1

FUNCTIONAL EQUATIONS IN THE THEORY OF DYNAMIC PROGRAMMING—III

MULTI-STAGE GAMES

By

Richard Bellman

P-676

13 May 1955
Rev. 14 March 1956

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA

Summary

~~In this~~ paper ~~we~~ establishs existence and uniqueness theorems for a class of functional equations occurring in the theory of multi-stage games.

CONTENTS

§1. Introduction.

In the preceding papers of this series [1], [3], we have studied two particular classes of functional equations which arise in the theory of dynamic programming. The first paper, [1], presented some results concerning the equation

$$(1) \qquad f(P) = \operatorname*{Max}_{Q} \left[R(P,Q)+h(P,Q)f(T(P,Q))\right],$$

(a detailed exposition is to be found in [2]), while the second paper of the series, [3], was devoted to a discussion of equations of the form

$$(2) \qquad \frac{dx_i}{dt} = \operatorname*{Max}_{Q} f_i(x_1,x_2,\ldots,x_n,t;Q), x_i(0)=c_i, i=1,2,\ldots n.$$

Equations of the first type arise from the study of discrete processes, while equations of the second type arise from the study of a certain class of continuous decision processes. Partial differential analogues of (2) arise from the calculus of variations, and from the theory of integral equations, cf [4],[5].

All of these equations pertain to one-person decision processes. In this paper, we shall study a class of functional equations arising from the theory of multi-stage games. The equation we shall use to illustrate our techniques is

$$(3) \qquad f(P,P') = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} \left[\int\int_{D(P,P')} \left[R(u,v)+h(u,v)f(T,T')\right]dG(u)dG'(v)\right]$$
$$= \operatorname*{Min}_{G'} \operatorname*{Max}_{G} \left[\ \ldots\ \right],$$

where $T = T(P,P';u,v)$, $T'=T'(P,P';u,v)$, and $G(u),G'(v)$ are distribution functions for u and v respectively over the allowable regions. To simplify the notation, we write $R(u,v)$ and $h(u,v)$, although we actually allow these functions to depend upon P and P'.

The precise restrictions we shall impose upon the functions which appear will be discussed below.

The earliest formulation of multi-stage games in terms of functional equations, of which we are aware, is contained in R. Bellman and J. La Salle, [11], and R. Bellman and D. Blackwell, [10], where "games of survival" are introduced. Following this, further studies of the existence and uniqueness of the solution, together with properties of the solution, are contained in M. Peisakoff [13], and Bellman, [6]. The first paper stating some general existence and uniqueness theorems for other classes of multi-stage games is that of L. Shapley, [15]. Since then, a number of papers have appeared on the subject of multi-stage games. The subject has attracted a great deal of attention, and deservedly so, since the theory of multi-stage games constitutes a natural extension of the Von Neumann-Morgenstern theory. In some sense, we may even consider the multi-stage process as basic, giving rise to the single-stage theory as a limiting case corresponding to a "steady state". This is a clear inference from the Brown-Von Neumann iterative solution of games. A concept of this type is useful in discussing the play of n—person games and non—zero sum games. For an application of this idea, see [8], where the idea is applied in a heuristic fashion.

We shall begin our discussion in the following section with the description of a multi-stage game arising from the study of two-person allocation processes. The "principle of optimality", [7], will be used to show that we may reduce the study of the N-stage process to the study of a certain system of recurrence relations. We shall then consider, in turn, the corresponding process involving an unbounded number of stages, the process where the interaction is stochastic rather than deterministic, and finally, some time-dependent cases. In this way we shall be led to consider the equation in (3).

Using this equation as our model, we shall turn to a discussion of existence and uniqueness, under various hypotheses concerning the coefficient functions. Our proofs will depend upon the method of successive approximations, and a lemma which exploits the quasi-linear aspect of the functional equation.

We shall demonstrate that the strategies determined by the functional equation are effective, and consider the stability of the solution under changes in $R(u,v)$.

The method we shall employ in presenting the results sketched above is also applicable to the one-sided functional equation

$$(4) \qquad f(P,P') = \operatorname*{Max}_{u} \operatorname*{Min}_{v} \left[R(u,v)+h(u,v)f(T,T')\right],$$

as was pointed out to us by W. Fleming. We shall also use a cruder method to treat this problem.

Finally, we shall discuss some other classes of multi-stage games, such as "games of survival" and "pursuit games", which lead to functional equations amenable to the same analysis.

§2. Description of a Multi-Stage Game

Let us now describe in detail the multi-stage game we wish to analyze. Two players, whom we may rather prosaically designate by A and B, possessing, respectively, resources which we may represent as M-dimensional vectors, P and P', are engaged in a multi-stage process carried on in the following manner. At the beginning of each stage of an N-stage process, A allocates a certain quantity of his resources, a vector u, and B a certain quantity of his resources, a vector v; this will be represented symbolically by the notation $0 \leq u \leq P$, $0 \leq v \leq P'$.

As a result of this allocation, there are two consequences. A receives a pay-off of $R(u,v)$, a scalar function, and B a pay-off of $-R(u,v)$. Furthermore, their resources are altered, P is transformed into $T(P,P';u,v)$ and P' becomes $T'(P,P';u,v)$. The process now continues in the same fashion for (N-1) additional stages.

The total return to A of the N-stage process is assumed to be additive,

$$(1) \quad R_N = R_N(u,u_1,\ldots,u_{N-1};v,v_1,\ldots,v_{N-1}) = R(u,v)+R(u_1,v_1)+\ldots+R(u_{N-1},v_{N-1}),$$

There are two ways of treating the N-stage process. We can either consider the N-stage game as a single-stage game of complicated type, requiring a choice of the set of vectors $(u,u_1,\ldots,u_{N-1})$ by A, and the set $(v,v_1,\ldots,v_{N-1})$ by B, where the choice of u_k and v_k depends upon the choice of $u,u_1,\ldots,u_{k-1},v,v_1,\ldots,v_{k-1}$, or we can use the functional equation approach of dynamic programming, [6], [7], and thus reduce the dimensions of the process. For the case of unbounded processes, or processes involving stochastic interaction of u and v, which we shall discuss below, the recurrence relation technique seems to be the only feasible one, while in the case of finite deterministic processes, this technique is simpler analytically, conceptually, and computationally.

