JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 3, 102–116 (1961)

The Convergence Problem For Differential Games

WENDELL H. FLEMING*

Mathematics Department, Brown University, Providence, R. I.

Submitted by Richard Bellman

I. INTRODUCTION

The subject of differential games was begun by Isaacs [9], who was concerned with problems of continuous pursuit. In such problems there are two players — a pursuer and an evader — each of whom maintains continual surveillance of the other. The pursuer wishes to choose tactics to minimize, for example, the time to capture his opponent, and the evader to maximize this time. Another important class of differential games arose from the study of tactical warfare models. Several problems of this type have been explicitly solved by Berkovitz and Dresher [5].

When the game has only one player it becomes a maximization or minimization problem of the calculus of variations. This may then be treated either by classical techniques of by the functional equation method of Bellman [1]. Still another source for differential games seems to be the study of optimal control problems in which the system to be controlled is subject to unknown random disturbances. A few remarks concerning this case are made in Section VIII.

In a differential game a continuum of moves for each player is envisioned, and each move at time t is made with the knowledge of all previous moves. Profound difficulties are involved even in the precise formulation of such a game. Hence it is natural to start instead with a corresponding sequence of games with discrete time, the time Δ_n between successive moves tending to 0 as n tends to ∞. The *convergence problem* is to show that the value V_n of the nth game tends to a limit V as n tends to ∞. V then is the value of an (appropriately defined) limiting differential game.

Scarf [10] and the author [8] treated certain aspects of the convergence problem; however, both had to make the drastic assumption that

* Consultant, Mathematics Division, The RAND Corporation, Santa Monica, California.

a certain partial differential equation for V has a smooth solution. In the present paper we drop this assumption and impose instead restrictions on the form of the functions f and g which, together with the initial conditions and the constraints, specify the differential game. We are able to settle the convergence problem only in certain cases; an obvious direction for further research is to reduce the assumptions we have had to impose. When there is only one player the convergence problem is easier, and was settled by Bellman [2].

There are two methods for trying to solve differential games. The first method consists in solving recursively in time the difference equation [(2.7) below] for V_n. The experience of Bellman and Dreyfus [3] with optimization problems of this type seems applicable here. In the second method one writes down formally a first order partial differential equation [(2.8) below] analogous to the Hamilton-Jacobi equation, which V must satisfy. A solution is built up by constructing a field of characteristics (i.e., extremals) of this equation. Unlike the situation for calculus of variations, a single extremal does not describe a solution to the game with given initial data. A solution must prescribe a strategy for both players for every possible position of the game. Isaacs enriched the theory by many examples solved with this technique, and indicated a general method. The use of fields was clarified by Berkovitz and Fleming [6], and is being extended to very general problems of the Bolza type by Berkovitz [4]. However, the effectiveness of the method is hampered by various technical difficulties, including the fact that the partial derivatives of V can have discontinuities whose locations are not known in advance.

II. Games of Prescribed Duration

In this paper we shall consider only games of bounded duration. To begin with let us suppose that the duration T of the game does not depend on the strategies chosen by the players. Let t denote a point of the time interval $(0, T)$, x a point in a euclidean space of dimension $r \geqslant 1$, and y, z points of compact, convex sets Y, Z respectively. The vector x will have the role of a positional variable, y that of a control variable for player I, and z that of a control variable for player II. Let $f(x, y, z)$ and $g(x, y, z)$ be continuous functions, f being real valued and $g = (g_1, \ldots, g_r)$ having values in r-space. It is assumed that f and g satisfy a uniform Lipschitz condition in x, namely, there is a positive constant K such that for all x, x', y, z:

$$|f(x, y, z) - f(x', y, z)| \leqslant K|x - x'|, \tag{2.1}$$

$$|g(x, y, z) - g(x', y, z)| \leqslant K|x - x'|. \tag{2.2}$$

To define the time-discrete version of the game we fix $T_0 > 0$ and let

$$\Delta_n = 2^{-n} T_0, \qquad n = 1, 2, \ldots, \tag{2.3}$$

$$T = j\Delta_n, \qquad j = 1, 2, \ldots, 2^n. \tag{2.4}$$

An initial position $x = x_1$ is given. The game has j moves; and at move $i = 1, \ldots, j$ player I chooses y_i from Y and II chooses z_i from Z simultaneously. Both know all previous moves, and in particular the position x_i of the game. The new position is given by

$$x_{i+1} = x_i + \Delta_n g(x_i, y_i, z_i). \tag{2.5}$$

The total payoff to player I is

$$\sum_{i=1}^{j} \Delta_n f(x_i, y_i, z_i) + V^0(x_{j+1}), \tag{2.6}$$

where V^0 is a given function describing the worth of the final position x_{j+1}. We suppose that $V^0(x)$ satisfies a Lipschitz condition on any bounded set. In continuous form (2.5) would become a differential equation and (2.6) an integral from 0 to T. Assumption (2.2) guarantees that $x_1, \ldots, x_{j+1}$ remain uniformly bounded as long as x_1 is restricted to a given bounded set and T remains in the basic time interval $(0, T_0)$.

