AD 74903

Armed Services Technical Information Agency

Reproduced by
DOCUMENT SERVICE CENTER
KNOTT BUILDING, DAYTON, 2, OHIO

UNCLASSIFIED

FORCE
Project RAND

AN INTRODUCTION TO THE THEORY OF DYNAMIC PROGRAMMING

RICHARD BELLMAN

June, 1953

R-245

REPRODUCIBLE COPY

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

PROJECT RAND

AN INTRODUCTION TO THE THEORY OF DYNAMIC PROGRAMMING

RICHARD BELLMAN

June, 1953

R-245

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

PREFACE

The purpose of this study is to provide an introduction to a class of mathematical techniques required to treat the problems that arise in the planning of multistage processes, many of which are of day-to-day importance to military and other government workers. These are programming problems, to use the terminology currently popular, and I have introduced the adjective "dynamic" to indicate that they are problems in which time plays an important role and in which the order of performance of operations is all-important. This differentiation is not merely one of nomenclature, but is definitely conceptual, and we shall see that, properly interpreted, it furnishes us with a powerful mathematical tool with which to treat these problems.

The multistage processes in which we are interested are composed of sequences of operations in which the outcome of the preceding operations may be used to guide the course of the future ones. There are two types of operations that we can distinguish immediately: those in which the outcome is completely determined, and those in which the outcome is predictable on the basis of a probability distribution. Depending on the point of view, either type may be considered to be an approximation to the reality represented by the other. Although we shall see that mathematically the two viewpoints are not far apart, in any practical situation the two philosophies may clash violently.

Any realistic treatment of investment and replacement theory, of scientific sampling and testing, of learning theory, of industrial production problems—to mention only a few areas of importance—must involve to a greater or lesser extent problems of dynamic programming. From this it follows that however important planning has been in the past in the face of the riddles of an uncertain future, it must inevitably assume a role of greater and greater importance as an increasing population with increasing technological demands faces the challenge of a world with shrinking resources.

The theory has particular relevance to government planning, ranging in scope from the study of actual operations to questions of the procurement and replacement of equipment and to problems of the training and allocation of personnel.

Since most of the problems that arise are of an entirely novel type frequently offering formidable mathematical difficulties, we shall restrict ourselves to a consideration of the simplest problems possessing a germ of reality in order not to obscure by extraneous analytic and algebraic complications the techniques we employ.

The realistic problems that confront the theory of dynamic programming are in order of complexity on a par with the three-body problem of classical dynamics, whereas the theory painfully scrambles to solve problems on a level with that of the motion of a freely falling particle. Nonetheless, there is no cause for discouragement. Consider the case of the nuclear physicist. In attempting to explain the behavior of heavy atoms, he is forced to treat of an n-body problem infinitely more complex than the above only-partially-solved astronomical problem. Nevertheless, by combining the exact results of the one-body problem and the two-body problem with the results of experiment and observation, he is able to construct an imposing theoretical structure, albeit one with an occasional blind alley or barred window, that is amazingly useful in predicting and explaining experimental results.

Similarly, in discussing the exceedingly involved planning problems of economic life, further complicated by sociological and psychological problems, we must combine in a skillful fashion the exact results of the simple models with the intuitive theory derived from experience. The fashion in which this is to be done is beyond the power of the mathematician to describe. It must be realized that however elegant the mathematical theory, however consistent and economical its axioms, eventually the point of meta-mathematics will be reached at which someone will have to say, "I prefer this theory."

In order not to increase unduly the size of this study, I have been forced to omit any mention of a number of important and interesting investigations and to include only a part of the results known concerning the topics included.

To begin with, I have not included any treatment of the mathematical theory of learning as formulated by R. Bush and F. Mosteller, jointly, and by M. Flood. Extensive results in this field have been obtained by T. Harris, H. N. Shapiro, and the author, and, independently, together with generalizations, by S. Karlin.

Nor have I included results recently obtained by S. Johnson and S. Karlin concerning processes in which the distribution of outcomes is only partially known. These are problems of great importance in statistical applications and arise in other connections as well. A description of problems of this type will be found in an expository paper by H. Robbins.

Because of the difficulty of adequately summarizing his results in any brief space, no mention has been made of the extensive theory of pursuit games created by R. Isaacs. These games are related to the games of survival briefly discussed in Chapter 6. Both types of games belong to the general class of multistage games, which has not been touched upon here, although there are many interesting results known concerning these games, as, for example, the results of D. Blackwell and the author concerning games of bluffing and elimination of randomization and the related results of A. Dvoretzky, H. Wald, and J. Wolfowitz.

Finally, I have not included the recent investigations of I. Glicksberg, O. Gross, and myself concerning the important and novel variational problems that arise in connection with problems of economic and mechanical control.

In connection with the computational aspects of the theory of dynamic programming, I have not discussed any applications of the "simplex" method of G. Dantzig that has proved of such great value in the theory of linear programming and yields the solution of many important classes of dynamic programming problems.

It is a pleasure to acknowledge my indebtedness to a number of sources: First, to the von Neumann theory of games, as developed by J. von Neumann, O. Morgenstern, and others, which shows how to treat by mathematical analysis vast classes of problems formerly thought far out of the reach of the mathematician—and relegated, therefore, to the limbo of imponderables—and, simultaneously, to the Wald theory of sequential analysis, as developed by A. Wald, D. Blackwell, A. Girshick, J. Wolfowitz, and others, which shows the vast economy of effort that may be effected by the proper consideration of

multistage testing processes; second, to a number of colleagues and friends who have discussed various aspects of the theory with me and have contributed greatly to its clarification and growth.

In particular, the last section of Chapter 6 is taken verbatim from an unpublished paper by M. Peisakoff. Section 2.12 is from an unpublished paper by S. Karlin and H. N. Shapiro, as is also Section 3.12, while Section 3.11 is based on a personal communication from H. N. Shapiro. A partial solution of the problem in Section 3.11 had previously been given by O. Gross, using a different approach. The solution of Eq. (3.1) was obtained while collaborating with M. Shiffman; the solution of Eq. (5.45) was obtained in collaboration with D. Blackwell; and the formulation in mathematical terms of games of survival was obtained in collaboration with J. LaSalle.

The optimal inventory problem mentioned in Chapter 1 and discussed briefly in Chapter 4 was first studied by K. Arrow, T. Harris, and J. Marschak. Following this, an extensive treatment, together with many generalizations, was given by A. Dvoretzky, J. Kiefer, and J. Wolfowitz.

I should like to thank Oliver Gross, who read the final manuscript through with great care and made a large number of valuable suggestions and corrections.

Finally, I should like to record a particular debt of gratitude to O. Helmer and E. W. Paxson, who early appreciated the importance of the study of multistage processes and, in addition to furnishing a large number of stimulating problems arising naturally in various important applications, constantly encouraged me in my researches.

SUMMARY

Dynamic programming is a mathematical theory devoted to the study of multistage processes. The multistage processes discussed in this report are composed of sequences of operations in which the outcome of those preceding may be used to guide the course of future ones. Operations of both deterministic and stochastic types are discussed.

After an introductory chapter, in which a number of representative problems are investigated, and a succeeding chapter, in which some general mathematical results are obtained, the remainder of the report is devoted to the study of equations of particular types that arise in various applications of the theory.

CONTENTS

CHAPTER 1
FUNDAMENTAL CONCEPTS

1.1. Introduction

We propose in this chapter to discuss a number of representative problems in the theory of dynamic programming, emphasizing their conceptual and analytic aspects. In place of any discussion of an abstract type, which at this stage would necessarily remain rather vague, we shall begin the chapter by posing a number of simple prototypes of general problems that fall within the domain of our theory. Following this we shall present various mathematical approaches to these problems and introduce the reader to the functional equation technique which we shall employ throughout most of the study. Since both the class of problems we shall encounter and the techniques we shall employ possess certain features of novelty, we shall not hesitate to be repetitious to a certain extent, feeling that, in an introductory work, sins of repetition are of lesser magnitude than sins of omission.

1.2. Some Problems

PROBLEM 1.1. We are given a quantity $x > 0$ that may be divided into two parts, y and $x - y$. From y we obtain a return of $g(y)$; and from $(x - y)$, a return $h(x - y)$. In so doing we expend a certain amount of our original resources and are left with a new quantity, $ay + b(x - y)$, $0 < a, b < 1$, with which to continue the process. How does one proceed so as to maximize the total return obtained in a finite, or unbounded, number of stages?

PROBLEM 1.2. We are given a quantity $x > 0$ that is to be utilized to accomplish a certain task. If an amount y, where $0 \leq y \leq x$, is used on any single attempt, the probability of success is $a(y)$. If the task is not accomplished on the first try, we continue with the new quantity $x - y$. How does one proceed in order to maximize the over-all probability of success?

PROBLEM 1.3. We are informed that a particle is in either state 0 or 1, and we are given initially the probability x that it is in state 1. Use of the operation A will reduce this probability to ax, where a is some positive constant less than 1, whereas operation L, which consists in observing the particle, will tell us definitely which state it is in. If it is desired to transform the particle into state 0 in a minimum time, what is the optimal procedure?

PROBLEM 1.4. At each stage of sequence of actions we are allowed our choice of one of two actions. The first has associated a probability p_1 of gaining one unit, a probability p_2 of gaining two units, and a probability p_3 of terminating the process. The second has a similar set of probabilities p'_1, p'_2, p'_3. What sequence of choices maximizes the probability of attaining at least n units before the process is terminated?

PROBLEM 1.5. We are fortunate enough to possess two gold mines, A and B, the first of which possesses an amount x of gold, while the second possesses an amount y. If the only gold-mining machine we have is used in A, there is a probability p_1 that r per cent of the gold there will be brought up safely, the machine still being usable, and a probability $(1 - p_1)$ that the machine will be damaged, will mine no gold, and will be of no further use. Similarly, mine B has the probabilities q_1 and $(1 - q_1)$ associated with it. How does one proceed in order to maximize the total amount of gold obtained before the machine is defunct?

PROBLEM 1.6. Let us consider the above problem in the case in which we know only the expected amounts of gold in each mine and the expected amount mined each time, without being able to observe the results of individual operations.

PROBLEM 1.7. Two players, A and B, the first possessing x dollars and the second possessing y dollars, play a modified coin-tossing game described by the matrix

$$M = \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}.$$

Assuming that each player is motivated by a desire to ruin the other, how does each play?

1.3. Enumerative Solutions—Deterministic Case

Having posed the problems, let us now consider what we shall accept as an answer. Clearly, what we desire is a rule that, when given the initial parameters, the allowable operations, and their outcomes, yields a sequence of actions which achieves, or attempts to achieve, a designated goal in a fashion that is optimal in some previously specified sense.

One way of obtaining this rule is to list all possible rules, calculate the effect of each, and then choose the one best suited for our purposes. This method we call the "enumerative" method.

Let us pursue this technique in the case of Problem 1.1, which is deterministic, beginning with the case in which exactly N operations are permitted. Let $y_1, y_2, \cdots, y_N$ be the sequence of choices. The total return will be

$$J(y_1, y_2, \cdots, y_N) = \sum_{i=1}^{N} g(y_i) + \sum_{i=1}^{N} h(x_i - y_i), \tag{1.1}$$

where the variables are constrained by the conditions

$$\begin{aligned}
&\text{(a)} \qquad & 0 &\le y_i \le x_i, \\
&\text{(b)} \qquad & x_1 &= x, \\
& & x_2 &= ay_1 + b(x_1 - y_1), \\
& & &\cdots\cdots\cdots\cdots\cdots\cdots \\
& & x_N &= ay_{N-1} + b(x_{N-1} - y_{N-1}).
\end{aligned} \tag{1.2}$$

The problem is now to maximize J subject to the above constraints. Since several of the y_i may be 0 or x_i (end points of the allowable intervals), any naïve application of calculus is somewhat hazardous.

For N infinite, which is to say that an unbounded number of operations are permitted, the problem is one involving an infinite number of variables, and rather more discussion is required. Let us observe that the case of infinite N, which is meaningless in any practical situation, possesses a very important invariance property from the mathematical point of view, since after any finite number of stages there still remains a process with an infinite number of stages. This fact will be of great utility in our subsequent discussion.

1.4. Enumerative Solutions—Stochastic Case

In Problems 1.1 through 1.7 the outcome of any action is indeterminate, specified only by a distribution function, which we take to be known. Problems of a second order of difficulty, overlapping the domain of sequential analysis, are those in which the distribution function is only partially known. Third-order problems would perhaps be those in which it is not known whether or not a distribution function exists. We see from this brief listing that it is possible to construct a hierarchy of problems ranging from the blissful state of complete determinacy to the inferno of utter ignorance. In this introductory treatment we shall consider only first-order problems.

In order to understand what an enumerative solution of a stochastic decision problem involves, let us discusss Problem 1.4, considering the simple case in which only two stages are allowed.

In the general case in which N stages are allowed, we require $2 \cdot 4^{N-1}$ listings in order to enumerate all possible rules. If N is infinite, which is to say that the process is allowed to continue until it terminates of itself, the number of possible rules is non-enumerable.* This fact will make any direct application of the enumerative method somewhat tedious of execution.

The possible sequences of choices may be illustrated graphically by means of a tree, as shown in Fig. 1.1.

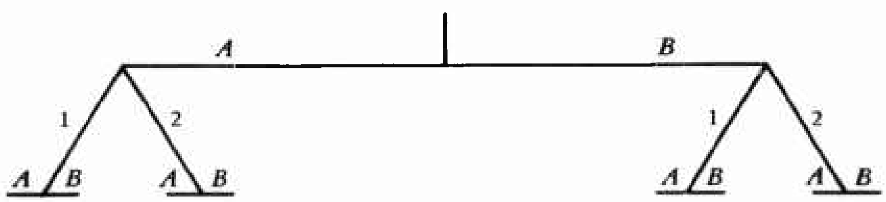

Fig. 1.1

The eight possible rules are

$$\begin{array}{cccc} A(1)A, & A(2)A, & A(1)B, & A(2)B, \\ B(1)A, & B(2)A, & B(1)B, & B(2)B, \end{array} \tag{1.3}$$

* We recall that an infinite process which allows one of two choices at each stage yields a set of possible sequences that may be put into 1 — 1 correspondence with the dyadic expressions of the real numbers in [0, 1].

where, for example, $A(i)B$ means that A is chosen first, i units are obtained, and B is then chosen. If the process terminates with an initial choice of A or B, there is, of course, no further need for decision.

We now require a method for comparing the outcomes of different rules. Since we are dealing with stochastic sequences, let us use the metric of probability theory and consider the expected return. In general, let us note, it is not the expected return that is important, but rather the expected value of some function of the total return. In the case of Problem 1.4, this function has the form, if R is the return,

$$f(R) = \begin{cases} 1, & R \geq n, \\ 0, & R < n, \end{cases} \tag{1.4}$$

since the expected value of $f(R)$ is precisely the probability that $R \geq n$.

It is now not difficult to calculate the desired expected value and to compare the eight numbers obtained in this way to obtain the optimal policy. Although feasible for small N, this technique is impossible of execution for N of even moderate size.

