MATH/AMATH 509

OPTIMAL CONTROL

Lecture notes by R. T. Rockafellar (1990)

Part 1

INTRODUCTORY IDEAS

1 States and Controls

In mathematical modeling the concept of a "state" is very general, but it typically refers to a specification of the quantities that fully describe, at a particular moment in time, the system being modeled. For our purposes here, this will mean the specification of a vector $x = (x_1, \ldots, x_n)$ in the space R^n.

For example, in describing the solar system on an elementary level in terms of only the nine planets, we might have x consist of three position coordinates and three velocity coordinates for each planet. This is because the specification of these coordinates at any particular time is supposed to determine by the laws of classical mechanics all future positions and velocities of the planets through the effects of gravity. A state of the solar system would then be a vector in R^{54}.

The state of an economy, like that of the United States, might be identified with a vector having thousands of components. These components could give the current amounts of amounts of employment in various sectors, the current inventories of various goods, interest rates for various types of borrowing and lending, and so forth.

Our interest is drawn not just to one fixed state, but to the way that states may change or evolve in time. We usually want to think therefore of a varying state $x(t) = (x_1(t), \ldots, x_n(t))$, which can be pictured as a moving point in R^n. Fundamentally then we focus on the study of a function from some real interval $[t_0, t_1]$ to R^n, which is called a *state trajectory*.

This trajectory could be denoted again simply by the symbol x, but to avoid confusion over the point of view being taken at any moment, whether that of a single vector or a vector-valued function, it is often helpful to write $x(\cdot)$ in the latter case. Derivatives of $x(t)$ with respect to t, when they exist, are denoted by $\dot{x}(t)$, $\ddot{x}(t)$, ...

The evolution of states is very often governed by differential equations with initial conditions, such as

$$\dot{x}(t) = f(t, x(t)) \text{ with } x(t_0) = a_0,$$

$$\ddot{x}(t) = f(t, x(t), \dot{x}(t)) \text{ with } x(t_0) = a_0,\ \dot{x}(t_0) = \dot{a}_0,$$

which are said to be of first order, second order, and so forth. Under fairly general assumptions, which will be discussed later, such an equation with initial conditions determines a unique trajectory $x(\cdot)$. Of course this kind of evolution is utterly deterministic. Possibilities for controlling the evolution can be introduced, though, as will be seen presently.

Example: An object moved by forces. Let $x(t)$ denote the position of the object in R^3. Let m be its mass, and let the force vector acting at time t and position $x(t)$ be $\varphi(t, x(t))$. Then, according to Newton's Law, the motion of the object will be governed by

$$\ddot{x}(t) = m^{-1}\varphi(t, x(t)), \text{ with } x(t_0) = a_0, \; \dot{x}(t_0) = \dot{a}_0.$$

The initial position a_0 and initial velocity $\dot{a}_0$ will thus determine the position $x(t)$ at all times $t > t_0$.

It may be possible, however, for us to alter the magnitude or direction of the force in order to influence how the object moves. Let us model this capability by saying that the force exerted at time t depends not only on t and the current position, but on a vector $u(t) \in R^d$ whose components are parameter values that we may specify as functions of t. For instance, these parameters might represent the settings on various levers or dials that operate a piece of equipment. The differential equation then takes the form

$$\ddot{x}(t) = m^{-1}\varphi(t, x(t), u(t)) \text{ with } x(t_0) = a_0, \; \dot{x}(t_0) = \dot{a}_0.$$

We choose the function $u(\cdot)$, and a certain trajectory $x(\cdot)$ results—depending of course on the initial position a_0 and velocity $\dot{a}_0$, which likewise could be subject to our choice.

Control of general dynamical systems. Any system governed by an ordinary differential equation can be expanded to a control system by introducing additional variables—called *control* variables in contrast to the given *state* variables. The control variables are the components of a vector $u = (u_1, \ldots, u_d)$ in a control space R^d. The differential equation then takes a form such as

$$\dot{x}(t) = f(t, x(t), u(t)) \text{ with } x(t_0) = a_0,$$

$$\ddot{x}(t) = f(t, x(t), \dot{x}(t), u(t)) \text{ with } x(t_0) = a_0, \; \dot{x}(t_0) = \dot{a}_0,$$

or perhaps something of still higher order. Luckily it's not necessary to develop the theory with all these different orders of differential equations running parallel to each other, which would be very burdensome. First order covers everything, because of a standard trick.

Recall that for the second-order equation just written we can achieve a first-order formulation by introducing a new state vector $y(t) = (y_1(t), y_2(t))$, where y_1 stands for x and y_2 for $\dot{x}$. In this manner we get an equivalent equation in R^{2n}:

$$\dot{y}(t) = g(t, y(t), u(t)) \text{ with } y(t_0) = b_0,$$

where $b_0 = (a_0, \dot{a}_0)$ and

$$g(t, y, u) = g(t, y_1, y_2, u) = (y_2, f(t, y_1, u)).$$

In the case of the object moved by forces, the new state vector would be comprised of the position and velocity vectors of the object and would be an element of R^6.

For ordinary differential equations of order higher than 2, the reduction to the first-order case follows a similar pattern.

Constraints. Working with a control system

$$\dot{x}(t) = f(t, x(t), u(t)) \text{ with } x(t_0) = a_0,$$

we may want to achieve a certain specified state a_1 at a later time time t_1: $x(t_1) = a_1$. This would be a *terminal* constraint. More generally such a constraint could take the form $x(t_1) \in E_1$, or even $(t_1, x(t_1)) \in E_1$ in the case of variable t_1, where the set E_1 might be defined by a collection of equations or inequalities.

There could be restrictions on the control parameter values at our disposal, as represented by *control* constraints:

$$u(t) \in U, \text{ or } u(t) \in U(t), \text{ or } u(t) \in U(t, x(t)).$$

A simple example of the first case can be found in the control of forces by the equipment that generates them. Each control variable u_i might be imagined as ranging over an interval which corresponds, say, to the possible positions of a certain lever or the settings of a certain dial. The intervals can all be normalized to $[0, 1]$ by a change of scale. Then $U = [0, 1] \times \ldots [0, 1]$.