Let us now make some assumptions of continuity. We shall take $R(u,v)$ and $h(u,v)$ to be a continuous function of u, v, P and P' over any finite $(P,P';u,v)$- region, and similarly for $T(P,P';u,v)$, $T'(P,P';u,v)$ to be continuous functions of P,P',u and v. The case where P,P',u,v, $R(u,v),T,T'$, assume only finite sets of values is also interesting and may be treated by the same general techniques.

The value of the N-stage game described above is given by the expression

$$(2) \qquad v_N = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} \left[\int\int R_N dG(u,u_1,\ldots,u_{N-1}) dG'(v,v_1,\ldots,v_{N-1}) \right]$$

$$= \operatorname*{Min}_{G'} \operatorname*{Max}_{G} \left[\ \ldots \ \right],$$

where G and G' are distribution functions over regions of quite complicated form defined by the inequalities

$$(3) \qquad \begin{aligned} &0 \le u \le P, \quad 0 \le v \le P', \\ &0 \le u_1 \le T, \quad 0 \le v_1 \le T', \\ &\vdots \\ &0 \le u_{N-1} \le T_{N-1}, \quad 0 \le v_{N-1} \le T'_{N-1}. \end{aligned}$$

Note that T and T' depend upon P,P', u and v, T_1,T_1' depend upon P,P',u,v,u_1,v_1, and so on.

Observing that v_N depends upon P and P', the initial states, and only upon these quantities, let us define the sequence of functions, $\{f_N(P,P')\}$, by means of the relation

$$(4) \qquad f_N(P,P') = v_N, \quad N=1,2,\ldots$$

§3. <u>The Principle of Optimality</u>.

In [7], we enunciated a principle which yields the functional equations of the theory of dynamic programming, namely the <u>Principle of Optimality</u>. <u>An optimal policy has the property that</u>

<u>whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decisions.</u>

Applying this to the multi-stage game, we obtain the following recurrence relation

$$(1) \qquad f_1(P,P') = \underset{G}{\text{Max}}\ \underset{G'}{\text{Min}} \int\limits_{\substack{0\le u\le P\\ 0\le v\le P'}}\int R(u,v)\,dG(u)\,dG'(v) = \underset{G'}{\text{Min}}\ \underset{G}{\text{Max}}\ [\ \dots\ ]$$

$$f_{N+1}(P,P') = \underset{G}{\text{Max}}\ \underset{G'}{\text{Min}} \Big[\int\limits_{\substack{0\le u\le P\\ 0\le v\le P'}}\int [R(u,v)+f_N(T,T')]\,dG(u)\,dG'(v)\Big]$$

$$= \underset{G'}{\text{Min}}\ \underset{G}{\text{Max}}\ [\ \dots\ ].$$

That the above principle is valid for one-person processes where we are attempting to maximize a return or minimize cost is clear by contradiction. Since its validity may not be as obvious for game processes, let us present a brief proof for the sake of completeness.

The recurrence relation in (1) provides a sequence, not necessarily unique, of pairs of distribution functions, $\{G_N(u,P,P'), G'_N(v,P,P')\}$ which furnish the sequence $\{f_N(P,P')\}$. In order to show that the function $f_N(P,P')$ is actually the value of the N-stage game, it is sufficient to show that A can guarantee an expected return of $f_N(P,P')$ if he chooses u at the first stage of an N-stage process in accordance with the distribution function $G_N(u,P,P')$, when the states of A and B are described by P and P', respectively, and similarly that B can guarantee an expected loss of not more than $-f_N(P,P')$.

To demonstrate this, consider the one-person N-stage process in which A employs the fixed strategy represented by the sequence of distribution functions, $\{G_k(u,P,P')\}$, $k=1,2,\ldots,N$, and B attempts to minimize A's expected N-stage return. It is sufficient to consider this process, since any other policy employed by B yields a larger expected return for A. Let

(2) $w_N(P,P')$ = N-stage expected return to A when A employs the fixed strategy $\{G_N(u,P,P')\}$, B employs a minimizing strategy, and A and B are in the initial states P and P'.

Then we have the recurrence relations

(3)
$$w_1(P,P') = \operatorname*{Inf}_{G'} \int_{0\leq v\leq P'} \left[\int_{0\leq u\leq P} R(u,v)\,dG(u,P,P')\right] dG'(v),$$
$$w_{N+1}(P,P') = \operatorname*{Inf}_{G'} \int_{0\leq v\leq P'} \left[\int_{0\leq u\leq P} [R(u,v)+w_N(T,T')]\,dG_{N+1}(u,P,P')\right] dG'(v)$$

upon employing the principle of optimality for the one-person process.

Considering the origin of the function F_1, we see that the minimum in the relation for $w_1(P,P')$ in (3) is attained by the function $G'=G_1'$, not uniquely in general. Hence,

(4) $w_1(P,P') = v_1(P,P')$.

Since $w_1\equiv v_1$, the relation for w_2 yields in the same way the fact that $w_2\equiv v_2$, and thus, inductively, we see that

(5) $w_N(P,P') \equiv v_N(P,P')$.

In precisely the same way we show that if B employs the strategy $\{G'_N(v,P,P')\}$, A cannot obtain more than $v_N(P,P')$. Hence $v_N(P,P')$ is the value of the N-stage game.

§4. Related Classes of Games.

Proceeding formally for the moment, without regard to the existence of the quantity we define, let us consider the unbounded process. Define

(1) $f(P,P')$ = the value of the infinite stage process to A when A has P initially, B has P', and both players employ optimal strategies.

If $f(P,P')$ exists, it satisfies the equation

$$(2) \qquad f(P,P') = \underset{G}{\text{Max}}\ \underset{G'}{\text{Min}} \left[\int_{0\le u\le P} \int_{0\le v\le P'} [R(u,v)+f(T,T')]dG(u)dG'(v) \right]$$

$$= \underset{G'}{\text{Min}}\ \underset{G}{\text{Max}} \left[\quad \cdots \quad \right],$$

provided that Max-Min = Min-Max. The legitimacy of this will be discussed in the following sections, under suitable assumptions.

