

## SOME NEW TECHNIQUES IN THE DYNAMIC PROGRAMMING SOLUTION OF VARIATIONAL PROBLEMS\*

BY

RICHARD BELLMAN

*The Rand Corporation, Santa Monica, Cal.*

**Summary.** In previous papers, it has been shown that the functional equation technique of dynamic programming may be applied to yield the numerical solution of a wide class of variational problems of the type occurring in mathematical physics, engineering, and economics.

It was seen that the numerical solution of a problem involving  $N$  state variables depended upon the computation of sequences of functions of  $N$  variables. This fact made the method routine only for the case where  $N = 1$  or  $2$ , with grave difficulties arising in the general case.

In this paper, it is indicated how this difficulty can be overcome for a large class of problems in which the underlying equations and criterion function are linear, although the restraints on the forcing functions may be non-linear, corresponding say to energy considerations.

The same methods are applicable to other classes of linear equations, and, in particular, to differential-difference equations, arising from time-lag problems and to various classes of partial differential equations. These problems could not previously be treated by dynamic programming techniques in any usable fashion.

Finally, it is briefly indicated how the method of successive approximations may be combined with the foregoing techniques to reduce general variational problems with non-linear equations and criterion function to sequences of problems that can be solved numerically by means of sequences of functions of one variable. There are a number of interesting and difficult convergence questions associated with this program; these, however, are not discussed here.

**1. Introduction.** A variational problem that is encountered in many areas of pure and applied mathematics is that of determining the minimum or maximum of a functional of the form

$$J(y) = \int_0^T F(x_1, x_2, \dots, x_N; y_1, y_2, \dots, y_M) dt \quad (1.1)$$

over all functions  $y_1, y_2, \dots, y_M$ , connected to the  $x_i$  by means of the relations

$$\frac{dx_i}{dt} = G_i(x_1, x_2, \dots, x_N; y_1, y_2, \dots, y_M), \quad x_i(0) = c_i, \quad i = 1, 2, \dots, N, \quad (1.2)$$

and satisfying constraints of the form

$$\begin{aligned} (a) \quad & x_i(T) = b_i, \quad i = 1, 2, \dots, k, \\ (b) \quad & R_j(x; y) \leq 0, \quad j = 1, 2, \dots, m. \end{aligned} \quad (1.3)$$

For a variety of reasons, which preliminary mathematical study discloses quite readily, this problem, to the degree of generality stated above, presents formidable analytic

---

\*Received Sept. 23, 1957.

difficulties. Not only do these difficulties arise in connection with the explicit solution of the problem, see [9] and [10], but also in connection with the apparently more modest demand for a numerical solution.

We have discussed elsewhere, [1, 2, 3, 4], various applications of the functional equation technique of dynamic programming to the numerical solution of classes of problems of the foregoing type. In this paper, we wish to indicate some recent developments which greatly enlarge the scope of the methods utilized in the references cited above. These new developments, combined with the classical tool of successive approximations, enable us to attack systematically large classes of problems formerly far beyond our powers.

We shall begin our discussion with a terminal control problem for a linear system with constant coefficients. The problem we consider is that of maximizing a functional of the form

$$J(y) = H[x_1(T), x_2(T), \dots, x_k(T)] \quad (1.4)$$

over all functions  $y_i$  related to the  $x_i$  by means of the linear differential equations

$$\frac{dx_i}{dt} = \sum_{j=1}^N a_{ij}x_j + \sum_{j=1}^N b_{ij}y_j, \quad x_i(0) = c_i, \quad i = 1, 2, \dots, N, \quad (1.5)$$

and satisfying constraints of the following type:

$$\begin{aligned} (a) \quad & m_i \leq y_i(t) \leq m'_i, \quad 0 \leq t \leq T, \quad i = 1, 2, \dots, N, \\ (b) \quad & \int_0^T K_j(y_1, y_2, \dots, y_N) dt \leq c_j, \quad j = 1, 2, \dots, l. \end{aligned} \quad (1.6)$$

Whereas the techniques previously given enable us to convert this problem into one involving the sequential computation of functions of  $N$  variables, the linearity of the defining equation in (1.5) permits, as we shall see, a transformation of the problem into one involving sequences of functions of  $k$  variables, where  $k$  is as in (1.4).