The general third case of control constraints, where $u(t) \in U(t, x(t))$, is said to involve *partial feedback*. The trajectory $x(\cdot)$ depends on the chosen control function $u(\cdot)$, but the control vector $u(t)$ at time t is restricted to a set that depends in part on the current state $x(t)$.

An example of such forced partial feedback in control can be seen in the modeling of an inventory system, where the components of $x(t)$ represent the supplies of various goods being kept in a storage facility, say, while the components of $u(t)$ describe activities like production, consumption or maintenance of these goods. The extent of the activities could be affected by the current supplies. For instance, if the capacity for storing a particular type of good is completely used up, then for the time being no more of that good can be added to the inventory. Another possibility is that storage of a particular good requires a lot of maintenance effort, and this restricts the amount of effort that can be devoted to processing a different kind of good into or out of inventory. Budget limits could place joint restrictions on $x(t)$ and $u(t)$ too.

Besides control constraints there could be *state* constraints of the form

$$x(t) \in X, \text{ or } x(t) \in X(t).$$

Controllability. An obvious question under all these circumstances is whether there even exists a control function $u(\cdot)$ which, along with the corresponding trajectory $x(\cdot)$ obtained from the differential equation and initial condition, satisfies all the given constraints. This is the issue of *controllability*. A substantial part of control theory is devoted to the study of controllability. One hopes, of course, to have posed the situation and model in such a way that controllability is present.

Then, however, a second issue arises. Usually there will be more than one control function for which all the conditions are met, and we have to decide which such function should be selected. An opportunity therefore arises for optimization: We can try to choose the $u(\cdot)$ that is "best" according to some criterion.

Objectives. One simple concept of optimality in the case of a terminal constraint like $(t_1, x(t_1)) \in E_1$ is to have this condition be achieved with the lowest possible value of $t_1 > t_0$. Such would be a *minimum time* problem of optimal control. More generally one could try to choose the control function to minimize an expression of the form

$$\int_{t_0}^{t_1} f_0(t, x(t), u(t))dt + k(t_1, x(t_1)).$$

Either of the functions f_0 or k in this expression could vanish. The case where f_0 vanishes and $k(t_1, x(t_1)) = t_1$ is the one of minimum time, as already described. In some applications the integral term could be related to energy, more specifically to the amount of work performed. Or it could represent accumulating "costs," in contrast to the final cost given by the term at time t_1.

Control theory as a branch of optimization. The framework we have arrived at is one where a set $\mathcal{U}$ of control functions $u(\cdot)$ is specified by various constraints, and over this set $\mathcal{U}$, if it is nonempty, a certain "cost" expression is to be minimized. (It would equally be possible to speak of maximizing something, of course, but minimization is adopted as the standard formulation.) In this we have a typical problem of optimization, with its questions of *feasibility* and *optimality.* Feasibility is studied more specially under the heading of controllability. Optimality is studied in terms of finding useful characterizations of the control functions and trajectories that achieve minimal "cost."

Optimal control problems are distinguished from other kinds of optimization problems, however, in several important respects. Chief among these is the fact that the set $\mathcal{U}$ over which the minimization takes place is a set of *functions of time,* normally continuous time. Further, the constraints defining this set involve a differential equation. We are concerned therefore with *infinite-dimensional* optimization.

The specification of the particular *function space* in which to work becomes an important consideration in such a context. Mathematical technicalities take on a much larger role than in other branches of optimization such as nonlinear programming.

Note that the question of whether we are free to choose the terminal time t_1 as part of the specification of the control function $u(\cdot)$, or whether t_1 is fixed for us in advance, poses a serious complication in the choice of function space. Most of the standard function spaces are linear spaces (vector spaces) involving functions over a *fixed* interval. Two control functions defined over different time intervals can't be added together and therefore can't ordinarily be regarded as belonging to the same standard space. Fortunately it is possible to develop much of the theory as if t_1 were fixed and then by various mathematical devices derive from it the results desired for problems in which t_1 is not fixed.

2 Open–Loop Versus Closed–Loop Control

The overview of control theory given so far is one-sided. It neglects one of the most crucial ideas in the subject, that of *feedback.* We have mentioned that situations may arise where the control vectors available at a particular time may be subject to restrictions that depend on the current state. More important, though, is that we may wish to develop the solution to an optimal control problem in the form of a *feedback law*:

$$u(t) = c(t, x(t)).$$

Such a law represents a way of automatically executing the solution to a control problem by means of a mechanism that engineers could build into a system. Instead of imagining that we first turn our computers to minimize a cost expression over a set $\mathcal{U}$ of control functions to get an optimal $u(\cdot)$, and then plug this $u(\cdot)$ into the control equation to get

$$\dot{x}(t) = g(t, x(t)) \text{ with } g(t, x) = f(t, x, u(t)),$$

we imagine replacing the given control system by a deterministic system

$$\dot{x}(t) = g(t, x(t)) \text{ with } g(t, x) = f(t, x, c(t, x)).$$

This needs some explanation. From the usual optimization point of view we would be focusing on a single problem where the initial time t_0 and state a are fixed. The solution to the problem would be a control function $u(\cdot)$, not necessarily unique, which is calculated in advance and then "implemented." We would make no input during the implementation itself; the system would just follow the trajectory we have prescribed.

There are difficulties with such an approach in practice. It only tells us what to do for a single choice of initial time and state. If we were to change our minds, or try to solve slightly different problem a little later, we would have to carry out the entire computation all over again. Most seriously, the approach relies too heavily on the deterministic framework and doesn't allow us to make corrections based on random events or errors that might intervene during implementation. It does not seem to be the right way to be thinking about controlling an actual complex system in real time.

Quite possibly we don't know with absolute certainty what the state of the system is at any given moment. The presumed initial state may have been obtained from measurements that are subject to inaccuracies. If after following the implementation of a control $u(\cdot)$ for a while we see that the system doesn't seem to be doing what it is supposed to, shouldn't we be able to take some recourse, and if so, how?

Our need for recourse increases when the system is not totally cut off from the world and may have to react to outside events that have not been included in the model. For example, we may be trying to keep a telescope or antenna focused on a particular target. A sudden gust of wind or a minor earth tremor could knock it out of line. A control system would then come into play in order to restore it to the proper position. This would be activated with the displaced position of the telescope as the initial state. But while the restoration is going on, another gust of wind could come along.