Let us, however, observe briefly how more general processes can give rise to various extensions of (2). If we allow the process to be time-dependent in the sense that the return from the k^{th} stage, as well as the transformations T and T', depends upon k, in place of the function defined by (1), we must consider the sequence of functions

(3) $f(P,P';k)$ = the value to A of the infinite process beginning at the k^{th} stage when A possesses P at this stage and B possesses P', and both employ optimal strategies.

This sequence satisfies the recurrence relation

$$(4) \qquad f(P,P';k) = \underset{G}{\text{Max}}\ \underset{G'}{\text{Min}} \left[\int_{0\le u\le P} \int_{0\le v\le P'} [R(u,v,k)+f(T_k,T'_k;k+1)]dg(u)dG'(v) \right.$$

$$= \underset{G}{\text{Min}}\ \underset{G'}{\text{Max}} \left[\ \cdots \ \right].$$

Let us now complicate the process to a further degree. We have assumed in the above formulation that the interaction between the players was perfectly determined once u and v were chosen. It is interesting occasionally to consider more general processes in which a choice of u and v merely determine a distribution of outcomes, denoted by the function $K_k(z,t,t';u,v)$, which depends upon the stage, where z is the value of $R(u,v)$, t the value of T, and t' the value of T'. Then (4) is replaced by

$$(5) \qquad f(P,P';k) = \underset{G}{\mathrm{Max}}\ \underset{G'}{\mathrm{Min}} \left[\int\limits_{0 \le u \le P} \int\limits_{0 \le v \le P'} \left[\int \left[z+f(t,t';k+1) \right] dK_k(z,t,t';u,v) \right] dG(u)\ dG'(v) \right]$$

$$= \underset{G'}{\mathrm{Min}}\ \underset{G}{\mathrm{Max}} \left[\ \dots \ \right].$$

Finally, let us consider the case where we are not interested in the sum of the returns, but in some nonlinear function of the total return. A particularly important example is the probability of achieving a return of at least R_o. This is the expected value of the function defined by

$$(6) \qquad \phi(u) = 0, \quad 0 \le u < R_o,$$
$$= 1, \quad u \ge R_o.$$

Another interesting utility function is e^{aR}.

To describe the general non-linear situation, we must introduce an additional state variable, a, the return obtained by A from the previous stages of the process. Defining $f(P,P',a;k)$ essentially as in (3), we obtain the functional equation

$$(7) \qquad f(P,P'a;k) = \underset{G}{\mathrm{Max}}\ \underset{G'}{\mathrm{Min}} \left[\int\limits_{0 \le u \le P} \int\limits_{0 \le v \le P'} \left[\int f(t,t',a+z;k+1)\, dK_k(z,t,t';u,v) \right] dG(u)\, dG'(v) \right]$$

$$= \underset{G'}{\mathrm{Min}}\ \underset{G}{\mathrm{Max}} \left[\ \dots \ \right]$$

We shall not consider any of these more complicated functional equations since the basic approach is the same in all cases, despite the fact that the analytic details increase in complexity.

A genuinely new class of functional equations emerges from the study of "learning processes", [14], [9]. Here it is assumed that the distribution function $K(z,t,t';u,v)$ exists, but is not completely known. In the course of carrying out the process, additional information is obtained concerning K. The problem is once again to maximize the expected value of the total return. It is clear, of course, that we are encroaching upon the domain of sequential analysis. We shall consider the functional equations obtained in this way in a subsequent work.

The methods we employ here, combined with those used to establish the results of [3], can be used in very much the same form to treat the nonlinear differential equations

$$(11) \qquad \frac{dx_i}{dt} = \underset{G}{\mathrm{Max}}\ \underset{G'}{\mathrm{Min}} \left[\int\int f_i(x,t;u,v)dG(u)dG'(v) \right]$$

$$= \underset{G'}{\mathrm{Min}}\ \underset{G}{\mathrm{Max}} \left[\ \ldots \ \right], \quad i=1,2,\ldots,n,$$

and similar types of integro-differential equations, cf. [16], [17].

§5. Statement of Principal Results.

Before stating our results, let us introduce some notation and definitions. We shall take P and P' to be n- and n'-dimensional vectors defined over regions D and D' respectively, each containing the origin in its respective space. For all values of u, v, P and P', the transformed vectors $T(P,P';u,v)$, $T'(P,P';u,v)$, are required to lie within these same domains, where u and v are k and k'-dimensional choice vectors respectively, constrained to domains S and S' which, in general, depend upon P and P'. Since we are dealing with shrinking transformations, there is no loss in assuming D and D' to be finite.

In each space, let us introduce the norm, $||P||$, equal to the sum of the absolute values of the components of P,

$$(1) \qquad ||P|| = \sum_{i=1}^{n} |P_i|,$$

$$||P'|| = \sum_{i=1}^{n'} |P'_i|.$$

Actually, these need not be the same norms, and in some situations, it might be useful to consider norms molded to the structure of the functional equations arising, rather than standard norms of the above type.

The functional equation we shall consider is

$$(2) \qquad f(P,P') = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} \Big[\int\int_{\substack{u \in S(P,P') \\ v \in S(S'(P,P')}} \big[R(P,P';u,v) + h(P,P',u,v) f(T,T') \big] dG(u) dG'(v)$$

$$= \operatorname*{Min}_{G'} \operatorname*{Max}_{G} [\ \dots\ ],$$

where

$$(3) \qquad T = T(P,P';u,v),$$

$$T' = T'(P,P';u,v)$$

To simplify our notation, let us represent the operator appearing within brackets in equation (2) by $T(P,P';f;G,G')$, so that the above functional equations take the form

$$(4) \qquad f(P,P') = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} T(P,P';f;G,G')$$

$$= \operatorname*{Min}_{G'} \operatorname*{Max}_{G} T(P,P';f;G,G').$$

Whether one wishes to call this one functional equation or a pair of linked functional equations appears to be a matter of taste.