We shall see, subsequently, that the enumerative method possesses theoretical value in some cases and computational value in others. In general, however, it is inferior to the method we shall employ throughout most of the study.

1.5. Enumerative Approach—II

The problems above lead to a complicated enumeration of cases because of the fact that a policy consists not merely in a selection of choices of A or B, but actually in a selection coupled with actual occurrences. Hence, in place of the four policies AA, AB, BA, BB for the two-stage process, we have the eight policies of the form $A(i)B$, $B(j)A$ to consider.

In Problem 1.7, in which the results of an individual choice cannot be ascertained, we need consider, in the two-stage process, only the four choices AA, AB, BA, BB. Let us observe that a policy such as AB is to be interpreted to mean that B is used on the second trial, if the machine survives the first trial.

It is interesting to note that analytically there will be no difference between (1) the above problem, in which we do not know the precise outcome of any individual action, (2) a similar problem in which we do observe the effect of each choice, provided that the effects are of sufficiently simple type, and (3) a similar completely deterministic problem.

The enumerative approach here leads to a very interesting geometric treatment of the problem, which we shall present later in Chapter 3.

1.6. The Functional Equation Approach

Let us begin by observing that the problems posed above have the following features in common:

1. The state of the system is described by a small set of parameters.
2. The effect of a decision is to transform this set of parameters into a similar set.

3. The past history of the system is of no importance in determining future actions, a Markovian property.

We have purposely left this description rather vague, since we feel it is the spirit of the problem rather than the letter that is significant. It is extremely important to realize that one cannot axiomatize mathematical formulation and legislate away ingenuity. In some problems the state variables are forced on one; in others there is a choice, and the mathematical solution will stand or fall depending on the choice that is made. Experience alone helps in the setting up of useful mathematical models.

In addition to the above facts, we require the following simple *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.*

We shall now apply this principle to obtain functional equations whose solutions will yield the optimal strategies.

PROBLEM 1.8. Let us set

$$f(x) = \text{total return obtained using an optimal policy of allocation of resources at each stage, where an unlimited number of operations is permitted.} \tag{1.5}$$

If the initial allocation is y and $x - y$, the return from this division will be $g(y) + h(x - y)$, with $ay + b(x - y)$ remaining to continue the process. From the definition of $f(x)$, paying heed to our fundamental principle, above, it follows that the total return from $ay + b(x - y)$ will be $f[ay + b(x - y)]$. Consequently, the total return derived from an initial allocation of y and $x - y$ will be

$$R(y) = g(y) + h(x - y) + f[ay + b(x - y)]. \tag{1.6}$$

The maximum return will be obtained if y is chosen to maximize $R(y)$. Since this maximum return is, by definition, $f(x)$, we obtain the functional equation

$$f(x) = \operatorname*{Max}_{0 \le y \le x} \{g(y) + h(x - y) + f[ay + b(x - y)]\}. \tag{1.7}$$

Since we have no a priori assurance that $f(x)$ is continuous, even if g and h are, it is better to write

$$f(x) = \operatorname*{Sup}_{0 \le y \le x} \{g(y) + h(x - y) + f[ay + b(x - y)]\} \tag{1.8}$$

and then to prove, under certain assumptions, that the supremum is actually attained.

For the N-stage process, we have, using an obvious notation and taking g and h to be continuous,

$$\begin{aligned} f_1(x) &= \operatorname*{Max}_{0 \le y \le x} [g(y) + h(x - y)], \\ f_N(x) &= \operatorname*{Max}_{0 \le y \le x} \{g(y) + h(x - y) \\ &\qquad + f_{N-1}[ay + b(x - y)]\}, \qquad N = 2, 3, \cdots. \end{aligned} \tag{1.9}$$

We now see the advantage of the mathematical fiction of taking N infinite. In place of the sequence of functions, $\{f_N(x)\}$, given by (1.9), we have one function $f(x)$ satisfying (1.8). There is, naturally, a close connection between the sequence $\{f_N(x)\}$ and $f(x)$, which we shall subsequently exploit.

Having a functional equation for $f(x)$ that under certain simple natural conditions determines $f(x)$ uniquely, as we shall see, the question naturally arises as to how this function is to be used to determine an optimal policy. Turning to (1.7) we see that y is the quantity, or a quantity, that maximizes $g(y) + h(x - y) + f[ay + b(x - y)]$ in $[0, x]$. This quantity y is a known function of x if $f(x)$ is once found.

It is clear then that there is an equivalence between the optimal policy and the solution of the functional equation. We shall subsequently discuss this in more detail.

PROBLEM 1.9. Let us set, in similar fashion,

$$f(x) = \text{over-all probability of success using an optimal procedure.} \tag{1.10}$$

If we use an amount y on the first try, our probability of success is $a(y)$. If we fail on the first try, an occurrence with probability $[1 - a(y)]$, we use an optimal policy starting with the residual amount $x - y$. Hence, $f(x)$ satisfies the relation

$$f(x) = \operatorname*{Max}_{0 \le y \le x} \{a(y) + [1 - a(y)]f(x - y)\}. \tag{1.11}$$

The problem is much simpler mathematically if we consider the probability of failure rather than the probability of success.

PROBLEM 1.10. Let

$$f(x) = \text{expected time, using an optimal procedure, to transform the particle into state 0, if the probability that it is initially in state 1 is } x. \tag{1.12}$$

If we observe the system, we find it in state 1 with probability x and continue with that knowledge; whereas, if we find it in state 0, the process terminates. Hence, if $f_L(x)$ denotes the expected time spent if we observe on the first move, we have

$$f_L(x) = 1 + xf(1). \tag{1.13}$$

On the other hand, if we act, we have

$$f_A(x) = 1 + f(ax). \tag{1.14}$$

Combining these two results, we see that

$$\begin{aligned} f(x) &= \operatorname{Min}\begin{bmatrix} 1 + xf(1) \\ 1 + f(ax) \end{bmatrix}, \qquad 0 < x \le 1, \\ f(0) &= 0. \end{aligned} \tag{1.15}$$

PROBLEM 1.11. Let

$$f(n) = \text{probability of obtaining at least } n, \text{ using an optimal procedure.} \tag{1.16}$$

Enumerating the possibilities relative to each choice, we obtain

$$f(n) = \text{Max}\begin{Bmatrix} p_1f(n-1) + p_2f(n-2) \\ p_1f(n-1) + p_2f(n-2) \end{Bmatrix}, \qquad n \geq 2. \tag{1.17}$$

The reasoning behind this equation is as follows: If one obtains k on the first step, one continues so as to maximize the probability of obtaining at least $n - k$ on the following steps. For $n = 1$, we have the same equation with the convention that $f(-k) = 1$, $k \geq 0$.

PROBLEM 1.12. Let

$$f(x, y) = \text{expected amount of gold obtained using an optimal sequence of choices.} \tag{1.18}$$

If choice A is made, we have

$$f_A(x, y) = p_1\{r_1x + f[(1 - r_1)x, y]\}; \tag{1.19}$$

whereas, if choice B is made, we obtain

$$f_B(x, y) = p_2\{s_1y + f[x, (1 - s_1)y]\}, \tag{1.20}$$

where $r_1 = r/100$, $s_1 = s/100$.

Hence,

$$f(x, y) = \text{Max}\begin{bmatrix} A\colon\ p_1\{r_1x + f[(1 - r_1)x, y]\} \\ B\colon\ p_2\{s_1y + f[x, (1 - s_1)y]\} \end{bmatrix}, \qquad x, y \geq 0. \tag{1.21}$$

PROBLEM 1.12′. Let us consider the same situation in which it is desired to maximize not the expected value of the total return, R, but the expected value of $\varphi(R)$, where ϕ is a given function.

In this case it is necessary to introduce another state variable, namely z, the amount already mined. If we set

$$f(x, y, z) = \exp \phi(R), \text{ starting with an amount } z, \text{ using an optimal policy}, \tag{1.22}$$

we obtain for f the functional equation

$$f(x, y, z) = \text{Max}\begin{bmatrix} A\colon\ p_1f[(1 - r_1)x, y, z + r_1x] + (1 - p_1)\phi(z) \\ B\colon\ p_2f[x, (1 - s_1)y, z + s_1y] + (1 - p_2)\phi(z) \end{bmatrix}, \qquad x, y \geq 0,$$
$$f(0, 0, z) = \phi(z). \tag{1.23}$$

PROBLEM 1.13. Let

$$f(x, y) = \text{expected amount of gold obtained using an optimal sequence of choices when } A \text{ has expected amount } x \text{ and } B \text{ has expected amount } y. \tag{1.24}$$

If choice A is made, we have

$$f_A(x, y) = p_1\{r_1x + f[(1 - r_1)x, y]\}, \tag{1.25}$$

while choice B yields

$$f_B(x, y) = p_2\{s_1 y + f[x, (1 - s_1)y]\}. \tag{1.26}$$

Hence,

$$f(x, y) = \text{Max}\begin{bmatrix} A\colon\ p_1\{r_1 x + f[(1 - r_1)x, y]\} \\ B\colon\ p_2\{s_1 y + f[x, (1 - s_1)y]\}\end{bmatrix}, \qquad x, y \geq 0. \tag{1.27}$$

We see, as noted above, that the two problems, 1.12 and 1.13, yield the same functional equation, although quite different in structure. This will only be true in the simplest versions of Problem 1.12.

Problem 1.14. Since the total amount of money in the game remains constant, c, it is sufficient to specify the amount of money, x, held by the first player. Let

$$\begin{aligned} f(x) = {} & \text{probability that } B \text{ is ruined before } A \text{ when } A \text{ has } x \text{ and } B \text{ has } c - x \\ & \text{and when both sides use optimal strategies.} \end{aligned} \tag{1.28}$$

If A uses the strategy $p = (p_1, p_2)$, where p_1 and p_2 denote, respectively, the frequencies with which the first and second rows of M are played, and if B uses the strategy $q = (q_1, q_2)$, the frequencies with which B chooses the columns, we obtain, for $0 < x < c$,

$$f(x) = p_1 q_1 f(x + 1) + p_1 q_2 f(x - 1) + p_2 q_1 f(x - 2) + p_2 q_2 f(x + 2). \tag{1.29}$$

Let us denote the right-hand side of this equation by $T[p, q, f(x)]$. If both sides play optimally, we have

$$\begin{aligned} f(x) &= \underset{p}{\text{Max}}\ \underset{q}{\text{Min}}\ T[p, q, f(x)] = \underset{q}{\text{Min}}\ \underset{p}{\text{Max}}\ T[p, q, f(x)], \qquad 0 < x < c, \\ f(0) &= 0, \qquad x \leq 0, \\ f(x) &= 1, \qquad x \geq c. \end{aligned} \tag{1.30}$$

This is equivalent to saying that $f(x)$ is the value of the game whose payoff matrix is

$$\begin{pmatrix} f(x + 1) & f(x - 1) \\ f(x - 2) & f(x + 2) \end{pmatrix}. \tag{1.27}$$

1.7. Discussion

It is important to observe that in all these problems the functions—the solutions of the functional equations—are essentially secondary, since the optimal procedures are the items of primary interest. Actually the two are equivalent, since a procedure defines a function, and, conversely, a solution of the functional equation defines a procedure. Frequently the procedure is quite easy to describe, whereas the function is quite complicated. From the point of view of application, the function yields little or no immediate information as to the structure of an optimal procedure, whereas the individual steps in the process may illustrate some valuable principles that may be applied in heuristic fashion to the more complicated problems which frequently and almost maliciously defy exact analysis.

The plan of this study is first to formulate some general mathematical problems of the type discussed above with the concomitant existence and uniqueness theorems, and then to discuss a number of interesting simple representatives of the general problem.

It is not too difficult to subsume the problems we discuss under a more abstract framework. However, it is important to postpone this inundation until a large number of individual problems have been formulated and solved, since certain indigenous features of each problem in its native setting will facilitate its solution. It is further important that we consider problems that arise naturally in the external world, since, in general, it is only these that we can expect to possess solutions with simple and easily discernible structures. Particularly in a theory involving non-linear functional equations are the signposts of nature most valuable in preventing us from wandering desolately in the trackless wilderness of existence and uniqueness theorems.

1.8. General Mathematical Formulation

Let p be a point in an abstract space, a "phase space"; let $f(p)$ be a function of p, whose values lie in another abstract space; and let T_k be a set of operators applicable to f. The general class of functional equations in which we are interested has the form

$$f(p) = \operatorname*{Max}_{k} [T_k(f)]. \tag{1.31}$$

Minimization problems are included in this formulation, since they may be converted into maximization problems by a simple change of sign. The index k may run over a finite, infinite denumerable, or non-denumerable set.

The simplest examples of such operators are furnished by the class

$$T_k(f) = g_k(p) + \sum_{l} b_{kl}(p) f(s_{kl}p), \tag{1.32}$$

where p is a point in n-dimensional space, $g_k(p)$, $b_{kl}(p)$ are scalar functions, and s_{kl} is a point transformation. Examples of equations connected with such transformations are (1.7), (1.11), (1.15), (1.17), (1.21), and (1.27). We shall consider only equations of this type in this study, except in Chapter 4, where some simple integral operators appear.

Problems leading to equations of this type arise from both deterministic and stochastic models, with slight differences of form, as we can see upon comparing Eq. (1.7), which arises from a deterministic model, with Eq. (1.11), which arises from a stochastic model. These slight differences force us to use different techniques in establishing existence and uniqueness theorems.

In this section, devoted to a mathematical formulation of the problems occurring in one phase of dynamic programming, we shall discuss various types of general problems that lead to the diverse classes of functional equations we shall consider in the remainder of the report.

1.8.1. Deterministic Investment Problems. At each stage of a sequence of operations we are permitted to divide our resources, of total amount x, into k parts $x_1, x_2, \cdots, x_k$, where $x_i \geq 0$ and $x_1 + x_2 + \cdots + x_k = x$. The return from this partition is given by $a(x_1, x_2, \cdots, x_k)$, and a total quantity $b(x_1, x_2, \cdots, x_k)$ will be available to continue, repeating the process. If we define

$$f(x) = \text{total return obtained using an optimal policy, with an unlimited number of operations,} \tag{1.33}$$

we obtain the functional equation

$$f(x) = \operatorname*{Max}_{x_i} \{a(x_1, x_2, \cdots, x_n) + f[b(x_1, x_2, \cdots, x_n)]\}, \tag{1.34}$$

where the x_i are subjected to the conditions $x_i \geq 0_1$, $x_1 + x_2 + \cdots + x_n = x$. The boundary condition is $f(0) = 0$, assuming, naturally, that $a(0) = 0$, $b(0) = 0$.

