufficient second order condition is then

$$
G_y\left(\theta, y^{\theta}\right) = \frac{\partial}{\partial y} \left\{ y\left(1 - T'(y)\right) + \frac{y}{\theta} \frac{U_l\left(y - T(y), \frac{y}{\theta}\right)}{U_c\left(y - T(y), \frac{y}{\theta}\right)} \right\}\Bigg|_{y=y^{\theta}} < 0, \tag{34}
$$

where we define

$$
G(\theta, y) := y\left(1 - T'(y)\right) + \frac{y}{\theta} \frac{U_l\left(y - T(y), \frac{y}{\theta}\right)}{U_c\left(y - T(y), \frac{y}{\theta}\right)}.
$$

The first order condition implies $G\left(\theta, y^{\theta}\right) = 0$, and if we totally differentiate this equation with respect to θ, we find

$$0 = \frac{d}{d\theta} G\left(\theta, y^{\theta}\right) = G_{\theta}\left(\theta, y^{\theta}\right) + G_{y}\left(\theta, y^{\theta}\right) \dot{y}^{\theta}.$$

After rearranging, we have the equality

$$G_{y}\left(\theta, y^{\theta}\right) \dot{y}^{\theta} = -G_{\theta}\left(\theta, y^{\theta}\right).$$

Direct calculation of the right side yields

$$G_{y}\left(\theta, y^{\theta}\right) \dot{y}^{\theta} = \frac{y^{\theta}}{\theta^{2}} \frac{\partial}{\partial l} \left\{ l \frac{U_{l}(c,l)}{U_{c}(c,l)} \right\} \Bigg|_{c = y^{\theta} - T\left(y^{\theta}\right), l = \frac{y^{\theta}}{\theta}}.$$

The Spence-Mirrlees condition (33) implies that the right side is strictly negative, so the assumption $\dot{y}^{\theta} > 0$ implies the sufficient second order condition (34).

Under our differentiability assumptions, the evolution equation (32) can be shown to be equivalent to the agent's first order condition via the Envelope Theorem. We have thus shown that under the Spence-Mirrlees condition (33), any direct mechanism $\left(V^{\theta}, l^{\theta}\right)$ that satisfies the evolution equation (32) and is such that y^{θ} is strictly increasing in θ must be locally incentive-compatible. However, it is also easy to see that global incentive-compatibility also holds: Simply compute the difference

$$U\left(c\left(V^{\theta'}, l^{\theta'}\right), \frac{\theta'}{\theta} l^{\theta'}\right) - U\left(c\left(V^{\theta}, l^{\theta}\right), l^{\theta}\right)$$

using the Fundamental Theorem of Calculus (our differentiability assumptions imply that V^{θ} and l^{θ} are at least differentiable), and make use of the Spence-Mirrlees condition (33).

In the interest of proceeding to the derivation of optimality conditions, we will not prove that the evolution equation (32) and increasing effective labor are sufficient for incentive-compatibility under the Spence-Mirrlees condition (33) in the full generality needed here (i.e., when the tax T is simply continuous.) However, this result does hold (see [5]), and we will use it below. We can then reformulate the government's problem (26) as an optimal control problem:

$$\sup_{V,l} \int_{\underline{\theta}}^{\bar{\theta}} G\left(V^{\theta}\right) f(\theta)\, d\theta \tag{35}$$

subject to

$$V^{\theta} = V^{\underline{\theta}} + \int_{\underline{\theta}}^{\theta} -\frac{l^s}{s} U_l\left(c\left(V^s, l^s\right), l^s\right) ds \quad \theta \in \left[\underline{\theta}, \bar{\theta}\right], \tag{36}$$

$$\int_{\underline{\theta}}^{\bar{\theta}} E - \left(\theta l^{\theta} - c\left(V^{\theta}, l^{\theta}\right)\right) f(\theta)\, d\theta \leq 0, \tag{37}$$

$$c\left(V^{\theta}, l^{\theta}\right) - \theta l^{\theta} \leq 0 \qquad \theta \in \left[\underline{\theta}, \bar{\theta}\right], \tag{38}$$

$$\theta \mapsto \theta l^{\theta} \text{ increasing.} \tag{39}$$

Here we take $V : \left[\underline{\theta}, \bar{\theta}\right] \to \mathbb{R}$ to be absolutely continuous and $l : \left[\underline{\theta}, \bar{\theta}\right] \to [0, 1)$ to be piecewise continuous. It is standard to relax this problem by removing the non-negativity constraint (38) and the second-order condition (39). (These are verified ex post in simulations.) We then arrive at the relaxed control problem:

$$\sup_{V,l} \int_{\underline{\theta}}^{\bar{\theta}} G\left(V^{\theta}\right) f(\theta)\, d\theta \tag{40}$$

subject to

$$V^{\theta} = V^{\underline{\theta}} + \int_{\underline{\theta}}^{\theta} -\frac{l^s}{s} U_l\left(c\left(V^s, l^s\right), l^s\right) ds \quad \theta \in \left[\underline{\theta}, \bar{\theta}\right], \tag{41}$$

$$\int_{\underline{\theta}}^{\bar{\theta}} E - \left(\theta l^{\theta} - c\left(V^{\theta}, l^{\theta}\right)\right) f(\theta)\, d\theta \leq 0. \tag{42}$$

This problem is (nearly) in the form of the control problems considered in §2 and §3, and in particular we will be able to apply Pontryagin's Maximum Principle to derive optimality conditions.

4.3 General Necessary Conditions

The following theorem describes the necessary conditions for a local optimum to the government's relaxed control problem (40):

Theorem 5. *Let (V_*, l_*) be a local optimum for the relaxed problem (40) that is not an extremal* *of the constraint set. Then* $\exists\varphi : \left[\underline{\theta}, \bar{\theta}\right] \to \mathbb{R}$ *absolutely continuous and* $\exists\psi \in \mathbb{R}$ *such that*[7]

$$V_*^\theta = V_*^{\underline{\theta}} + \int_{\underline{\theta}}^{\theta} -\frac{l_*^s}{s} U_l^s ds \quad \forall\theta \in \left[\underline{\theta}, \bar{\theta}\right], \tag{43}$$

$$0 = \int_{\underline{\theta}}^{\bar{\theta}} E - \left(\theta l_*^\theta - c\left(V_*^\theta, l_*^\theta\right)\right) f(\theta)\, d\theta, \tag{44}$$

$$0 = \psi\left(\frac{U_l^*}{U_c^*} + \theta\right) f(\theta) + \frac{\varphi^\theta}{\theta}\left[-l_*^\theta \frac{U_l^*}{U_c^*} U_{cl}^\theta + l_*^\theta U_{ll}^\theta + U_l^\theta\right] \quad \text{if } l_*^\theta > 0, \tag{45}$$

$$0 = \int_{\underline{\theta}}^{\bar{\theta}} \left[G'\left(V_*^s\right) + \frac{\psi}{U_c^s}\right] f(s) \exp\left(-\int_{\underline{\theta}}^{s} \frac{l_*^t}{t} \frac{U_{cl}^t}{U_c^t} dt\right) ds, \tag{46}$$

where

$$\varphi^\theta = \int_{\theta}^{\bar{\theta}} \left[G'\left(V_*^s\right) + \frac{\psi}{U_c^s}\right] f(s) \exp\left(-\int_{\theta}^{s} \frac{l_*^t}{t} \frac{U_{cl}^t}{U_c^t} dt\right) ds.$$

Before presenting the proof, a quick note: Equations (43) and (45) should be viewed as determining the functions V_* and l_*, given ψ and $V_*^{\underline{\theta}}$. The conditions (44) and (46) then determine ψ and $V_*^{\underline{\theta}}$, respectively. Thus the set of necessary conditions (43-46), in principle, fully characterize the solution to the government's relaxed control problem (40). Finally, note that equations (45) and (46) correspond to equations (27) and (26) in [5], respectively.

Proof. Despite its interpretation as a static optimal taxation problem, (40) has the same formal structure as the Basic Control Problem, where θ is the "time" variable, V is the trajectory, and l is the control function. (The resource constraint (42) is clearly binding, so it can be considered an equality constraint.) In particular, we can derive necessary conditions for a local optimum using Pontryagin's Maximum Principle (Theorem 2).

Introduce a new "state" variable for the resource constraint,

$$B^\theta := \int_{\underline{\theta}}^{\theta} E - \left(\theta l^\theta - c\left(V^\theta, l^\theta\right)\right) f(\theta)\, d\theta.$$

With this additional state variable and the binding resource constraint (42), note that the target set $S \subseteq \mathbb{R}^2$ for the appended state vector (V, B) is now given by $S = \mathbb{R} \times \{0\}$, i.e., the horizontal

[7] To simplify notation, we write U_x^θ for the x^{th} partial derivative of U, evaluated at the optimum $\left(c\left(V_*^\theta, l_*^\theta\right), l_*^\theta\right)$.

axis in the V-B plane. Define the Hamiltonian

$$H(V,B,l,\varphi,\psi,\lambda_0,\theta) := -\lambda_0 G(V) f(\theta) - \varphi \frac{l}{\theta} U_l(c(V,l),l) + \psi(E - \theta l + c(V,l)) f(\theta).$$