The result we wish to prove is

Theorem 1. *Consider the above equation,* (4), *under the following assumptions.*

(5) (a) *The functions* $R(P,P';u,v)$, $h(P,P';u,v)$, $T(P,P';u,v)$ *and* $T(P,P';u,v)$ *are continuous functions of* P *and* P', u *and* v, *in any bounded regions of the variables.*

(b) *The choice domains,* $S(P,P')$, $S'(P,P')$, *vary continuously with* P *and* P'.

(c) T *and* T' *are shrinking transformations,* i.e.

$$\operatorname*{Max}_{\substack{u \in S \\ v \in S'}} \left(\| T(P,P';u,v) \| + \| T'(P,P';u,v) \| \right) \leq k(\|P\| + \|P'\|)$$

where k *is a fixed constant less than* 1.

(d) *Let*

$$w(c) = \operatorname*{Max}_{\|P\| + \|P'\| \leq c} \left(\operatorname*{Max}_{\substack{u \in S, \\ v \in S'}} |R(u,v)| \right)^{*}$$

Then $\displaystyle\sum_{n=1}^{\infty} w(k^n c) < \infty$.

(e) $$\operatorname*{Max}_{u,v,P,P'} \| h(P,P';u,v) \| \leq 1.$$

If the above conditions are satisfied, we can assert that there is a unique solution of (4) *within the class of functions* $f(P,P')$

---

*For simplicity, let us suppress the P and P' in $R(P,P';u,v)$.

which are continuous for all finite P and P' and vanish when P' and P' are both null vectors.

This solution may be found by the method of successive approximations,

$$(6)\qquad f_0(P,P') = \underset{G}{\text{Max}}\ \underset{G'}{\text{Min}} \left[\iint_{\substack{u\in S,\\ v\in S'}} R(u,v)dG(u)dG'(v) \right],$$

$$= \underset{G'}{\text{Min}}\ \underset{G}{\text{Max}} \left[\ \ldots\ \right],$$

$$f_{n+1}(P,P') = \underset{G}{\text{Max}}\ \underset{G'}{\text{Min}}\ T(P,P';f_n;G,G'),$$

$$= \underset{G'}{\text{Min}}\ \underset{G}{\text{Max}}\ T(P,P';f_n;G,G'), \qquad n \geq 0,$$

with $f(P,P') = \lim_{n \to \infty} f_n(P,P')$ in any bounded domain of the (P,P') space.

We shall further demonstr ;e

Theorem 2. Under the hypotheses of Theorem 1, a set of functions, (G(u),G'(v)), furnished by the functional equation constitute a set of optimal strategies for A and B respectively in the multi-stage game described in the preceding sections.

§6. Lemma.

Let us present a simple but extremely useful inequality which exhibits the quasi-linearity of the transformation

$$(1)\qquad L(f) = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} T(P,P';\, f;\, G,\, G') = \operatorname*{Min}_{G'} \operatorname*{Max}_{G} T$$

Lemma 1.* Let

$$(2)\qquad L(f) = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} \left[\int_{u \in S} \int_{v \in S'} \left[R(u,v) + h(P,P';u,v)\, f(T,T') \right] dG(u)\, dG'(v) \right]$$

$$= \operatorname*{Min}_{G'} \operatorname*{Max}_{G} \left[\;\cdots\; \right],$$

$$L_1(F) = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} \left[\int_{u \in S} \int_{v \in S'} \left[R_1(u,v) + h(P,P';u,v) F(T,T') \right] dG(u)\, dG'(v) \right].$$

$$= \operatorname*{Min}_{G'} \operatorname*{Max}_{G} \left[\;\cdots\; \right].$$

Then

$$(3)\qquad |L(f) - L_1(F)| \le \operatorname*{Max}_{G} \operatorname*{Max}_{G'} \left[\int_{u \in S} \int_{v \in S'} \left| |R(u,v) - R_1(u,v)| + |h(P,P';u,v)|\, |f(T,T') - F(T,T')| \right| dg(u)\, dG'(v) \right].$$

Proof: Let us write

$$(4)\qquad L(f) = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} T(P,P';f;G,G') = \operatorname*{Min}_{G'} \operatorname*{Max}_{G}$$

$$L_1(F) = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} T_1(P,P';F;G,G') = \operatorname*{Min}_{G'} \operatorname*{Max}_{G}$$

Let (G_1, G_1') be a pair of functions yielding the value $L(f)$, and (G_2, G_2') be a pair of functions yielding the value $L_1(F)$. Then, by virtue of the saddle-point property, we have the following chain of equalities and inequalities

*It is assumed that max-min = min-max for each transformation. A similar result holds for the one-sided max-min operator; see § 14.

$$(5)\qquad L(f) = T(P,P';f;G_1,G_1') \geq T(P,P';f;G_2,G_1')$$
$$\leq T(P,P';f;G_1,G_2'),$$
$$L_1(F) = T_1(P,P';F;G_2,G_2') \geq T_1(P,P';F;C_1,G_2')$$
$$\leq T_1(P,P';F;G_2,G_1')$$

Combining these inequalities we have

$$(6)\qquad L(f)-L_1(F) \geq T(P,P';f;G_2,G_1') - T_1(P,P';F;G_2,G_1')$$
$$\leq T(P,P';f,G_1,G_2') - T_1(P,P';F;G_1,G_2')$$

The inequality in (6) yields

$$(8)\qquad L(f)-L_1(F) \geq \int_{u\in D}\int_{v\in D'}\Big[R(u,v)-R_1(u,v)+h(P,P';u,v)\big[f(T,T')-F(T,T')\big]\Big]\,dG_2(u)dG_1'(v)$$
$$\leq \int_{u\in D}\int_{v\in D'}\Big[R(u,v)-R_1(u,v) + h(P,P';u,v)\big[f(T,T')-F(T,T')\big]\Big]\,dG_1(u)dG_2'(v).$$

Using the fact that $a \leq c \leq b$ implies $|c| \leq \text{Max}(|a|, |b|)$, we obtain from (8) the further inequality

$$(9)\qquad |L(f)-L_1(F)| \leq \text{Max}\left\{\left|\int_{u\in D}\int_{v\in D'}\Big[|R(u,v)-R_1(u,v)| + |h(P,P';u,v)|\,|f(T,T')-F(T,T')|\Big]\,dG_2(u)dG_1'(v)\right|,\right.$$
$$\left.\left[\int_{u\in D}\int_{v\in D'}\Big[|R(u,v)-R_1(u,v)| + |h(P,P';u,v)|\,|f(T,T')-F(T,T')|\Big]\,dG_1(u)dG_2'(v)\right]\right\},$$

from which (3) follows immediately.