1.8.2. Stochastic Investment Problems—The Gold-mining Problem. There are n sources of profit having respective total yields $x_1, x_2, \cdots, x_n$. We are allowed an unbounded number of operations and a choice of one of a set of possible actions on each operation. Associated with the kth operation there is a distribution function of returns:

$$p_{ik} = \text{probability of a return of } \sum_j a_{ijk} x_j, \tag{1.35}$$

where we assume that $\sum_i p_{ik} \leq c_1 < 1$ for all k. This means that associated with each choice of an action there is a non-zero probability of not being able to continue the sequence of operations. If this kth operation yields a return of $\sum a_{ijk} x_j$, the remaining total yields are now $x_j(1 - a_{ijk})$, $j = 1, 2, \cdots, n$. Hence, if we set

$$f(x_1, x_2, \cdots, x_n) = \text{expected yield employing an optimal policy,} \tag{1.36}$$

we obtain for f the functional equation

$$f(x_1, x_2, \cdots, x_n) = \operatorname*{Max}_{k} \left[\sum_{i=1}^{n} p_{ik} \left\{ \sum_j a_{ijk} x_j + f[x_1(1 - a_{i1k}), \cdots, x_n(1 - a_{ink})] \right\} \right]. \tag{1.37}$$

The simplest example of this is furnished by the equation

$$f(x_1, x_2) = \operatorname*{Max}_{k} \begin{bmatrix} p_1\{r_1 x_1 + f[(1 - r_1)x_1, x_2]\} \\ p_2\{r_2 x_2 + f[x_1, (1 - r_2)x_2]\} \end{bmatrix}. \tag{1.38}$$

Here p_i is the probability of receiving $r_i x_i$ and being allowed to continue the operations.

1.8.3. A Testing Problem. A system is known to be in some one of $N + 1$ different states, which we denote by $0, 1, 2, \cdots, N$, with an initial probability $\{p_k\}$ that it is in the kth state. By means of the following operations, we wish to transform it into a given state, which may as well be 0, with the certainty that it is there in a minimum time:

L: We observe the actual state of the system and proceed with that knowledge. This requires a time t_L.

A: We perform an operation A_l that converts the original probability distribution $\{p_k\}$ into a new distribution $\{p_{kl}\}$. This operation consumes a time t_l.

Let us set

$$p = \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_N \end{bmatrix}, \qquad T_l p = \begin{bmatrix} p_{0l} \\ p_{1l} \\ \vdots \\ p_{Nl} \end{bmatrix}, \qquad x_k = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad k = 0, 1, 2, \cdots, N, \tag{1.39}$$

where the "1" occurs in the kth place and define

$$f(p) = f(p_1, p_2, \cdots, p_n) = \text{expected time required to have the system in state 0 with certainty, using an optimal policy.} \tag{1.40}$$

Then f satisfies the equation

$$\begin{aligned} f(p) &= \text{Min} \begin{Bmatrix} L\colon\ t_L + \sum_{k=0}^{N} p_k f(x_k) \\ A_l\colon\ t_l + f(T_l p) \end{Bmatrix}, \qquad p \neq x_0, \\ f(x_0) &= 0. \end{aligned} \tag{1.41}$$

There is a natural discontinuity at x_0, since for $p \neq x_0$, no matter how close it may be, we must look or act, either of which consumes a certain non-zero time.

1.8.4. A Production Problem. Let us suppose that we are given initial amounts $x_1, x_2, \cdots, x_n$ of substances $A_1, A_2, \cdots, A_n$ with the knowledge that at each stage of a sequence of operations each substance may be used to produce both more of itself and more of the other substances. If it is desired to maximize the amount of one given substance we possess at the end of a fixed number of stages, a question arises as to the allocation of resources at each stage.

Let us consider a simple case in which there are only two substances, A and B, and in which a quantity x of A yields $c_1 x$ of A, if used to produce A, and $c_2 x$ of B, if used to produce B. Similarly y of B produces $d_1 y$ of A and $d_2 y$ of B. Assuming that at each stage we are allowed only these operations:

$$T_1\colon \begin{matrix} A \to A \\ B \to A \end{matrix}, \qquad T_2\colon \begin{matrix} A \to A \\ B \to B \end{matrix}, \qquad T_3\colon \begin{matrix} A \to B \\ B \to A \end{matrix}, \qquad T_4\colon \begin{matrix} A \to B \\ B \to B \end{matrix}, \tag{1.42}$$

we see that if the initial amounts of A and B are x and y, respectively, at the end of the first stage the results of the various operations will be

$$\begin{aligned} T_1\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} c_1 x + d_1 y \\ 0 \end{pmatrix}, \qquad & T_2\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} c_1 x \\ d_2 y \end{pmatrix}, \\ T_3\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} d_1 y \\ c_2 x \end{pmatrix}, \qquad & T_4\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ c_2 x + d_4 y \end{pmatrix}. \end{aligned} \tag{1.43}$$

The effects of various choices T_1, T_2, T_3, T_4 are equivalent to matrix operations:

$$T_1 = \begin{pmatrix} c_1 & d_1 \\ 0 & 0 \end{pmatrix}, \quad T_2 = \begin{pmatrix} c_1 & 0 \\ 0 & d_2 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 0 & c_2 \\ d_1 & 0 \end{pmatrix}, \quad T_4 = \begin{pmatrix} 0 & 0 \\ c_2 & d_2 \end{pmatrix}. \tag{1.44}$$

If, given the initial vector x_0 whose components are x and y, we wish to maximize the final amount of A, the problem is that of choosing the sequence of matrices $T_{i_1}, T_{i_2}, \cdots, T_{i_n}$ so that the x-component of

$$x_n = \left(\prod_{k=1}^{n} T_{i_k}\right) x_0 \tag{1.45}$$

will be a maximum.

In general we may wish to maximize some linear combination of the final amounts of the various commodities, which is to say, the inner product of x with a fixed vector c, (x, c). If we define

$$\phi_n(x) = (x_n, c) = \text{value of the inner product obtained using an optimal } n\text{-stage procedure,} \tag{1.46}$$

we obtain the functional equation

$$\phi_n(x) = \operatorname*{Max}_k \phi_{n-1}(T_k x, c), \qquad n \geq 1. \tag{1.47}$$

This problem is complicated by a lack of invariance in time, i.e., $n \to n - 1$. If an invariant formulation is desired, we may be able to obtain this in certain cases by using the following device: We may suppose that we are performing these operations in order to meet some contingency that has a certain probability of occurring between stages of the sequence of operations.

Let

$\phi(x) = \phi(x_1, x_2) =$ probability that the contingency can be successfully met with current quantities x_1 of A and x_2 of B.

$f(x) = f(x_1, x_2) =$ probability that the contingency can be successfully met whenever it occurs, given initial amounts of x_1 of A and x_2 of B and using an optimal allocation policy.

$p =$ probability that the contingency occurs between two stages. (1.48)

Then

$$f(x) = p\phi(x) + (1 - p) \operatorname*{Max}_k f(T_k x). \tag{1.49}$$

1.8.5. An Investment Problem. In the previous problems we have been dealing with expected values and the maximization or minimization of these values. Frequently, however, the actual purpose of a program is not so much to maximize the expected value of a critical variable as it is to maximize the probability that this variable is above a certain level. Consequently, it is of interest to compare the optimal strategy derived from using the expected-value criterion with the strategy corresponding to the more realistic criterion.

There is an additional reason for studying alternative criteria. It is possible that the strategy which yields the maximum expected value will have an undesirably large variance associated with it which makes its use quite risky. On the other hand, it is to be expected that the strategy associated with the criterion of maximum probability will automatically possess a smaller variance.

Let us consider a situation in which there is an infinite sequence of operations to be performed. At each stage we have a choice of one of the operations $A_1, A_2, \cdots, A_k$. If the ith operation is chosen, there is a probability p_{ij}, $1 \leq j \leq r$, of receiving an amount j, with $\sum_{j=1}^{r} p_{ij} < 1$ for all i, and a probability of terminating the sequence of operations equal to the remaining probability $1 - \sum_{j=1}^{r} p_{ij}$.

If we define

$$f(n) = \text{probability of obtaining a return greater than or equal to } n, \tag{1.50}$$

then clearly $f(n)$ satisfies the equation

$$f(n) = \operatorname*{Max}_{i} \left[\sum_{j=1}^{r} p_{ij} f(n-j) \right], \tag{1.51}$$

since, if one obtains j on the first operation, one continues so as to maximize the probability of obtaining at least $n - j$.

This is the general case of Problem 1.8.4, discussed previously.

1.8.6. An Optimal Inventory Problem. Let us assume that we have a quantity, x, of merchandise on hand, and that there is a probability $\phi(y)\,dy$ that at some specified time we shall be called on to deliver a quantity y of this merchandise. To meet this potential demand, we may order an additional quantity, z, of merchandise at a cost of $g(z)$. If the demand, y, exceeds the total quantity, $x + z$, the request for merchandise is satisfied as far as possible, and a penalty cost of M is levied. Assuming that this situation repeats itself indefinitely, and that future costs are discounted at a fixed rate, a, determine the ordering policy which minimizes the over-all expected cost.

Let

$$f(x) = \text{expected total cost using an optimal ordering policy.} \tag{1.52}$$

If z is ordered initially, the total expected cost will be

$$T(z, f) = g(z) + a\left\{ [M + u(0)] \int_{x+z}^{\infty} \phi(y)\,dy + \int_0^{x+z} f(x + z - y)\phi(y)\,dy \right\}. \tag{1.53}$$

Since z is to be chosen to minimize the total cost, we have, for our functional equation,

$$f(x) = \operatorname*{Inf}_{z \geq 0} [T(z, f)]. \tag{1.54}$$

CHAPTER 2

EXISTENCE AND UNIQUENESS THEOREMS

2.1. Introduction

In this chapter we shall discuss the questions of the existence and uniqueness of the solutions of the various functional equations formulated in response to the problems posed in Chapter 1.

We shall first show how that general factotum—the method of successive approximations—yields, under assumptions that are natural to the problem, existence and uniqueness theorems together with information concerning the dependence of the solution on the variables and parameters in the equation.

Using these facts, we shall turn to a discussion of the rigorous concept of a solution and then to various questions concerning computational and approximate methods. This last is of great practical importance, since the non-linearity of the equations reduces the number that may be resolved purely by analytic means to a woeful handful.

2.2. The Equation $f(p) = \operatorname{Max}_k [g_k(p) + h_k(p)f(T_kp)]$

In Section 1.6 we encountered the equation

$$f(x, y) = \operatorname{Max} \begin{bmatrix} p_1\{r_1x + f[(1 - r_1)x, y]\} \\ p_2\{s_1y + f[x, (1 - s_1)y]\} \end{bmatrix},$$

in connection with a gold-mining problem.

This is a special case of the more general equation

$$f(p) = \operatorname*{Max}_{1 \le k \le n} \left[g_k(p) + \sum_{i=1}^{M} h_{ij}(p) f(T_{ik}p) \right],$$

where p is a point in N-dimensional space, E_N, and $T_{ik}p$ is a transformation taking p into another point in E_N. To simplify the notation we shall assume throughout that $M = 1$. It will be quickly seen that this is no essential restriction.

Our first result is

THEOREM 2.1. *Consider the equation*

$$f(p) = \operatorname*{Max}_{1 \le k \le n} [g_k(p) + h_k(p) f(T_kp)], \tag{2.1}$$

where we assume that

(a) *The point p is restricted to a region R with the property that $p \in R$ implies that $T_kp \in R$,*

$$\text{(b)} \qquad 0 \le g_k(p) \le c_1 \qquad \textit{for } p \in R,$$
$$\text{(c)} \qquad 0 \le h_k(p) \le c_2 < 1 \qquad \textit{for } p \in R. \tag{2.2}$$

Under these conditions there is a unique bounded solution to (2.1).

As we shall see below, conditions (2.2b) and (2.2c) could be replaced by $|g_k| \le c_1$, $|h_k| \le c_2 < 1$ without affecting the validity of the final result. In most applications, however, (2.2b) and (2.2c) will be realized, since h_k represents a probability and g_k an expected gain. Our first application of successive approximations will rely heavily upon (2.2b) and (2.2c).

There are several ways of applying the method of successive approximations that are distinct not only analytically, but also conceptually.

The first takes its origin in the viewpoint that an infinite process is only sensibly defined as a limit of a finite process. We consider, then, that at first we are allowed only n stages. If we define

$$f_n(p) = \text{return obtained using an optimal policy when at most } n \text{ stages are allowed,} \tag{2.3}$$

we obtain the recurrence relations

$$f_0(p) = \operatorname*{Max}_k \, [g_k(p)],$$
$$f_{n+1}(p) = \operatorname*{Max}_k \, [g_k(p) + h_k(p) f_n(T_k p)], \qquad n = 0, 1, 2, \cdots. \tag{2.4}$$

Let us show that the sequence $\{f_n(p)\}$ converges to a solution of (2.1). Since $g_k, h_k \ge 0$, it is clear that $f_1(p) \ge f_0(p)$ for all p in R. From this it follows inductively that $f_{n+1} \ge f_n \ge \cdots f_1 \ge f_0 \ge 0$. If we set $u_n = \operatorname{Sup}_R f_n(p)$, we obtain from (2.4), using (2.2a),

$$u_{n+1} \le c_1 + c_2 u_n, \tag{2.5}$$

which shows that $u_n \le c_1/(1 - c_2)$, $n = 0, 1, 2, \cdots$.

It follows, then, that for each $p \in R$, the sequence $\{f_n(p)\}$ converges to a function that we call $f(p)$. It remains to demonstrate that $f(p)$ is actually a solution of (2.1).

Turning to (2.4), we see that the monotone character of f_n yields

$$f_{n+1}(p) \le \operatorname*{Max}_k \, [g_k(p) + h_k(p) f(T_k p)], \tag{2.6}$$

whence

$$f(p) \le \operatorname*{Max}_k \, [g_k(p) + h_k(p) f(T_k p)]. \tag{2.7}$$

Similarly, from (2.4),

$$f(p) \ge \operatorname*{Max}_k \, [g_k(p) + h_k(p) f_n(T_k p)], \tag{2.8}$$

whence

$$f(p) \geq \operatorname*{Max}_k \left[g_k(p) + h_k(p) f(T_k p)\right]. \tag{2.9}$$

Comparing (2.7) and (2.9), we see that we must have equality.

Let us now demonstrate the uniqueness of this solution, $f(p)$. Let $F(p)$ be another bounded solution satisfying (2.1):

$$F(p) = \operatorname*{Max}_k \left[g_k(p) + h_k(p) F(T_k p)\right]. \tag{2.10}$$

Let $k = k(p)$ be an index which yields the maximum of $g_k(p) + h_k(p) f(T_k p)$, and let $m = m(p)$ be a corresponding index for $g_k(p) + h_k(p) F(T_k p)$. Then by virtue of the maximum property we have

$$\begin{aligned} f(p) &= g_k(p) + h_k(p) f(T_k p) \geq g_m(p) + h_m(p) f(T_m p), \\ F(p) &= g_m(p) + h_m(p) F(T_m p) \geq g_k(p) + h_k(p) F(T_k p). \end{aligned} \tag{2.11}$$

Hence,

$$\begin{aligned} f(p) - F(p) &\geq h_m(p)\left[f(T_m p) - F(T_m p)\right] \\ &\leq h_k(p)\left[f(T_k p) - F(T_k p)\right], \end{aligned} \tag{2.12}$$

which yields the result

$$|f(p) - F(p)| \leq \operatorname{Max} \begin{Bmatrix} h_m(p)\,|f(T_m p) - F(T_m p)| \\ h_k(p)\,|f(T_k p) - F(T_k p)| \end{Bmatrix}. \tag{2.13}$$

If we define

$$S = \operatorname*{Sup}_R |f(p) - F(p)|, \tag{2.14}$$

we have for a p for which $|f(p) - F(p)| \geq S - \epsilon$, ϵ small, from (2.13),

$$S - \epsilon \leq c_2 S. \tag{2.15}$$

Since $c_2 < 1$, this leads to a contradiction for ϵ sufficiently small, unless $S = 0$. This completes the proof of the uniqueness.

