Rivista Mat. Univ. Parma 6, 193–213 (1955)

RICHARD BELLMAN (*)

Dynamic programming and a new formalism in the Calculus of variations – I. (**)

§ 1. – Introduction.

In a series of recent papers, see [1] and [3] where bibliographies may be found, we have investigated the theoretical aspects of multistage decision processes of both deterministic and stochastic type. Since these are programming problems of non-static type we have coined the name « dynamic programming » to describe them.

A fundamental tool in our investigations is the use of functional equations. The purpose of this paper is to show how this general method may be applied to the calculus of variations regarded as a multi-stage decision process of continuous type. Consistent with this initial aim of exposition we shall restrict ourselves only to the formalism of the method and bypass all questions of rigor in this first paper of the series. Actually it is not difficult, using classical results in the calculus of variations, to justify our results and procedures as we shall show in the second paper. However, we feel that this justification would at the moment obscure the scene with analytic foliage.

We furthermore restrict ourselves to one-dimensional problems. A treatment of multi-dimensional problems would seem to involve functionals and functional analysis. A hint of this appears in our corresponding treatment of integral equation theory, [4].

By use of functional equations we shall obtain partial differential equations for the extremum values of integrals as functions of certain state variables. Using these equations we can obtain successive approximations to the extre-

(*) Address: The RAND Corporation, 1700 Main St., Santa Monica, California, U.S.A.

(**) Received May 20, 1955.

mum value and extremal curves, and monotone convergence. We shall illustrate these points by a consideration of the problem of determining the maximum of

$$J(y) = \int_0^T F(x, y)\, dt \tag{1}$$

subject to

$$\frac{dx}{dt} = G(x, y), \qquad x(0) = c, \tag{2}$$

and the eigenvalue problem associated with

$$u'' + \lambda^2 \Phi(t) u = 0, \qquad u(0) = u(1) = 0\,. \tag{3}$$

§ 2. – The Calculus of variations as a continuous decision process.

Before showing why we may regard various problems in the calculus of variations as continuous decision processes, let us discuss the concept of a decision process itself.

Let S be a spce of some type, N-dimensional Euclidean or, as more frequently occurs, a function space and P a typical point in this space. Let $\{ T(P, Q) \}$, where Q belongs to another space R, be a set of transformations of point in S into points in S.

Furthermore let D be a domain in S with the property that $P \in D$ implies $T(P, Q) \in D$ for all $Q \in R$. We shall occasionally refer to P as a *state variable* below.

A choice of Q is a choice of a transformation and we call this choice a *decision*. A sequence of decisions, i.e. a sequence of Q's, $\{ Q_k \}$, discrete or continuous, we call a *policy*. Each policy yields a corresponding sequence of P's, $\{ P_k \}$, which in many cases will be a set of stochastic variables. Considering for the moment an N-stage process, one where we make N decisions, let $F(P)$ be some function defined for $P \in D$, and assume that we are to choose these N decisions so as to maximize $F(P_N)$. If $\{ P_k \}$ is a stochastic sequence, then $F(P)$ will be an expected value. This maximum will be a function of P, the initial point, and N the number of states. We may then write

$$f_N(P) = \underset{\{Q_k\}}{\text{Max}}\, F(P_N)\,. \tag{1}$$

A policy which yields $f_N(P)$ we call an *optimal policy*. By a solution to a decision process we mean the determination of all optimal policies.

It is clear that an optimal policy is characterized by the following intuitively obvious:

Principle of Optimality. *An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.*

Using this principle, we see that the members of the sequence $\{ f_N(P) \}$ satisfies the functional recurrence relation

$$f_{N+1}(P) = \operatorname*{Max}_{Q} f_N(T(P, Q)). \tag{2}$$

If we have a continuous process where the discrete sequence 1, 2, ..., is replaced by a continuous parameter T, then (2) becomes

$$f(P, S + T) = \operatorname*{Max}_{D[0, S]} f(T_D(P), T), \tag{3}$$

where $T_D(P)$ is the state at time S due to decisions over the interval $[0, S]$, and we maximize over all policies over this interval. A more detailed discussion may be found in [**1**] and [**3**].

Let us now show that we can transform a general class of problem in the calculus of variations into continuous decision processes. In some cases, the transformation is almost immediate; in others it requires the introduction of additional state variables.