Since current digital computers do not allow the storage of functions of more than two variables in any feasible way, this is a very important reduction if  $k$  is equal to one or two. As it turns out, a number of significant problems arising in the fields of economic and engineering control may be formulated in these terms.

If the function  $H(x_1, x_2, \dots, x_k)$  is linear,

$$H(x_1, x_2, \dots, x_k) = \sum_{i=1}^k \alpha_i x_i, \quad (1.7)$$

the problem may be still further reduced to a computation involving sequences of functions of one variable, regardless of the values of  $k$  and  $N$ . The same result holds for the case where it is desired to maximize a linear functional of the form

$$J = \int_0^T \left[ \sum_{i=1}^N \rho_i x_i(t) \right] dt. \quad (1.8)$$

In exactly the same way, we can treat discrete variational problems where difference equations replace differential equations. As a matter of fact, these techniques were developed in connection with a discrete problem, the "caterer" problem, [5].

As we shall see, the same methods are applicable to the treatment of variational

questions involving time lags and retarded control, see [7]. In simple cases involving only one time delay, the equation corresponding to (1.5) is

$$\frac{dx_i}{dt} = \sum_{j=1}^N a_{ij}x_j(t) + \sum_{j=1}^N \alpha_{ij}x_j(t - \delta) + \sum_{j=1}^N b_{ij}y_j(t), \quad (1.9)$$

$$x_i(t) = c_i(t), \quad 0 \leq t \leq \delta, \quad i = 1, 2, \dots, N.$$

More complicated types of hereditary processes and various classes of partial differential equations may also be treated by means of the techniques we present below. Formerly, problems of this nature could not be treated computationally by dynamic programming techniques because of their dependence upon functionals rather than functions.

As noted above, the results discussed in the foregoing paragraphs depend in an essential manner upon the linearity of the equations defining the process. To extend these techniques to cover more general processes, we turn to that general factotum of analysis, the method of successive approximations. Without entering into any of the questions of rigor, we indicate briefly how a variety of apparently multi-dimensional problems can be reduced to sequences of problems involving functions of one variable.

Although a large number of interesting and important questions concerning convergence, rapidity of convergence, stability, and so on, arise from these investigations, we shall postpone any investigation of these matters until a later date.

Similar techniques can be used to treat processes involving random effects. These will also be discussed subsequently.

**2. Dynamic programming.** In order to appreciate the improvement in technique afforded by the methods we present here, let us sketch briefly the direct approach of dynamic programming to variational problems of the aforementioned variety.

To treat the general question posed in (1.1)-(1.3), we define a function of  $N$  variables,  $f(c_1, c_2, \dots, c_N; T)$ , by means of the relation

$$f(c_1, c_2, \dots, c_N; T) = \max_y J(y), \quad (2.1)$$

where the maximum is taken over all functions  $y_i(t)$  satisfying (1.3). Under appropriate assumptions concerning the continuous dependence of the maximizing functions upon the initial values  $c_i$  and the upper limit  $T$ , we obtain for the function  $f$  a non-linear partial differential equation

$$\frac{\partial f}{\partial T} = \max_{v_i} \left[ F(c, v) + \sum_{i=1}^N G_i(c, v) \frac{\partial f}{\partial c_i} \right] \quad (2.2)$$

$$f(c_1, c_2, \dots, c_N; 0) \equiv 0.$$

The maximization in the above expression is over quantities  $v_i$  satisfying the constraints

$$(a) \quad m_i \leq v_i \leq m'_i, \quad i = 1, 2, \dots, N, \quad (2.3)$$

$$(b) \quad R_i(c, v) \leq 0.$$

If  $F$  and  $G$  satisfy suitable differentiability conditions, this non-linear partial differential equation leads, in the absence of constraints, to the usual Euler equations, via the method of characteristics, see [1, 2].