By induction on j one shows that the game has a value, denoted by $V_n(x, T)$, satisfying the functional equation.

$$\begin{gathered} V_n(x, T) = \operatorname*{val}_{y,z} [\Delta_n f(x, y, z) + V_n(x + \Delta_n g(x, y, z), T - \Delta_n)], \\ V_n(x, 0) = V^0(x), \qquad n = 1, 2, \ldots . \end{gathered} \tag{2.7}$$

The symbol val denotes the value of the game over $Y \times Z$ for fixed (x, T) which has the expression in brackets as payoff. Its solution gives the optimal strategies for the first move of the game (2.6) with initial state x and duration T.

The continuous analogue of (2.7) is a partial differential equation for a function $V(x, T)$:

$$\begin{gathered} \frac{\partial V}{\partial T} = \operatorname*{val}_{y,z} [f(x, y, z) + (\operatorname{grad} V) \cdot g(x, y, z)], \\ V(x, 0) = V^0(x), \end{gathered} \tag{2.8}$$

where grad V is the gradient in x. Unlike (2.7), this equation is derived only in a formal way; and there is no theorem guaranteeing that (2.8) has a solution except when a field of extremals has been constructed.

III. Majorant and Minorant Games

Besides the game just described it will be convenient to introduce majorant and minorant games which result if the information pattern is biased in favor of one player or the other. The minorant game is defined exactly as before, except that at each move $i = 1, \ldots, j$ player II also knows y_i before he chooses z_i; i.e., I must commit himself first at each move. The value of the game is denoted by $V_n^-(x, T)$. With this information pattern mixed strategies are irrelevant and (2.7) is replaced by:

$$
\begin{aligned}
V_n^-(x, T) &= \max_y \min_z \,[\Delta_n f(x, y, z) + V_n^-(x + \Delta_n g(x, y, z), T - \Delta_n)], \\
V_n^-(x, 0) &= V^0(x), \qquad n = 1, 2, \ldots .
\end{aligned} \tag{3.1$^-$}
$$

For the majorant game, II must commit himself first at each move. The value V_n^+ satisfies:

$$
\begin{aligned}
V_n^+(x, T) &= \min_z \max_y \,[\Delta_n f(x, y, z) + V_n^+(x + \Delta_n g(x, y, z), T - \Delta_n)], \\
V_n^+(x, 0) &= V^0(x), \qquad n = 1, 2, \ldots .
\end{aligned} \tag{3.1$^+$}
$$

It is easily seen that

$$
V_n^- \leqslant V_n \leqslant V_n^+. \tag{3.2}
$$

We shall look for conditions under which V_n^-, V_n and V_n^+ all tend to the same limit V as $n \to \infty$.

As an aid to comparing a kth game with an nth game we introduce functions U_{nk}^- and U_{nk}^+ defined inductively as follows:

$$
U_{nk}^-(x, T) = \max_{y_1} \min_{z_1} \max_{y_2} \ldots \min_{z_m} \Omega^- \tag{3.3$^-$}
$$

$$
U_{nk}^+(x, T) = \min_{z_1} \max_{y_1} \min_{z_2} \ldots \max_{y_m} \Omega^+, \tag{3.3$^+$}
$$

$$
U_{nk}^-(x, 0) = U_{nk}^+(x, 0) = V^0(x),
$$

where

$$
k \leqslant n, \qquad m = 2^{n-k},
$$

$$
T = j\Delta_k, \qquad j = 1, \ldots, 2^k,
$$

$$
\Omega^\pm = \sum_{i=1}^{m} \Delta_n f(x, y_i, z_i) + U_{nk}^\pm\left(x + \sum_{i=1}^{m} \Delta_n g(x, y_i, z_i), T - \Delta_k\right)
$$

and each y_i [or z_i] can be a function of all those z_p [or y_p] which preceed it in (3.3). By iterating (3.1) m times we get an equation differing from (3.3) only in the fact that $f(x_i, y_i, z_i)$ and $g(x_i, y_i, z_i)$ appear instead of $f(x, y_i, z_i)$ and $g(x, y_i, z_i)$ in the expression corresponding to $\Omega^{\pm}$. This suggests the result of Lemma 2 below.