Consider the problem of maximizing the functional

$$J(x) = \int_0^T F(x, x', t)\, dt \tag{4}$$

The classical approach is to obtain variational conditions which characterize a local maximum, $\bar{x}$, by considering $J(\bar{x} + \varepsilon x)$ for small ε and suitably chosen x. This is modelled after the finite dimensional approach, and considers the extremal curve, $\bar{x}$, as a point in function space.

Alternatively, we can consider the choice of $\bar{x}(t)$ as a continuous decision process where we must choose our continuation along $\bar{x}(t)$ at each time t. That is to say at each time t we determine $d\bar{x}(t)/dt$ which in turn determines the continuation of the curve.

In geometric terms we are determining an extremal as an envelope of tangents. This is the approach of the theory of dynamic programming as contrasted with the classical approach which views a curve as a locus of points. Our emphasis then is upon determining the « im kleinem » continuation from a generic position rather than the « im grossem » continuation from a fixed initial point.

Since we are exploiting the duality of Euclidean space, we might expect that this duality should manifest itself analytically by means of the theory of characteristics of partial differential equations, and this is indeed so.

It seems reasonable to expect also that the combination of both approaches should yield more than just one or the other, and this also seems to be so. Particularly the concept of a policy, and the further concept of « approximation in policy space », seems to be fruitful.

For those who are interested in seeing further applications of the theory of dynamic programming, we recommend the papers [**1**], [**3**]. The results contained in this paper were announced in two Notes [**5**], [**6**].

§ 3. – Applications – I.

As an application of the concepts which we have discussed above let us consider the problem of determining the maximum of

$$J(y) = \int_0^\infty F(x,\ y)\, dt, \tag{1}$$

subject to the constraint

$$\frac{dx}{dt} = G(x,\ y), \qquad x(0) = c\,. \tag{2}$$

We set

$$\operatorname*{Max}_y J(y) = f(c)\,. \tag{3}$$

Then, we have

$$f(c) = \operatorname*{Max}_{D[0,\,s]} \left[\int_0^s F(x,\ y)\ dt + \int_s^\infty F(x,\ y)\ dt\right], \tag{4}$$

where $D[0,\ s]$ is a choice of $y(t)$ over the interval $[0,\ s]$.

Employing the principle of optimality cited above, (4) is equivalent to

$$(5) \qquad f(c) = \underset{D[0,s]}{\mathrm{Max}} \left[\int_0^s F(x, y)\, dt + f(c(s)) \right].$$

Let us now assume that $f(c)$ has a continuous derivative and that $y(t)$ is continuous in t. Then, (5) yields

$$(6) \qquad f(c) = \underset{y(0)}{\mathrm{Max}} \left[s\, F(c, y(0)) + f(c) + s\, G(c, y(0))\, f'(c) + o(s) \right],$$

as $s \to 0$ using the consequence of (2), $c(s) = c + s\, G(c, y(0)) + o(s)$. The limiting relation is

$$(7) \qquad 0 = \underset{v}{\mathrm{Max}} \left[F(c, v) + G(c, v)\, f'(c) \right],$$

where we have set for typographical convenience $v = y(0)$. The maximum occurs at the values of v determined by

$$(8) \qquad 0 = F_v + G_v\, f'(c).$$

Hence, for the determination of $v = v(c)$ and $f'(c)$ we have the two simultaneous equations

$$(9) \qquad 0 = F + G\, f'(c), \qquad 0 = F_v + G_v\, f'(c).$$

These yield, upon elimination of $f'(c)$,

$$(10) \qquad \begin{vmatrix} F(c, v) & G(c, v) \\ F_v(c, v) & G_v(c, v) \end{vmatrix} = 0$$

as the equation determining v as function of c. This function yields $y(0)$ as a function of $x(0)$ for any initial point or generally, y as a function of x for any t. To determine x and y as functions of t, we return to (2) and use the equation

$$(11) \qquad \frac{dx}{dt} = G(x, y(x)), \qquad x(0) = c,$$

to determine $x(t)$ and thus $y(t)$.

It is interesting to observe that x and y as functions of t may be quite complicated, yet y may be simple function of x. Thus the solution in terms of optimal policies may be much more meaningful and informative. Essentially, y as a function of x is an instrinsic equation of the curve, much as the definition of a circle as the locus of points equidistant from a given point is an intrinsic definition.

