MEMORANDUM
RM-3814-PR
NOVEMBER 1963

ASYMPTOTIC CONTROL THEORY

Richard Bellman and Richard Bucy

PREPARED FOR:
UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM
RM-3814-PR
NOVEMBER 1963

ASYMPTOTIC CONTROL THEORY

Richard Bellman and Richard Bucy

This research is sponsored by the United States Air Force under Project RAND—contract No. AF 49(638)-700 monitored by the Directorate of Development Planning, Deputy Chief of Staff, Research and Development, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force.

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA

PREFACE

Part of the Project RAND research program consists of basic supporting studies in mathematics. The mathematical research presented here concerns control theory and in particular the relation between optimal policies for control over a finite and infinite time interval.

SUMMARY

The purpose of this paper is to describe some general problems concerning the asymptotic behavior of control processes as the time-interval becomes infinite, to present some partial results in the general case, and to provide a detailed analysis of a one-dimensional control process.

CONTENTS

ASYMPTOTIC CONTROL THEORY

by

Richard Bellman[1] and Richard Bucy[2]

1. INTRODUCTION.

In recent years the mathematical theory of control has received an increasing amount of attention. New theories have been developed and older theories have been refined and extended [1,2,3,4,5,6].

In this paper, we wish to initiate discussion of a problem in the calculus of variations which has not had the attention due it in the classical literature. The problem is concerned with the asymptotic behavior of the solution of a variational problem as the time interval becomes infinite. From the standpoint of control theory, and more generally from the standpoint of dynamic programming, this is a very natural type of behavior to study. In many significant cases, the "steady-state" policy is simpler conceptually, analytically and computationally.

We shall consider the minimization of the functional

$$J(u) = \frac{1}{2}\int_0^T (u^2 + L(x))dt \tag{1.1}$$

This research is sponsored by the United States Air Force under Project RAND—Contract No. AF 49(638)-700.

---

(1) RAND Corporation, Santa Monica, California.

---

(2) Consultant RAND Corporation, currently at RIAS, Baltimore, Maryland.

over all functions u where

(1.2) $\dot{x} = f(x) + u, \quad x(0) = c.$

Let $V(c,T) = \min_u J(u)$. For finite and sufficiently small T the classical calculus of variations, or dynamic programming applies, under certain reasonable assumptions on L and f. We shall be interested, however, in the following questions:

(1) When does the problem for infinite T make sense?

(2) When it does, are the optimal motions and policies for finite T the limits of the corresponding optimal motions and policies for finite T?

(3) What is the effect of using steady-state optimal policy for the finite problem?

This is an example of what we mean by <u>asymptotic control theory</u>.

For example, if $f = 0$ and $L = x^2 + \frac{1}{2} x^4$ the problem is that of minimizing the functional

(1.3) $$J(u) = \int_0^T [\dot{x}^2 + x^2 + \frac{1}{2} x^4] dt$$

over all C^1 curves for which $x(0) = c$. The Euler equation is

(1.4) $$\ddot{x} - x - x^3 = 0,$$

subject to the two-point boundary conditions

$$(1.5) \qquad x(0) = c, \quad \dot{x}(T) = 0.$$

Establishing the existence and uniqueness of solutions of (1.4) and determining the asymptotic behavior as $T \to \infty$ is analogous to the classical problem of Poincaré-Lyapunov [5], but materially more difficult because of the two-point boundary-value condition.

We shall first, using quite general arguments, show that $V(T,c)$ is monotone increasing as a function of T, and uniformly bounded under mild restrictions concerning $L(x)$. Taking advantage of the fact that the Euler equation possesses a first integral, we can analyze the behavior of the solution in detail as $T \to \infty$.

This analysis shows that the formal asymptotic series obtained from the partial differential equation

$$(1.6) \qquad V_T = \min_u \left[\tfrac{1}{2}\left(u^2 + L(x)\right) + V_x(ax + u)\right],$$

an equation derived from dynamic programming considerations which yields the Hamilton-Jacobi equation relevant to the variational problem when $f(x) = ax$([1] and [12]), is an actual asymptotic series for $V(c,T)$. This corresponds to the result easily derived in the case where the integrand in (1.1) is merely quadratic in x and u.