It is not difficult to see that

$$U_{nk}^{-} \leqslant U_{nk}^{+}. \tag{3.4}$$

LEMMA 1. *On any bounded set in (x, T) space each of the functions V_n^-, V_n^+, U_{nk}^-, U_{nk}^+ satisfies a Lipschitz condition, with Lipschitz constant independent of k and n.*

To illustrate the method of proof it is enough to consider V_n^-. It suffices to show a uniform Lipschitz condition in x for fixed T and in T for fixed x. For a minorant game of duration T, a pair of strategies consists of a choice of $y_1, \ldots, y_j, z_1, \ldots, z_j$, where y_i is a function of $z_1, \ldots, z_{i-1}$ and z_i a function of $y_1, \ldots, y_i$. Let $x = x_1$ and $x' = x_1'$ be two initial positions, and x_i, x_i' the corresponding later positions which result from this pair of strategies. By a standard estimate in the theory of differential equations,

$$|x_i - x_i'| \leqslant |x - x'| \, e^{KT}, \qquad i = 1, \ldots, j+1,$$

where K is as in (2.2). If L is a Lipschitz constant for V^0, then the payoff starting from x differs from that starting from x' by at most $(KT + L)\, e^{KT}$ times $|x - x'|$. Since this is true for any pair of strategies,

$$|V_n^-(x, T) - V_n^-(x', T)| \leqslant (KT + L)\, e^{KT} \, |x - x'|.$$

Let M be a bound for $|f(w, y, z)|$ and $|g(w, y, z)|$, for all y in Y, z in Z, and all w which are possible positions for a game with (x, T) in the given bounded set. For given initial position x, a pair of strategies for a game of duration T induces by truncation strategies for any game of duration $T' < T$. The respective final positions are distant no more than $M(T - T')$. We find therefore that

$$|V^-(x, T) - V^-(x, T')| \leqslant M(1 + L)(T - T').$$

LEMMA 2. *On any bounded set in (x, T) space there exists a constant Q such that*

$$|U_{nk}^-(x, T) - V_n^-(x, T)| \leqslant QT\Delta_k,$$

$$|U_{nk}^+(x, T) - V_n^+(x, T)| \leqslant QT\Delta_k,$$

for any $k = 1, 2, \ldots, n \geqslant k$, *and*

$$T = j\Delta_k, \qquad j = 1, \ldots, 2^k.$$

This lemma is clearly true for $T = 0$. One proceeds by induction on j, and makes use of Lemma 1 together with the following estimates:

$$|x_i - x| \leqslant M\Delta_k, \qquad i = 1, \ldots, m, \quad x = x_1,$$

$$\sum_{i=1}^{m} \Delta_n |f(x_i, y_i, z_i) - f(x, y_i, z_i)| \leqslant KM\Delta_k^2,$$

$$\sum_{i=1}^{m} \Delta_n |g(x_i, y_i, z_i) - g(x, y_i, z_i)| \leqslant KM\Delta_k^2.$$

IV. Convergence for the Majorant and Minorant Games

Throughout the rest of the paper we make the following assumption:

(a) f is concave in y for each fixed x and z, and convex in z for each fixed x and y. The function g is bilinear in y and z for each fixed x.

Linearity here means, of course, linearity with respect to convex linear combinations. With assumption (a) the "infinitesimal" game $f + (\operatorname{grad} V) \cdot g$ over $Y \times Z$ appearing in (2.8) has a solution in pure strategies. This suggests that the discrepancy occasioned by neglecting mixed strategies in the time-discrete version (2.7) tends to 0 as $n \to \infty$. This conjecture will be proved in the next section, under an additional assumption (b). In practically all known examples of pursuit or tactical games (a) holds.

We continue to consider only games with duration T a dyadic rational number times T_0.

Lemma 3. $U_{nk}^{-} \geqslant V_k^{-}$ *and* $U_{nk}^{+} \leqslant V_k^{+}$.