If $G(x, y)$ is taken to be negative for all values of x and y, equation (7) is susceptible to a very interesting interpretation. We have

$$0 \geqslant F(c, v) + G(c, v) f'(c) \tag{12}$$

for all v with equality for at least one value $v = v(c)$. Hence

$$f'(c) \geqslant F(c, v)/\{-G(c, v)\} \tag{13}$$

with equality for at least one value, which implies

$$f'(c) = \operatorname*{Max}_{v} [F(c, v)/\{-G(c, v)\}]. \tag{14}$$

If we think as the original process as an economic process, then the interpretation of (14) is that at each time we proceed so as to maximize immediate « gain » over immediate « cost. ».

It may be verified that (10) is a first integral of the EULER equation derived used the classical variational techniques, and (14) is again equivalent to the EULER equation.

Applying the same technique to the two-dimensional problem of maximizing

$$J(y_1, y_2) = \int_0^\infty F(x_1, x_2, y_1, y_2) \, dt \tag{15}$$

subject to the constraints

$$\begin{cases} \dfrac{dx_1}{dt} = G(x_1, x_2, y_1, y_2), & x_1(0) = c_1, \\ \dfrac{dx_2}{dt} = H(x_1, x_2, y_1, y_2), & x_2(0) = c_2, \end{cases} \tag{16}$$

we obtain the equation

$$(17)\qquad \begin{vmatrix} F(c_1,\ c_2,\ u,\ v) & G(c_1,\ c_2,\ u,\ v) & H(c_1,\ c_2,\ u,\ v) \\ F_u & G_u & H_u \\ F_v & G_v & H_v \end{vmatrix} = 0,$$

connecting $u = x_1(0)$ and $v = x_2(0)$.

It is again easy to verify that (17) is a consequence of the two EULER equations. We have not been able to the present to obtain another equation which will yield expressions for u and v in terms of c_1 and c_2 and thus solve the original problem without the intervention of differential equations.

§ 4. – Applications - II.

Let us now consider the more general problem of maximizing

$$(1)\qquad J(y) = \int_0^T F(x,\ y,\ t)\ \mathrm{d}t,$$

subject to the constraint

$$(2)\qquad \frac{\mathrm{d}x}{\mathrm{d}t} = G(x,\ y,\ t), \qquad x(0) = c\,.$$

We imbed this problem within the problem of maximizing

$$(3)\qquad J(y) = \int_a^T F(x,\ y,\ t)\ \mathrm{d}t,$$

subject to the constraint

$$(4)\qquad \frac{\mathrm{d}x}{\mathrm{d}t} = G(x,\ y,\ t), \qquad x(a) = c\,,$$

Keeping T fixed we may write

$$(5)\qquad \operatorname*{Max}_{y}\ J(y) = f(a,\ c).$$

The functional equation for f is

$$f(a,\ c) = \underset{D[a,\, a+S]}{\mathrm{Max}} \left[\int_a^{a+S} F(x,\ y,\ t)\ dt + f(a+S,\ c(s)) \right], \tag{6}$$

which yields, following the same formal procedure as before, in the limit as $S \to 0$,

$$0 = \underset{v}{\mathrm{Max}} \left[F(c,\ v,\ a) + f_a + G(c,\ v,\ a)\, f_c \right], \tag{7}$$

where $v = v(a,\ c)$ is the value of $y(0)$.

From (7) we obtain the two equations

$$\begin{cases} 0 = F(c,\ v,\ a) + f_a + G(c,\ v,\ a)\, f_c, \\ 0 = F_v \qquad\qquad\quad + G_v\, f_c\,. \end{cases} \tag{8}$$

Solving for f_c and f_a, we obtain

$$\begin{cases} f_c = -F_v/G_v = P(c,\ v,\ a), \\ f_a = (FG_v - F_vG)/G_v = Q(c,\ v,\ a), \end{cases} \tag{9}$$

with P and Q known functions of c, v and a. Equating f_{ca} and f_{ac}, we obtain the first order partial differential equation for v,

$$P_v\, v_a + P_a = Q_v\, v_c + Q_c\,. \tag{10}$$

It may again be verified readily that the characteristics of this equation are equivalent to the EULER equation obtained from (3) and (4). This result carries over to the more general multi-dimensional situation. Here, however, the proof of the equivalence is much more complicated, see OSBORN **[8]**.