Random events and errors aren't the only thing to worry about. Our entire model of the system being controlled may be an oversimplification, and indeed it almost should be: That is the nature of mathematical modelling and one of its stunning virtues. Tremendous success can be achieved by working with just a few key features in a situation and neglecting the rest. Mathematics is, after all, the science of abstraction. But we hope at the same time to be protected in some degree from the effects of oversimplification. We can't fall into the trap of taking our model *too* seriously, where practical consequences are concerned.

It is in this setting that we can appreciate the importance of having a control law that does not require us, each time something happens or seems threatening to go wrong, to turn on our whole optimal-control-finding computer program again and grind out an updated control function $u(\cdot)$ that is henceforth supposed to be optimal for all time. Instead we can just have a mechanism that reads the current time t and state $x(t)$ of the system from various data inputs and sensors, and on the basis of this transmits an appropriate control vector $u(t)$ to the controlling equipment. If a disturbance affects the system, no matter: The mechanism will perceive the new state and alter the control vector accordingly.

In such a case it is customary to speak of *closed–loop* control. The feedback loop from states to controls is "closed." The other case, where the optimal control function is determined in advance and then merely plugged in, is accordingly termed *open–loop* control. No feedback at all occurs from states to controls during execution; the feedback loop is "open."

A great advantage of the closed–loop point of view is that it suggests ways of proceeding even in the absence of full knowledge or treatment of what might be optimal. One can experiment with various control laws $u(t) = c(t, x(t))$ to see how well they work. Through experience, one can select a law that operates more or less satisfactorily and is reasonably easy to implement. It is no wonder, therefore, that in closed–loop theory there is a large emphasis on so-called *suboptimal* control, and that this branch of the subject is much more heavily drawn upon by engineers than open–loop theory. One of the most significant facts learned through applications of control has been that relatively elementary control laws can serve well in obtaining acceptable performance of complex systems.

We shall see, however, that closed–loop control depends very much in concept and insight on open–loop control, which is therefore more fundamental. It is by studying the optimality rules in open–loop control that we can discover the kinds of control laws that lead to the best performance and correctly account for the presence of various constraints.

Tracking problems with penalties. These ideas can understood more fully in the context of tracking problems in engineering. In such problems one desires to guide a system "as well as one can" along a pre-specified trajectory $\bar{x}(\cdot)$. Suppose the control system has the *autonomous* form

$$\dot{x}(t) = f(x(t), u(t)) \text{ with } u(t) \in U, x(t) \in X.$$

(Autonomous in referring to differential equations means time-independent: here f, U, and X don't themselves depend on t.) Let T denote the future time marking the end of the period over which we are trying to track the trajectory $\bar{x}(\cdot)$. For reasons soon to become apparent, let τ rather than t_0 denote the time at which we wish to start the tracking. The control system might not have the capability of tracking the trajectory $\bar{x}(\cdot)$ exactly, and anyway we may have to start from some point that is not on the trajectory, namely $x(\tau) = a$. (This may be due to a disturbance, as already explained.) What we can attempt to do in this case is to choose the control function $u(\cdot)$ to minimize a *penalty cost* expression

$$\int_{\tau}^{T} [p(x(t) - \bar{x}(t)) + q(u(t))]dt,$$

where $q(\cdot)$ gives the cost associated with the choice of control vector and $p(\cdot)$ assigns a positive penalty to all vectors different from 0. In our earlier description of optimal control, this would correspond of course to choosing

$$f_0(t, x, u) = p(x - \bar{x}(t)) + q(u), \quad l \equiv 0.$$

In particular one might have

$$p(x(t) - \bar{x}(t)) = \alpha |x(t) - \bar{x}(t)|^2,$$

where $|\cdot|$ denotes the euclidean norm and $\alpha > 0$ is a weighting coefficient.

This sounds merely like open–loop control, but the crucial idea is that we consider the problem not just for one initial pair (τ, a), but for all such possible pairs. In solving the problem for a particular pair (τ, a), we can expect to obtain a certain control function $u_{\tau,a}$ over the interval $[\tau, T]$. Let the initial control vector specified by this function be denoted by $c(\tau, a)$; thus $c(\tau, a) = u_{\tau,a}(\tau)$. We then have in principle a control law that should work at all times and in all states: $u(\tau) = c(\tau, x(\tau))$, where notation can now be changed from τ to t.

No more than a heuristic introduction to such a way of thinking is intended at this stage. There are many serious technical questions involved. Among them: Does the control function $u_{\tau,a}$ even exist? What if it is not unique? What will be the mathematical nature of the function $c(\cdot, \cdot)$ derived in this way, and can it be legitimately substituted in the given differential equation? We will address such questions later.

Nonetheless some lessons are apparent. The most important of these is that the form of the closed–loop feedback law obtained through such considerations depends on the optimality conditions for the open–loop control problems solved to get the functions $u_{\tau,a}(\cdot)$.

At present, most applications of control theory to tracking problems assume that the functions p and q in the penalty cost expression are *quadratic*, the differential equation is *linear*, and that constraints on controls and states are not present at all: $U = R^d$, and $X = R^n$. This is due to the fact that optimal control laws can readily be calculated for this more elementary case by what amounts to the method just described. Constraints on controls and states truly are present, of course, in most situations, so this is an illustration of mathematical convenience getting the better of reality.

Mathematical models don't have to be completely realistic, or we would never achieve anything in applying mathematics. The main test is whether a model provides good insights and useful, reasonably reliable results when used judiciously. Still, we should not lose sight of the route by which a particular methodology has been reached. By following the route more closely at a later time, and with the advantage or newer theoretical ideas, it may be possible to come up with significant improvements. This is certainly one of the motivations for the continued study of open–loop control in association with serious applications of control, even if on the surface it may appear less practical than closed–loop control.

Restoration of equilibrium. The most important case of tracking problems occurs when $\bar{x}(t) \equiv \bar{x}_0$, a point representing a state of equilibrium. We wish to stay in this state of equilibrium and could do so, once we are there, with the control $u(t) \equiv 0$: We suppose we have $f(\bar{x}_0, 0) = 0$, so that the trajectory $x(t) \equiv \bar{x}_0$ does correspond to the control $u(t) \equiv 0$ in the differential equation $\dot{x}(t) = f(t, x(t), u(t))$.