In general, in questions of greatest interest in applications, this reduction from partial differential equations to ordinary differential equations does not occur due to the presence of various constraints of a physical nature.

**3. Computational aspects.** Since the analytic solution of problems of the type discussed in the foregoing sections is only rarely achieved, we turn to the study of computational techniques. Our aim initially is to provide a numerical algorithm, based upon the use of a digital computer, which resolves questions of this nature in a routine fashion without the use of an extensive analysis strongly dependent upon the particular analytic structure of the functions occurring. In practice, simplifications can always be made, taking advantage of the structure of the particular problem under consideration. As we shall see below, it is only by proceeding in this manner that we can make significant advances. General methods, however, are always useful, if only as a court of last resort.

To determine  $f(c_1, c_2, \dots, c_N, T)$ , we can either use (2.2) and any of a number of standard techniques for the numerical solution of partial differential equations of this type, or, as has turned out to be preferable, we can go over to a discrete version of the original continuous process. Analytically, this means that the original differential equations are replaced by difference equations. Thus, (1.2) becomes

$$\begin{aligned} x_i(t + \delta) &= x_i(t) + \delta G_i[x_1(t), x_2(t), \dots, x_N(t); y_1(t), y_2(t), \dots, y_M(t)], \\ x_i(0) &= c_i, \quad i = 1, 2, \dots, N, \end{aligned} \quad (3.1)$$

with  $t$  assuming only the values  $0, \delta, 2\delta, \dots$ .

The non-linear partial differential equation is then replaced by the non-linear recurrence relation

$$\begin{aligned} f(c, T) &= \max [\delta F(c, v) + f(c_1 + \delta G_1(c, v), \dots, c_N + \delta G_N(c, v); T - \delta)], \\ f(c, 0) &= 0. \end{aligned} \quad (3.2)$$

In order to carry out the indicated process, we must be able to tabulate functions of  $N$  variables. This means that at the present time the approach outlined above is only feasible if  $N = 1$  or  $2$ . For  $N > 2$ , the memory requirements become prohibitive.

It is necessary then to develop some new techniques if we wish to utilize dynamic programming to solve large scale problems of the type arising in the engineering and economic spheres.

**4. Preliminaries on linear systems.** In this section we shall mention some well-known results concerning the solution of vector-matrix systems of linear differential equations. These will be utilized in what follows. Proofs of the results cited here may be found in [6].

The linear system of (1.5) may be written, using an obvious vector-matrix notation, in the form

$$\frac{dx}{dt} = Ax + By, \quad x(0) = c. \quad (4.1)$$

Consider first the case in which  $A$  is constant. The solution of (1) may then be written in the form

$$x = \exp(At)c + \int_0^t \exp[A(t-s)]By(s) ds. \quad (4.2)$$

If  $A = A(t)$ , a matrix dependent upon  $t$ , then  $x$  may be written in the form

$$x = X(t)c + \int_0^t X(t)X^{-1}(s)By(s) ds, \quad (4.3)$$

where  $X(t)$  is the matrix solution of

$$\frac{dX}{dt} = A(t)X, \quad X(0) = I. \tag{4.4}$$

We shall utilize these representations in a crucial manner below.