The value of v at $a = T$ is obtained by maximizing $F(c,\ v,\ T)$ as a function of v.

§ 5. – Constraints – I.

In many problems of interest in physical, economic, engineering and direct analytic applications, a forthright application of the classical variational technique is not possible because free variations are not always permissibile. Problems of this kind occur in many physical situations when we impose realistic bounds on velocity, acceleration or radius of curvature, and similarly in mathematical economics when we impose bounds on rate of allocation, and so forth.

Let us consider a simple problem of this type. We wish to maximize

$$J(y) = \int_0^T F(x, \; y) \, \mathrm{d}t, \tag{1}$$

subject to the constraints

$$\begin{cases} \text{(a)} \quad \dfrac{\mathrm{d}x}{\mathrm{d}t} = G(x, \; y), \qquad x(0) = c, \\ \text{(b)} \quad 0 \leqslant y \leqslant x \, . \end{cases} \tag{2}$$

Let

$$f(c, \; T) = \operatorname*{Max}_{y} J(y) \, . \tag{3}$$

Then, as above, f satisfies the equation

$$f_T = \operatorname*{Max}_{0 \leq v \leq c} \left[F(c, \; v) + G(c, \; v) \; f_c\right]. \tag{4}$$

We have in another place, [7], discussed this problem and obtained the structure of the solution under various assumptions concerning F and G, using both the classical method and the techniques of dynamic programming.

Let us proceed formally here to illustrate how the equation in (4) may be used to obtain an over-all view of the solution. We shall assume that $F(x, y)$ and $G(x, y)$ are *concave* functions of y. Then $F(c, v) + G(c, v) \; f_c$ is a concave function of v provided we assume that $f_c \geqslant 0$. This will be true if $F(x, y)$ and $G(x, y)$ are monotone increasing in x.

Since a concave function has a unique maximum which must occur at $v = 0$, c or a point in between, we see that the solution will have intervals where $y = x$ followed by intervals where $0 < y < x$ followed by intervals where $y = 0$ or x, and so on. An interval where $y = 0$ cannot be followed directly by or follow directly one where $y = x$. If we wish to obtain more information concerning the structure of the solution, assume that f_c is monotone increasing in T. In particular, assume that $f_c \to \infty$ as $T \to \infty$, and also that $F_v < 0$, $G_v > 0$, for all c.

At $T = 0$, $F_v < 0$ and $v = 0$ is the maximum value. As T increases, $F(c, v) + G(c, v)$ achieves a maximum inside the interval $[0, c]$. As T gets still larger $G_v f_c + F_v$ becomes positive for all v in $[0, c]$. Hence the maximum stays at $v = c$.

In (x, y) space this means that the solution has the structure

$$(5)\qquad \begin{cases} \text{(a)} & y = x, \qquad T \geqslant T_1(c), \\ \text{(b)} & 0 < y < c, \qquad T_2(c) < T < T_1(c), \\ \text{(c)} & y = 0, \qquad 0 \leqslant T \leqslant T_2(c). \end{cases}$$

As we have mentioned above, rigorous details may be found in [7].

§ 6. – Constraints – II.

Let us now consider the same problem with the additional constraint

$$(1)\qquad \int_a^T y \, dt \leqslant m .$$

Then

$$(2)\qquad \operatorname*{Max}_y J(y) = f(a, c, m) .$$

The functional equation for f is now

$$(3)\qquad f_T = \operatorname*{Max}_{0 \leq v \leq c} [F(c, v) + G(c, v) f_c - v f_m] .$$

Here the analysis is more difficult and we have not as yet investigated any particular problems using this technique. We have, in the meantime, developped a new method, combining the functional equation approach whith LAGRANGE multipliers, wich is particularly applicable to problems of this type.

§ 7. – Successive approximations.

Returning to an equation such as (3.5), it is tempting to envisage the use of successive approximations in solving this equation. If we choose an initial function $f_0(c)$ and define

$$(1)\qquad f_1(c) = \operatorname*{Max}_{D[0, S]} \left[\int_0^S F(x, y) \, dt + f_0(c(S)) \right],$$

we see that in the limit as $S \to 0$ we must have $f_1(c) = f_0(c)$.