The question then is what to do if we should find ourselves out of equilibrium at time τ in some state $a \neq \bar{x}_0$. We look for an appropriate control law that will optimally, or at least suboptimally, get us back into equilibrium.

In this situation it is especially apparent that, while optimality would be nice, the matter of being able to move reliably back toward the equilibrium state from wherever we are could in itself be far more important. It is the *stability* of the system that is at issue. We seek a control law guaranteeing that once a disturbance has intervened and the system is momentarily out of equilibrium, it won't just slip into worse and worse states until a breakdown occurs. An example would be the control of mechanisms such as wing flaps on an aircraft in flight.

An optimal control law that in principle did the best possible job of restoring equilibrium would be worthless if, in practice, it were so delicate to implement that it might fail as a result of small errors. But too much emphasis on optimality is only potential one source of trouble. Any control law is bound to be based on a mathematical model of the system being controlled, and as already mentioned, this might be quite crude. While this is often a virtue, it can also lead to danger if we become too complacent about it. The convenience of deriving nice control laws by using quadratic penalties and linearized dynamics, neglecting constraints on states and controls, is wonderful, but it might afford *too local* an approximation in some cases. The control scheme could then go awry if disturbances are larger than anticipated.

A final comment at this point about equilibrium problems is that the choice of a particular terminal time T seems rather arbitrary. For this reason it often makes sense pass to the case where $T = \infty$. One speaks then of an *infinite horizon* control problem. The study of such problems requires still greater battles with mathematical technicalities, but somewhat paradoxically it often leads to control laws that are simpler and more agreeable in the end.

Closely connected to this is the idea of trying to follow from any disturbed state the *minimum time* trajectory for getting back to equilibrium, instead of a trajectory that minimizes a particular penalty expression. Such an approach too frees us from having to specify a finite horizon T.

3 Variational Problems

As mathematical theories go, optimal control theory is relatively new. It came into its own in the late 1950's and 1960's. Yet it is deeply connected with a far older subject which has been one of the most fertile in all the history of mathematics, namely, the *calculus of variations.*

The calculus of variations seems at first sight to have an entirely different orientation, and indeed the concept of "control" never occurs in its confines. It does concern the optimization of trajectories in R^n, however. These trajectories, because they are not usually visualized as arising from differential equations, are often referred to as *arcs* instead. (The word "arc" is a sort of synonym for "parametrized curve.")

Problems in the calculus of variations can have many forms, but the most fundamental is that of a so-called *problem of Lagrange*: Minimize

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

over all arcs $x(\cdot)$ (with prescribed differentiability properties) that satisfy the endpoint conditions

$$x(t_0) = a_0 \text{ and } x(t_1) = a_1.$$

A vastly more general form is that of a *problem of Bolza*: Minimize

$$\int_{t_0}^{t_1} L_0(t, x(t), \dot{x}(t))dt + l_0(t_0, x(t_0), t_1, x(t_1))$$

subject to constraints

$$L_i(t, x(t), \dot{x}(t)) \begin{cases} \leq 0 & \text{for } i = 1, \ldots, r, \\ = 0 & \text{for } i = r+1, \ldots, r_0, \end{cases}$$

$$l_j(t_0, x(t_0), t_1, x(t_1)) \begin{cases} \leq 0 & \text{for } j = 1, \ldots, s, \\ = 0 & \text{for } j = s+1, \ldots, s_0. \end{cases}$$

Note that endpoint constraints in this formulation do cover the simpler ones for a problem of Lagrange, which could be expressed as a system of $2n$ equations on the coordinates of the vectors $a_0 = (a_{01}, \ldots, a_{0n})$ and $a_1 = (a_{11}, \ldots, a_{1n})$, namely

$$x_j(t_0) - a_{0j} = 0 \text{ and } x_j(t_1) - a_{1j} = 0 \text{ for } j = 1, \ldots, n.$$

The constraints in a problem of Bolza can also be viewed more abstractly in the form

$$(t, x(t), \dot{x}(t)) \in D \text{ and } (t_0, x(t_0), t_1, x(t_1)) \in E$$

for certain sets D and E.

Example: A hanging cable. Classical problems in the calculus of variations often use t just as a general parameter, not necessarily interpreted as time. For example, there is the problem of a hanging cable, fixed at both ends and having a given length. What is the shape of the curve it assumes, this being the curve that minimizes the potential energy? This can be set up in various ways, but one of them is to use t as the arc length parameter along the cable, which for simplicity we can suppose not to be stretchable. The constraints only concern fixed endpoints of the cable, so the problem is one of Lagrange in which the t interval is fixed. More generally, though, there might be obstructions that prevent the cable from hanging freely. Perhaps it has to drape partly over a large rock. In this case constraints must be imposed on the positions $x(t)$ that can be taken on, and the problem becomes one of Bolza. If the endpoints of the cable are not completely fixed, e.g. because they are fastened to poles that might bend, one would get a problem in which the endpoints just have to satisfy certain equations that characterize the range of locations that they could taken on. If the cable is stretchable, the t interval might be modeled as variable too.

Calculus of variations as a branch of optimal control. Every problem in the calculus of variations can be regarded mathematically as a special case of a problem in optimal control. All we need to do is introduce the trivial differential equation

$$\dot{x}(t) = u(t), \text{ i.e. take } f(t, x, u) \equiv u.$$

Thus in the Bolza format we minimize

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

subject to this differential equation and

$$(t, x(t), u(t)) \in D \text{ and } (t_0, x(t_0), t_1, x(t_1)) \in E.$$

The condition $(t, x(t), u(t)) \in D$ could be rewritten notationally as $u(t) \in U(t, x(t))$ for a certain choice of the set $U(t, x(t))$.

Optimal control as a branch of the calculus of variations. What is less apparent, but will be quite important for us theoretically, is the converse fact: *Every problem of optimal control can be formulated as a seemingly unconstrained problem of Bolza in the calculus of variations* —provided that one is willing to admit functions L and l that are extended-real-valued, i.e. can take on not only real values but ∞. Technically this is more complicated to achieve with rigor, but a rough sketch of the idea at this early stage will nonetheless be helpful. It will serve to underline the essential unity of control theory and the calculus of variations and also to point the way toward some of the technical developments that will later have to be undertaken.