It is easy to make the modifications required to obtain the analogous result for the case where Max Min is replaced by Sup Inf.

§7. Existence and Uniqueness.

We can now proceed to a proof of Theorem 1.

Let

(1) $$f_o(P,P') = \underset{G}{\mathrm{Max}} \ \underset{G'}{\mathrm{Min}} \left[\int_{u\in S} \int_{v\in S'} R(u,v)\,dG(u)\,dG'(v) \right],$$

$$= \underset{G'}{\mathrm{Min}} \ \underset{G}{\mathrm{Max}} \left[\quad \cdots \quad \right]$$

and

(2) $$f_{n+1}(P,P') = \underset{G}{\mathrm{Max}} \ \underset{G'}{\mathrm{Min}} \ T(P,P';f_n;\ G,G') = \mathrm{Min}\ \mathrm{Max}\ T,$$

where T is defined as in (4.2) and (4.4).

By virtue of our assumptions concerning the coefficient functions, and the domains, S and S', we can assert the existence of the saddlepoint in (1), and the continuity of $f_o(P,P')$. Inductively, then, all the $f_n(P,P')$ exist and are continuous for all finite P and P'.

Let us now show that the sequence $\{f_n\}$ converges uniformly in any finite portion of the (P,P')-regions. Using Lemma 1 we obtain the inequality

(3) $$|f_{n+1}(P,P')-f_n(P,P')| \leq \underset{G}{\mathrm{Max}} \ \underset{G'}{\mathrm{Max}} \left[\int \int |f_n(T,T')-f_{n-1}(T,T')| \, dG(u)\,dG'(v) \right], \quad n=2,3,\ldots .$$

Define the new sequence

(4) $$u_{n+1}(c) = \underset{||P||+||P'||\leq c}{\mathrm{Max}} |f_{n+1}(P,P')-f_n(P,P')|.$$

Then (3) yields, using the assumption of (4a) of §3,

(5) $$u_{n+1}(c) \leq u_n(kc), \quad n=2,3,\ldots ,$$

Also, we have

(6) $$|f_2(P,P')-f_o(P,P')| \le \underset{G}{\text{Max}} \ \underset{G'}{\text{Max}} \int\int |R(u,v)|dG(u)dG'(v),$$

whence

(7) $$u_2(c) \le w(c).$$

Using our assumption that $\Sigma\, w(k^n c) < \infty$, we see that the series $\sum_n \left[f_{n+1}(P,P')-f_n(P,P')\right]$ converges uniformly in any finite region. Hence $f_n(P,P')$ converges uniformly to a function $f(P,P')$ which satisfies the original functional equation.

This completes the proof of existence. Let us now turn to a proof of uniqueness. Let $F(P,P')$ be another solution which is continuous at $P=0, P'=0$, and bounded in any finite region. We see that $F(P,P')$ is then actually continuous for all finite P and P', although this fact is not necessary for our proof. It does simplify it a bit since we can replace Sup—Inf by Max–Min.

We then have the two equations

(8) $$F(P,P') = \underset{G}{\text{Max}} \ \underset{G'}{\text{Min}} \ T(P,P';F;G,G')$$

$$f(P,P') = \underset{G}{\text{Max}} \ \underset{G'}{\text{Min}} \ T(P,P';f;G,G').$$

Applying Lemma 1, we see that

(9) $$|F(P,P')-f(P,P')| \le \underset{G}{\text{Max}} \ \underset{G'}{\text{Max}} \left[\int_{u\in S}\int_{v\in S'} |F(T,T')-f(T,T')|dGdG'.\right.$$

Let

(10) $$\Delta(c) = \underset{||P||+||P'||\le c}{\text{Max}} |F(P,P')-f(P,P')|.$$

Then (9) yields the relation

(11) $$\Delta(c) \le \Delta(kc),$$

which, upon iteration, yields $\Delta(c) \leq \Delta(k^n c)$, $n=1,2,\ldots,$. Since F and f are both continuous at $P = 0$, $P' = 0$, and have the common value 0 there, we see that $\Delta(k^n c) \to 0$ as $n \to \infty$. Hence $\Delta(c) = 0$ and $F \equiv f$.

This completes the proof of Theorem 1.

§8. Successive Approximations in General

The sequence of approximations, $\{f_n(P,P')\}$, we used to construct the function $f(P,P')$ was precisely that obtained from the sequence of values associated with the finite n-stage processeses. This is actually not the best sequence to use if we are interested only in the infinite stage process. As we have pointed out elsewhere, [4], [7], approximation in "policy space", here "strategy space", is in many ways a more natural and more important type of approximation.

To justify this and other types of approximations we require

Theorem 3. *Under the assumptions of Theorem 1, the sequence defined by*

$$(1) \qquad f_{n+1}(P,P') = \underset{G}{\mathrm{Max}}\ \underset{G'}{\mathrm{Min}}\ T(P,P';f_n;G,G'),\quad n=0,1,\ldots$$

$$= \underset{G'}{\mathrm{Min}}\ \underset{G}{\mathrm{Max}}$$

Converges to the solution of (5.3) for any initial function $f_0(P,P')$ which is continuous in any finite part of the (P,P')-domain, and equal to zero at $P=0, P'=0$.

The proof is precisely the same as that given above.

§ 9. Effectiveness of Solution.

We have established existence and uniqueness of the functional equation we derived under the assumption that the infinite process possessed a value for each player. The question now arises as to whether the functional equation actually yields sufficient information to allow each player to obtain this value. If so, we say that the solution is effective.