At first sight this would seem to render the use of successive approximations impossible. The answer is that we must approximate in *policy space* rather than *function space*. What this means is that we must concentrate on the function $v(c, T)$ rather than the function $f(c, T)$.

Let us, to illustrate this point, discuss the problem of maximizing

$$J(y) = \int_0^T F(x, y)\, dt \tag{2}$$

subject to $dx/dt = G(x, y)$, $x(0) = c$. We choose an initial approximation $v_0 = v_0(c, T)$, which is equivalent to $y_0 = y_0(x, T - t)$. Using this value of y_0, we compute x_0 by means of the differential equation

$$\frac{dx_0}{dt} = G(x, y_0(x_0, T - t)), \qquad x_0(0) = c. \tag{3}$$

Having x_0 and y_0 we compute

$$f_0(c, T) = \int_0^T F(x_0, y_0)\, dt. \tag{4}$$

This function, f_0, satisfies the partial differential equation

$$f_{0T} = F(c, v_0) + G(c, v_0) f_{0c}. \tag{5}$$

To determine the next approximation to an extremal y, an « optimal policy », we determine $v_1(c, t)$ as a function which maximizes

$$F(c, v) + G(c, v)\, f_{0c}, \tag{6}$$

Let us for the moment assume that v_1 is unique. Using v_1 we compute x_1 and f_1 as above. Having obtained f_1 we compute v_2 as the function $v = v_2(c, T)$ which maximizes

$$F(c, v) + G(c, v)\, f_{1c}, \tag{7}$$

and continue in this way, deriving a sequence of approximations to f, $\{ f_n \}$, and a sequence of approximations to v, $\{ v_n \}$.

§ 8. – Monotone approximations.

Let us now show that this sequence of approximations is *monotone* increasing. We shall in the second paper of this series discuss the convergence of the process. This fact has important theoretical and computational advantages. We have

$$(1)\qquad \begin{cases} f_{1T} = F(c,\, v_1) + G(c,\, v_1)\, f_{1c}, \\ f_{0T} = F(c,\, v_0) + G(c,\, v_0)\, f_{0c} \leqslant F(c,\, v_1) + G(c,\, v_1)\, f_{0c}\,. \end{cases}$$

Hence

$$(2)\qquad f_{1T} - f_{0T} \geqslant G(c,\, v_1)\ (f_{1c} - f_{0c}),$$

and from this follows that $f_1 \geqslant f_0$.

§ 9. – Eigenvalue problems.

Let us now turn to the problem of ascertaining the values of λ which permit a non-trivial solution of the equation

$$(1)\qquad u'' + \lambda^2\, \Phi(t)\, u = 0, \qquad u(0) = u(1) = 0,$$

to exist.

Under light conditions upon $\Phi(t)$, this is equivalent to the problem of determining the relative minima of $\int_0^1 u'^2\, dt$ subject to the constraints

$$(2)\qquad \int_0^1 \Phi(t)\ u^2\ dt = 1, \qquad u(0) = u(1) = 0,$$

or, conversely, that of determining the relative maxima of $\int_0^1 \Phi(t) u^2\ dt$ subject to the constraints

$$(3)\qquad \int_0^1 u'^2\, dt = 1, \qquad u(0) = u(1) = 0\,.$$

In this form the problem is quite similar to the the problems we have considered above.

There is, however, one major difference. In the previous problems, it was immediately evident what the state variables were. In this problem, as we proceed along an extremal, the type of problem changes since the condition $u(0) = 0$ is violated immediately.

Consequently, we must :mbed this problem in a class of problems which possess the requisite invariance properties. There are several ways of doing this.

The first method we shall employ is that of determining the minimum of

$$J(u) = \int_a^T u'^2 \, dt. \tag{4}$$

Subject to the constraints

$$\begin{cases} \text{(a)} \quad \int_a^T \Phi(t) u^2 \, dt + k \int_a^T (T - t)\Phi(t) \, dt = 1, \\ \text{(b)} \quad u(a) = u(T) = 0. \end{cases} \tag{5}$$

As we shall show below this will provide us with an invariant formulation. The state parameters are a and k, keeping T fixed. Let us write

$$\operatorname*{Min}_u J(u) = f(a, \ k). \tag{6}$$

We shall derive a partial differential equation for f below which will be nonlinear. Nevertheless it will be useful for computational purposes, and can be utilized to obtain power series in k.