Let us first consider not a control problem, but a problem of Bolza with constraints expressed in abstract form, as above. Such a problem can be rewritten as a seemingly *unconstrained* problem of Bolza, as follows. We define functions $\bar{L}$ and $\bar{l}$ by

$$L(t,x,v) = \begin{cases} L_0(t,x,v) & \text{if } (t,x,v) \in D \\ \infty & \text{if } (t,x,v) \notin D \end{cases}$$

$$l(t_0,x_0,t_1,x_1) = \begin{cases} l_0(t_0,x_0,t_1,x_1) & \text{if } (t_0,x_0,t_1,x_1) \in E \\ \infty & \text{if } (t_0,x_0,t_1,x_1) \notin E. \end{cases}$$

We then consider the problem of minimizing the expression

$$\int_{t_0}^{t_1} L(t,x(t),\dot{x}(t))dt + l(t_0,x(t_0),t_1,x(t_1))$$

with no constraints at all imposed explicitly on the arc $x(\cdot)$.

In fact this reformulated problem is equivalent to the given problem of Bolza and merely has the constraints represented implicitly instead of explicitly. We can't fully justify this without getting prematurely into too many technical details, but the crucial idea is that of representing constraints by imposing infinite penalties when they are violated. In the reformulated problem, we have no interest in arcs $x(\cdot)$ for which the expression being minimized has the value ∞. Arcs for which the expression is not ∞ must, however, under the right technical assumptions and interpretation, satisfy

$$L(t,x(t),\dot{x}(t)) < \infty \text{ and } l(t_0,x(t_0),t_1,x(t_1)) < \infty.$$

This means of course that such arcs satisfy the desired constraints

$$(t,x(t),u(t)) \in D \text{ and } (t_0,x(t_0),t_1,x(t_1)) \in E.$$

Now let us look at an optimal control problem where the expression

$$\int_{t_0}^{t_1} f_0(t, x(t), u(t))dt + k(t_1, x(t_1)).$$

is to be minimized subject to constraints

$$u(t) \in U(t, x(t)), \quad x(t) \in X(t), \quad (t_1, x(t_1)) \in E_1,$$

and of course

$$\dot{x}(t) = f(t, x(t), u(t)) \text{ with } x(t_0) = a_0.$$

Let D be the set of all $(t, x, v) \in R \times R^n \times R^n$ such that $x \in X(t)$ and there exists a vector u satisfying

$$u \in U(t, x) \text{ and } f(t, x, u) = v.$$

For each choice of $(t, x, v) \in D$, let $L(t, x, v)$ be the minimum value of $f(t, x, u)$ over the set of u vectors just described, but for $(t, x, v) \notin D$ define $L(t, x, v) = \infty$. Thus $L(t, x(t), \dot{x}(t))$ will be the cheapest cost at which the velocity $\dot{x}(t)$ can be achieved by some choice of control vector at time t, if it can be achieved at all, with an infinite penalty if it can't. Similarly define the function l by

$$l(t_0, x_0, t_1, x_1) \begin{cases} k(t_1, x_1) & \text{if } (t_0, x_0, t_1, x_1) \in R \times \{a_0\} \times E_1 \\ \infty & \text{otherwise.} \end{cases}$$

Our claim (heuristic for now) is that under these definitions the expression

$$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt + l(t_0, x(t_1), t_1, x(t_1))$$

will take the value ∞ unless the arc $x(\cdot)$ is such that

$$x(t_0) = a_0, \quad (t_1, x(t_1)) \in E_1, \quad x(t) \in X(t),$$

and there exists a function $u(\cdot)$ with $u(t) \in U(t, x(t))$ and $f(t, x(t), u(t)) = \dot{x}(t)$. Moreover then the expression in question is the lowest value of

$$\int_{t_0}^{t_1} f_0(t, x(t), u(t))dt + k(t_1, x(t_1))$$

over all control functions $u(\cdot)$ corresponding in this way to $x(\cdot)$. Thus we have a problem of Bolza whose minimizing arcs $x(\cdot)$ are the optimal arcs for the given control problem. From any such arc $x(\cdot)$ it is possible to recover the corresponding optimal control function $u(\cdot)$.

Nonsmooth analysis. The seemingly unconstrained problems of Bolza that furnish this bridge between calculus of variations problems in $x(\cdot)$ and control problems in $u(\cdot)$ are far from being the kind covered by classical theory. In the past, only *finite* functions L and l were admitted, whereas these can take on ∞. Moreover they need not even be continuous functions, much less differentiable. Classical methods that rely on taking derivatives of L and l can't be applied.

All is not hopeless, though. There is now in existence a version of calculus that can effectively be used in handling such functions L and l. It is called *nonsmooth analysis.* ("Smoothness" is a synonym for continuous differentiability.) This form of analysis works with generalized concepts of continuity and differentiability that remain viable even when function values can jump to ∞. It makes possible the derivation of very general optimality conditions that reduce to well known conditions when applied to standard types of problems.

By means of nonsmooth analysis, a *neoclassical* theory of variational problems can be built up which emulates classical theory in its fundamental concepts but at the same time provides a solid framework for problems in optimal control and other modern problems as well. The development of this theory is a matter of current research, but we shall at least be able to see the main outlines that are emerging. The new theoretical view offers fresh insights into many important issues.

4 The Right Spaces of Functions

It may at first seem natural, in dealing with the control equation

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

to assume that $x(\cdot)$ is differentiable and $u(\cdot)$ is continuous. In much of mathematics, especially applied mathematics, people are accustomed to taking for granted that such properties are physically reasonable and are there for the asking. To this way of thinking, based heavily on the traditions and successes of classical calculus, there is hardly any point in tampering with such fundamental assumptions, other than as a purely academic experiment. Real systems do behave "smoothly," if for no other reason than inherent frictions and lags in response that prevent true discontinuities in performance, so why focus on anything less? Anyway, how could the control equation even be interpreted if $x(\cdot)$ were not differentiable?

In adopting such an attitude, however, one would naively be missing an essential feature of control models, and more generally of any mathematical models involving optimization. When we set up a problem in terms of an optimization model, we certainly have in mind that the problem ought to have a solution. But whether it does have a solution or not is an issue that can't merely be entrusted to our intuition. It depends on technical details about which we need to exercise caution. Wishful thinking about the properties we imagine a solution ought to have could lead us to an inappropriate formulation that excludes the solution we should be looking for. Optimization problems can have unexpected complications.