**5. Terminal control—General non-linear criterion.** Let us now turn to the problem of maximizing a given function  $H[x_1(T), x_2(T), \dots, x_k(T)]$  of the terminal state of the system over all control functions  $y_i(t)$  which are related to the  $x_i$  by means of the linear equation in (4.1), and which are subject to the constraints

$$\begin{aligned} \text{(a)} \quad & m_i \leq y_i(t) \leq m'_i, \quad 0 \leq t \leq T, \quad i = 1, 2, \dots, N, \\ \text{(b)} \quad & \int_0^T G(y_1, y_2, \dots, y_N) dt \leq k. \end{aligned} \tag{5.1}$$

We wish to show that the numerical solution of a problem of this type can be made to depend upon a sequence of functions of  $k$  variables, rather than upon sequences of functions of  $N$  variables. We shall consider first the case where  $A = (a_{ij})$  is a constant matrix.

We begin with the linear representation of (4.2), which yields a set of equations

$$x_i(t) = z_i(t) + \int_0^t \left[ \sum_{j=1}^N x_{ij}(t-s)y_j(s) \right] ds, \quad i = 1, 2, \dots, N, \tag{5.2}$$

where  $z_i(t)$  is the  $i$ th component of  $\exp(At)c$ , and  $X(t) = [x_{ij}(t)]$ .

The problem we wish to consider may then be cast in the form of maximizing a functional of the type

$$H\left(u_1 + \int_0^T \left[ \sum_{i=1}^N x_{1i}(T-s)y_i(s) \right] ds, \dots, u_k + \int_0^T \left[ \sum_{i=1}^N x_{ki}(T-s)y_i(s) \right] ds \right), \tag{5.3}$$

where  $u_1, u_2, \dots, u_k$  are given quantities, over all functions  $y_1, y_2, \dots, y_N$  satisfying the constraints (5.1a) and (5.1b).

Let us then consider the sequence of functions  $f(u_1, u_2, \dots, u_k; T)$ , implicitly dependent upon  $\lambda$ , defined as follows.

$$\begin{aligned} f(u_1, u_2, \dots, u_k; T) = \max_{\nu} & \left[ H\left(u_1 + \int_0^T [\dots] ds, \dots, u_k + \int_0^T [\dots] ds \right) \right. \\ & \left. - \lambda \int_0^T G(y_1, y_2, \dots, y_N) ds \right], \end{aligned} \tag{5.4}$$

where the functions  $y_i(t)$  are now constrained only by (5.1a).

A motivation and discussion of this use of the Lagrange multiplier may be found in [8], and a numerical example in [11]. The value of the method lies in the fact that it enables us to reduce multi-dimensional problems, where the dimensionality is reckoned in the dynamic programming sense, to sequences of lower dimensional problems.

To obtain a functional equation for  $f(u_1, u_2, \dots, u_k; T)$ , we proceed as follows. Suppose that the values of  $y_1(t), y_2(t), \dots, y_N(t)$  have been determined over  $[0, \delta]$ .

Then we may write

$$\begin{aligned}
 & H\left(u_1 + \int_0^T [\dots] ds, \dots, u_k + \int_0^T [\dots] ds\right) - \lambda \int_0^T G(y_1, y_2, \dots, y_N) ds \\
 &= H\left(u_1 + \int_0^\delta [\dots] ds + \int_\delta^T [\dots] ds, \dots, u_k + \int_0^\delta [\dots] ds + \int_\delta^T [\dots] ds\right) \\
 &\quad - \lambda \int_0^\delta G(y_1, y_2, \dots, y_N) ds - \lambda \int_\delta^T G(y_1, y_2, \dots, y_N) ds \tag{5.5} \\
 &= H\left(u_1 + \int_0^\delta [\dots] ds + \int_0^{T-\delta} \left[ \sum_{i=1}^N x_{i,i}(T - \delta - s)y_i(s + \delta) \right] ds, \dots\right) \\
 &\quad - \lambda \int_0^\delta G(y_1, y_2, \dots, y_N) ds - \lambda \int_0^{T-\delta} G[y_1(s + \delta), \dots, y_N(s + \delta)] ds.
 \end{aligned}$$

The principle of optimality, see [1], then yields the functional equation

$$\begin{aligned}
 f(u_1, u_2, \dots, u_k; T) = \max_{v \in |0, \delta|} & \left[ -\lambda \int_0^\delta G(y_1, y_2, \dots, y_N) ds \right. \\
 & \left. + f\left(u_1 + \int_0^\delta \left[ \sum_{i=1}^N x_{i,i}(T - s)y_i(s) \right] ds, \dots\right) \right]. \tag{5.6}
 \end{aligned}$$

The maximum is now taken over all functions  $y_i(s)$  defined over  $0 \leq s \leq \delta$ , and satisfying the constraints  $m'_i \leq y_i(s) \leq m_i$  in  $|0, \delta|$ .