The second application of successive approximations proceeds upon the basis that the physical origin of the equation is of no interest. We choose, consequently, an arbitrary non-negative function, $f_0(p)$, uniformly bounded over R, as our first approximation.

The recurrence relation is now

$$f_{n+1}(p) = \operatorname*{Max}_k \left[g_k(p) + h_k(p) f_n(T_k p)\right], \qquad n = 0, 1, 2, \cdots. \tag{2.16}$$

To show that the sequence $\{f_n(p)\}$ converges, we use the device of (2.11), above. Let $k = k(p)$ denote an index that furnishes the maximum for f_{n-1}, and let $m = m(p)$ denote a corresponding index for f_n. Proceeding as in (2.12), we obtain, for $n \geq 1$,

$$|f_{n+1}(p) - f_n(p)| \leq \operatorname{Max} \begin{Bmatrix} c_2\,|f_n(T_m p) - f_{n-1}(T_m p)| \\ c_2\,|f_n(T_k p) - f_{n-1}(T_{k-1} p)| \end{Bmatrix}. \tag{2.17}$$

If we define

$$u_n = \sup_R |f_n(p) - f_{n-1}(p)|, \qquad n = 1, 2, \cdots, \tag{2.18}$$

we obtain, from (2.17), $u_{n+1} \leq c_2 u_n$. Hence, if $c_3 = \text{Sup}_R f_0(p)$, we obtain finally the inequality $u_{n+1} \leq c_2^{n+1} c_3$. This shows that the series

$$\sum_{n=0}^{\infty} (f_{n+1} - f_n) \tag{2.19}$$

converges uniformly in R, and thus that the sequence $\{f_n\}$ converges uniformly in R to a function $f(p)$.

It follows from this that $f(p)$ will be continuous if each $f_n(p)$ is a continuous function of p. This will be true if $g_k(p)$ and $h_k(p)$ are continuous in p, and if $f_0(p)$ is chosen to be continuous.

We have thus demonstrated

THEOREM 2.2. *Under the conditions*

(a) $g_k(p)$ *is a continuous function of p in R,*

(b) $h_k(p)$ *is a continuous function of p in R, and* $|h_k(p)| \leq c_2 < 1$, (2.20)

the solution of (2.1) *is a continuous function of p.*

Furthermore, if $g_k(p)$ *and* $h_k(p)$ *are continuous functions of a set of parameters, q,* $f(p)$ *will be a continuous function of these parameters.*

In Section 2.7, below, devoted to approximate and computational techniques, we shall show that a combination of the above two ideas can be used in many cases to furnish quite useful initial approximations.

2.3. The Equation $f(x) = \text{Max}\,\{a(x_1, x_2, \cdots, x_n) + f[b(x_1, x_2, \cdots, x_n)]\}$

The existence and uniqueness theorem in the previous section does not apply to the equation

$$f(x) = \max_{0 \leq y \leq x} \{g(y) + h(x-y) + f[ay + b(x-y)]\}, \tag{2.21}$$

encountered in Section 1.6 in connection with an investment problem. To remedy this, we shall prove

THEOREM 2.3. *Consider the equation*

$$f(x) = \max_R \{a(x_1, x_2, \cdots, x_n) + f[b(x_1, x_2, \cdots, x_n)]\}, \tag{2.22}$$

where $R = R(x)$ *is defined by* $x_k \geq 0$, $\sum_{k=1}^n x_k = x$.

If

(a) $a(x_1, x_2, \cdots, x_n)$ *is continuous over* $R(x)$ *for* $0 \leq x \leq x_0$ *and non-negative,* $a(0, 0, \cdots, 0) = 0$,

(b) *$b(x_1, x_2, \cdots, x_n)$ is continuous and non-negative over R, and*

$$b(x_1, x_2, \cdots, x_n) \leq c \sum_{1}^{n} x_k, \qquad 0 < c < 1,$$

in $R(x_0)$,

(c) $\sum_{l=0}^{\infty} b(c^l x_0) < \infty$, *where*

$$b(x) = \max_{0<y<x} [\max_{R(y)} a(x_1, x_2, \cdots, x_n)], \tag{2.23}$$

there is a unique continuous solution to (2.22) *for which* $f(0) = 0$, *for* $0 \leq x \leq x_0$.

PROOF. Let $f_0(x)$ be the value obtained by choosing $x_1 = x$, $x_2 = x_3 = \cdots x_n = 0$ repeatedly. Then

$$f_0(x) = a(x) + a[b(x)] + \cdots, \tag{2.24}$$

where we have set $a(x) = a(x, 0, \cdots, 0)$, $b(x) = b(x, 0, \cdots, 0)$. The series on the right is majorized by

$$\sum_{l=0}^{\infty} b(c^l x_0),$$

and hence converges uniformly.

Define

$$f_{n+1} = \max_{R(x)} \{a(x_1, x_2, \cdots, x_N) + f_n[b(x_1, x_2, \cdots, x_N)]\}. \tag{2.25}$$

From the definition of f_0, it follows that $f_1 \geq f_0$, and hence that $f_{n+1} \geq f_n$. Let $M_n(x) = \max_{0 \leq y \leq x} f_n(y)$. Then, from (2.25),

$$M_{n+1}(x) \leq b(x) + M_n(cx), \tag{2.26}$$

whence $M_n(x) \leq \sum_{l=0}^{\infty} b(c^l x)$. Therefore, $f_n(x)$ converges to a function $f(x)$ for all x in $[0, x_0]$, which, as above, is readily seen to be a solution.

The technique utilized above is readily adapted to show the uniqueness of a continuous solution, f. If g is another continuous solution, we obtain, for a pair of points $(x_1, x_2, \cdots, x_n)$, $(y_1, y_2, \cdots, y_n)$ which yield the respective maxima,

$$\begin{aligned} f(x) &= a(x_1, x_2, \cdots, x_N) + f[b(x_1, x_2, \cdots, x_N)] \\ &\geq a(y_1, y_2, \cdots, y_N) + f[b(y_1, y_2, \cdots, y_N)], \\ g(x) &= a(y_1, y_2, \cdots, y_N) + g[b(y_1, y_2, \cdots, y_N)] \\ &\geq a(x_1, x_2, \cdots, x_N) + g[b(x_1, x_2, \cdots, x_N)], \end{aligned} \tag{2.27}$$

whence

$$\begin{aligned} |f(x) - g(x)| \leq \max \{&|f[b(x_1, x_2, \cdots, x_N)] - g[b(x_1, x_2, \cdots, x_N)]|, \\ &|f[b(y_1, y_2, \cdots, y_N)] - g[b(y_1, y_2, \cdots, y_N)]|\}. \end{aligned} \tag{2.28}$$

Let

$$M(x) = \operatorname*{Max}_{0 \le y \le x} |f(y) - g(y)|. \tag{2.29}$$

Then, from (2.28),

$$M(x) \le M(cx) \le M(c^n x), \qquad n = 1, 2, \cdots. \tag{2.30}$$

Since $M(x)$ is continuous and $M(0) = 0$, we obtain $M(x) \le 0$, which means that $M(x)$ is identically zero. This completes the proof.

Let us note that all we have actually proved at this point is that there is a unique bounded solution to

$$\begin{aligned} f(x) &= \operatorname*{Sup}_{R} \{a(x_1, x_2, \cdots, x_n) + f[b(x_1, x_2, \cdots, x_n)]\}, \\ f(0) &= 0, \end{aligned} \tag{2.31}$$

since, at the moment, we have no assurance that the maximum is assumed. The simplest way to ensure that the maximum is assumed is to prove that f itself is continuous. This fact and the corresponding results concerning continuity as a function of parameters may be readily derived by using the modification of the method in Section 2.2, Eqs. (2.17) through (2.19), given above.

2.4. The Equation $f(p) = \text{Min}\,[1 + \sum_{k=0}^{n} p_k f(x_k), 1 + f(T_l p)]$

Let us now turn our attention to the more complicated functional equation involved in the testing problem discussed in Section 1.8.3:

$$\begin{aligned} f(p) &= \text{Min}\left\{1 + \sum_{k=0}^{n} p_k f(x_k), \operatorname*{Min}_{l}\,[1 + f(T_l p)]\right\}, \\ f(x_0) &= 0, \end{aligned} \tag{2.32}$$

where l runs over the set $l = 1, 2, \cdots, M$. Here, we set

$$p = \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \end{bmatrix}, \qquad T_l p = \begin{bmatrix} p_{0l} \\ p_{1l} \\ \vdots \\ p_{nl} \end{bmatrix}, \qquad x_k = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{2.33}$$

where $p_{kl} = p_{kl}(p_0, p_1, \cdots, p_n)$, the "1" occurring in the kth place, and take $f(p)$ to be a scalar function of p. That $f(x_0) = 0$ is a consequence of the fact that x_0 is the desired state and that no action is taken when $p = x_0$.

We shall prove

THEOREM 2.4. *If for each transformation T_l it is true that*

$$\sum_{k=1}^{n} p_{kl} \le c_1 \sum_{k=1}^{n} p_k, \qquad 0 < c_1 < 1, \tag{2.34}$$

for all p such that $\sum_{k=0}^{n} p_k = 1$, $p_k \ge 0$, then there exists a unique bounded positive solution to Eq. (2.32).

PROOF. We shall once more employ the method of successive approximations. Consider the procedure S_1 represented by $LT_1LT_1\cdots$, which means that we look, then act, if p is not x_0; look, act; and so on, repeatedly. Similarly, define S_2 by $T_1LT_1L\cdots$. It is clear from simple considerations of the probability theory that the expected times $F_1(p)$, $F_2(p)$ required to transform p into x_0 with certainty are finite for both strategies.

To calculate $F_1(p) = f_{S_1}(p)$ and $F_2(p) = f_{S_2}(p)$, we employ the equations

$$\begin{aligned} F_1(p) &= 1 + F_2(T_1p), \qquad p \ne x_0, \\ F_2(p) &= 1 + \sum_{k=0}^{n} p_{kl}F_1(x_k), \qquad p \ne x_0. \end{aligned} \tag{2.35}$$

Hence,

$$F_1(x_l) = 2 + \sum_{k=1}^{n} p_{kl}F_1(x_k), \qquad l = 1, 2, \cdots, n. \tag{2.36}$$

Since $0 \le \sum_{k=1}^{n} p_{kl}(x_l) < c_1 < 1$, the determinant of the system does not vanish and this system has a unique solution, necessarily positive, as we see by solving iteratively. Having determined $F_1(x_k)$, we readily determine $F_1(p)$ and $F_2(p)$.

Now define

$$\begin{aligned} f_0(p) &= \operatorname{Min}\,[F_1(p), \quad F_2(p)], \\ f_{n+1}(p) &= \operatorname{Min}\begin{bmatrix} 1 + \sum_{k=1}^{n} p_{kl}f_n(x_k) \\ 1 + f_n(T_1p) \end{bmatrix}, \\ f_{n+1}(x_0) &= 0. \end{aligned} \tag{2.37}$$

Considering the related policies or the equations themselves, it is clear that $f_1(p) \le f_0(p)$, whence $f_0 \ge f_1 \ge f_2 \cdots \ge f_n \ge 1$ for $p \ne x_0$. Hence $f_n(p)$ converges monotonically to a function $f(p)$ which, as in the previous proofs, may be shown readily to satisfy the functional equation.

The uniqueness proof is rather more complicated. Let f and g be two positive, bounded solutions of (2.32). We prove first a lemma.

LEMMA 2.1. $\operatorname{Sup}_{p} |f(p) - g(p)| = \operatorname{Max}_{k} |f(x_k) - g(x_k)|$.

PROOF. The inequality

$$\operatorname{Max}_{k} |f(x_k) - g(x_k)| \le \operatorname{Sup}_{p} |f(p) - g(p)| \tag{2.38}$$

is clear. To demonstrate the reverse inequality, we consider four cases:

$$
\begin{aligned}
&\text{(a)} && f(p) = 1 + \sum_{k=0}^{n} p_k f(x_k), \\
& && g(p) = 1 + \sum_{k=0}^{n} p_k g(x_k), \\
&\text{(b)} && f(p) = 1 + \sum_{k=0}^{n} p_k f(x_k), \\
& && g(p) = 1 + g(T_1 p), \\
&\text{(c)} && f(p) = 1 + f(T_1 p), \\
& && g(p) = 1 + \sum_{k=1}^{n} p_k g(x_k), \\
&\text{(d)} && f(p) = 1 + f(T_1 p), \\
& && g(p) = 1 + g(T_1 \cdot p).
\end{aligned} \tag{2.39}
$$

Consider first the case corresponding to (a). We have

$$f(p) - g(p) = \sum_{k=0}^{n} p_k[f(x_k) - g(x_k)], \tag{2.40}$$

whence

$$|f(p) - g(p)| \leq \operatorname*{Max}_k |f(x_k) - g(x_k)|. \tag{2.41}$$

Therefore, for all p for which (a) holds, the assertion of the lemma is correct. The equations of (2.39a) will hold whenever p is close enough to x_0, since $f(p), g(p) \geq 1$ for all $p \neq x_0$, whereas $1 + \sum_{k=1}^{n} p_k f(x_k)$ and $1 + \sum_{k=1}^{n} p_k g(x_k)$ are close to 1 for p close to x_0. Thus, $1 + f(T_1 p)$ and $1 + g(T_1 p)$ will be dominated by the observation moves for p close to x_0.

This is an important point, since the crux of our proof is the fact that (2.39a) will always occur after a finite number of moves, by virtue of Eq. (2.34).

Now consider Eq. (2.39b). We have

$$
\begin{aligned}
f(p) &= 1 + \sum_{k=0}^{n} p_k f(x_k) \geq 1 + f(T_1 p), \\
g(p) &= 1 + g(T_1 p) \geq 1 + \sum_{k=0}^{n} p_k g(x_k).
\end{aligned} \tag{2.42}
$$

Hence,

$$
\begin{aligned}
|f(p) - g(p)| \leq \operatorname{Max} [&\operatorname*{Max}_k |f(x_k) - g(x_k)|, \\
&\operatorname*{Sup}_p |f(T_1 p) - g(T_1 p)|],
\end{aligned} \tag{2.43}
$$

and similarly for (2.39c). For (2.39d) we obtain

$$\sup_p |f(p) - g(p)| \le \operatorname{Max}\,[\sup |f(T_l p) - g(T_l p)|,\\ \sup |f(T_{l'} p) - g(T_{l'} p)|]. \tag{2.44}$$

We now iterate these inequalities. For any fixed p, $T_{l_1}, T_{l_2}, \cdots, T_{l_r} p$ will be in the region governed by (2.39a) for r large enough. Consequently, we obtain

$$\sup_p |f(p) - g(p)| \le \operatorname{Max}_k |f(x_k) - g(x_k)|. \tag{2.45}$$

This completes the proof of the lemma. It remains to show that $\operatorname{Max}_k |f(x_k) - g(x_k)| = 0$. Let K be an index at which the maximum is assumed. It follows from the functional equation for f and g that

$$\begin{aligned} f(x_K) &= 1 + f(T_l x_K), \qquad l = l(K),\\ g(x_K) &= 1 + g(T_{l'} x_K), \qquad l' = l'(K), \end{aligned} \tag{2.46}$$

and that

$$\begin{aligned} f(x_K) &= 1 + f(T_l x_K) \ge 1 + f(T_{l'} x_K),\\ g(x_K) &= 1 + g(T_{l'} x_K) \ge 1 + g(T_l x_K). \end{aligned} \tag{2.47}$$

If both inequalities are proper, we obtain

$$|f(x_K) - g(x_K)| < \operatorname{Max}\,[\,|f(T_l x_K) - g(T_l x_K)|,\\ |f(T_{l'} x_K) - g(T_{l'} x_K)|\,] \le \sup_p |f(p) - g(p)|, \tag{2.48}$$

which contradicts (2.45). Thus, for l or l' we have

$$\begin{aligned} f(x_K) &= 1 + f(T_l x_K),\\ g(x_K) &= 1 + g(T_l x_K). \end{aligned} \tag{2.49}$$

This means that the first moves can be the same.