Proof. Choose an element of y^0 of Y such that

$$V_k^{-}(x, T) = \min_z \,[\Delta_k f(x, y^0, z) + V_k^{-}(x + \Delta_k g(x, y^0, z), T - \Delta_k)];$$

and $z_1, \ldots, z_m$ such that z_i is a function of $y_1, \ldots, y_i$ and

$$\sum_{i=1}^{m} \Delta_n f(x, y_i, z_i) + U_{nk}^{-}\left(x + \sum_{i=1}^{m} \Delta_n g(x, y_i, z_i), T - \Delta_k\right) \leqslant U_{nk}^{-}(x, T)$$

for any admissible choices of $y_1, \ldots, y_m$. In particular, take $y_i = y^0$; and set

$$z_i^0 = z_i(y^0, \ldots, y^0), \qquad i = 1, \ldots, m,$$

$$z^0 = \frac{1}{m}(z_1^0 + \ldots + z_m^0).$$

The lemma is true for $T = 0$. Proceeding inductively, we may assume it true for time $T - \Delta_k$ and all initial states x. Then

$$V_k^-(x, T) \leqslant \Delta_k f(x, y^0, z^0) + V_k^-(x + \Delta_k g(x, y^0, z_i^0), T - \Delta_k).$$

Using assumption (a), and the fact that $\Delta_n = m^{-1}\Delta_k$, the right side is no more than

$$\Delta_n \sum_{i=1}^{m} f(x, y^0, z_i^0) + V_k^-\left(x + \Delta_n \sum_{i=1}^{m} g(x, y^0, z_i^0), T - \Delta_k\right)$$

Then using the induction hypothesis,

$$V_k^-(x, T) \leqslant \Delta_n \sum_{i=1}^{n} f(x, y^0, z_i^0) + U_{nk}^-\left(x + \Delta_n \sum_{i=1}^{m} g(x, y^0, z_i^0), T - \Delta_k\right)$$

$$\leqslant U_{nk}^-(x, T).$$

The inequality $U_{nk}^+ \leqslant V_k^+$ is proved in the same way.

THEOREM 1. *If* (a) *holds, then* $V_n^+(x, T)$ *and* $V_n^-(x, T)$ *converge to limits* $V^+(x, T)$ *and* $V^-(x, T)$ *as* $n \to \infty$, *with* $V^-(x, T) \leqslant V^+(x, T)$. *The convergence is uniform on bounded sets.*

PROOF. From Lemmas 2 and 3

$$V_n^-(x, T) \geqslant V_k^-(x, T) - QT\Delta_k$$

for $n \geqslant k$ and $T = j\Delta_k$. Suppose that for some (x, T) the sequence of numbers $V_n^-(x, T)$ had two limit points a and b with $a < b$. For some k_0

$$2QT\Delta_{k_0} < b - a.$$

Choose $k \geqslant k_0$ such that $V_k^-(x, T)$ is close to b and $n > k$ such that $V_n^-(x, T)$ is close to a. This leads to a contradiction. Thus $V_n^-(x, T)$ tends to a limit. Similarly $V_n^+(x, T)$ tends to a limit. By Ascoli's theorem and Lemma 1 the convergence is uniform on bounded sets.

Transplantation

The basic idea of the above proof was that an optimal first move y^0 for I in the kth minorant game can be "transplanted" to define the first m moves $y^0, \ldots, y^0$ which are nearly optimal in the nth minorant game. Player I loses little (no more than $QT\Delta_k$ according to Lemma 2) by observing the position of the nth game every Δ_k time units instead of every Δ_n time units. Similarly for II in the majorant game. We do not know whether nearly optimal transplantation is possible for I in the majorant game and II in the minorant game except when (b) defined below is satisfied.

Transplantation from k to $k+1$ was used by Bellman in [2].

V. Convergence of V_n

Since $V_n^- \leqslant V_n^+$ it suffices in view of Theorem 1 to show that $V_n^+ - V_n^-$ tends to 0, or by Lemma 2 that $U_{nk}^+ - U_{nk}^-$ tends to 0 for each fixed k as $n \to \infty$. We are at present able to do this only under the additional assumption:

(b) The functions f and g have the form

$$f(x, y, z) = a(x, y) + b(x, z),$$

$$g(x, y, z) = c(x, y) + d(x, z).$$

In many examples f is a function of x only, and in fact often $f = 0$ (terminal payoff games). When this is the case, (b) requires that each player's contribution at any move i to the change $x_{i+1} - x_i$ in the position of the game is not affected by his opponent's choice at the ith move.