For computational purposes, we may use the approximate relation

$$\begin{aligned}
 f(u_1, u_2, \dots, u_k; T) = \max_v & \left[ -\lambda \delta G(v_1, v_2, \dots, v_N) \right. \\
 & \left. + f\left[u_1 + \delta \sum_{i=1}^N x_{i,i}(T)v_i, \dots\right] \right] \tag{5.7}
 \end{aligned}$$

or we may start with a discrete version of the original process.

We have thus reduced the numerical solution of the variational problem to the determination of a sequence of functions of  $k$  variables. If  $k = 1$  or  $2$ , we have a feasible method of solution.

**6. Terminal control—Quadratic criterion.** Let us note that the problem of minimizing the functional

$$Q[x_1(T), x_2(T), \dots, x_N(T)] \tag{6.1}$$

over all functions  $y$  satisfying the relations

$$\frac{dx}{dt} = A(t)x + B(t)y, \quad x(0) = c, \tag{6.2}$$

can be reduced to the solution of systems of linear equations, if  $Q$  is a quadratic form in the  $x_i(T)$ ; see [9].

The same holds if we add to  $Q$  a quadratic function of the form

$$\int_0^T P(x_1, x_2, \dots, x_N; y_1, y_2, \dots, y_N) dt, \tag{6.3}$$

where  $P$  is quadratic in its arguments.

Problems of this type may also be very simply treated by means of the formalism of dynamic programming, a matter we will discuss elsewhere.

**7. Terminal control—Variable coefficients.** It is important, in connection with our subsequent discussion of the use of successive approximations, to consider the same problem for the case where  $A$  is a variable matrix. Let us consider then the situation where the equation governing the process has the form

$$\frac{dx}{dt} = A(t)x + B(t)y + \phi(t), \quad x(0) = c. \tag{7.1}$$

As we know, the solution of this equation is given by the expression

$$x = x(t)c + \int_0^t x(t)x^{-1}(s)B(s)y(s) ds + \int_0^t x(t)x^{-1}(s)\phi(s) ds. \tag{7.2}$$

Hence the components of  $x(T)$  have the form

$$x_i(T) = u_i + \int_0^T \left[ \sum_{j=1}^N x_{ij}(T, s)y_j(s) \right] ds, \tag{7.3}$$

where the  $u_i$  are independent of  $y$ .

In order to take account of the non-stationarity of the process, we count time backwards. In place of noting the time at which the process ends, we single out the time at which it begins. Fixing  $T$ , we consider the function  $f(u_1, u_2, \dots, u_k; r)$  defined by the relation

$$f(u_1, u_2, \dots, u_k; r) = \max_{v \{r, T\}} \left[ H \left[ u_1 + \int_r^T \left[ \sum_{i=1}^N w_{1i}(T, s)y_i(s) ds, \dots \right] - \lambda \int_r^T G(y_1, y_2, \dots, y_N) ds \right] \right]. \tag{7.4}$$

Arguing as in the preceding section, we see that  $f$  satisfies the relation

$$f(u_1, u_2, \dots, u_k; r) = \max_{v \{r, r+\delta\}} \left[ -\lambda \int_r^{r+\delta} G(y_1, y_2, \dots, y_N) ds + f \left( u_1 + \int_r^{r+\delta} [\dots] ds, \dots, u_k + \int_r^{r+\delta} [\dots] ds \right) \right] \tag{7.5}$$

For computational purposes, this reduces to

$$f(u_1, u_2, \dots, u_k; r) = \max_v \left[ -\lambda \delta G(v_1, v_2, \dots, v_N) + f \left[ u_1 + \delta \sum_{i=1}^N w_{1i}(T, r)v_i, \dots \right] \right], \tag{7.6}$$

with  $f(u_1, u_2, \dots, u_k; T) = 0$ . Here  $r = n\delta, \dots, T = M\delta$ .