Rightly or wrongly we may be confident that the real situation we are attempting to model does yield a solution, but this is almost irrelevant as far as our model is concerned, because we can't be completely sure that the model does a very good job of capturing the real situation. If we are so confident about a solution existing, then we should look to this as a test of the validity of our model: We should check to see that the problem we have formulated is guaranteed *by mathematics* to have a solution, under the assumptions we have imposed. If not, then we have failed to do a good job of mathematical modeling and may have overlooked some crucial features. Here, by the way, lies the important function of *existence proofs* in mathematics, which is sometimes unappreciated.

Example: An emergency ride. An elementary control problem will illustrate this point. An ambulance picks up an injured man to take him to the hospital a short distance away. He needs to be delivered there as soon as possible. Let us suppose the road to the hospital is straight, and there is no other traffic. We can think of the ambulance as an object of mass m represented by a point in one-dimensional space. Its motion is to be controlled by forces so that, starting at rest, it will proceed from the pick up location to the origin (the hospital) and stop there. As explained in Section 1, the state of the ambulance at any time needs to be modeled in this case as a vector in R^2, the first coordinate giving the ambulance's position and the second coordinate its velocity. The ambulance must therefore be brought from an initial state of the form $(b, 0)$ to the terminal state $(0, 0)$. The question is how to do so as quickly as possible.

The control equation will be based on Newton's Law of motion. We could go through the details of setting it up, but for the purpose at hand this is not really necessary. It is clear that the control possibilities are essentially of two kinds:

$$\begin{aligned} u_1(t) &= \text{how much the driver presses on the ambulance's accelerator,} \\ u_2(t) &= \text{how much the driver presses on the ambulance's brakes.} \end{aligned}$$

We can model the control therefore as a vector $u(t) \in [0,1] \times [0,1]$.

In this simplified situation the solution to the control problem is obvious. The driver should begin by pressing the accelerator all the way down and keep it that way until a certain critical moment, when he switches over to pressing the brakes all the way down. The switch-over is timed so that the ambulance comes to a halt exactly at the hospital. Mathematically this means that the optimal control function will have the form

$$u(t) = \begin{cases} (1,0) & \text{for } t < \tau, \\ (0,1) & \text{for } t \geq \tau, \end{cases}$$

where τ is the switch-over time. The solution $u(\cdot)$ is thus a *discontinuous* function; it is of a type known as a "bang-bang" control, because it makes use only of the extreme possibilities in the control set $U = [0,1] \times [0,1]$. Because of the discontinuities in the control, the corresponding state trajectory will exhibit discontinuities in its derivatives, and at these discontinuities certain derivatives will fail to exist, at least in the two-sided sense.

We could, of course, adopt the view that these discontinuities are an artifact caused by oversimplification. Truly the driver can't *instantaneously* shift his foot from the accelerator to the brakes, so the optimal control we have found is not implementable. Our model doesn't therefore conform to reality. But what would be the alternative? Should we try to incorporate a complete and accurate model of the motions of the driver's foot? Should we just introduce arbitrary "smoothing coefficients" into the mathematics? If the latter, how could we claim that our model was still anything other than a convenient mathematical approximation?

As long as we are going to introduce a mathematical approximation anyway—which is inherent in the nature of modeling, as already discussed—we may as well choose one that gets to the heart of the situation, which in this case is in fact an underlying tendency to discontinuity. Although the optimal control furnished by such a model may not actually be implementable, it can serve as an ideal toward which we may strive in implementation.

Piecewise continuous controls—no panacea. Examples like the preceding suggest that we might be wise in allowing for *piecewise continuous* control functions $u(\cdot)$ in the equation

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

Assuming that f itself is continuous, we would then have $\dot{x}(\cdot)$ piecewise continuous. Thus $x(\cdot)$ would be piecewise differentiable; there could be jumps in the derivative that cause abrupt "corners" on the trajectory.

The need for at least this much generalization was long ago recognized even in the calculus of variations, where arcs with discontinuous derivatives are known to occur as solutions to some problems of Lagrange in which the integrand L is infinitely many times differentiable. Most textbooks on optimal control proceed in terms of piecewise continuous controls. This approach works out less simply than one might suppose, however, and it still leaves the resulting theory vulnerable to serious questions.

The biggest question is whether solutions can be guaranteed to exist in the class of piecewise continuous controls any more than in the class of continuous controls. Without this, not much has been gained. A control problem can be thought of as a kind of machine for which the input consists of the data defining the problem, such as the function f in the dynamics, the control set U, and so forth, and the output is the solution or set of solutions. It is one thing to make assumptions about the input, which can be verified, and another in the case of the output, which is not known in advance.

Many attempts at applying control theory exhibit a fundamental flaw in logic. It is assumed that the formulated problem has a solution $u(\cdot)$ in the class of piecewise continuous control functions, and certain mathematical properties that are known to characterize a solution, if it happens to be in this class, are then invoked. After some work, a piecewise continuous $u(\cdot)$ is identified that does have these properties, and it may even be the only such control function. The conclusion is then made that this function $u(\cdot)$ does indeed solve the problem—but this is erroneous. The reasoning would be correct if we did know that the problem had a solution that belonged to the class of piecewise continuous functions, and this was the only $u(\cdot)$ in that class with the necessary properties. But we don't usually know that.

Again we see the importance of having an appropriate existence theorem before proceeding to invoke conditions on a solution. Unfortunately, the class of piecewise continuous functions is not well disposed to existence theory, because of certain difficulties about taking limits.

The term "piecewise" refers to *finitely many* pieces. Trouble arises because in proving existence of solutions one generally needs to establish some "compactness" property: A sequence of approximate solutions will have a subsequence converging to a true solution. But in a sequence of piecewise continuous functions, successive functions may involve more and more pieces, so that the limit, in whatever sense it is to be interpreted, won't just have finitely many pieces.

The concept of piecewise continuous functions has other subtleties as well, and we should pause here to give a really rigorous definition. A function $u(\cdot) : [t_0, t_1] \to R^d$ is *piecewise continuous* if the interval $[t_0, t_1]$ can be expressed as the union of finitely many *closed* subintervals I_k, on each of which there is a continuous function $u_k(\cdot)$ agreeing with $u(\cdot)$ except possibly at the endpoints of I_k. ¿From this it follows that a piecewise continuous function has right and left limits at every point, and these are the same except at possibly a finite number of points.