In the concluding section, we shall mention some open and apparently quite difficult questions in connection with asymptotic behavior and give some references to analogous results obtained for dynamic programming processes by Kalman and Bucy [6], Beckwith [7], Iglehart [8], Freimer [9], and Bellman [10].

2. MONOTONICITY AND BOUNDEDNESS.

Let us introduce the function

$$(2.1) \qquad V(c,T) = \min_u J(u),$$

(with the assumption that $f(x) = ax$). Let $x(t,T)$, $u(t,T)$ represent the functions that furnish the minimum of $J(u)$ under the assumption that $L(x)$ is a nonnegative entire function of x. In most processes of interest $L(x)$ is a polynomial in x.

Since

$$(2.2) \qquad \begin{aligned} V(c,T+\Delta) &= \int_0^T + \int_T^{T+\Delta} \\ &\geq \min_u \int_0^T + \int_T^{T+\Delta} \\ &\geq V(c,T) + \int_T^{T+\Delta} \end{aligned}$$

we see that $V(c,T)$ is monotone increasing in T.

To show uniform boundedness in T, for fixed c, let us choose an appropriate control policy, say

$$(2.3)\qquad \begin{aligned} u &= 0 && \text{when } a < 0, \\ u &= -2ax && \text{when } a > 0, \\ u &= -x && \text{when } a = 0. \end{aligned}$$

In each case, we see that $u = ce^{-bt}$ with b positive. Hence,

$$(2.4)\qquad J(u) = \int_0^T [O(e^{-2bt}) + L(ce^{-bt})]\,dt.$$

Under the assumption that $L(x) = O(x)$ as $x \to 0$, the integral is uniformly bounded as $T \to \infty$.

Having established boundedness and monotonicity as $T \to \infty$, we can assert convergence,

$$(2.5)\qquad V(c,T) \to V(c)$$

as $T \to \infty$.

It is not settled, however, whether or not the states $X(t,T)$ and the policies $u(t,T)$ converge as $T \to \infty$. The foregoing argument extends to quite general situations, but leaves unanswered the interesting and important questions concerning the convergence of policies.

3. DETAILED ANALYSIS.

We will be interested in an explicit solution to the partial differential equation

(3.1) $$V_T = \frac{1}{2} L(c) + acV_c - \frac{1}{2} V_c^2$$

subject to the boundary conditions

(3.2) $$V(T,c)|_{T=0} = 0, \quad a < 0,$$

$$V(T,c)|_{T=0} = ac^2, \quad a > 0.$$

As is well known ([12]) existence of a sufficiently smooth solution to (3.1) is a sufficient condition for the variational problem (1.1) to have a solution. The equation of (3.1) is (1.6) with the minimization carried out.

It will be assumed that L satisfies the following conditions:

(3.3)
(1) L is even, and positive,
(2) L and L_x are continuous and increasing for positive x,
(3) $L(x) = O(|x|)$ as $|x| \to 0$,
(4) L is analytic.

Now the Cauchy-Kowalewski theorem implies (3.1) has a unique local analytic solution ([11]).

With the aim of solving (3.1) we introduce the function $y(c,T)$ which corresponds physically to the final state of the controlled system along an optimal trajectory initiating at $(c,0)$ and ending at (y,T). The following lemma shows that y is well defined for

$c > 0$. The case $c < 0$ is similar. When $L(c) = c^2 + c^4$, y will be defined by an elliptical integral of the first kind.