Consider now the situation for second moves. Using the same argument, we see that the second moves, i.e., the equations for $f(T_l x_K)$, $g(T_l x_K)$, can be the same, and so on, by induction.

Let $p_n = p_n(x_K)$ be the distribution achieved after n moves, where the $(n+1)$th move puts x_K into the region governed by (2.39a). The same argument as that above shows that both f and g may be put into this situation at the same move. Then

$$\begin{aligned} f(x_K) &= (n+1) + \sum_{k=0}^{n} p_{kn} f(x_k),\\ g(x_K) &= (n+1) + \sum_{k=0}^{n} p_{kn} g(x_k). \end{aligned} \tag{2.50}$$

Therefore,

$$f(x_K) - g(x_K) = \sum_{k=0}^{n} p_{kn}[f(x_k) - g(x_k)]. \tag{2.51}$$

By assumption, $p_{0n} > 0$. Therefore,

$$|f(x_K) - g(x_K)| \leq (1 - p_{0n})\,|f(x_K) - g(x_K)| \tag{2.52}$$

implies that $|f(x_K) - g(x_K)| = 0$. This implies that $f(p) - g(p) \equiv 0$, and completes our proof.

2.5. The Optimal Inventory Equation

Let us now treat the equation that occurs in connection with the optimal inventory problem discussed in Section 1.8.6,

$$u(x) = \operatorname*{Inf}_{y \geq x} \Big[g(y - x) + a\{u(0)[1 - F(y)] + \int_y^\infty \Psi(v, y)\, dF(v) + \int_0^y u(y - v)\, dF(v)\}\Big], \tag{2.53}$$

where we shall assume that

(a) $g(0) = 0, \quad g(y) \geq 0 \quad$ for $y \geq 0$,

(b) $\int_0^\infty dF = 1, \quad dF \geq 0$,

(c) $0 < a < 1$,

(d) $0 \leq \int_y^\infty \Psi(v, y)\, dF(v) \leq c_1 < \infty \quad$ for all $y > 0$. (2.54)

Under these assumptions we shall prove

THEOREM 2.5. *There is a unique uniformly bounded solution to* (2.53).*

PROOF. We shall, as before, employ the method of successive approximations. Let

$$u_0(x) = a\Big\{u_0(0)[1 - F(x)] + \int_x^\infty \Psi(v, x)\, dF(v) + \int_0^x u_0(x - v)\, dF(v)\Big\},$$

$$u_{n+1}(x) = \operatorname*{Inf}_{y \geq x} \Big[g(y - x) + a\Big\{u_n(0)[1 - F(y)] + \int_y^\infty \Psi(v, y)\, dF(v) + \int_0^y u_n(y - v)\, dF(v)\Big\}\Big]. \tag{2.55}$$

* This result is contained, along with a multitude of other results, in the paper of Dvoretzky, Kiefer, and Wolfowitz referred to in the Preface. The method of proof given here is, however, different from theirs.

Setting $x = 0$, we obtain

$$u_0(0) = \frac{a\left[\int_0^\infty \Psi(v, 0)\, dF(v)\right]}{(1 - a)}. \tag{2.56}$$

Substituting in (2.55), we obtain an equation, the "renewal" equation, for $u_0(x)$, which may be readily solved by Laplace Transform techniques or by simple iteration, since $0 < a < 1$, and $\int_0^\infty dF = 1$, $dF \geq 0$.

Referring to (2.55) again, we see that

$$u_1(x) = \operatorname*{Inf}_{y \geq x} [g(y - x) + u_0(y)]. \tag{2.57}$$

Since

$$\operatorname*{Inf}_{y \geq x} [g(y - x) + u_0(y)] \leq g(0) + u_0(x) = u_0(x), \tag{2.58}$$

we obtain the important result that $u_1(x) \leq u_0(x)$. From this it is immediately clear, using the recurrence relation in (2.55), that $u_2 \leq u_1$, and that, generally,

$$u_0 \geq u_1 \geq \cdots \geq u_n \cdots > 0. \tag{2.59}$$

It follows that for all $x \geq 0$, the sequence $\{u_n(x)\}$ converges to a uniformly bounded function $u(x)$. Using (2.55) again, it is clear that $u(x)$ satisfies (2.53).

It is not difficult to use the methods of the previous sections to show that the convergence is actually geometric, i.e., $|u(x) - u_n(x)| \leq c_2 a^n$, for some $c_2 > 0$.

To establish uniqueness we proceed as before. Let us assume that there are two solutions, u and w, both bounded in any finite interval $[0, x]$. Let $y = y(x)$ and $z = z(x)$ be two decision functions that yield values of $u(x)$ and $w(x)$, respectively, within ϵ_1 and ϵ_2 of the actual infima, where ϵ_1 and ϵ_2 are small positive quantities.

Then we have

$$\begin{aligned} u(x) &= T(u, y) + \epsilon_1 \leq T(u, z) + \epsilon_3, \\ w(x) &= T(w, z) + \epsilon_2 \leq T(w, y) + \epsilon_4, \end{aligned} \tag{2.60}$$

where ϵ_3 and ϵ_4 are again small quantities, and $T(u, y)$ is an abbreviation for

$$\begin{aligned} g(y - x) + a\Big\{u(0)[1 - F(y)] &+ \int_y^\infty \Psi(v, y)\, dF(v) \\ &+ \int_0^y u(y - v)\, dF(v)\Big\}. \end{aligned} \tag{2.61}$$

We have then

$$\begin{aligned} u(x) - w(x) &\leq \epsilon_3 - \epsilon_2 + T(u, z) - T(w, z) \\ &\geq \epsilon_1 - \epsilon_4 + T(u, y) - T(w, y). \end{aligned} \tag{2.62}$$

Since

$$T(u, z) - T(w, z) = a \int_0^z [u(z - v) - w(z - v)] \, dF(v) + a[u(0) - w(0)][1 - F(z)], \tag{2.63}$$

the inequalities in (2.61) yield

$$\begin{aligned} |u(x) - w(x)| \leq \text{Max} \Big\{ & \epsilon_5 + a \int_0^z |u(z - v) - w(z - v)| \, dF(v) \\ & + a|u(0) - w(0)|[1 - F(z)], \\ & \epsilon_5 + a \int_0^y |u(y - v) - w(y - v)| \, dF(v) \\ & + a|u(0) - w(0)|[1 - F(y)] \Big\}. \end{aligned} \tag{2.64}$$

Let x be chosen to be a point at which $|u(x) - w(x)|$ is within ϵ_6 of its supremum, d. Then, since

$$\begin{aligned} & \int_0^z |u(z - v) - w(z - v)| \, dF(v) + |u(0) - w(0)|[1 - F(z)] \\ & \leq d \int_0^z dF(v) + d[1 - F(z)] = d, \end{aligned} \tag{2.65}$$

we obtain, from (2.63),

$$d - \epsilon_6 \leq \epsilon_5 + ad, \tag{2.66}$$

which yields $d(1 - a) \leq \epsilon_5 + \epsilon_6$. Since $1 - a > 0$ and ϵ_5 and ϵ_6 may be chosen arbitrarily small, d must necessarily be zero.

2.6. Definition of a Solution

We have shown that the functional equation in (2.1) possesses a unique bounded solution. It is clear that this function defines a strategy S, since the first choice will be A_k, where k is the index that maximizes $g_k(p) + h_k(p)f(T_kp)$, the second choice being similarly determined by the expression for $f(T_kp)$, and so on.

The question arises as to whether or not this is actually an optimal strategy, and, if so, as to whether or not it is unique. That it is an optimal strategy we see by the following argument: Suppose that we have another prescription S_a for determining an optimal yield. This prescription defines a function $\phi(p)$ that must satisfy the same functional equation, (2.1), because if it did not, it would not possess the necessary optimal continuation policy. Hence, $\phi(p) = f(p)$. Since S yields $f(p)$, we see that no other policy S_a is preferable.

It is not necessarily true that S is unique. This arises from the fact that for various p's several choices may be equivalent, although the continuations from equivalent choices will be quite different. We shall subsequently meet a very simple example of this. We observe, however, that the functional equation permits us to obtain all optimal strategies.

Another question that arises is one as to the precise a priori definition of $f(p)$. The definitions given in the Introduction are loose, since it is not clear that the required optimal procedures exist.

There are several alternative procedures, corresponding to the techniques of successive approximation that we employed.

We may define, by ukase, $f(p)$ to be the solution of our functional equation, using the existence and uniqueness of the solution as a *de facto* justification. Or we may define $f_N(p)$, unambiguously, as the maximum return when only N stages are allowed, and set

$$f(p) = \lim_{N \to \infty} f_N(p) \tag{2.67}$$

whenever this limit exists. Or we may ambitiously consider the space of all sequences of decisions $S = A_1^{a_1} A_2^{a_2} \cdots$, remembering that the exponents, a_i, are, in general, random variables depending on the pattern of events and not merely fixed in advance, and define

$$f(p) = \operatorname*{Max}_S f_S(p), \tag{2.68}$$

when it exists. In general it will be clear that $\operatorname{Sup}_S f_S(p)$ exists, and an essential part of the problem will be to show that the maximum is actually attained. That the maximum is actually attained may be demonstrated by use of the functional equations or abstract topological techniques, which we shall not present here.

From the mathematical standpoint it would seem that (2.67) is a preferable definition, since it furnishes a stronger hold on $f(p)$ than does (2.68). However, since the two definitions lead to the same function, it is actually convenience that decides which to treat as fundamental in any particular problem.

2.7. Approximate and Computational Methods

At this point it must be confessed that in the theory of dynamic programming, as in most other theories treating of the physical world, the majority of the functional equations that arise will be resolutely, if impartially, insoluble by analytic means, as far as explicit solutions are concerned. Consequently, the theoretical and practical development of the theory requires that efficient and readily applicable approximate and computational methods be developed.

In theory there is only one method that may be used to approximate the solution of a functional equation, namely, the solution of an approximate functional equation. In practice the variants of this technique differ greatly.

Let us write our functional equation in the form

$$f = T(f, P), \tag{2.69}$$

where f represents the unknown function, T is the transformation induced upon f by the physical process, and P is a quantity representing various parameters that occur, constants and functions.

The method of successive approximations in its usual guise relies on solving the following approximate equation

$$f_{n+1} = T(f_n, P), \tag{2.70}$$

where f_0 is a suitable first guess at the solution. In more refined applications, (2.70) is replaced by

$$f_{n+1} - R(f_{n+1}) = T(f_n, P) - R(f_n), \tag{2.71}$$

where $R(f)$ is a transformation so chosen as to force f_n to possess certain desired properties or so chosen as to increase the rapidity of convergence.

In place of the above approach, we may consider the equation

$$f = T(f, P'), \tag{2.72}$$

where P' represents a modified set of parameters so chosen that the solution of (2.72) may be obtained in a simple way. Thus, for example, in the theory of differential equations, the treatment of non-linear and linear equations with variable coefficients depends to an enormous extent on the fortunate circumstance that linear equations with constant coefficients are explicitly solvable in terms of exponentials. Similarly, in the consideration of the functional equations occurring in the theory of dynamic programming, any results that may be obtained under the simplifying hypotheses of linearity, convexity, and so on are extremely important, insofar as they furnish guides to the behavior of the actual solutions. The justification, from the larger point of view, of searching for complete solutions of simplified equations lies precisely in the hope and expectation of using these special solutions as approximations—not only quantitatively, but also qualitatively—to the solutions of the more realistic and complicated equations.

The functional equations that we treat of afford yet a third approach, which arises from the duality between function space and strategy space. In place of the original class of transformations, we may consider a subclass obtained by restricting the permissible choices. Thus, for example, in place of infinite processes, we may consider finite processes; in place of three-choice processes, we may consider two-choice processes.

Employing the optimal policies for the simplified model, we obtain approximate policies for the larger model. A computation will then yield an approximate solution to the functional equation. It is clear mathematically and intuitively that if the method of successive approximations is now employed, the convergence will be monotone, an important fact from the computational standpoint.

The essential idea behind the preceding method is that we obtain suitable first approximations most readily by approximating in the strategy space rather than in the function space. It is in this fashion that the physical process generating the functional equation can best be exploited, and that experience and intuition gained by solving simpler problems can be most efficiently utilized.

2.8. A Geometric Technique

Let us now describe an interesting approach particularly applicable to a certain class of problems of deterministic type, in particular to equations of the form

$$f(x, y) = \underset{1 \leq i \leq M}{\text{Max}} \, [A_i \colon P_i(x, y) + c_i f(r_i x, s_i y)], \tag{2.73}$$

$0 < r_i, s_i, c_i < 1$, where $P_i(x, y)$ is a homogeneous polynomial in x and y. Since the solution will be homogeneous in x and y, it is sufficient to consider only values of x and y for which $x + y = 1$.

For any given x and y we may write

$$\bar{x} = \frac{x}{x + y}, \qquad \bar{y} = \frac{y}{x + y}, \qquad f(\bar{x}, \bar{y}) = \frac{f(x, y)}{(x + y)^k}. \tag{2.74}$$

Any strategy in (2.73) has the form

$$S = A_1^{a_1} A_2^{a_2} \cdots. \tag{2.75}$$

Employing this strategy, we obtain

$$f_S(x, y) = P_1(x, y) + \cdots, \tag{2.76}$$

a homogeneous function of x and y. This function $f_S(x, y)$ may now be regarded as a function of one variable, $\bar{x}$, $0 \leq \bar{x} \leq 1$, $f_S(x, y) = f_S(\bar{x})$. To each strategy, S, in consequence, corresponds a continuous curve $f_S(\bar{x})$.

If from all these curves we now form the envelope, above, we obtain a new curve,

$$E(\bar{x}) = \underset{S}{\overline{\text{Env}}} \, f_S(\bar{x}), \tag{2.77}$$

which must necessarily be $f(\bar{x}, \bar{y}) = f(\bar{x})$.