THEOREM 2. *If both* (a) *and* (b) *hold, then*

$$V^-(x, T) = V^+(x, T) = \lim_{n \to \infty} V_n(x, T).$$

PROOF. It suffices to show that, for fixed k and T of the form $j\Delta_k$,

$$\lim_{n \to \infty} [U_{nk}^+(x, T) - U_{nk}^-(x, T)] = 0 \tag{5.1}$$

uniformly for x in any bounded set. This is true for $T = 0$ since then $U_{nk}^+ = U_{nk}^- = V^0$. We proceed by induction on j, and suppose the statement true for $T - \Delta_k$.

Let

$$z_1^0(y_1), z_2^0(y_1, y_2), \ldots, z_m^0(y_1, \ldots, y_m)$$

be optimal in the minorant problem U_{nk}^-, and

$$y_1{}^0(z_1),\ y_2{}^0(z_1, z_2), \ldots, y_m{}^0(z_1, \ldots, z_m)$$

optimal in the majorant problem U_{nk}^+. Set

$$z_1{}^1 = \text{arbitrary},$$

$$z_i{}^1(y_1, \ldots, y_{i-1}) = z_{i-1}^0(y_1, \ldots, y_{i-1}), \qquad i = 2, \ldots, m.$$

If in the majorant problem I uses the optimal choices $y_1{}^0, \ldots, y_m{}^0$ and II uses $z_1{}^1, \ldots, z_m{}^1$,

$$U_{nk}^+(x, T) \leqslant \Delta_n R + U_{nk}^+(x + \Delta_n S, T - \Delta_k), \tag{5.2}$$

where

$$R = a(x, y_1{}^0) + b(x, z_1{}^1) + \ldots + a(x, y_m{}^0) + b(x, z_m{}^1),$$

$$S = c(x, y_1{}^0) + d(x, z_1{}^1) + \ldots + c(x, y_m{}^0) + d(x, z_m{}^1).$$

For the minorant problem, let

$$y_i{}^1(z_1, \ldots, z_{i-1}) = y_i{}^0(z_1{}^1, z_1, \ldots, z_{i-1}), \qquad i = 1, \ldots, m,$$

$$R' = a(x, y_1{}^1) + b(x, z_1{}^0) + \ldots + a(x, y_m{}^1) + b(x, z_m{}^0),$$

$$S' = c(x, y_1{}^1) + d(x, z_1{}^0) + \ldots + c(x, y_m{}^1) + d(x, z_m{}^0).$$

Then

$$\Delta_n R' + U_{nk}^-(x + \Delta_n S', T - \Delta_k) \leqslant U_{nk}^-(x, T). \tag{5.3}$$

However,

$$y_1{}^1 = y_1{}^0(z_1{}^1), z_1{}^0(y_1{}^1) = z_2{}^1(y_1{}^0), y_2{}^1(z_1{}^0) = y_2{}^0(z_1{}^1, z_2{}^1),$$

$$\ldots, y_m{}^1(z_1{}^0, \ldots, z_{m-1}^0) = y_m{}^0(z_1{}^1, \ldots, z_m{}^1),$$

$$R - R' = b(x, z_1{}^1) - b(x, z_m{}^0),$$

$$S - S' = d(x, z_1{}^1) - d(x, z_m{}^0).$$

Let N be a bound for $|b|$ and $|d|$, and P a Lipschitz constant for U_{nk}^+. From (5.2) and (5.3)

$$U_{nk}^+(x, T) \leqslant \Delta_n R' + U_{nk}^+(x + \Delta_n S', T - \Delta_k) + 2N\Delta_n + 2NP\Delta_n$$

$$0 \leqslant U_{nk}^+(x, T) - U_{nk}^-(x, T) \leqslant U_{nk}^+(x + \Delta_n S', T - \Delta_k) - \tag{5.4}$$

$$U_{nk}^-(x + \Delta_n S', T - \Delta_k) + 2N(1 + P)\Delta_n.$$

As n tends to infinity the right side of (5.4) tends to 0 uniformly for x in any bounded set. Hence the same is true in (5.1), which proves Theorem 2.

VI. Time-Continuous Form

Several definitions have been proposed for an analogue with continuous time of the game described by (2.5) and (2.6). The relationship between these definitions is not known in general. The one in [4] and [6] requires the existence of a field of extremals. The definition we shall give here is a variant of the one in [8]; and is in terms of time-discrete games with the time between successive moves not specified in advance. We again consider times only of the form a dyadic rational number times the basic time interval T_0. It is merely a technical exercise to remove this restriction.