Lemma 1. Suppose $c > 0$, and assume conditions (3.3) are fulfilled, then for every K, $\infty > K > 0$, there exists a unique $0 < y \leq c$ such that

$$(3.4) \qquad I(y) = \int_y^c \frac{dx}{\sqrt{a^2x^2 + L(x) - L(y)}} = K.$$

Proof. Since it is clearly continuous, elementary bounding of $I(y)$ shows it takes on all finite positive values as y ranges over $(0,c]$. To show uniqueness assume for some finite positive K that there exist y_1 and y_2, $y_1 > 0 > y_2 > 0$ where both satisfy (3.4), then

$$(3.5) \qquad \int_{y_1}^c \frac{dx}{\sqrt{a^2x^2 + L(x) - L(y_1)}} = \int_{y_2}^c \frac{dx}{\sqrt{a^2x^2 + L(x) - L(y_2)}} .$$

But (3.5) implies

$$(3.6) \qquad \int_{y_2}^{c-\Delta} \frac{dx}{\sqrt{a^2(x+\Delta)^2 + L(x+\Delta) - L(y_1)}}$$

$$> \int_{y_2}^{c-\Delta} \frac{dx}{\sqrt{a^2x^2 + L(x) - L(y_2)}}$$

where $\Delta = y_1 - y_2 > 0$. It suffices to show (3.7) to contradict (3.6):

$$(3.7) \qquad 2a^2x\Delta + a^2\Delta^2 > L(x) - L(x+\Delta) + L(y_1) - L(y_2).$$

Consider the right-hand side of (3.7) for $x \geqq y_1$. The mean value theorem gives

$$L(x+\Delta) - L(x) = L_x(\varphi)\Delta \geqq L_x(y_1)\Delta,$$

$$\varphi \in (x, x+\Delta),$$

$$L(y_1) - L(y_2) = L_x(\theta)\Delta \leqq L_x(y_1)\Delta,$$

$$\theta \in (y_2, y_1),$$

and for these x (3.7) is satisfied. Then for $y_2 < x < y_1$

$$L(x+\Delta) - L(y_1) = L_x(\delta)(x - y_2) \geqq L_x(y_1)(x - y_2),$$

$$L(x) - L(y_2) = L_x(\gamma)(x - y_2) \leqq L_x(x)(x - y_2)$$

$$\leqq L_x(y_1)(x - y_2),$$

and (3.7) is satisfied for all $x \in (y_2, c - \Delta)$.

A closer examination of the previous lemma shows $I(y)$ decreases strictly as y increases through $(0,c]$ and hence $\partial/\partial y\, I(y) < 0$. Defining $y(c,T)$ as that value of y which satisfies

$$(3.8) \qquad \int_y^c \frac{dx}{\sqrt{a^2x^2 + L(x) - L(y)}} = T,$$

it is easy to see by the implicit function theorem that y is C^1 in c and T on $(0,c] \times [0,\infty)$. Further, y is analytic! The following lemma characterizes the behavior of y as T tends to infinity.

<u>Lemma 2</u>. <u>Under the assumptions</u> (3.3) <u>and</u> $a \neq 0$, y <u>defined by</u> (3.8) <u>tends exponentially to zero as</u> $T \to \infty$.

<u>Proof</u>.

$$0 < T = \int_y^c \frac{dx}{\sqrt{a^2x^2 + L(x) - L(y)}}$$

$$\geq \int_y^c \frac{dx}{|a|x} = \frac{1}{|a|} \ln \frac{c}{y}$$

or

$$0 \leq y \leq ce^{-|a|T}.$$

Now we will characterize the solution of (3.1) subject to (3.2).

* Theorem 1. Equation (3.1) has the following analytic solutions in the regions $c > 0$ and $c < 0$ under assumptions (3.3). For $a < 0$

$$(3.9)\quad V(T,c) = \begin{cases} \int_y^c a\xi + \sqrt{a^2\xi^2+L(\xi)-L(y)}\,d\xi + T\,\frac{L(y)}{2}, & c > 0, \\ 0, & c = 0 \\ \int_y^c a\xi - \sqrt{a^2\xi^2+L(\xi)-L(y)}\,d\xi + T\,\frac{L(y)}{2}, & c < 0, \end{cases}$$

while for $a > 0$

$$(3.10)\quad V(T,c) = \begin{cases} \int_y^c a\xi + \sqrt{a^2\xi^2+L(\xi)-L(y)}\,d\xi + T\,\frac{L(y)}{2} + ay^2, & c > 0, \\ 0, & c = 0, \\ \int_y^c a\xi - \sqrt{a^2\xi^2+L(\xi)-L(y)}\,d\xi + T\,\frac{L(y)}{2} + ay^2, & c < 0, \end{cases}$$