Although in general the envelope will be difficult to obtain explicitly, various qualitative features of the solution may often be obtained readily. An example of the application of this technique will be given in Section 3.12.

In the important case in which the $P_i(x, y)$ are linear functions of x and y, the envelope curve is convex, a result of great utility in the application of the method.

CHAPTER 3

THE GOLD-MINING EQUATION

3.1. Introduction

In this chapter we turn to a more detailed study of the "gold-mining" equation, beginning with the simplest representative,

$$f(x, y) = \text{Max}\begin{Bmatrix} A: & p_1[x + f(0, y)] + p_2[c_1x + f(c_2x, y)] \\ B: & q_1[y + f(x, 0)] + q_2[d_1y + f(x, d_2y)] \end{Bmatrix}, \tag{3.1}$$

where $x, y \geq 0$, and the constants that appear are subject to the following conditions:

$$\text{(a)} \qquad 0 < p_1, p_2, q_1, q_2 < 1, \qquad p_1 + p_2 < 1, \qquad q_1 + q_2 < 1,$$

and

$$\text{(b)} \qquad 0 < c_1, d_1 < 1, \qquad c_1 + c_2 = 1, \qquad d_1 + d_2 = 1. \tag{3.2}$$

The origin of this equation was discussed in Problem 1.5, page 2, and the required existence and uniqueness theorems were given in Chapter 2.

We shall begin by presenting a solution to (3.1) and also some generalizations. We shall then consider the equation

$$f(x, y, a) = \text{Max}\begin{bmatrix} A: & p_1f(0, y, a + x) + p_2f(c_2x, y, a + c_1x) + p_3\phi(a) \\ B: & q_1f(x, 0, a + y) + q_2f(x, d_2y, a + d_1y) + q_3\phi(a) \end{bmatrix}, \tag{3.3}$$

$x, y, a \geq 0$, with $f(0, 0, a) = \phi(a)$, which arises when we use as a criterion function $\phi(z)$ in place of z, where z is the total yield.

This equation may be solved explicitly in the case in which $\phi(z) = z$, as above, and in the case in which $\phi(z) = e^{bz}$. The asymptotic form of the solution for large x and y will be given in this latter case.

After this we shall discuss briefly some extensions of (3.1) that are at present obdurately resisting analytic solution.

Turning from this analytic treatment, we shall then present an interesting geometric treatment of (3.1), using the ideas of Section 2.8.

3.2. The Solution of Equation (3.1)

The purpose of this section is to provide an introduction to the analytic techniques we shall employ throughout the chapter and to demonstrate

THEOREM 3.1. *Consider the functional equation*

$$f(x, y) = \text{Max} \begin{bmatrix} A: \sum_{k=1}^{N} p_k[c_k x + f(c_k' x, y)] \\ B: \sum_{k=1}^{N} q_k[d_k y + f(x, d_k' y)] \end{bmatrix}, \tag{3.4}$$

where

$$\begin{aligned} &\text{(a)} \qquad p_k \geq 0, \quad q_k \geq 0, \quad \sum_{k=1}^{N} p_k, \sum_{k=1}^{N} q_k < 1, \\ &\text{(b)} \qquad 1 \geq c_k, \quad d_k \geq 0, \quad c_k' + c_k = d_k' + d_k = 1, \\ &\text{(c)} \qquad x, y \geq 0. \end{aligned} \tag{3.5}$$

The optimal choice of operations is the following: If

$$\frac{\sum_{k=1}^{N} p_k c_k}{1 - \sum_{k=1}^{N} p_k} x > \frac{\sum_{k=1}^{N} q_k d_k}{1 - \sum_{k=1}^{N} q_k} y, \tag{3.6}$$

choose A; if the reverse inequality holds, choose B. In case of equality, either choice is satisfactory.

To simplify the notation and the algebra, let us consider first the simpler form of (3.4) given by (3.1). As noted above, we already know from Chapter 2 that there is a unique solution to this equation. Let us turn, then, to a discussion of some of the simpler properties of $f(x, y)$. Since $p_1 + p_2 < 1$, $q_1 + q_2 < 1$, it follows that $f(0, 0) = 0$. From the fact that $f(kx, ky)$ and $kf(x, y)$ satisfy the same equation for $k \geq 0$, it follows that $f(kx, ky) = kf(x, y)$, for $k \geq 0$. Setting $y = 0$ and using $f(c_2 x, 0) = c_2 f(x, 0)$, we obtain

$$\begin{aligned} f(x, 0) &= \text{Max} \begin{bmatrix} A: & (p_1 + p_2 c_1)x + p_2 c_2 f(x, 0) \\ B: & (q_1 + q_2) f(x, 0) \end{bmatrix} \\ &= (p_1 + p_2 c_1)x + p_2 c_2 f(x, 0), \end{aligned} \tag{3.7}$$

whence

$$f(x, 0) = \frac{(p_1 + p_2 c_1)x}{(1 - p_2 c_2)}, \tag{3.8}$$

and, similarly,

$$f(0, y) = \frac{(q_1 + q_2 d_1)y}{(1 - q_2 d_2)}. \tag{3.9}$$

These results are, of course, obvious if we consider the process generating the function. On these grounds we should also suspect that A would be employed whenever y was sufficiently small compared with x. This fact follows from the continuity of

$f(x, y)$ (compare Section 2.2), since the inequality

$$f(x, y) > (q_1 + q_2d_1)y + q_1f(x, 0) + q_2f(x, d_2y) \tag{3.10}$$

must hold for small positive $y = y(x)$, for $x > 0$, seeing that it is valid for $y = 0$.

It follows then that there are two regions, close to the x and y axes, in which the optimal choices are, respectively, A and B, whenever (x, y) is contained in either of these regions, as shown in Fig. 3.1.

It is reasonable to suppose that the solution has the form shown in Fig. 3.2. The meaning of Fig. 3.2 is that A is employed whenever (x, y) is in R_A, the region between the x-axis and L, and B is employed in the complementary region. On the line L either A or B may be used.

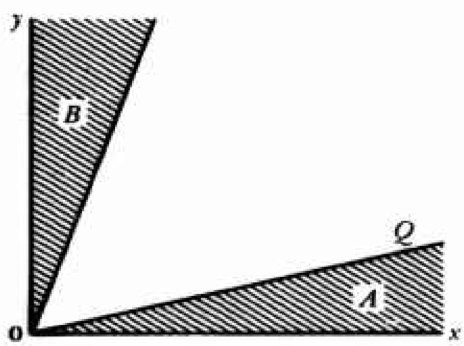

Fig. 3.1

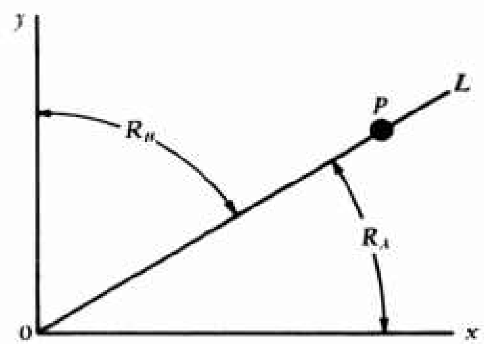

Fig. 3.2

That the boundary curve, if it exists, must be a straight line follows from the homogeneity of $f(x, y)$. Assuming that the solution has this form, we shall show that the equation of L may be calculated from the fact that it is an indifference curve. By this term we mean that for points (x, y) on the curve, the value of the function $f(x, y)$ is the same whether we employ A or B.

Observe that the effect of employing A is always to drive P into R_B, whereas the use of B sends P into R_A. Consequently, if A is used at P, the next choice, in an optimal policy, must be B, and vice versa if B is used.

This alone would not be sufficient to determine L, were it not for another fact. Since the operations A and B operate on x and y alone, there will be a certain symmetry in the results obtained by using A and then B, or B and then A, which plays a decisive role in the solution.

Let us now do a small amount of computing. Using the values of $f(x, 0)$ and $f(0, y)$ obtained above, we have

$$f(x, y) = \text{Max}\begin{bmatrix} A\colon\ (p_1 + p_2c_1)x + \dfrac{p_1(q_1 + q_2d_1)y}{1 - q_2d_2} + p_2f(c_2x, y) \\ B\colon\ (q_1 + q_2d_1)y + \dfrac{q_1(p_1 + p_2c_1)x}{1 - p_2c_2} + q_2f(x, d_2y) \end{bmatrix}. \tag{3.11}$$

To simplify the notation, let us denote the coefficients of x and y in the above equation by α_1, α_2 in A and by β_1, β_2 in B. If we employ A, we obtain, using an obvious notation,

$$f_A(x, y) = \alpha_1 x + \alpha_2 y + p_2 f(c_2 x, y). \tag{3.12}$$

Following this by B, we have

$$f_{AB}(x, y) = (\alpha_1 + \beta_1 p_2 c_2)x + (\alpha_2 + p_2\beta_2)y + p_2 q_2 f(c_2 x, d_2 y). \tag{3.13}$$

Similarly, the result of B and then A is

$$f_{BA}(x, y) = (\beta_1 + q_2\alpha_1)x + (\beta_2 + q_2\alpha_2 d_2)y + p_2 q_2 f(c_2 x, d_2 y). \tag{3.14}$$

If (x, y) lies upon L, we must have $f_{AB} = f_{BA}$. Equating the two expressions, we observe that the unknown function $f(c_2 x, d_2 y)$ disappears. Consequently, we obtain for L the equation

$$[\alpha_1(1 - q_2) + \beta_1(p_2 c_2 - 1)]x = [\alpha_2(q_2 d_2 - 1) + \beta_2(1 - p_2)]y. \tag{3.15}$$

Using the precise values of α_1, β_1, α_2, β_2 as given by (3.11), we finally obtain, as the equation of L,

$$\frac{(p_1 + p_2 c_1)x}{1 - p_1 - p_2} = \frac{(q_1 + q_2 d_1)y}{1 - q_1 - q_2}. \tag{3.16}$$

This is a remarkably simple equation, since, as we observe, the coefficient of x depends only on the A operation, while the coefficient of y depends only on the B operation. Furthermore, each coefficient admits of a very simple interpretation as the ratio of the expected yield of the operation to the probability of termination of the process.

Let us insert a word of warning: Although this elegant result holds for some generalizations of the functional equation, it does not hold in general, as we shall subsequently see.

Let us now prove that the solution actually has this simple form. To make the previous argument rigorous, we observe that below L, the procedure consisting of A, B, and an optimal continuation is superior to B, A, and an optimal continuation, and that the reverse is true above L. Referring to Fig. 3.1, let Q be a point above the known A-region and far enough below L so that any outcome of a B-choice transforms $Q(x, y)$ into the known A-region.

To show that A is used at Q, we argue by contradiction. Suppose that B were used; then the next choice would necessarily be A. However, we have seen, above, that below L, the procedure consisting of B, A, and an optimal continuation is inferior to A, B, and an optimal continuation. Hence, A is used at Q. It is clear that we may continue this argument until we have demonstrated that the region between L and the x-axis is an A-region. Similarly, starting from the known B-region, we may demonstrate that the region above L is a B-region.

We have carried through the proof for the simplest case of (3.4). There is no difficulty in verifying that the argument is general.

Geometrically, the pattern is as follows. When (x, y) is in R_A, A is employed until the resultant point is in R_B, at which time B is employed until the point is again in R_A, and so on.

3.3. A Generalization

There is no difficulty in extending the above analysis to the following n-dimensional equation

$$f(x_1, x_2, \cdots, x_n) = \operatorname*{Max}_i \left\{ \sum_{k=1}^{K} p_{ik}[c_{ik}x_i + f(x_1, x_2, \cdots, c'_{ik}x_i, \cdots, x_n)] \right\}, \tag{3.17}$$

where

$$\begin{aligned} &\text{(a)} \quad p_{ik} \geq 0, \quad \sum_{k=1}^{K} p_{ik} < 1, \quad i = 1, 2, \cdots, n, \\ &\text{(b)} \quad 1 \geq c_{ik} \geq 0, \quad c_{ik} + c'_{ik} = 1, \\ &\text{(c)} \quad x_i \geq 0. \end{aligned} \tag{3.18}$$

The decision functions are again the ratios of expected gain to probability of termination, namely,

$$D_i(x_i) = \frac{\sum_k p_{ik}c_{ik}}{1 - \sum_k p_{ik}} x_i. \tag{3.19}$$

If Max $D_i(x_i)$ is attained for $i = L$, then the Lth choice is made unless there is equality, in which case any one of the maximizing choices is optimal.

3.4. The Form of $f(x, y)$

Having obtained a very simple characterization of the optimal policy, let us now turn our attention to the function $f(x, y)$. In general, no simple analytic representation will exist. If, however, we consider Eq. (3.1), which we write again as

$$f(x, y) = \operatorname{Max} \begin{bmatrix} \alpha_1 x + \alpha_2 y + p_2 f(c_2 x, y) \\ \beta_1 x + \beta_2 y + q_2 f(x, d_2 y) \end{bmatrix}, \tag{3.20}$$

we shall show that if c_2 and d_2 are connected by a relation of the type $c_2^m = d_2^n$, m and n being positive integers, we shall obtain piecewise linear representations for $f(x, y)$.

It is sufficient, in order to illustrate the technique, to consider the simplest case, $c_2 = d_2$.

Let (x, y) be a point in the A-region. If A is applied, either (x, y) goes into $(0, y)$, in which case B is used continually thereafter, or it is transformed into $(c_2 x, y)$, which may be in either an A- or a B-region. Let L_1 be the line that is transformed into L when (x, y) goes into (c_2x, y), let L_2 be the line transformed into L_1, and so on. Similarly, let M_1 be the line transformed into L when (x, y) goes into (x, d_2y), and so on. In the sector LOL_1, A is used first, followed by B, as shown in Fig. 3.3.

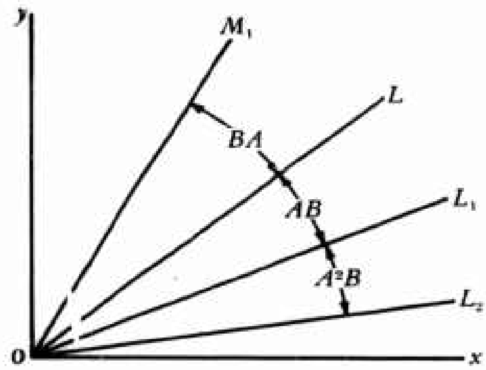

Fig. 3.3

Hence, for (x, y) in this sector we obtain

$$\begin{aligned} f(x, y) &= \alpha_1 x + \alpha_2 y + p_2 f(c_2x, y) \\ &= \alpha_1 x + \alpha_2 y + p_2(\beta_1 c_2 x + \beta_2 y) + p_2 q_2 f(c_2 x, c_2 y) \\ &= (\alpha_1 + p_2\beta_1 c_2)x + (\alpha_2 + p_2\beta_2)y + p_2q_2c_2 f(x, y). \end{aligned} \tag{3.21}$$

This yields

$$f(x, y) = \frac{(\alpha_1 + p_2\beta_1c_2)x + (\alpha_2 + p_2\beta_2)y}{1 - p_2q_2c_2} \tag{3.22}$$

for (x, y) in LOL_1. Similarly, we obtain a linear expression for f in LOM_1. Having obtained the representations in these sectors, it is clear that we obtain linear expressions in L_1OL_2, etc.