A strategy for player I consists in choosing a positive integer k_1 and functions $u_1(x),\ldots, u_s(x)$, $s = 2^{k_1}$, with values in Y. Similarly II chooses k_2 and $v_1(x),\ldots, v_t(x)$, $t = 2^{k_2}$, with values in Z. Let $n = \max(k_1, k_2)$. The payoff is calculated as in the nth time discrete game, using (2.5) and (2.6). If $k_1 = n$, then I chooses $y_i = u_i(x_i)$, $i = 1, 2, \ldots$. If $k_1 < n$, then his strategy is defined by transplantation: for $i = 1, 2, \ldots$, $y_i = u_p(x_p)$ where $p \leqslant i$ is the largest integer such that $p\Delta_n$ is of the form $j\Delta_{k_1}$. Similarly for II.

THEOREM 3. *If both* (a) *and* (b) *hold, then* $\lim_{n\to\infty} V_n(x, T)$ *is the value of the time-continuous game.*

PROOF. Let

$$V(x, T) = \lim_{n\to\infty} V_n(x, T).$$

Define $u_i(x)$ such that, for $T = i\Delta_k$, $i = 1,\ldots, s$,

$$V_k^-(x, T) = \min_z \left[\Delta_k f(x, u_i(x), z) + V_k^-(x + \Delta_k g(x, u_i(x), z), T - \Delta_k)\right]$$

where $k = k_1$ is to be chosen presently. By the proofs for Lemmas 2 and 3, this strategy for I yields against any strategy for II at least $V_k^-(x, T) - Q\Delta_k T$. By Theorem 2 given $\varepsilon > 0$ and (x, T) we may choose k_1 such that T is of the form $j\Delta_{k_1}$ and

$$V(x, T) - \varepsilon < V_{k_1}^-(x, T) - Q\Delta_{k_1} T.$$

Similarly, II has a strategy against which I can never get more than $V(x, T) + \varepsilon$. Since ε is arbitrary this proves Theorem 3.

VII. Problems of Pursuit Type

One of the most important types of differential games with finite, but not prescribed, duration is the following. Let the position vector x move in a region R; and let the payoff be the time to reach the boundary B from a given initial position x_1 and an initial time t_1. Isaacs pursuit games [9] are of this type, and also the problem of minimum time for an airplane to climb to a given altitude [11].

Let us suppose that B is a smooth manifold. We require that R be connected, but B may have several components. We again fix $T_0 > 0$ and for $n = 1, 2, \ldots$ consider time-discrete games with time Δ_n between successive moves. In this problem $f = 0$, and as before

$$x_{i+1} = x_i + \Delta_n g(x_i, y_i, z_i), \qquad i = 1, 2, \ldots . \tag{7.1}$$

The payoff shall be T if the polygonal path joining successively $x_1, x_2, \ldots$ first meets B at time $t_1 + T$. To avoid the possibility that play does not end we agree that if B is not reached before time T_0, then the payoff is

$$T_0 - t_1 + V^0(x_{N+1}), \qquad \Delta_n N = T_0 - t_1, \tag{7.2}$$

where $V^0(x)$ is a given nonnegative function such that $V^0(x) = 0$ for x in B and V^0 satisfies a Lipschitz condition on any bounded set.

In order to apply the analysis of Sections III, IV, and V we need to impose a condition to insure that, from any position sufficiently near B, player II can bring the game to an end in a short time by forcing the position to B. The condition we assume is:

(c) There is a constant $q > 0$ such that for every x in B there exists $z(x)$ such that, for all y,

$$g[x, y, z(x)] \cdot \nu(x) \geqslant q,$$

where $\nu(x)$ is the exterior unit normal to B at x.

Let $V_n(x, s)$ be the value of the game with initial conditions $x_1 = x$, $t_1 = T_0 - s$, where $s = j\Delta_n$, $j = 1, 2, \ldots, 2^n$, and x is in R. We have

$$\begin{aligned} V_n(x, s) &= 0 \quad \text{for } x \text{ in } B, \\ V_n(x, 0) &= V^0(x). \end{aligned} \tag{7.3}$$

The analogue of the recursive relation (2.7) for V_n is.

$$V_n(x, s) = \operatorname*{val}_{y,z} \theta(y, z), \tag{7.4}$$

where

$$\theta(y, z) = \begin{cases} \tau_1 \Delta_n, & \text{if} \quad \tau_1 \leqslant 1, \\ \Delta_n + V_n(x + \Delta_n g(x, y, z), s - \Delta_n), & \text{if} \quad \tau_1 \geqslant 1, \end{cases}$$

where τ_1 is the smallest nonnegative value of τ such that $x + \tau\Delta_n g(x, y, z)$ is in B. Using (c) one shows by induction on j that the value V_n exists, is a continuous function of s, and that (7.4) holds.