In order to obtain a useful approximation to the solution of this equation, we can if we wish consider the simpler problem of determining the minimum of $\int_a^T u'^2 \, dt$ subject to the constraints

$$\begin{cases} \text{(a)} \quad \int_0^T u^2 \, dt + k \int_a^T (T - t) \, u(t) \, dt = 1, \\ \text{(b)} \quad u(a) = u(T) = 0 \, . \end{cases} \tag{7}$$

For a close to T we may write, for $a \leqslant t \leqslant T$,

(8) $$\Phi(t) \cong \Phi(T),$$

and so obtain (7) with $k' = k\, \Phi(T)$. A better approximation would be to set

(9) $$\Phi(t) \cong \Phi(T) + (t - T)\ \Phi'(T),$$

a device used by LANGER in connection with the *WKB* method. Although the coefficients are no longer constant, the problem with $k = 0$ is soluble in terms of BESSEL functions of order 1/3.

§ 10. – The approximation of (9.7).

Let us now consider the approximation of (9·7) above. We write [1]

(1) $$f(k,\ T) = \operatorname*{Min}_u \int_0^T u'^2\ dt\ .$$

To obtain a functional equation for f we write, following the method we have employed above,

(2) $$f(k,\ S + T) = \int_0^S u'^2\,dt + \int_S^{S+T} u'^2\,dt\,,$$

for an extremal u. We now wish to express the second integral in terms of k^* and T, where $k^* = k^*(k,\ S,\ u)$.

Since we are interested principally in the partial differential equation for f, we shall restrain ourselves to the case where S is small. From (2) we obtain

(3) $$f(k,\ S + T) = S\, u'^2(0) + \int_S^{S+T} u'^2\,dt + o(S)\,.$$

---

[1] We are now keeping the lower limit fixed and using the upper limit as a state variable.

We must now perform a change of variable, transforming u into a variable v which is zero at $t = S$ and $t = S + T$.

We set

$$v(t) = u(t) - S\,u'(0)\,(S + T - t)/t\,. \tag{4}$$

Then

$$\begin{cases} v(S) = u(S) - Su'(0) + o(S), \\ v(S + T) = u(S + T) = 0\,. \end{cases} \tag{5}$$

Hence to terms in $o(S)$,

$$v(S) = v(S + T) = 0\,. \tag{6}$$

The expression for $u'(t)$ is

$$u'(t) = v'(t) - S\,u'(0)/T, \tag{7}$$

which yields

$$\int_S^{S+T} u'^2\,dt = \int_S^{S+T} v'^2\,dt - \frac{2Su'(0)}{T}\int_S^{S+T} v'\,dt + o(S) = \int_S^{S+T} v'^2\,dt + o(S). \tag{8}$$

Combining this with (3), we may write

$$f(k,\; S + T) = S\,u'^2(0) + \int_S^{S+T} v'^2\,dt + o(S)\,. \tag{9}$$

It remains to convert the constraint of (9·7a) into one involving v. We have

$$\int_S^{S+T} u^2\,dt + k\int_S^{S+T} (S + T - t)\,u\,dt = \tag{10}$$

$$= 1 - \int_0^S u^2\,dt - k\int_0^S (S + T - t)u\,dt = 1 + o(S)\,.$$

Replacing $u(t)$ by its expression in terms of $v(t)$, the above equation yields, after some slight simplification,

$$(11)\quad \int_S^{S+T} v^2\,dt + \left(k + \frac{2Su'(0)}{T}\right)\int_S^{S+T}(S+T-t)\,v\,(t)\,dt = 1 - kS\,u'(0)\,T^2/3 + o(S)\,.$$

To normalize this relation so that it will have the form of the original constraint, we set

$$(12)\qquad v(t) = w(t)\left\{1 - k\,Su'(0)\,T^2/6\right\}.$$

The relation in (11) then becomes

$$(13)\quad \int_S^{S+T} w^2\,dt + \left(k + \frac{2Su'(0)}{T} - \frac{Su'(0)k^2T^2}{6}\right)\int_S^{S+T}(S+T-t)\,w(t)\,dt = 1 + o(S),$$

with

$$(14)\qquad w(S) = w(S+T) = o(S).$$

Furthermore

$$(15)\quad f(k,\ S+T) = S\,u'^2(0) + \left\{1 - k\,S\,u'(0)T^2/6\right\}\int_S^{S+T} w'\,dt + o(S).$$

§ 11. – The functional equation for $f(k,\ T)$.