(Note the sign of the first term: We are solving a maximization problem, so the "loss" function l in the corresponding minimization problem is given by $-G(V) f(\theta)$.) We assume that the local optimum (V_*, l_*) is not an extremal of the constraint set, so we can normalize the abnormal multiplier λ_0 to -1. The Maximum Principle then implies that $\exists \varphi, \psi : \left[\underline{\theta}, \bar{\theta}\right] \to \mathbb{R}$ absolutely continuous such that

1. V_*, B_*, φ, and ψ follow the evolution equations

$$\begin{aligned}
\dot{V}_*^\theta &= H_\varphi = -\frac{l_*^\theta}{\theta} U_l\left(c\left(V_*^\theta, l_*^\theta\right), l_*^\theta\right), \\
\dot{B}_* &= H_\psi = \left(E - \theta l_*^\theta + c\left(V_*^\theta, l_*^\theta\right)\right) f(\theta), \\
\dot{\varphi}^\theta &= -H_V \\
&= -G'\left(V_*^\theta\right) f(\theta) + \varphi^\theta \frac{l_*^\theta}{\theta} U_{cl}\left(c\left(V_*^\theta, l_*^\theta\right), l_*^\theta\right) c_V\left(V_*^\theta, l_*^\theta\right) - \psi c_V\left(V_*^\theta, l_*^\theta\right) f(\theta), \\
\dot{\psi}^\theta &= -H_B = 0;
\end{aligned} \tag{47}$$

2. V_*, B_*, φ, and ψ satisfy the boundary conditions

$$\begin{aligned}
&B_*^{\underline{\theta}} = B_*^{\bar{\theta}} = 0, \\
&\left(\varphi^{\underline{\theta}}, \psi^{\underline{\theta}}\right) \perp T_{\left(V_*^{\underline{\theta}}, B_*^{\underline{\theta}}\right)}(\mathbb{R} \times \{0\}), \\
&\left(\varphi^{\bar{\theta}}, \psi^{\bar{\theta}}\right) \perp T_{\left(V_*^{\bar{\theta}}, B_*^{\bar{\theta}}\right)}(\mathbb{R} \times \{0\}).
\end{aligned} \tag{48}$$

3. l_* maximizes the Hamiltonian pointwise:

$$H\left(V_*^\theta, B_*^\theta, l_*^\theta, \varphi, \psi, -1, \theta\right) \geq H\left(V_*^\theta, B_*^\theta, l, \varphi, \psi, -1, \theta\right) \quad \forall \theta \in \left[\underline{\theta}, \bar{\theta}\right], \forall l \in [0,1). \tag{49}$$

All of these optimality conditions follow directly from Theorem 2), with the exception of the boundary condition

$$\left(\varphi^{\underline{\theta}}, \psi^{\underline{\theta}}\right) \perp T_{\left(V_*^{\underline{\theta}}, B_*^{\underline{\theta}}\right)}(\mathbb{R} \times \{0\}).$$

This is because the Basic Control Problem (7) considered in §3 has a fixed initial state, while the government's relaxed control problem (40) allows $V^{\underline{\theta}}$ to vary freely. Since we must have

$B^{\underline{\theta}} = 0$, we can construct an initial "target set" for the appended state by constraining $\left(V^{\underline{\theta}}, B^{\underline{\theta}}\right)$ to $\mathbb{R} \times \{0\}$. By the same argument as for the boundary conditions at the terminal time, the costate vector $\left(\varphi^{\underline{\theta}}, \psi^{\underline{\theta}}\right)$ must be orthogonal to the tangent space of this target set at $\left(V^{\underline{\theta}}, B^{\underline{\theta}}\right)$.

To derive interpretable necessary conditions for (V_*, l_*), we must simplify the Maximum Principle conditions (47-49). Regarding the boundary conditions (48), the simple fact that the tangent space of $\mathbb{R} \times \{0\}$ at any point is again $\mathbb{R} \times \{0\}$ implies that the costate conditions are equivalent to

$$\varphi^{\underline{\theta}} = \varphi^{\bar{\theta}} = 0. \tag{50}$$

Regarding the evolution equations (47), we first use the relation (31) to compute

$$c_V\left(V_*^{\theta}, l_*^{\theta}\right) = \frac{1}{U_c\left(c\left(V_*^{\theta}, l_*^{\theta}\right), l_*^{\theta}\right)}.$$