Note that the particular value assigned to a piecewise continuous function at a "jump" point does not much matter. Functions that differ only in such respects can, for most purposes, be regarded as the "same." Thus in an important sense a piecewise continuous function is not just a single function but a sort of equivalence class of functions.

We are led by this and the cited difficulties about working with piecewise continuous functions to search for a better class of functions, and we are prepared to accept that for such function a kind of equivalence could well play a role too. We find just the right class in the so-called *measurable* functions.

Measurable Functions. The term "measurable" is unfortunate as an appellation for functions. It comes from measure theory, of course, but measurable functions attract interest for reasons of their own. It is not necessary to study measure theory to see their usefulness.

In what follows, we suppose the reader has some acquaintance already with measurable functions and related concepts, because otherwise we would be obliged to go into much more detail about them. Nonetheless we sketch a more or less independent way of developing the subject, as motivated by the applications we intend to make. This approach would not be suitable for abstract levels of measure theory, but it does provide insights—sadly neglected in most texts—about the nature of measurable functions of a single real variable.

The fundamental notion is that of a *negligible* subset S of R. By definition this means that for any $\varepsilon > 0$, no matter how small, it is possible to find a sequence of intervals I_k of lengths l_k such that $S \subset \bigcup_{k=1}^{\infty}$ and $\Sigma_{k=1}^{\infty} l_k < \varepsilon$.

In particular if S contains only countably many points, e.g. if S is finite or empty, then S is negligible. It can be shown on the basis of the definition that *the union of countably many negligible sets is still negligible.*

Two functions from $[t_0, t_1]$ to R^d are regarded as *equivalent* if they differ only on a negligible subset of $[t_0, t_1]$. This is often expressed by saying that they *agree a.e.*, where the abbreviation "a.e." stands for "almost everywhere" and is always to be interpreted as referring to an exceptional set that is negligible. For example, piecewise continuous functions that differ only at a finite number of jump points agree a.e. and are thus equivalent in the sense just introduced.

A sequence of functions $u^\nu(\cdot)$ (we shall typically use ν, the Greek letter "nu," as the index for sequences) is said to converge *pointwise a.e.* on $[t_0, t_1]$ to a function $u(\cdot)$ if

$$\lim_{\nu \to \infty} u^\nu(t) = u(t) \text{ for almost every } t \in [t_0, t_1].$$

For example, the functions $u^\nu(t) = \sin^\nu 2\pi\nu t$ on $[0, 1]$ converge pointwise a.e. to the function $u(t) \equiv 0$, but there are countably many points t where convergence does not take place.

Now we come to the main characterization of measurable functions for our context—this could actually be used as a definition of measurability. *A function $u(\cdot) : [t_0, t_1] \to R^d$ is measurable if and only if there is a sequence of continuous functions $u^\nu(\cdot)$ that converges to $u(\cdot)$ pointwise a.e.* Moreover this characterization remains true if "continuous" is replaced by "piecewise continuous" or by "piecewise constant"—piecewise constant functions are (vector-valued) step functions.

We shall not furnish a proof of this characterization here. (It follows from a well known theorem of Lusin.) Rather we adopt it as the criterion for measurability that we use in demontrating things about the concept. In terms of this criterion, for example, one can readily verify the following standard facts:

- If a sequence of measurable functions $u^\nu(\cdot)$ converges pointwise a.e. to a function $u(\cdot)$, then $u(\cdot)$ is measurable.

- A vector-valued function is measurable if and only if each of its real-valued component functions is measurable.

- Sums and scalar multiples of measurable functions are measurable. In the real-valued rather than vector-valued case, the same is true also for products and quotients (where nonzero).

- The pointwise supremum or infimum of a finite (or countable) collection of measurable real-valued functions is again measurable.

- If the interval $[t_0, t_1]$ is expressed as the union of finitely many (or even countably many) disjoint subintervals I_k, and the function $u(\cdot)$ on $[t_0, t_1]$ is defined by $u(t) = u_k(t)$ for all $t \in I_k$, where each $u_k(\cdot)$ is measurable, then $u(\cdot)$ is measurable.

- If $v(t) = h(u(t))$ where $u(\cdot)$ is measurable and h is *continuous*, then $v(\cdot)$ is measurable.

In the last assertion, it is important that h be continuous and not just measurable. Otherwise the conclusion could fail. A more general composition result, involving $v(t) = h(t, u(t))$, will be proved in Section 1 of Part 2.

Integration. Although we shall not say much about integration theory, a few reminders about the role of measurability may be helpful. For the most part, it is enough to speak of real-valued and extended-real-valued functions, because vector-valued functions can always be broken down into their real-valued component functions.

The main thing is that the definition of an integral $\int_{t_0}^{t_1} f(t)dt$ requires, as a precondition, that $f(t)$ be measurable in t. Measurability is normally not enough by itself, though, for the integral to be well defined. One case where it *is* enough is the case of $f(t) \geq 0$ for almost every t. Then the integral is a uniquely determined number in $[0, \infty]$, and it is 0 if and only if actually $f(t) = 0$ for almost every t. More generally, if f is not nonnegative but satisfies $f(t) \geq -\rho(t)$ for almost every t, where ρ is a nonnegative function with $\int_{t_0}^{t_1} \rho(t)dt < \infty$; then the integral of f is a well defined number in $(-\infty, \infty]$. Similarly, of course, if $f(t) \leq \rho(t)$ for almost every t, then the integral of f is a value in $[-\infty, \infty)$. If actually $|f(t)| \leq \rho(t)$ for such a function ρ, then the integral of f is necessarily finite. In this case, f is said to be a *summable* function. Thus f is summable if and only if $\int_{t_0}^{t_1} |f(t)|dt < \infty$. The integral of f is the same as the integral of any function equivalent to f.