3.5. The Problem for a Finite Number of Stages

Let us now consider the problem that arises when only a finite number of stages are allowed. If we set

$$f_N(x, y) = \text{expected return using an optimal } N\text{-stage policy}, \tag{3.23}$$

then

$$\begin{aligned} f_1(x, y) &= \text{Max}\,[(p_1 + p_2c_1)x, (q_1 + q_2d_1)y], \\ f_{N+1}(x, y) &= \text{Max}\begin{Bmatrix} A\colon & p_1[x + f_N(0, y)] + p_2[c_1x + f_N(c_2x, y)] \\ B\colon & q_1[y + f_N(x, 0)] + q_2[d_1y + f_N(x, d_2y)] \end{Bmatrix}. \end{aligned} \tag{3.24}$$

We know from the results concerning existence and uniqueness in Section 2.2 that, as $N \to \infty$, $f_N(x, y) \to f(x, y)$. However, it is not reasonable to suspect that for each N the optimal policy will be that of $f(x, y)$. Furthermore, it is clear that, in general, the policies will not be the same for $N = 1$.

It does, however, follow from our previous argumentation that if for some N the decision regions of $f_N(x, y)$ and $f(x, y)$ coincide, they must do so for all larger N.

To show that the regions need not coincide for $N = 2$, consider the following simple example

$$f_{N+1}(x, y) = \mathrm{Max}\begin{bmatrix} \alpha x + p f_N(cx, y) \\ \beta y + q f_N(x, dy) \end{bmatrix}, \qquad N = 1, 2, \cdots, \tag{3.25}$$

where $\alpha, \beta > 0$, $0 < c, d < 1$, $0 < p, q < 1$. For $N = 1$, we have $f_1 = \mathrm{Max}\,[\alpha x, \beta y]$. We may take $\alpha = \beta$, since this is equivalent to changing the x or y scale. The boundary line for $N = 1$ is then $x = y$, which we call L_1. For $N = 2$, we consider the possible strategies AA, AB, BA, BB. We then have the following boundary curves:

$$\begin{aligned} L_1\colon\quad A &= B, & x &= y, \\ AA &= BA, & y &= (1 + pc - q)x, \\ BB &= BA, & y &= \frac{x}{d}, \\ AB &= AA, & y &= cx, \\ AB &= BA, & y &= \frac{(1-q)}{(1-p)}\,x. \end{aligned} \tag{3.26}$$

If $c \geq (1 - q)/(1 - p)$, the lines will have the relative positions shown in Fig. 3.4. It is clear then that for $N = 2$, the decision regions will be as shown in Fig. 3.5.

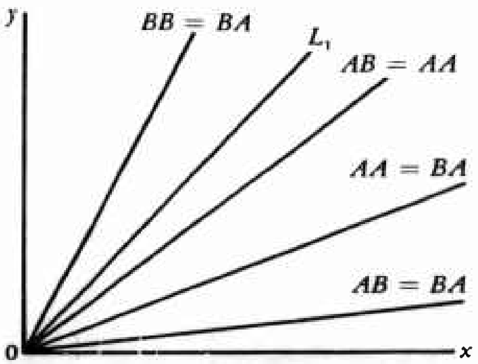

Fig. 3.4

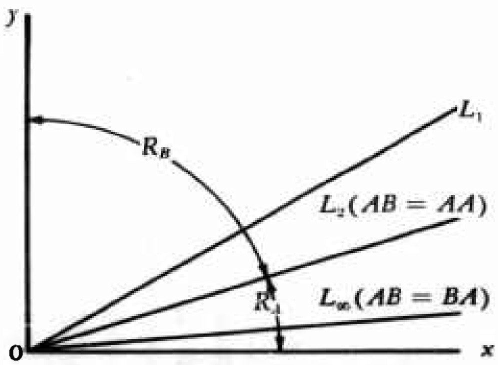

Fig. 3.5

Let us now show that decision regions for f_N converge toward that of f as $N \to \infty$, and that there will always be an N_0 with the property that for $N \geq N_0$ the regions will coincide.

The proof is very simple. Consider the situation for $N = 3$, as in Fig. 3.6.

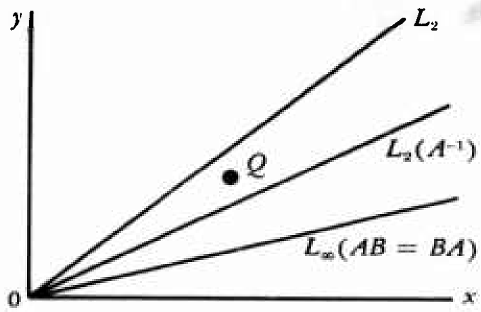

Fig. 3.6

Let $L_2(A^{-1})$ denote the line that is transformed into L_2 when (x, y) goes into (cx, y). Let Q be in the sector between L_2 and $L_2(A^{-1})$. If A is used at Q, then B is used next, since the transformed point is in the R_B-region for $N = 2$. If Q is above L_∞, we know that AB is inferior to BA, regardless of N, as a set of first two choices. Hence, B is used at Q. This shows us that the B-region for the N-stage process is at least that containing the sector bounded by the y-axis and $L_2(A^{-1})$. This process continues until $L_k(A^{-1})$, for some k, lies below L_∞, which must necessarily occur after some finite number of stages.

The argument is general and applies to the general equations discussed above. However, we cannot assert that the convergence is monotone, as we suspect, until we know more about the A- and B-regions for the N-stage process. It is probably true that there are two regions for each N, but this is a result that has only been demonstrated in the case of the simple equation (3.20).

To show this result, we use the fact that this equation arises from a model in which the results of an operation are known only as far as the expected outcome is concerned. Any N-stage policy has the form, therefore,

$$S_N = A^{a_1}B^{b_1} \cdots A^{a_k}B^{b_k}, \tag{3.27}$$

where the a_i and b_i are 0 or positive integers. There are now two cases: S_N is either equal to A^N or B^N, or it has the form $A^kB \cdots$ or $B^lA \cdots$, where $k, l < N$.

Referring to Fig. 3.6, consider a point Q above L_∞. If an optimal policy has the form $A^kB \cdots$, $k < N$, which may be written $A^{k-1}(AB) \cdots$, it may be improved by replacing AB with BA, since A iterated any number of times maintains Q above L_∞. It follows then that in the region above L_∞, either B is used first or A is used repeatedly; and, similarly, in the region below L_∞, either A is used first or B is used repeatedly.

Since A^N is clearly the optimal policy for points sufficiently close to the x-axis, and B^N is the optimal policy for points sufficiently near the y-axis, it follows from the analytic form of the yield for any S_N—an expression which is linear in x and y—that if

A^N is used at Q, it is used for all points below the line $0Q$, and similarly for B^N, "below" being replaced by "above."

It follows that there are always two regions, separated either by $AB = BA$ or by a line of more complicated form, if A^N or B^N are still dominant. For large N it is clear that A^N and B^N become less and less influential, so that eventually $AB = BA$ emerges as the sole dividing line.

3.6. A General Utility Function

We have in the previous sections considered only the case in which the utility of a total yield z was proportional to z. Let us now turn to the more interesting case in which the utility is measured by a function $\phi(z)$.

The non-linearity of $\phi(z)$ will, in general, require the introduction of a new state parameter—the quantity obtained as a result of the preceding operations. Denoting this quantity by a, we obtain the equation

$$f(x, y, a) = \text{Max}\begin{bmatrix} A\colon & p_1 f(0, y, a + x) + p_2 f(c_2 x, y, a + c_1 x) + p_3\phi(a) \\ B\colon & q_1 f(x, 0, a + y) + q_2 f(x, d_2 y, a + d_1 y) + q_3\phi(a) \end{bmatrix},$$
$$f(0, 0, a) = \phi(a), \tag{3.28}$$

as noted in Section 3.1.

This equation is more difficult to treat of than that occurring for $\phi(z) = z$, and we shall only be able to present its solution for certain classes of functions.

We have

$$f(0, y, a) = \text{Max}\begin{bmatrix} A\colon & p_1 f(0, y, a) + p_2 f(0, y, a) + p_3\phi(a) \\ B\colon & q_1 f(0, 0, a + y) + q_2 f(0, d_2 y, a + d_1 y) + q_3\phi(a) \end{bmatrix}. \tag{3.29}$$

Since $f(x, y, a) \geq f(0, 0, a) = \phi(a)$ for $x, y \geq a$, with strict inequality if x or y is positive, it follows, since $p_1 + p_2 + p_3 < 1$, that

$$f(0, y, a) = q_1\phi(a + y) + q_3\phi(a) + q_2 f(0, d_2 y, a + d_1 y), \tag{3.30}$$

and, similarly, that

$$f(x, 0, a) = p_1\phi(a + x) + p_3\phi(a) + p_2 f(c_2 x, 0, a + c_1 x). \tag{3.31}$$

For given ϕ, these equations may now be solved by iteration for the functions $f(0, y, a)$ and $f(x, 0, a)$.

Let us again proceed formally before turning to a justification of our operations. It is clear from the conservative nature of the processes involved that the quantity $x + y + a$ remains constant throughout the sequence of operations. Consequently, the effect of any choice is to transform a point in the region R: $x + y + a = c$, $x, y, a \geq 0$ into another point in the region, as shown in Fig. 3.7 on page 40.

The problem that confronts us is that of determining the set of points in R in which A is used and the set in which B is used. If we assume, as before, that these sets constitute connected regions having a boundary curve P, we may proceed to find the boundary as before, using the fact that the boundary is an indifference curve.

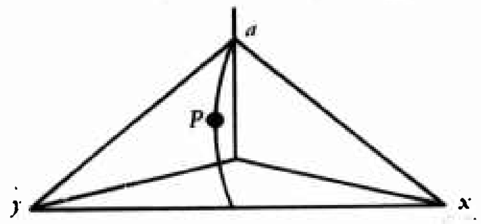

Fig. 3.7

However, we must assume more about the boundary curve than previously, where the fact that it was a straight line resulted in considerable simplification. Let us assume that the result of applying A to a point P on the boundary curve is to transform it into the B-region, and vice versa.

Having provided ourselves with a cushion of assumptions, let us now go through the calculations. If A is employed, we obtain

$$f(x, y, a) = p_1 f(0, y, a + x) + p_2 f(c_2 x, y, a + c_1 x) + p_3 \phi(a). \tag{3.32}$$

Employing B at $(0, y, a + x)$ and $(c_2 x, y, a + c_1 x)$, we obtain

$$\begin{aligned} f(x, y, a) = {} & p_1[q_1\phi(a + x + y) + q_2 f(0, d_2 y, a + x + d_1 y) \\ & + q_3\phi(a + x)] + p_2[q_1 f(c_2 x, 0, a + c_1 x + y) \\ & + q_2 f(c_2 x, d_2 y, a + c_1 x + d_1 y) \\ & + q_3 \phi(a + c_1 x)] + p_3\phi(a). \end{aligned} \tag{3.33}$$

A similar expression is obtained by using B and then A. Equating the two, we obtain, for the equation of the boundary curve,

$$\begin{aligned} & p_1 q_3 \phi(a + x) + p_2 q_3 \phi(a + c_1 x) + p_3 \phi(a) \\ & = q_1 p_3 \phi(a + y) + q_2 p_3 \phi(a + d_1 y) + q_3 \phi(a), \end{aligned} \tag{3.34}$$

which may be written

$$\begin{aligned} & p_1 q_3[\phi(a + x) - \phi(a)] + p_2 q_3[\phi(a + c_1 x) - \phi(a)] \\ & = q_1 p_3[\phi(a + y) - \phi(a)] + q_2 p_3[\phi(a + d_1 y) - \phi(a)]. \end{aligned} \tag{3.35}$$

In order to establish the result rigorously, we must ascertain whether or not the boundary curve has the desired transformation property.

What we actually require is

PROPERTY T. *If*

$$\begin{aligned} F(x, y, a) = {} & p_1 q_3[\phi(a + x) - \phi(a)] \\ & + p_2 q_3[\phi(a + c_1 x) - \phi(a)] - q_1 p_3[\phi(a + y) - \phi(a)] \\ & - q_2 p_3[\phi(a + d_1 y) - \phi(a)] \geq 0, \end{aligned} \tag{3.36}$$

then $F(c_2x, y, a + c_1x) \geq 0$. *If* $F(x, y, a) \leq 0$, *then* $F(x, d_2y, a + d_1y) \leq 0$.

Unfortunately it seems to be difficult to present any simple criterion which will ensure that a general utility function $\phi(z)$ will satisfy Property T. It is not difficult to show, for example, that $\phi(z) = z^2$ does not satisfy it for all values of p_i and q_i.

Let us now demonstrate

THEOREM 3.2. *If*

$$\begin{aligned} &\text{(a)} \qquad && \phi(z) \text{ is strictly increasing and continuous,} \\ & && \phi(z) \geq 0, \\ &\text{(b)} && \text{Property T is satisfied,} \end{aligned} \tag{3.37}$$

then the solution to (3.28) *is given by*

$$f(x, y, a) = p_1 f(0, y, a + x) + p_2 f(c_2x, y, a + c_1x) + p_3\phi(a) \tag{3.38}$$

for $F(x, y, a) \geq 0$, *and by*

$$f(x, y, a) = q_1 f(x, 0, a + y) + q_2 f(x, d_2y, a + d_1y) + q_3\phi(a), \tag{3.39}$$

for $F(x, y, a) \leq 0$.

The optimal policy is to apply A when $F(x, y, a) > 0$ *and B if* $F(x, y, a) < 0$. *When there is equality, it is a matter of indifference as to which choice is made.*

PROOF. The proof is carried through in two stages. First we show that there is a region in the plane $x + y + a = c$ where A is always used, namely, a region close to $y = 0$. Then we consider what happens at a point Q in the region defined by $F(x, y, a) \geq 0$ and $x + y + a = c$.

Let us assume for the moment that we have already established the existence of a region where A is always used. If B is used at Q, it follows from Property T that the transformed point is again in the same region. It cannot be true that B is used repeatedly, if $x > 0$, since eventually the y coordinate will be so small that the point will be in the A-region. Hence, if at Q an optimal policy employs B for the first k choices, the sequence of moves has the form

$$S = BB \cdots (k \text{ times}) \cdots BA. \tag{3.40}$$

On the basis of Property T, we are still in the region $F(x, y, a) \geq 0$, $x + y + a = c$ after employing B $(k - 1)$ times. The next two moves, B and then A, cannot be optimal, however, since the region is defined by the property that AB plus optimal continuation is superior to BA plus optimal continuation. This shows that at Q, move B cannot be used first in an optimal policy.

It remains then to establish the existence of the A-region mentioned above. Since $f(x, y, a) > \phi(a)$ for $x, y \geq 0$ and one at least positive, it follows that

$$\begin{aligned} &p_1 f(0, y, a + x) + p_2 f(c_2x, y, a + c_1x) + p_3\phi(a) \\ &\quad > q_1 f(d, 0, a + y) + q_2 f(x, d_2y, a + d_1y) + q_3\phi(a), \end{aligned} \tag{3.41}$$

which holds at $y = 0$, must by virtue of the continuity of the functions involved, for any $x > 0$, hold for some interval $0 \leq y \leq y(x, a)$.