We are now ready to derive the functional equation for $f(k,\ T)$.

Combining (10.15) with (10.13) and (10.14), we see that we have

$$(1)\quad f(k,\ S+T) = S\,u'^2(0) + \left(1 - \frac{k\,Su'(0)\,T^2}{6}\right) f\left(k + \frac{2S\,u'(0)}{T} - \frac{S\,u'(0)\,k^2T^2}{6},\ T\right) + o(S)\,.$$

Since $u'(0)$ is to be chosen to minimize $f(k, S+T)$, we obtain the basic functional relation

(2) $$f(k, S+T) = \\ = \operatorname*{Min}_v \left[S v^2 + \left(1 - \frac{k S v T^2}{6}\right) f\left(k + \frac{2S v}{T} - \frac{S v k^2 T^2}{6}, T\right)\right] + o(S),$$

where we have $v = u'(0) = v(k, T)$.

Expanding both sides in powers of S, this yields in the limit as $S \to 0$,

(3) $$f_T = \operatorname*{Min}_v \left[v^2 + v \left\{ \left(\frac{2}{T} - \frac{k^2 T^2}{6}\right) f_k - \frac{k T^2}{6} f \right\}\right].$$

The minimum is assumed for

(4) $$v = -\frac{1}{2}\left[\left(\frac{2}{T} - \frac{k^2 T^2}{6}\right) f_k - \frac{k T^2 f}{6}\right].$$

Substituting in (3), we see that f satisfies the nonlinear partial differential equation

(5) $$f_T = -\frac{1}{4}\left[\left(\frac{2}{T} - \frac{k^2 T^2}{6}\right) f_k - \frac{k T^2 f}{6}\right]^2.$$

§ 12. - Power series in k.

For $k = 0$, the characteristic functions are

(1) $$u = \sqrt{\frac{2}{T}} \sin(2\pi m/T) \qquad (m = 1, 2, \ldots).$$

yielding as the minimum of $\int_0^T u'^2 \, dt$ the value

(2) $$J_0 = 4\mu^2/T^2.$$

Let us then attempt to find a power series for $f(k, T)$ in k,

(3) $$f(k, T) = J_0 + kJ_1 + k^2 J_2 + \ldots.$$

Substituting in (11·5), we may determine the coefficients recurrently. In particular

$$(4)\qquad \begin{cases} J_1 = \pi\sqrt{8}/\sqrt{T}, \\ J_2 = \left(\dfrac{\pi^2}{6} + \dfrac{1}{8}\right) T . \end{cases}$$

Continuing in this way we can obtain an excellent approximation to $f(k, T)$ for small k and T. Having the solution for small T we can use numerical integration to determine it for larger values. These results are useful in connection with the approsimation $\Phi(a) \cong \Phi(T)$ discussed above.

§ 13. – The functional equation in the general case.

Using the same methods as above we may show that the function $f(a, k)$ as defined by [2]

$$(1)\qquad f(a, k) = \operatorname*{Max}_u \left[\int_a^1 \Phi(t)u^2\,dt + k\int_a^1 (1-t)\Phi(t)u\,dt\right],$$

subject to the constraints

$$(2)\qquad \begin{cases} \text{(a)} \quad \displaystyle\int_a^1 u'^2\,dt = 1, \\ \text{(b)} \quad u(a) = u(1) = 0, \end{cases}$$

satisfies the nonlinear partial differential equation

$$(3)\qquad f_a = \left(\frac{2 + k\,G(a)}{1-a}\right)^2 / (k\,f_a - f),$$

where

$$(4)\qquad G(a) = \int_a^1 (1-t)^2\,\Phi(t)\,dt .$$

The advantage of considering this formulation rather than the one above resides in the fact that $f(a, k) \to 0$ as $a \to 1$.

---

[2] We have now reverted to use of the lower limit.

§ 14. – The higher eigenvalues.

There are several methods by which the problem of determining functional equations for the higher eigenvalues may be approached, based upon min-max characterizations, or upon variational representations involving restraints. We shall discuss these questions at a later date.