A fundamental convergence result for integrals is provided by Fatou's Lemma:

> *If the extended-real-valued measurable functions $f^\nu(\cdot)$ satisfy $f^\nu(t) \geq g(t)$ for almost every t, where $g(\cdot)$ is summable, then*
>
> $$\liminf_{\nu\to\infty} \int_{t_0}^{t_1} f^\nu(t)dt \geq \int_{t_0}^{t_1} \liminf_{\nu\to\infty} f^\nu(t)dt.$$

Likewise, if $f^\nu(t) \leq h(t)$ for almost every t, where $h(\cdot)$ is summable, then

$$\limsup_{\nu\to\infty} \int_{t_0}^{t_1} f^\nu(t)dt \leq \int_{t_0}^{t_1} \limsup_{\nu\to\infty} f^\nu(t)dt.$$

By combining these two statements in the case where actually $\lim_{\nu\to\infty} f^\nu(t) = f(t)$ for almost every t, one obtain the convergence of $\int_{t_0}^{t_1} f^\nu(t)dt$ to $\int_{t_0}^{t_1} f(t)dt$. This is the Lebesgue dominated convergence theorem. It can be stated as follows in the general vector-valued case.

If a sequence of measurable functions $w^\nu(\cdot)$ converges pointwise a.e. to a function $w(\cdot)$ and satisfies $|w^\nu(t)| \leq \rho(t)$ a.e. for all ν, where $\rho(\cdot)$ is summable, then

$$\lim_{\nu\to\infty} \int_{t_0}^{t_1} w^\nu(t)dt = \int_{t_0}^{t_1} w(t)dt.$$

This theorem, in conjunction with the characterization of measurable functions as pointwise a.e. limits of continuous (or piecewise continuous, or piecewise constant) functions, tells us by the way that the integral of a measurable function, when it exists, can be obtained as the limit of the classically defined integrals of approximating functions.

Function spaces. If $u(t)$ is measurable in t, then so is $|u(t)|$, where $|\cdot|$ denotes the absolute value when $u(t)$ is a real number and the Euclidean norm when $u(t)$ is a vector. This is a special case of the composition of a measurable function with a continuous function. A function $u(\cdot) : [t_0, t_1] \to R^d$ is said to be *essentially bounded* if there is a number ρ, with $0 \leq \rho < \infty$, such that $|u(t)| \leq \rho$ for almost every $t \in [t_0, t_1]$. The lowest such number ρ (it exists) is denoted by $||u(\cdot)||_\infty$.

The space of all essentially bounded functions $u(\cdot) : [t_0, t_1] \to R^d$ is denoted by $\mathcal{L}_d^\infty[t_0, t_1]$. It is the main space we shall be interested in for control functions. It contains all continuous and piecewise continuous function, in particular. It is a Banach space under the norm $||\cdot||_\infty$, which means of course that it is a linear space (closed under addition and scalar multiplication) in which any sequence of elements $u^\nu(\cdot)$ that is a Cauchy sequence converges to some element $u(\cdot)$:

$$||u^\nu(\cdot) - u(\cdot)||_\infty \to 0 \text{ as } \nu \to \infty.$$

(The sequence is *Cauchy* if for every $\varepsilon > 0$ there is an index ν_0 such that for all $\nu, \nu' > \nu_0$, one has $||u^{\nu'}(\cdot) - u^\nu(\cdot)||_\infty < \varepsilon$.)

Other standard spaces, of course, are the Lebesgue spaces $\mathcal{L}_d^p[t_0, t_1]$ for $1 \leq p < \infty$, which consist of the functions $u(\cdot) \to R^d$ having $\int_{t_0}^{t_1} |u(t)|^p dt < \infty$. Each of these is a Banach space with norm $||u(\cdot)||_p = [\int_{t_0}^{t_1} |u(t)|^p dt]^{1/p}$. In particular, the space $\mathcal{L}_d^1[t_0, t_1]$ consists of all the summable functions $u(\cdot) \to R^d$. Strictly speaking, these spaces consist of equivalence classes of functions that agree almost everywhere, rather than of individual functions;

We shall denote by $\mathcal{A}_n^p[t_0, t_1]$ the space of functions $x(\cdot) : [t_0, t_1] \to R^n$ such that

$$x(t) = a + \int_{t_0}^{t} v(\tau) d\tau \text{ for some } v(\cdot) \in \mathcal{L}_n^p \text{ and } a \in R^n.$$

Such a function is necessarily continuous with $x(t_0) = a$ and has the property that for almost every t, $\dot{x}(t)$ exists and equals $v(t)$.

The functions in $\mathcal{A}_n^1[t_0, t_1]$ are the so-called *absolutely continuous* functions, whereas the ones in $\mathcal{A}_n^\infty[t_0, t_1]$ are known to be the *Lipschitz continuous* functions $x(\cdot)$. The latter functions, which will concern us the most in the context of control, are characterized by the existence of a constant $\lambda \geq 0$ such that

$$|x(t') - x(t)| \leq \lambda |t' - t| \text{ for all } t, t' \in [t_0, t_1].$$

In fact, λ has this property if and only in $\lambda \geq ||\dot{x}(\cdot)||_\infty$.

The norm on $\mathcal{A}_n^1[t_0, t_1]$ is taken to be

$$|||x(\cdot)|||_1 = |x(t_0)| + ||\dot{x}(\cdot)||_1,$$

whereas the one on $\mathcal{A}_n^\infty[t_0, t_1]$ is

$$|||x(\cdot)|||_\infty = \max\{|x(t_0)|, ||\dot{x}(\cdot)||_\infty\}.$$

More generally, on $\mathcal{A}_n^p[t_0, t_1]$ for $1 < p < \infty$ we would have

$$|||x(\cdot)|||_p = [|x(t_0)|^p + ||\dot{x}(\cdot)||_p^p]^{1/p}.$$

In each case the norm gives us a Banach space.

The space of *continuous* functions $x(\cdot) : [t_0, t_1] \to R^n$ will be denoted by $\mathcal{C}_n[t_0, t_1]$. Obviously

$$\mathcal{A}_n^\infty[t_0, t_1] \subset \mathcal{C}_n[t_0, t_1] \subset \mathcal{L}_n^\infty[t_0, t_1].$$

The norm on $\mathcal{C}_n[t_0, t_1]$ is the same as the one on $\mathcal{L}_n^\infty[t_0, t_1]$, namely $|| \cdot ||_\infty$, and under this $\mathcal{C}_n[t_0, t_1]$ too is a Banach space.