**8. Terminal control—Linear criterion.** Let us now consider the case where  $H$  is a linear function, which is to say, we consider the problem of maximizing the inner product  $[s(T), a]$  where  $a$  is a given vector. To simplify the notation, let us consider only the case where  $A$  is a constant matrix.

Using the representation for  $x(t)$  given in (4.2), we see that

$$[x(T), a] = (\exp(AT)c, a) + \left[ \int_0^T \exp[A(T-s)]y(s) ds, a \right]. \quad (8.1)$$

Neglecting the term  $[\exp(AT)c, a]$ , which is independent of  $y$ , we have the problem of maximizing

$$J(y) = \left[ \int_0^T \exp[A(T-s)]y(s) ds, a \right] \quad (8.2)$$

over all  $y_i$  satisfying the constraints

$$\begin{aligned} \text{(a)} \quad m_i &\leq y_i \leq m'_i, & 0 \leq t \leq T, & \quad i = 1, 2, \dots, N, \\ \text{(b)} \quad \int_0^T G(y_1, y_2, \dots, y_N) ds &\leq k. \end{aligned} \quad (8.3)$$

Introduce the function

$$f(k, T) = \max_y J(y), \quad (8.4)$$

where the maximum is over all functions  $y_i$  satisfying (8.3a) and (8.3b), and  $k$  is as in (8.3b).

It is easy reasoning as above to see that  $f(k, T)$  satisfies the equation

$$\begin{aligned} f(k, T) = \max_{\nu \in [0, \delta]} &\left[ \left( \int_0^\delta \exp[A(T-s)]y ds, a \right) \right. \\ &\left. + f \left[ k - \int_0^\delta G(y_1, y_2, \dots, y_N) ds, T - \delta \right] \right], \quad f(k, 0) \equiv 0. \end{aligned} \quad (8.5)$$

For computational purposes, we can use the relation

$$f(k, T) = \max_\nu [\delta(\exp(AT)v, a) + f(k - \delta G(v_1, v_2, \dots, v_N), T - \delta)]. \quad (8.6)$$

We see then that in the case where the underlying equation is linear, and there is only one constraint of the form in (8.3b), we can compute the solution using sequences of functions of one variable.

If there are two constraints of the type

$$\int_0^T G_i(y_1, y_2, \dots, y_N) ds \leq k_i, \quad i = 1, 2. \quad (8.7)$$

we introduce a Lagrange multiplier and consider the problem of maximizing

$$[x(T), a] - \lambda \int_0^T G_1(y_1, y_2, \dots, y_N) dt. \quad (8.8)$$

For each value of  $\lambda$  we have a one-dimensional problem. As the parameter  $\lambda$  is varied, we range over a set of values of the constraint  $\int_0^T G_1(y_1, y_2, \dots, y_N) ds$ .

In some cases, if the constraint is that given in (8.3b), we can solve the problem analytically, see [9, 10], and thus eliminate all computational aspects. It can easily



happen that a direct numerical solution for a range of values of  $k$  and  $c$  consumes less time and effort than a solution based upon an explicit analytic expression.

**9. Time-lag processes.** As the preceding sections show, the success of the method presented in the foregoing sections rested upon the superposition principle. Given an equation of the form

$$\begin{aligned} L(x) &= y, \\ x_{t=0} &= c, \end{aligned} \tag{9.1}$$

we were able to write the solution in the form

$$x = X(t)c + T(y), \tag{9.2}$$

thus avoiding any interaction between the initial state and the forcing function.