To show effectiveness, under the hypotheses of Theorem 1, we must show that if A uses a distribution function $G(u) = G(u;P,P')$ obtained from a pair (G,G') which yield the min-max, then, regardless of what B may do, we can guarantee a return of at least $f(P,P')$ to A.

Employing this fixed strategy, A's return will be, at worst, determined by the solution of the functional equation

$$(1) \qquad F(P,P') = \underset{G'}{\text{Min}} \left[\int_{u \in S} \int_{v \in S'} \left[R(u,v) + h(P,P';u,v)F(T,T') \right] dG(u)dG'(v) \right].$$

It is easy to show, using the techniques of the preceding sections, cf [2], where these equations are treated in detail, that this equation has a unique continuous solution which is zero at $P=0, P'=0$. Furthermore, the solution of this equation may be obtained as the limit of the sequence defined by

$$(2) \qquad F(P,P') = \underset{G'}{\text{Min}} \left[\int_{u \in S} \int_{v \in S'} R(u,v)dG(u)dG'(v) \right],$$

$$F_{n+1}(P,P') = \underset{G'}{\text{Min}} \left[\int_{u \in S} \int_{v \in S'} [R(u,v) + h(P,P';u,v)F_n(T,T')dG(u)dG'(v) \right]$$

It is clear, from the derivation of $G(u)$, that $F \equiv f_1$. Hence, inductively, $F_{n+1} \equiv f_{n+1}$, as defined by (7.2). Thus

$$(3) \qquad F(P,P') = \lim_{n \to \infty} F_n = \lim_{n \to \infty} f_n = f(P,P').$$

This demonstrates the effectiveness of the solution in the continuous case.

§10. Stability of the Solution

An important aspect of any physical process is the dependence of measuring functions (in this case, the value of the game), upon the parameters and coefficient functions which determine the process. In general, we expect small changes in these quantities to result in small changes in the measuring functions. If this is true, with appropriate definitions of "smallness", we say the process is <u>stable</u>. A principle of wide validity is that physical processes are stable. Of course, since the mathematical transcription of any physical process is never precise, we cannot conclude immediately that the mathematical process, as defined by the equation, is stable. Actually, one of the most useful tests of the realism of a mathematical model of a physical process is that of stability.

Let us now establish

<u>Theorem</u> 4. <u>Let</u>

$$(1) \qquad \Delta(c) = \underset{||P||+||P'||\le c}{\mathrm{Max}} \ \underset{\substack{u\in S\\ v\in S'}}{\mathrm{Max}} |R(u,v)-R'(u,v)|.$$

<u>Then, under the hypotheses of Theorem</u> 1, <u>the solutions of</u>

$$(2) \qquad f(P,P') = \underset{G}{\mathrm{Max}}\ \underset{G'}{\mathrm{Min}} \left[\int\limits_{u\in S}\int\limits_{v\in S'} \left[R(u,v)+h(P,P';u,v)f(T,T')\right]dG\,dG'\right]$$

$$= \underset{G'}{\mathrm{Min}}\ \underset{G}{\mathrm{Max}} \left[\qquad \cdots \qquad \right],$$

$$F(P,P') = \underset{G}{\mathrm{Max}}\ \underset{G'}{\mathrm{Min}} \left[\int\limits_{u\in S}\int\limits_{v\in S'} R'(u,v)+h(P,P';u,v)F(T,T')\right]dG\,dG'\Big]$$

$$= \underset{G'}{\mathrm{Min}}\ \underset{G}{\mathrm{Max}} \left[\cdots \qquad \right]$$

satisfy the inequality

$$(3) \qquad |f(P,P')-F(P,P')| \leq \sum_{n=0}^{\infty} \Delta(k^n c).$$

Proof: Applying the Lemma of §3, we see that

$$(4) \qquad |f(P,P')-F(P,P')| \leq \operatorname*{Max}_{G} \operatorname*{Max}_{G'} \left[\int_{u \in S} \int_{v \in S'} \left[|R-R'| + |f(T,T')-F(T,T')| \right] dG dG' \right].$$

Iteration of this inequality yields the desired result.

§11. Further Results.

There are a number of different ways in which the results of Theorems 1 through 4 can be extended and generalized.

These results depended upon the fact that the transformation (T,T') was a shrinking transformation, in the sense explained above. An intuitive visualization of this is to consider (P,P') as representing the resources of each side. Then each play of the game diminishes the total resources available.

We can introduce a shrinking transformation in another way by imposing the condition that

$$(1) \qquad |h(P,P';u,v)| \leq k < 1.$$

A condition of this type can arise in two ways. First of all, it may represent the discounted value of future actions as contrasted with the present; secondly, it may represent a probability of survival in situations in which there is a non-zero probability of the termination of the process associated with every play of the game of §16. Both concepts are connected through the intermediary of prediction theory.

Having introduced (1), it is no longer necessary to assume that (T,T') is a shrinking transformation. We must, however, assume that (P,P') and transforms always lie within some fixed region. With these conditions, the analogues of Theorems 1 through 4 are readily obtained.

Furthermore, similar results may be obtained, using the same methods, for the generalized equations mentioned in §4, under various combinations of the above assumptions.

Equations satisfying either of the above conditions correspond to the equations of Types One and Two discussed in our paper on one-person processes, [2]. When we consider other types of processes, the analysis becomes more specialized and complicated, cf. [2], §10.

In a different direction, we may relax the restrictions of continuity which we have imposed and investigate the conditions under which we obtain solutions to equations of the form

$$f(P,P') = \operatorname*{Sup}_{G} \operatorname*{Inf}_{G'} T(P,P';f;G,G'),$$

$$= \operatorname*{Inf}_{G'} \operatorname*{Sup}_{G} T(P,P';f;G,G').$$

In these cases, we will obtain ϵ-effective strategies.

§ 2. Differentiability.

If we assume that under suitable assumptions of concavity and convexity the functional equation reduces to

(1) $$f(P,P') = \max_{u} \operatorname*{Min}_{v} \left| R(u,v)+f(T,T') \right|$$

then, in certain other fortunate cases, we can reduce the equation to a functional equation of more conventional form.