where y satisfies

$$\int_y^c \frac{d\xi}{\sqrt{a^2\xi^2 + L(\xi) - L(y)}} = T \quad \text{for} \quad c > 0,$$

$$-\int_y^c \frac{d\xi}{\sqrt{a^2\xi^2 + L(\xi) - L(y)}} = T, \quad c < 0,$$

---

*Replacing $a\xi$ and $(a\xi)^2$ by $f(\xi)$ and $f(\xi)^2$ with f continuous and ay^2 by $2\int_0^y f(\xi)d\xi$, provides a solution to the equation $V_T = L + V_x f(x) - \frac{1}{2} V_x{}^2$ which is local unless y is defined for all T. Further Haar's uniqueness theorem ([13]) is applicable and implies (3.9) for $c > 0$ is the unique C^1 solution of (3.1).

and y is the same sign as c. Further,

$$(3.11)\qquad V(\infty,c) = \lim_{T\to\infty} V(T,c) = \int_0^c a\xi + \sqrt{a^2\xi^2 + L(\xi)}\,d\xi, \quad c > 0,$$

$$V(\infty,c) = \lim_{T\to\infty} V_c(T,c) = ac + \sqrt{a^2c^2 + L(c)}, \qquad c > 0,$$

and the corresponding formulae for $c < 0$ and $c = 0$. Finally, the optimum control law is

$$(3.12)\qquad u^0(t) = -\,ax(t) - \operatorname{sgn} x(t)\,\sqrt{a^2x(t)^2 + L(x(t)) - L(y)}.$$

Proof. (3.9) and (3.10) follow from Lemma 1 and direct substitution. (3.11) follows since for fixed c, $V_c(c,T)$ are monotone in T and uniformly bounded while the explicit form follows from Lemma 2, (3.3) (3) and the dominated convergence theorem. (3.12) is just the principle of optimality.

Corollary 1. Under the previous assumptions, the value of the T-infinite case $V(\infty,c)$ satisfies

$$\tfrac{1}{2}L + V_c ac - \tfrac{1}{2}V_c^2 = 0$$

just (3.1) with $V_T = 0$.

4. ASYMPTOTIC BEHAVIOR.

As mentioned above, the principle of optimality yields the partial differential equation

$$(4.1)\qquad V_T = \tfrac{1}{2}L(c) + acV_c - \tfrac{1}{2}V_c^2.$$

It does not seem possible to obtain the asymptotic behavior of V, even formally, by means of a series of the form

$$(4.2) \qquad V = V_0(c) + V_1(c,T) + \cdots,$$

without some additional information concerning the analytic structure of V_1, e.g.,

$$(4.3) \qquad V_1(c,T) = V_1(c)u_1(T).$$

Here $V_0(c) = \lim_{T\to\infty} V(c,T)$.

We can, however, obtain an interesting bound for $V(c,\infty) - V(c,T)$ in the following fashion. Consider the expression

$$(4.4) \qquad V(c,\infty) = \min_u \int_0^\infty [x^2 + u^2 + L(x)]\,dt.$$

Let $u(t,T)$, $x(t,T)$ denote the minimizing set of functions for the interval $[0,T]$. Then, it is clear that

$$(4.5) \qquad V(c,\infty) \le \int_0^T [u(t,T)^2 + x(t,T)^2 + L(x(t,T))]\,dt + \int_T^\infty [\ \cdots\ ]\,dt,$$

where in the second integral our choice of u and x are constrained only by the condition $x(T) = x(T,T)$. Write $x(T,T) = x(c,T)$, the state of the system at time T starting in state c at time 0 associated with the finite variational process over $[0,T]$. Then (4.5) yields the inequality

$$V(c,\infty) \leq V(c,T) + V(x(c,T),\infty). \tag{4.6}$$

Hence, we can obtain an estimate of the difference between $V(c,\infty)$ and $V(c,T)$ if we obtain an estimate for $x(c,T)$ as $T \to \infty$.