§ 15. – Eigenvalue problems. Second approach.

In the first approach above we maintained the boundary conditions $u(a) = u(T) = 0$ at the expense of introducing the somewhat artificial constraint

$$\int_a^T \Phi(t)\, u^2\, dt + k \int_a^T (T - t)\, \Phi(t)\, u\, dt = 1. \tag{1}$$

In lieu of this, let us consider the problem of determining the minimum of $\int_a^T u'^2\, dt$ subject to the constraints

$$\begin{cases} \text{(a)} \quad \int_a^T \Phi(t)\, u^2\, dt = 1, \\ \text{(b)} \quad u(a) = k, \quad u(T) = 0\,. \end{cases} \tag{2}$$

The procedure is very much as above, and we derive for

$$f(a, k) = \operatorname*{Min}_u \int_a^T u'^2\, dt, \tag{3}$$

the nonlinear partial differential equation

$$f_T = - f_k^2/4 + k^2\, \Phi(a)\, f_k/2 - k^2\, \Phi(a)\, f\,. \tag{4}$$

§ 16. – Monotone convergence of eigenvalues.

Using the concept of approximation in policy space, we can now obtain a monotone sequence of approximations to the eigenvalues. Let us consider, for example, the problem of determining the maximum of

$$J(u) = \int_a^{a+T} \Phi(t)\, u^2\, dt + k \int_a^{a+T} (T + a - t)\Phi(t)\, u\, dt \tag{1}$$

subject to

$$(2)\quad \begin{cases} \text{(a)} & u(a) = u(a+T) = 0, \\ \text{(b)} & \int_a^{a+T} u^2\, dt = 1. \end{cases}$$

Using the above technique, we obtain for

$$(3)\qquad f(a,\ k,\ T) = \operatorname*{Max}_u\ J(u),$$

the partial differential equation

$$(4)\qquad f_T = \operatorname*{Max}_v\ [H(v,\ f)],$$

where

$$(5)\qquad H(v,\ f) = f_a + v\cdot\left[\frac{2f_k}{T} + G(a)\right] + v^2\left[\frac{kf_k}{2} - f\right],$$

with

$$(6)\qquad G(a) = \frac{k}{T}\int_a^{a+T} (T + a - t)^2\, \Phi(t)\, dt.$$

To obtain a sequence of successive approximations, we make our first approximation in policy space,

$$(7)\qquad v_0 = u_0'(a) = v_0(a,\ k,\ T)\,.$$

Using this function v_0 we compute f_0 as the solution of the partial differential equation

$$(8)\qquad f_{0T} = H(v_0,\ f_0), \qquad f_0(a,\ k,\ 0) = 0\,.$$

Now choose $v_1 = u_1'(a)$ as the function of v which maximizes $H(v,\ f_0)$,

$$(9)\qquad v_1 = \left\{ G(a) + 2f_{0k}/T \right\}/2(f_0 - kf_{0k}/2)\,.$$

Using this value of v_1 we compute f_1 as the solution of

$$f_{1r} = H(v_1, f_1), \qquad f_1(a, k, 0) = 0 . \tag{10}$$

and continue in this fashion.

The proof that $f_1 \leqslant f_0$ is now almost precisely as before.

Bibliography.

1. R. Bellman, *An introduction to the theory of dynamic programming*, Rand Monograph No. R-245, June 1953.

2. R. Bellman, *Dynamic programming of continuous processes*, Rand Monograph No. R-271, November 1954.

3. R. Bellman, *The theory of dynamic programming*, Bull. Amer. Math. Soc. **60**, 503-516 (1954).

4. R. Bellman, *Dynamic programming and a new formalism in the theory of integral equations*, Proc. Nat. Acad. Sci. U.S.A. **41**, 31-34 (1955).

5. R. Bellman, *Monotone convergence in dynamic programming and the Calculus of variations*, Proc. Nat. Acad. Sci. U.S.A. **40**, 1073-1075 (1954).

6. R. Bellman, *Dynamic programming and a new formalism in the Calculus of variations*, Proc. Nat. Acad. Sci. U.S.A. **40**, 231-235 (1954).

7. R. Bellman, W. Fleming and D. Widder, *Studies in the Calculus of variations*, Ann. Mat. Pura Appl., (to appear).

8. H. Osborn, *Euler equations and characteristics*, (to appear).