To illustrate the idea, consider the functional equation

(2) $$f(x) = \underset{0\le y\le x}{\text{Max}} \left[g(y)+h(x-y)+f(ay+b(x-y)) \right],$$

where $0<a,b<1$, which we have discussed in a number of papers, [2, 7].
If the maximum always occurs in the interior of the interval, and if we assume that g and h are differentiable, we obtain the two equations

(3) $$g'(y)-h'(x-y)+(a-b)f'(ay+b(x-y))=0$$
$$f'(x)=h'(x-y)+bf'(ay+b(x-y)),$$

which allow us to compute $f'(x)$ via a relatively simple recurrence relation.

Similarly, if we consider the equation

(4) $$f(x,y) = \underset{0\le u\le x}{\text{Max}}\ \underset{0\le v\le y}{\text{Min}} \left[R(u,v;x,y)+f(T(x,y,u,v),T'(x,y,u,v))\right]$$
$$= \underset{0\le u\le x}{\text{Min}}\ \underset{0\le v\le y}{\text{Max}} \left[\ldots\right]$$

where x and y are now scalars, and assume that the saddlepoint exists and is always inside the region for all non-negative x and y, we can reduce (4) to the set of simultaneous equations

$$R_u+T_uf_x(T,T')+T'_uf_y(T,T')=0,$$
(5) $$R_v+T_vf_x(T,T')+T_v'f_y(T,T')=0,$$
$$f_x=R_x+T_xf_x(T,T'),$$
$$f_y=R_y+T_yf_y(T,T').$$

§13. <u>One-Sided Min-Max</u>

Let us now consider the equation

(1) $$f(P,P')=\underset{v\in S'}{\text{Min}}\ \underset{u\in S}{\text{Max}} \left[R(u,v)+h(P,P';u,v)f(T,T')\right],$$

which arises from the allocation process described above if the second player is required to announce his choice of v before each play.

Using a technique due to Wendell Fleming, we can treat this equation in exactly the same fashion as those appearing in the previous sections. We begin by noting that for any function $R(u,v)$

$$(2) \qquad \underset{v \in S'}{\text{Min}} \; \underset{u \in S}{\text{Max}} \; R(u,v) = \underset{v \in S'}{\text{Min}} \; \underset{u(v) \in S}{\text{Max}} \; R(u,v),$$

where $u(v)$ is now a function of v which maximizes $R(u,v)$ for fixed v. Let $U(v)$ be this function.

Let V be a value of v which minimizes $R(U(v),v)$. Then we have the two inequalities

$$(4) \qquad \begin{aligned} R(U(V),V) &\leq R(U(v),v), \\ R(U(V),V) &\geq R(u(V),V), \end{aligned}$$

for any other admissible values of u and v.

Using these inequalities we readily obtain the analogue of the lemma given in §4 for equations of the above type. In this way we can establish the analogues of the previous theorems.

§15. An Alternative Approach.

Let us now show that there are alternative approaches which can dispense with our lemma, and rely instead directly upon the shrinking properties of the transformation. Let us, for simplicity, consider the most important case where $R(u,v)$ and $h(P,P';u,v)$ are non-negative.

Let, as above

$$(1) \qquad w(c) = \underset{||P||+||P'||\leq c}{\text{Max}} \; \underset{u \in S}{\text{Max}} \; \underset{v \in S'}{\text{Max}} \; R(u,v).$$

As usual, let us introduce the sequence $\{f_n(P,P')\}$, where

$$(2) \qquad \begin{aligned} f_1(P,P') &= \underset{v \in S'}{\text{Min}} \; \underset{u \in S}{\text{Max}} \; R(u,v) \\ f_{n+1}(P,P') &= \underset{v \in S'}{\text{Min}} \; \underset{u \in S}{\text{Max}} \left[R(u,v) + h(P,P';u,v) f_n(T,T') \right], \quad n = 1,2,\ldots \end{aligned}$$

Then

(3) $$f_1(P,P') \le f_2(P,P') \le \operatorname*{Min}_{v\in S'} \operatorname*{Max}_{u\in S} \left[R(u,v)+w(kc)\right]$$

$$\le f_1(P,P')+w(kc).$$

Continuing in this fashion, we see that

(4) $$f_2(P,P') \le f_3(P,P') \le \operatorname*{Min}_{v\in S'} \operatorname*{Max}_{u\in S} \left[R(u,v)+h(P,P';u,v)f_2(T,T')+ w(a_2c)\right]$$

$$\le f_2(P,P')+w(k^2c).$$

(5) $$f_n(P,P') \le f_{n+1}(P,P') \le f_n(P,P')+w(k^nc).$$

Thus we have uniform convergence to a solution, under our assumption that $\sum_{n=1}^{\infty} w(k^nc) < \infty$.

§15. Probability of Survival.

Let us now briefly consider some other classes of multi-stage games which lead to related classes of functional equations. To begin with, let us consider the allocation process discussed in §2 in which we assume that there is a probability $h(P,P';u,v)$, dependent upon P,P', u and v, that the process will terminate at the end of the particular stage. The functional equation governing the process is then the equation we have discussed in the preceding sections where $f(P,P')$ is the expected return to A. In the allocation process, h may be either 1, or a "discount factor" emphasizing the present value of a return as opposed to a future value.

§16. Games of Survival.

Associated with the previous concept of probability of survival is the class of games called "games of survival". Here both players are actuated by the desire to survive the other, with each play of the game involving either a diminuation of resources of one or both players, or an actual chance of elimination.

Here the basic functional equation is

$$(1)\qquad f(P,P')=\underset{G}{\mathrm{Max}}\ \underset{G'}{\mathrm{Min}}\left[\int\limits_{\substack{u\in D\\ v\in D'}}\int h(P,P';u,v)f(T,T')dG(u)dG'(v)\right]$$

$$=\underset{G'}{\mathrm{Min}}\ \underset{G}{\mathrm{Max}}\left[\quad\cdots\quad\right]$$

where $f(P,P')$ is now the probability that A survive B. The equation is valid only for $P,P'>0$ with the side conditions

$$(2)\qquad f(P,0)=1,\ P>0;$$

$$f(0,P')=0,\ P'>0,$$

$$f(0,0)=1/2,\ \text{(as a matter of convention).}$$

This equation is very much more difficult to treat than the foregoing equations in the cases of greatest interest where $h=1$ and T and T' are merely restricted to lie in the bounded regions containing P and P'.