Also important was the fact that  $T(y)$  had the form

$$T(y) = \int_0^t K(t, s)y(s) ds. \tag{9.3}$$

It follows from this analysis, that the methods of the preceding sections are applicable to situations in which the underlying equations are differential-difference equations of the form given in (1.9). For the case where the coefficients are constant, Laplace transform methods yield the requisite representation formulas, see [7]. For the case of variable coefficients, these representation formulas have been developed in a forthcoming paper by the author and K. Cooke.

**10. Heat conduction processes.** In the study of the control of thermal processes, and in connection with the recent field of nuclear reactor control, we encounter variational problems in which the underlying equation is

$$\begin{aligned} u_t - u_{xx} &= f(x, t), & 0 < x < 1, & \quad t > 0, \\ u(x, 0) &= c(x), & 0 \leq x \leq 1, \\ u(0, t) &= u(1, t) = 0, & t > 0. \end{aligned} \tag{10.1}$$

Since the solution of this equation possesses the requisite properties described in the preceding section, it follows that a number of variational problems involving this equation can be treated by means of the foregoing techniques. A detailed discussion will appear subsequently.

**11. Successive approximations—I.** Let us now briefly, without entering into any rigorous discussion, which as may be imagined is non-trivial, indicate how the method of successive approximations may be combined with the foregoing techniques so as to reduce the computational solution of general variational problems involving non-linear differential equations to sequences of computational problems involving functions of fewer than  $N$  variables.

An important point to emphasize is that modern digital computers enable us to use as single steps in a computational process the solutions of problems once considered formidable in their own right.

Consider the problem of maximizing

$$J(y) = H[x_1(T), x_2(T), \dots, x_k(T)] \tag{11.1}$$

over all  $y_i(t)$  subject to

$$\begin{aligned}
 & \text{(a) } \frac{dx_i}{dt} = G_i(x, y), \quad x_i(0) = c_i, \quad i = 1, 2, \dots, N, \\
 & \text{(b) } m_i \leq y_i \leq m'_i, \quad i = 1, 2, \dots, N, \\
 & \text{(c) } \int_0^T G(y_1, y_2, \dots, y_N) dt \leq k.
 \end{aligned}
 \tag{11.2}$$

Let  $y^0(t) = [y_1^0(t), y_2^0(t), \dots, y_N^0(t)]$  be an initial guess in policy space, satisfying (11.2b) and (11.2c), and let  $x^0(t) = [x_1^0(t), x_2^0(t), \dots, x_N^0(t)]$  be the set of  $x$ -values determined by (11.2a) when  $y(t)$  is replaced by  $y^0(t)$ .

Consider the new system of differential equations

$$\begin{aligned}
 \frac{dx_i}{dt} = G_i(x^0, y^0) + \sum_{i=1}^N (x_i - x_i^0) \frac{\partial G_i}{\partial x_i}(x^0, y^0) \\
 + \sum_{i=1}^N (y_i - y_i^0) \frac{\partial G_i}{\partial y_i}(x^0, y^0) \quad i = 1, 2, \dots, N.
 \end{aligned}
 \tag{11.3}$$

The new variational problem is that of maximizing  $J(y)$  over all  $y$  satisfying (11.2b), (11.2c), and related to  $x$  by means of (11.3).

Since the underlying system in (11.3) is linear, this variational problem may be treated in terms of functions of  $k$  variables. Let the maximizing  $y(t)$  be called  $y^1(t)$ . Proceeding as above, we determine  $x^1(t)$  using (11.2a), and consider the new approximating linear equation

$$\frac{dx_i}{dt} = G_i(x^1, y^1) + \sum_{i=1}^N (x_i - x_i^1) \frac{\partial G_i}{\partial x_i}(x^1, y^1) + \sum_{i=1}^N (y_i - y_i^1) \frac{\partial G_i}{\partial y_i}(x^1, y^1).
 \tag{11.4}$$

The new variational problem determines a vector  $y^2$  which enables us to determine a new state vector  $x^2$  by way of (11.2a). Continuing in this way, we obtain a sequence of vector functions  $\{y^k\}$ , and a sequence of state vectors  $\{x^k\}$ , which we hope converges to a solution to the original variational problem.