We also consider the majorant and minorant games, with values V_n^+, V_n^-, satisfying (7.4) with min max, max min in place of val.

THEOREM 4. *If* $g(x, y, z) = a(x) + b(x, y) + c(x, z)$, *where* b *and* c *are linear functions in* y *and* z *for each fixed* x, *and if* (c) *holds, then* V_n^+, V_n^-, *and* V_n *all tend to the same limit as* $n \to \infty$. *The convergence is uniform on bounded sets.*

The proof proceeds by modifying that for Theorems 1 and 2, and it would be repetitious to repeat the details. The auxiliary functions $U_{nk}^{\pm}$ are defined as in (3.3), except now

$$\Omega^{\pm} = \begin{cases} \tau_1 \Delta_k, & \text{if} \quad \tau_1 \leqslant 1, \\ \Delta_k + U_{nk}^{\pm}\left(x + \sum_{i=1}^{m} \Delta_n g(x, y_i, z_i), s - \Delta_k\right), & \text{if} \quad \tau_1 \geqslant 1. \end{cases}$$

The crucial estimate which must be added to reasoning in Sections III, IV, and V is

$$V_n^{\pm}(x, s) < Ad, \qquad n = 1, 2, \ldots,$$
$$U_{nk}^{\pm}(x, s) < Ad, \qquad n \geqslant k, \quad k = 1, 2, \ldots,$$

where A is the sum of $2q^{-1}$ and the Lipschitz constant of V^0, and d the distance from x to B. This estimate is valid for x in some neighborhood of B.

VIII. REMARKS ABOUT STOCHASTIC MAXIMIZATION PROBLEMS

Let us return to the situation in Section II and assume that f and g do not depend on z. For simplicity take x to be scalar. Suppose that the change δx in the position during a small time interval δt is influenced not only by the control term $g(x, y)\,\delta t$ but also by a random term. Specifically, we assume that

$$\delta x \approx g(x, y)\,\delta t + \mu(x)\,\delta t + \sigma r(\delta t)^{1/2}, \qquad x(0) = x_1, \tag{8.1}$$

where $\sigma \geqslant 0$, r is a normalized Gaussian random variable (mean $= 0$, variance $= 1$), and μ is a given function satisfying a Lipschitz condition. The random inputs during disjoint time intervals are to be independent. The problem is to maximize the expected value

$$\exp \int_0^T f(x, y)\,dt. \tag{8.2}$$

The choice of Gaussian random variables is not arbitrary, but is forced upon us if we want continuous sample functions [7, p. 420]. If y and σ are known functions of time, or functions of position and time such that $g[x, y(x, t)]$ and $\sigma(x, t)$ satisfy Lipschitz conditions in x, then there is a well defined stochastic process corresponding to (8.1) [7, p. 277].

Suppose, however, that it is known only that $0 \leqslant \sigma \leqslant c$, where c is a given positive constant. Let us take a conservative view, and turn the problem into a game against potentially hostile nature. In this game player I controls y and II controls σ. To describe the game rigorously one should discretize time. However, the analogue of (2.7) is an even less tractable recurrence relation for the value. It seems more interesting to proceed formally to find a partial differential equation for the value $W(x, T)$ of the (ill-defined) game with continuous time. We have formally

$$W(x, T) = \operatorname*{val}_{y,\sigma} \operatorname*{exp}_{r} W(x + \delta x, T - \delta t).$$

Expand $W(x + \delta x, T - \delta t)$ in Taylor series up to second order terms and take account of the fact that $\exp r = 0$, $\exp r^2 = 1$. We get

$$W_T = \operatorname*{val}_{y,\sigma} \left\{ \frac{\sigma^2}{2} W_{xx} + [g(x, y) + \mu(x)] W_x + f(x, y) \right\}. \tag{8.3}$$

Since y and σ appear separately,

$$\begin{aligned} W_T &= \frac{\sigma_0{}^2(x, t)}{2} W_{xx} + \mu(x) W_x + \max_y \, [g(x, y) W_x + f(x, y)], \\ W(x, 0) &= 0, \\ \sigma_0 &= c \qquad \text{if} \qquad W_{xx} < 0, \\ \sigma_0 &= 0 \qquad \text{if} \qquad W_{xx} > 0. \end{aligned} \tag{8.4}$$

In particular if the value W is convex in x, then it equals the result obtained ignoring random effects. There will not in general be a solution of (8.4) with continuous partial derivatives. However, (8.4) suggests a discrete recurrence relation which may be more tractable than the one obtained directly from the time-discrete process.