Particular results may be found in [6], Milnor and Shapley, [12]. This approach is occasionally useful in treating non-zero sum games, see [8].

§17. Games of Pursuit.

Finally, let us mention the very interesting and difficult problems connected with pursuit games. There is as yet no satisfactory theory of continuous pursuit games, which necessarily restricts us to a discussion of discrete games.

Let us assume that two players, A and B, are restricted to positions at the lattice points of the plane. A can move up to k units from his position in either horizontal or vertical direction, and B can move up to l units in the same manner. With both players required to move simultaneously, we are interested in the strategies which enable A to catch B, to catch B in minimum time, or to minimize some other payoff function.

If played on the unbounded region of the plane, it is not easy to determine when capture occurs, and the problem is not trivial for a bounded region either.

Let us set

(1) $f(P,P')$ = the time required for A to catch B when A is at the lattice point P, B is at the lattice point Q, and both players employ optimal strategies.

Then, without inquiring into the existence of our function, the equation satisfied by f is

(2) $$f(P,P') = 1 + \underset{G}{\mathrm{Min}}\ \underset{G'}{\mathrm{Max}}\left[\int\int f(P+e,P'+f)dG(e)dG'(f)\right],$$

$$= 1 + \underset{G'}{\mathrm{Max}}\ \underset{G}{\mathrm{Min}}\left[\ \ldots\ \right].$$

where G(e) and G(f) are distributions over the allowable vectors e and f. This equation holds prior to capture. At capture, the process terminates. Very little has been done in connection with establishing existence and uniqueness theorems for these equations. We shall discuss them in a subsequent paper.

§18. Colonel Blotto.

Let us now consider a well-known single-stage game which can profitably be considered to be a multi-stage process. Let A and B possess the scalar quantities x and y, both positive.

A divides x into a sum of N non-negative quantities $x_1, x_2, \ldots, x_N$, and B does likewise with y, so that we have

(1) $$x = x_1 + x_2 + \ldots + x_N, \quad x_i \geq 0$$
$$y = y_1 + y_2 + \ldots + y_N.$$

As a result of this allocation, A receives a pay-off of

(2) $$R_N(x_i, y_j) = \sum_{i=1}^{N} \max\ (x_i - y_i, 0),$$

and B a pay-off of $-R_N(x_i, y_j)$.

Let us define $f_N(x,y)$ to be the value of this game. Then we have the recurrence relations

$$(3) \qquad f_1(x,y) = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} \int_0^x \int_0^y \operatorname{Max}(x_1 - y_1, 0)\, dG(x_1)\, dG'(y_1),$$

$$= \operatorname*{Min}_{G'} \operatorname*{Max}_{G} \int_0^x \int_0^y \operatorname{Max}(x_1 - y_1, 0)\, dG(x_1)\, dG'(y_1),$$

$$f_{N+1}(x,y) = \operatorname*{Max}_{G} \operatorname*{Min}_{G'} \left[\int_0^x \int_0^y [\operatorname{Max}(x_1 - y_1, 0) + f_N(x - x_1, y - y_1)]\, dG(x_1)\, dG'(y_1) \right]$$

$$= \operatorname*{Min}_{G'} \operatorname*{Max}_{G} [\ \dots\ ]$$

This formulation facilitates both analytical and computational treatment.

BIBLIOGRAPHY

1. R. Bellman, Some Functional Equations in the Theory of Dynamic Programming, Proc. Nat. Acad. Sci., Vol. 39 (1953), pp. 1077-82.

2. ——————— Some Functional Equations in the Theory of Dynamic Programming—I, (Functions of Points and Point Transformations), Trans. Amer. Math. Soc.

3. ——————— Functional Equations in the Theory of Dynamic Programming—II (Nonlinear Differential Equations), Proc. Nat. Acad. Sci.

4. ——————— Dynamic Programming and a New Formalism in the Calculus of Variations, Proc. Nat. Acad. Sci., Vol. 40 (1954), pp. 231-5.

5. ——————— Dynamic Programming and a New Formalism in the Theory of Integral Equations, Proc. Nat. Acad. Sci., Vol. 41 (1955), pp. 31-4.

6. ——————— An Introduction to the Theory of Dynamic Programming, RAND Monograph R-245, 1953.

7. ——————— The Theory of Dynamic Programming, Bull. Amer. Math. Soc., Vol. 60 (1954), pp. 503-16.

8. ——————— Decision-Making in the Face of Uncertainty—II, Navy Quarterly of Logistics

9. ——————— A Problem in the Sequential Design of Experiments, Sankhya, (to appear).

10. R. Bellman and D. Blackwell, On a Particular Non-zero Sum Game, RM-250, RAND Corporation, 1949.

11. R. Bellman and J. La Salle, On Non-zero Sum Games and Stochastic Processes, RM-212, RAND Corporation, 1949.

12. J. W. Milnor and L. Shapley On Games of Survival, RAND Paper P-622, 1955.

13. M. Peisakoff More on Games of Survival, The RAND Corporation, RM-884, 1952

14. H. Robbins Some Aspects of the Sequential Design of Experiments, Bull. Amer. Math. Soc., Vol 58 (1952), pp. 527-536.

15. L. Shapley Stochastic Games, Proc. Nat. Acad. Sci., Vol. 39 (1953), pp. 1095-1100.

16. R. Bellman, Functional Equations in the Theory of Dynamic Programming, V, Positivity and Quasi-Linearity, Proc. Nat. Acad. Sci., Vol. 41 (1955), pp. 743-746.

17. R. Bellman, Functional Equations in the Theory of Dynamic Programming, II, Nonlinear Differential Equations, ibid, pp. 482-485.