We can then write the evolution equation for φ as

$$\dot{\varphi}^{\theta} = -\left[G'\left(V_*^{\theta}\right) + \frac{\psi}{U_c^{\theta}}\right] f(\theta) + \varphi^{\theta} \frac{l_*^{\theta}}{\theta} \frac{U_{cl}^{\theta}}{U_c^{\theta}}. \tag{51}$$

It should be recognized that this is simply a linear first-order differential equation in φ with variable coefficients. The general solution to this differential equation is given by

$$\varphi^{\theta} = \exp\left(-\int_{\theta}^{\bar{\theta}} \frac{l_*^s}{s} \frac{U_{cl}^s}{U_c^s} ds\right)\left(A + \int_{\theta}^{\bar{\theta}} \left[G'\left(V_*^s\right) + \frac{\psi}{U_c^s}\right] f(s) \exp\left(\int_{s}^{\bar{\theta}} \frac{l_*^t}{t} \frac{U_{cl}^t}{U_c^t} dt\right) ds\right).$$

Using the boundary condition $\varphi^{\bar{\theta}} = 0$, we can solve for the constant $A = 0$ and combine the exponential terms to find

$$\varphi^{\theta} = \int_{\theta}^{\bar{\theta}} \left[G'\left(V_*^s\right) + \frac{\psi}{U_c^s}\right] f(s) \exp\left(-\int_{\theta}^{s} \frac{l_*^t}{t} \frac{U_{cl}^t}{U_c^t} dt\right) ds. \tag{52}$$

The other boundary condition $\varphi^{\underline{\theta}} = 0$ then implies

$$0 = \int_{\underline{\theta}}^{\bar{\theta}} \left[G'\left(V_*^s\right) + \frac{\psi}{U_c^s}\right] f(s) \exp\left(-\int_{\underline{\theta}}^{s} \frac{l_*^t}{t} \frac{U_{cl}^t}{U_c^t} dt\right) ds. \tag{53}$$

To simplify the Hamiltonian maximization condition (49), note first that the Hamiltonian is differentiable in l when $l > 0$. For any θ such that $l_*^\theta > 0$, we then have the first order condition

$$\begin{aligned} 0 &= H_l\left(V_*^\theta, B_*^\theta, l_*^\theta, \varphi, \psi, -1, \theta\right) \\ &= \psi\left(c_l^* - \theta\right) f(\theta) - \frac{\varphi^\theta}{\theta}\left[l_*^\theta U_{cl}^\theta c_l^\theta + l_*^\theta U_{ll}^\theta + U_l^\theta\right]. \end{aligned} \tag{54}$$

Again using the relation (31), we can calculate

$$c_l^* = -\frac{U_l^*}{U_c^*}.$$

The first order condition (54) is then

$$0 = \psi\left(\frac{U_l^*}{U_c^*} + \theta\right) f(\theta) + \frac{\varphi^\theta}{\theta}\left[-l_*^\theta \frac{U_l^*}{U_c^*} U_{cl}^\theta + l_*^\theta U_{ll}^\theta + U_l^\theta\right]. \tag{55}$$

Equations (41), (42), (53), and (55) yield the conditions given in the statement of the theorem, and equation (52) gives the expression for φ. □

4.4 Example: Quasilinear Preferences

An illuminating special case of Mirrlees's general setting is found when agents are assumed to have quasilinear preferences of the form

$$U(c, l) = c - v(l). \tag{56}$$

Under this specification, it is possible to derive a formula for the marginal tax rate T' using the conditions (43-46). This task was first undertaken in [1], and our analysis below is substantially similar.

To fix the setting, we will assume preferences are given by (56), where $v : [0, 1) \to \mathbb{R}$ is strictly increasing, strictly convex, and continuously differentiable. It is also assumed that $G'(v(l)) \to \infty$ quickly enough as $l \to 0$ so that the non-negativity constraint on the tax policy can be neglected. Finally, it will be convenient to assume that the skill density f is strictly positive. The optimality conditions (43-46) then take the form

$$V_*^\theta = V_*^{\underline{\theta}} + \int_{\underline{\theta}}^{\theta} \frac{l_*^s}{s} v'\left(l_*^s\right) ds \quad \forall \theta \in \left[\underline{\theta}, \bar{\theta}\right], \tag{57}$$

$$0 = \int_{\underline{\theta}}^{\bar{\theta}} E - \left(\theta l_*^\theta - \left(V_*^\theta + v\left(l_*^\theta\right)\right)\right) f(\theta)\, d\theta, \tag{58}$$

$$0 = \psi\left(-v'\left(l_*^\theta\right) + \theta\right) f(\theta) - \frac{\varphi^\theta}{\theta}\left[l_*^\theta v''\left(l_*^\theta\right) + v'\left(l_*^\theta\right)\right] \quad \text{if } l_*^\theta > 0, \tag{59}$$