As another approach to the problem, let us take $\sigma = c$ and try to estimate for small values of c how much the maximum value for the process without randomization is degraded by the random effects. We make this estimate for the case when

$$g(x, y) = bx + \phi(y), \qquad \mu(x) = 0,$$

where b is a constant and $\phi(y)$, $f(x, y)$ are arbitrary.

For the deterministic maximum problem it suffices to consider controls y which are functions of time only. Let $y(t)$ be any such function, and

$$h(t) = \phi[y(t)].$$

Let $p(x, t)$ the probability density that the process given by (8.1) with $y = y(t)$ and $\sigma = c$ is in position x at time t. Then p satisfies the Fokker-Planck equation [7, p. 275].

$$\frac{c^2}{2} p_{xx} = p_t + \frac{\partial}{\partial x}[(bx + h)p].$$

Upon multiplying this equation by x and by x^2 and integrating by parts with respect to x, one finds that the mean $m_1(t)$ and second moment $m_2(t)$ satisfy the equations

$$\frac{dm_1}{dt} = bm_1 + h = g(m_1, y(t)), m_1(0) = x_1,$$

$$\frac{dm_2}{dt} = c^2 + 2bm_2 + 2hm_1, m_2(0) = x_1^2.$$

As one expects, $m_1(t)$ is the path $\bar{x}(t)$ obtained for the deterministic process with control $y(t)$. For the variance $v = m_2 - m_1^2$ we obtain

$$\frac{dv}{dt} = c^2 + 2bv, \qquad v(0) = 0,$$

from which if $b \neq 0$

$$v(t) = \frac{c^2}{2b}(e^{2bt} - 1).$$

Upon expanding $f(x, y)$ in Taylor series in x, up to second order terms, we obtain for small values of c

$$\exp \int_0^T f(x, y)\, dt \approx \int_0^T f(\bar{x}, y)\, dt + \frac{c^2}{4b} \int_0^T f_{xx}(\bar{x}, y)\, (e^{2bt} - 1)\, dt. \tag{8.5}$$

Since $y(t)$ is chosen arbitrarily, (8.5) suggests that if $f_{xx} > 0$ then the maximum expected value is at least the maximum for the deterministic process for small values of c. In general the degradation is no more than λc^2 where λ can be estimated in terms of T and bound for f_{xx}. The case $b = 0$ is entirely similar.

The author wishes to acknowledge a helpful conversation with T. E. Harris in connection with this section.

References

1. Bellman, R. "Dynamic Programming." Princeton Univ. Press, Princeton, New Jersey, 1957.
2. Bellman, R. Functional equations in the theory of dynamic programming, VI. *Ann. Math.* **65**, 215–223 (1957).
3. Bellman, R., and Dreyfus, S. Computational aspects of dynamic programming. The RAND Corporation Report R-352, 1960.
4. Berkovitz, L. D. A variational approach to a class of differential games. *In* "Contributions to the Theory of Games," Vol. IV. (to appear).
5. Berkovitz, L. D., and Dresher, M. A game-theory analysis of tactical air war. *Operations Research* **7**, 599–620 (1959).
6. Berkovitz, L. D., and Fleming, W. H. On differential games with an integral payoff *In* "Contributions to the Theory of Games," Vol. III. *Ann. Math. Studies* **39**, 413–435 (1957).
7. Doob, J. L. "Stochastic Processes," Wiley, New York, 1953.
8. Fleming, W. H. A note on differential games of prescribed duration. *In* "Contributions to the Theory of Games," Vol. III. *Ann. Math. Studies* No. 39, 407–416 (1957).
9. Isaacs, R. Differential games I, II, III, IV. The RAND Corporation Research Memoranda RM-1391, RM-1399, RM-1411, RM-1486, 1955.
10. Scarf, H. E. On differential games with survival payoff. *In* "Contributions to the Theory of Games," Vol. III. *Ann. Math. Studies* No. 39, 393–406 (1957).
11. Breakwell, J. V. The optimization of trajectories. *J. Soc. Ind. Appl. Math.* **7**, 215–247 (1959).
12. Bellman, R. On the foundations of a theory of stochastic variational processes. The RAND Corporation Paper P-1903, 1960.