**12. Successive approximations—II.** If the criterion function  $H[x_1(T), x_2(T), \dots, x_k(T)]$  is linear,

$$H(x_1, x_2, \dots, x_k) = (x, b),
 \tag{12.1}$$

then the approximation procedure outlined above yields a problem which at each step can be resolved computationally in terms of sequences of functions of one variable, and even in explicit analytic terms under favorable circumstances.

**13. Terminal control vs. general control.** We have placed considerable emphasis upon terminal control processes because of the fact that a simple transformation enables us to treat general control processes as terminal control processes. If

$$J(y) = \int_0^T F(x_1, x_2, \dots, x_N; y_1, y_2, \dots, y_N) dt,
 \tag{13.1}$$

the introduction of a new variable,  $x_{N+1}(t)$ , determined by the relation

$$\frac{dx_{N+1}}{dt} = F(x_1, x_2, \dots, x_N; y_1, y_2, \dots, y_N), \quad x_{N+1}(0) = 0,
 \tag{13.2}$$

yields a new problem in which we wish to maximize  $x_{N+1}(T)$ .

Similarly, if we have the criterion function  $G[x_1(T), x_2(T), \dots, x_N(T)]$ , a new variable,  $x_{N+1}(t)$ , determined by the relation

$$\frac{dx_{N+1}}{dt} = \sum_{i=1}^N \frac{\partial G}{\partial x_i} \frac{dx_i}{dt} = \sum_{i=1}^N \frac{\partial G}{\partial x_i} H_i(x, y), \quad (13.3)$$

once again reduces the variational problem to one of terminal control.

#### BIBLIOGRAPHY

1. R. Bellman, *Dynamic programming*, Princeton University Press, 1957
2. R. Bellman, *Dynamic programming and its application to variational problems in mathematical economics*, Proceedings Symposium on Calculus of Variations and its Applications, Am. Math. Soc., 1955
3. R. Bellman, *On the application of the theory of dynamic programming to the study of control processes*, Symposium on Non-linear circuit analysis, Polytechnic Institute of Brooklyn, vol. VI, 1956
4. R. Bellman, *Notes on the theory of control processes—I: On the minimum of maximum deviation*, Quart. Appl. Math. XIV, 419-423 (1957)
5. R. Bellman, *On a dynamic programming approach to the caterer problem—I*, Management Sci. 3, 270-278 (1957)
6. R. Bellman, *Stability theory of differential equations*, McGraw-Hill, 1953
7. R. Bellman, *A survey of the mathematical theory of time-lag, retarded control, and hereditary processes*, The RAND Corp. Rept. R-256, 1954
8. R. Bellman, *Dynamic programming and Lagrange multipliers*, Proc. Natl. Acad. Sci. 42, 767-769, (1956)
9. R. Bellman, I. Glicksberg and O. Gross, *On some variational problems occurring in the theory of dynamic programming*, Rendiconti del Circolo Matematico di Palermo, Serie II, Tomo III 1-35 (1954)
10. R. Bellman, W. H. Fleming, and D. V. Widder, *Variational problems with constraints*, Annali di Matematica, Serie IV, Tomo XLI, 301-323, 1956
11. S. Dreyfus, *Dynamic programming solution of allocation problems*, Techniques of Industrial Operations Research Seminar, Illinois Institute of Technology, June 1956
12. H. Osborn, *Euler equations and characteristics*, Chap. 7 of *Dynamic programming of continuous processes*, The RAND Corp., Rept. R-271, 1955