$$0 = \int_{\underline{\theta}}^{\bar{\theta}} \left[G'\left(V_*^s\right) + \psi\right] f(s)\, ds, \tag{60}$$

where $\varphi^\theta = \int_\theta^{\bar{\theta}} \left[G'\left(V_*^s\right) + \psi\right] f(s)\, ds$ and ψ is a constant. In particular, condition (59) can be written

$$\psi\left(\theta - v'\left(l_*^\theta\right)\right) f(\theta) = \frac{1}{\theta}\left[v'\left(l_*^\theta\right) + l_*^\theta v''\left(l_*^\theta\right)\right] \int_\theta^{\bar{\theta}} \left[G'\left(V_*^s\right) + \psi\right] f(s)\, ds \quad \forall \theta : l_*^\theta > 0. \tag{61}$$

Suppose that at the optimum (V_*, l_*), the sufficient secondary condition holds, and θl_*^θ is strictly increasing in θ. Then we can define a tax schedule T_* on the range of θl_*^θ by

$$T_*\left(\theta l_*^\theta\right) := \theta l_*^\theta - c\left(V_*^\theta, l_*^\theta\right).$$

Suppose also that this function T_* is differentiable. Then the first order condition from the agent's problem is

$$\theta\left(1 - T_*'\left(\theta l_*^\theta\right)\right) = v'\left(l_*^\theta\right). \tag{62}$$

Substituting this expression into the left side of (61) and dividing both sides by $\psi \neq 0$,

$$T_*'\left(\theta l_*^\theta\right) = \frac{1}{\theta^2}\left[v'\left(l_*^\theta\right) + l_*^\theta v''\left(l_*^\theta\right)\right] \frac{1}{f(\theta)\psi} \int_\theta^{\bar{\theta}} \left[G'\left(V_*^s\right) + \psi\right] f(s)\, ds \quad \forall \theta : l_*^\theta > 0. \tag{63}$$

If T_* is twice continuously differentiable, then using (62) it is straightforward to show that the skill elasticity of effective labor is given by

$$e(\theta) := \frac{\theta}{\theta l_*^\theta} \frac{d}{d\theta}\left(\theta l_*^\theta\right) = \frac{v'\left(l_*^\theta\right)}{l_*^\theta v''\left(l_*^\theta\right)}.$$

We can then divide both sides of (63) by $1 - T_*'\left(\theta l_*^\theta\right) = \frac{1}{\theta} v'\left(l_*^\theta\right)$ to find

$$\begin{aligned}
\frac{T_*'}{1-T_*'} &= \frac{1}{\theta}\left[1+\frac{l_*^\theta v''}{v'}\right]\frac{1}{f(\theta)\psi}\int_\theta^{\bar{\theta}}\left[G'\left(V_*^s\right)+\psi\right]f(s)f(s)\,ds \\
&= \left[1+\frac{1}{e(\theta)}\right]\left(\frac{1}{\psi}\int_\theta^{\bar{\theta}}\left[G'\left(V_*^s\right)+\psi\right]f(s)\,ds\right)\frac{1}{\theta f(\theta)} \\
&= \underbrace{\left[1+\frac{1}{e(\theta)}\right]}_{A(\theta)}\underbrace{\left(\frac{1}{\psi(1-F(\theta))}\int_\theta^{\bar{\theta}}\left[G'\left(V_*^s\right)+\psi\right]f(s)\,ds\right)}_{B(\theta)}\underbrace{\frac{1-F(\theta)}{\theta f(\theta)}}_{C(\theta)} \quad \forall\theta : l_*^\theta > 0. \quad (64)
\end{aligned}$$

Here $F(\theta) := \int_{\underline{\theta}}^{\theta} f(\theta)\,d\theta$. We have thus arrived at Diamond's well-known "ABC" formula for the marginal tax T_*' (see [1]).

References

[1] Diamond, P. (1998), "Optimal Income Taxation: An example with a U-Shaped Pattern of Optimal Marginal Tax Rates." American Economic Review, 88(1), 83-95.

[2] Evans, L.C. (2010) *Partial Differential Equations*. Providence: American Mathematical Society.

[3] Liberzon, D. (2012) *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton: Princeton University.

[4] Luenberger, D.G. (1969) *Optimization by Vector Space Methods*. New York: John Wiley & Sons.

[5] Mirrlees, J. (1971), "An Exploration in the Theory of Optimum Income Taxation." Review of Economic Studies, 38(2), 175–208.