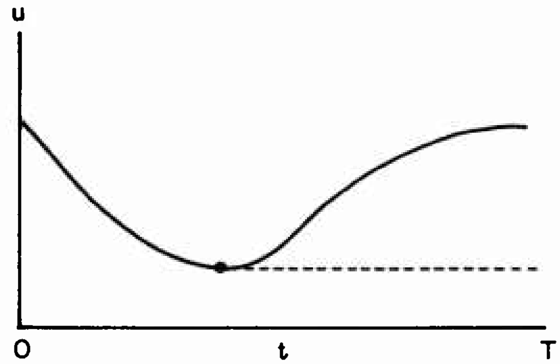
Observe that the estimate for $V(c,\infty)$ is readily obtained by using a convenient approximate policy of the type described in Sec. 2.

The estimate for $x(c,T)$ is not readily obtained in general. Let us indicate how elementary arguments yield the result for the problem of minimizing

$$J(x) = \int_0^T [\dot{x}^2 + x^2 + x^4]\,dt \tag{4.7}$$

where $x(0) = c$.

It is clear from the form of the integrand that if $c > 0$, then x is monotone decreasing. For, if as indicated below, x reached a turning point and started to increase, we could replace it by the dotted curve, obtaining obviously a smaller value of the integral:



The Euler equation is

$$(4.8) \qquad \ddot{x} - x - 2x^3 = 0, \quad x(0) = c, \quad \dot{x}(T) = 0.$$

If x decreases monotonically, the limit must be zero as $T \to \infty$. From the Poincaré-Lyapunov theorem, we know that all solutions of (4.8) which approach zero as $t \to \infty$ have an asymptotic expansion of the form $c_1e^{-t} + c_2e^{-2t} + \cdots$. Using this information in conjunction with the preceding results, we readily obtain an asymptotic series for $V(c,T)$ as $T \to \infty$.

5. FURTHER PROBLEMS.

The technique we have used here to obtain the asymptotic behavior of the state variables and the control variable is quite special and does not extend to the multidimensional case, to control processes with constraints, to more general control processes involving distributed parameters, to general stochastic control processes, or to adaptive control processes. Some partial results can be derived, but on the whole there appears to be a need for a development of some new techniques.

We feel that it is worthwhile in one case at least to show that the expected results actually hold.

For asymptotic results in dynamic programming for processes of quite different nature, see [7,8,9,10].

REFERENCES

1. Bellman, R., Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.
2. ———, Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, New Jersey, 1961.
3. Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, The Mathematical Theory of Optimal Processes, Interscience Publishers, New York, 1962.
4. Bellman, R., I. Glicksberg, and O. Gross, Some Aspects of the Mathematical Theory of Control Processes, The RAND Corporation, R-313, January 1958; Russian translation in Moscow, 1962.
5. Bellman, R., Stability Theory of Differential Equations, McGraw-Hill Book Company, Inc., New York, 1954.
6. Kalman, R. E., and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," ASME Journal of Basic Engineering, March 1961.
7. Beckwith, R. E., Analytic and Computational Aspects of Dynamic Programming Processes of High Dimension, Ph.D. Thesis, Purdue University, June 1959.
8. Iglehart, D., Ph.D. Thesis, Stanford University, 1960.
9. Freimer, M., A Dynamic Programming Approach to Adaptive Control Processes, Lincoln Lab. Report 54-2, 1959.
10. Bellman, R., "A Markovian Decision Process," J. Math. and Mech., Vol. 6, 1957, pp. 679-684.
11. Petrovsky, I. G., Partial Differential Equations, Interscience Publishers, New York, 1954.
12. Kalman, R. E., "The Theory of Optimal Control and the Calculus of Variations," Mathematical Optimization Techniques, University of California Press, Berkeley, California, 1963, pp. 309-331.
13. Courant, R., and D. Hilbert, Methods of Mathematical Physics, Interscience Publishers, New York, 1962.