3.7. The Exponential Utility Function

One way of obtaining utility functions that have the desired property, T, is to make the boundary equation independent of a. If we wish this to be true for all values of the parameters p_i and q_i, we must have

$$\phi(a + x) - \phi(x) = G(x)H(a), \tag{3.42}$$

which yields, using standard arguments, under the assumption of continuity,*

$$\begin{aligned} &\text{(a)} \qquad \phi(z) = mz + n, \qquad \text{or} \\ &\text{(b)} \qquad \phi(z) = ce^{bz}. \end{aligned} \tag{3.43}$$

We have already considered the first utility function; let us now consider the second.

The important property of these utility functions is that a policy which maximizes the expected value of $\phi(z)$ proceeds at each stage without regard for the amount already obtained, being dependent only on the remaining amount to be obtained.

If we set, for $b > 0$,

$$g(x, y) = \operatorname*{Max}_{P} \operatorname{Exp} (e^{bz}) \tag{3.44}$$

("Exp" denotes here "expected value," not "exponential"),

we obtain for g the functional equation

$$g(x, y) = \operatorname{Max} \begin{bmatrix} A\colon\ p_1 e^{bx} g(0, y) + p_2 e^{bc_1 x} g(c_2 x, y) + p_3 \\ B\colon\ q_1 e^{by} g(x, 0) + q_2 e^{bd_1 y} g(x, d_2 y) + q_3 \end{bmatrix}. \tag{3.45}$$

As a special case of Theorem 3.2, we obtain

THEOREM 3.3. *The solution of* (3.45) *is as follows: For*

$$\frac{p_1(e^{bx} - 1) + p_2(e^{bc_1 x} - 1)}{p_3} > \frac{q_1(e^{by} - 1) + q_2(e^{bd_1 y} - 1)}{q_3}, \tag{3.46}$$

use A; if the reverse inequality holds, employ B; if equal, either is applicable.

Observe that, as should be true, the limiting solution as $b \to 0$ is exactly that obtained from $\phi(z) \equiv z$.

3.8. Asymptotic Behavior of g(x, y)

We now turn to the problem of determining the asymptotic behavior of $g(x, y)$ as x and $y \to \infty$. We begin by deriving the asymptotic behavior of $g(x, 0)$ and $g(0, y)$. From the equation we obtain, for large x,

$$g(x, 0) = p_1 e^{bx} + p_3 + p_2 e^{bc_1 x} g(c_2 x, 0). \tag{3.47}$$

This equation may be solved by iteration:

$$g(x, 0) = (p_1 e^{bx} + p_3) + p_2 e^{bc_1 x}(p_3 + p_1 e^{bc_2 x}) + \cdots. \tag{3.48}$$

* This requirement of continuity can be considerably weakened.

To obtain the asymptotic behavior, however, we must proceed differently. Set

$$g(x, 0) = \frac{p_1 e^{bx}}{1 - p_2} + h(x)e^{bx}, \tag{3.49}$$

where h satisfies the equation

$$h(x) = p_3 e^{-bx} + p_2 h(c_2 x), \tag{3.50}$$

as we see by direct substitution. Although iteration yields

$$h(x) = p_3 e^{-bx} + p_2 p_3 e^{-bc_2 x} + \cdots, \tag{3.51}$$

the asymptotic behavior of $h(x)$ is still not apparent. We shall show that $h(x) = x^{-a}\Psi(x)[1 + o(1)]$ as $x \to \infty$, where $\Psi(x) = \Psi(c_2 x)$, $a = (\log 1/p_2)/(\log 1/c_2)$. To accomplish this, set $h(x) = k(x)x^{-a}$. Then k satisfies the simpler equation

$$k(x) - k(c_2 x) = p_3 x^a e^{-bx} = \phi(x). \tag{3.52}$$

The essential fact about ϕ that we shall use is that $\sum_{n=1}^{\infty} \phi(x/c_2^n)$ converges for each x. From (3.52) we have

$$k\left(\frac{x}{c_2^n}\right) - k\left(\frac{x}{c_2^{n-1}}\right) = \phi\left(\frac{x}{c_2^n}\right), \tag{3.53}$$

which yields

$$\lim_{n \to \infty} k\left(\frac{x}{c_2^n}\right) = k(x) + \sum_{n=1}^{\infty} \phi\left(\frac{x}{c_2^n}\right) = \Psi(x). \tag{3.54}$$

From the form of the limit function or from the equation for $k(x)$, we see that $\Psi(x) = \Psi(c_2 x)$ for all x. If then we write $y = x/c_2^n$ for $1 \leq x \leq 1/c_2$, we have

$$\begin{aligned} k(y) = k\left(\frac{x}{c_2^n}\right) &= \Psi(x)[1 + o(1)] = \Psi\left(\frac{x}{c_2^n}\right)[1 + o(1)] \\ &= \Psi(y)[1 + o(1)], \end{aligned} \tag{3.55}$$

as $y \to \infty$.

Collecting the previous results, we see that the asymptotic behavior of $g(x, 0)$ is given by

$$g(x, 0) = \frac{p_1}{1 - p_2} e^{bx} + \frac{e^{bx}\Psi(x)}{x^{a_1}} [1 + o(1)], \tag{3.56}$$

where

(a) $$\Psi(x) = \Psi(c_2 x),$$

(b) $$a_1 = \frac{\log \dfrac{1}{p_2}}{\log \dfrac{1}{c_2}}. \tag{3.57}$$

The corresponding result for $g(0, y)$ is

$$g(0, y) = \frac{q_1}{1 - q_2} e^{by} + \frac{e^{by}\zeta(y)}{y^{b_1}} [1 + o(1)], \tag{3.58}$$

where

(a) $$\zeta(y) = \zeta(d_2 y),$$

(b) $$b_1 = \frac{\log \frac{1}{q_2}}{\log \frac{1}{d_2}}. \tag{3.59}$$

Turning to the equation for $g(x, y)$ we have, for x and y large,

$$g(x, y) = \operatorname{Max} \begin{bmatrix} \frac{q_1 p_1}{1 - q_2} e^{b(x+y)} + p_2 e^{bc_1 x} g(c_2 x, y) + 0\left(\frac{e^{by}}{y^{b_1}}\right) \\ \frac{p_1 q_1}{1 - p_2} e^{b(x+y)} + q_2 e^{bd_1 y} g(x, d_2 y) + 0\left(\frac{e^{bx}}{x^{a_1}}\right) \end{bmatrix}. \tag{3.60}$$

Setting $h(x, y)e^{b(x+y)} = g(x, y)$, we obtain

$$h(x, y) = \operatorname{Max} \begin{bmatrix} \frac{q_1 p_1}{1 - q_2} + p_2 h(c_2 x, y) + 0\left(\frac{e^{-bx}}{y^{b_1}}\right) \\ \frac{p_1 q_1}{1 - p_2} + q_2 h(x, d_2 y) + 0\left(\frac{e^{-by}}{x^{a_1}}\right) \end{bmatrix}. \tag{3.61}$$

To simplify still further, we set $h(x, y) = \alpha + k(x, y)$, obtaining

$$\alpha + k(x, y) = \operatorname{Max} \begin{bmatrix} \frac{p_1 q_1}{1 - q_2} + p_2\alpha + p_2 k(c_2 x, y) + 0\left(\frac{e^{-bx}}{y^{b_1}}\right) \\ \frac{p_1 q_1}{1 - p_2} + q_2\alpha + q_2 k(x, d_2 y) + 0\left(\frac{e^{-by}}{x^{a_1}}\right) \end{bmatrix}. \tag{3.62}$$

If α is chosen to be the common solution of

$$\alpha = \frac{p_1 q_1}{1 - q_2} + p_2\alpha = \frac{p_1 q_1}{1 - p_2} + q_2\alpha, \tag{3.63}$$

namely, $p_1 q_1/(1 - p_2)(1 - q_2)$, (3.62) simplifies to

$$k(x, y) = \operatorname{Max} \begin{bmatrix} p_2 k(c_2 x, y) + 0\left(\frac{e^{-bx}}{y^{b_1}}\right) \\ q_2 k(x, d_2 y) + 0\left(\frac{e^{-by}}{x^{a_1}}\right) \end{bmatrix}. \tag{3.64}$$

To estimate $k(x, y)$, we use the fact that the solution may be obtained by means of successive approximations:

$$k_{n+1}(x, y) = \text{Max}\begin{bmatrix} p_2 k_n(c_2 x, y) + 0\left(\dfrac{e^{-bx}}{y^{b_1}}\right) \\ q_2 k_n(x, d_2 y) + 0\left(\dfrac{e^{-by}}{x^{a_1}}\right) \end{bmatrix}, \qquad k_0(x, y) = \frac{1}{x^r + y^r}, \tag{3.65}$$

considering, for our purposes, only values of x and y greater than 1. The exponent r will be chosen in a moment.

If we have an inequality of the type $k_n(x, y) \leq u_n/(x^r + y^r)$, u_n being a constant, which inequality is certainly valid for $n = 0$, we obtain

$$k_{n+1}(x, y) \leq \text{Max}\begin{bmatrix} \dfrac{p_2 u_n}{c_2^r(x^r + y^r)} + 0\left(\dfrac{e^{-bx}}{y^{b_1}}\right) \\ \dfrac{q_2 u_n}{d_2^r(x^r + y^r)} + 0\left(\dfrac{e^{-by}}{x^{a_1}}\right) \end{bmatrix}. \tag{3.66}$$

Choose r so that $p_2/c_2^r \leq \frac{1}{2}$, $q_2/d_2^r \leq \frac{1}{2}$. Since $a_1, b_1 > r$, we see, since $x^r e^{-bx} \leq d_r$ for all x, that $e^{-bx}/y^{a_1} \leq d_r/x^r y^r \leq d_r/(x^r + y^r)$, for $x, y \geq 1$. Hence, we have

$$k_{n+1}(x, y) \leq \text{Max}\begin{bmatrix} \dfrac{\frac{1}{2}u_n}{x^r + y^r} + \dfrac{a_2}{x^r + y^r} \\ \dfrac{\frac{1}{2}u_n}{x^r + y^r} + \dfrac{a_2}{x^r + y^r} \end{bmatrix}, \tag{3.67}$$

for some constant a_2. If we take $u_{n+1} = \frac{1}{2}(u_n + a_2)$, the inequality is preserved for u_{n+1}. Since u_n as defined by the recurrence relation is uniformly bounded, we obtain, in the limit, $k(x, y) \leq a_3/(x^r + y^r)$.

Knowing the form of the function, we readily obtain the optimal policy, deriving in this case the slightly paradoxical result that, asymptotically, as x and $y \to \infty$, it makes no difference which move is made first.

Collecting the above results, we obtain

$$g(x, y) = \frac{e^{b(x+y)} p_1 q_1}{(1 - p_2)(1 - q_2)} + 0\left(\frac{e^{b(x+y)}}{x^r + y^r}\right). \tag{3.68}$$

3.9. A More General Problem

We have, in the previous sections, considered the equations resulting from situations in which two choices are available at each stage. Let us now discuss a three-choice problem, as represented by the functional equation

$$f(x, y) = \text{Max}\begin{bmatrix} A\colon\ p_1[r_1 x + f(s_1 x, y)] \\ B\colon\ p_2[r_2 y + f(x, s_2 y)] \\ C\colon\ p_3[r_3 x + r_4 y + f(s_3 x, s_4 y)] \end{bmatrix}, \tag{3.69}$$

where

$$0 < p_1, p_2, p_3 < 1, \qquad 0 < r_i < 1, \qquad r_i + s_i = 1, \qquad i = 1, 2, 3, 4.$$

It might be suspected, on the basis of the previous results, that there will always exist three sectors, R_A, R_B, R_C, as pictured in Fig. 3.8, which determine the choice of the A, B, or C moves. Unfortunately, this is not true for general values of the parameters, since it has been shown that there is an equation of the form of (3.69) for which the decision regions are as shown in Fig. 3.9.

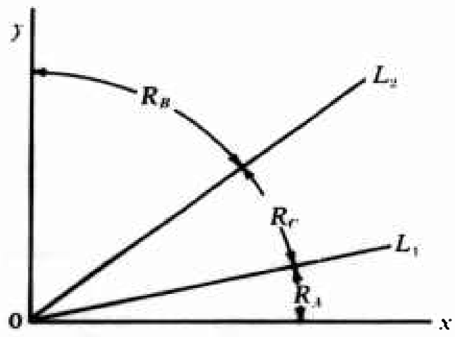

Fig. 3.8

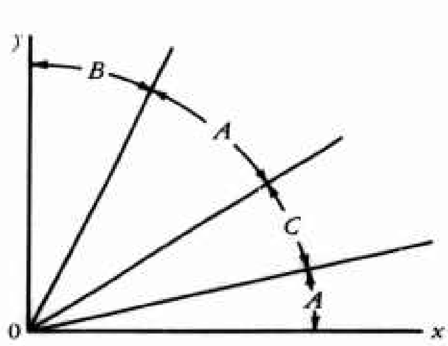

Fig. 3.9

From the fact that there exists a problem whose solution involves four decision regions, it follows immediately that the general solution of the multidecision problem cannot have the simple form of the solution in Section 3.2.

At the present time, although little is known about the general solution of the k-choice analogue of (3.69), it seems fairly certain that its general solution will possess a complicated and extremely unintuitive structure. It is not even known whether or not there is always a finite number of regions for any particular equation, and, if so, whether this number can be arbitrarily large or must be bounded by a number depending on k.

We shall illustrate a number of partially successful approaches by considering the two equations of special form

$$f(x, y) = \operatorname{Max} \begin{bmatrix} A\colon\ p_1[r_1 x + f(s_1 x, y)] \\ B\colon\ p_2[r_1 y + f(x, s_1 y)] \\ C\colon\ p_3[sx + sy + f(tx, ty)] \end{bmatrix} \tag{3.70}$$

and

$$f(x, y) = \operatorname{Max} \begin{bmatrix} A\colon\ x + f(ax, by) \\ B\colon\ y + f(cy, dx) \end{bmatrix}. \tag{3.71}$$

Before turning to a discussion of these equations, let us note that equations of this type also arise in connection with testing problems of an interesting type.

Consider the simplest version, in which we are given the information that a ball is in one of N boxes, and the a priori probability, p_k, that it is in the kth box. Assuming that each observation consumes one unit of time, it seems intuitively clear that we look in the most likely box first, in order to minimize the expected time required to find the ball. Note, however, that, for the case of two boxes, if we are merely interested in determining which box contains the ball, it makes no difference which box we examine first. If, however, we want to obtain the ball or to observe it, then it is best to examine the most likely spot.

Let us now consider the more general situation in which observation of the kth box consumes time t_k, and in which there is a probability q_k that if the kth box is observed, one is unable to examine its contents or to obtain them.

THEOREM 3.4. *If we wish to obtain the ball, the optimal policy is to examine the box for which*

$$\frac{p_k(1 - q_k)}{t_k} \tag{3.72}$$

is a maximum.

If we wish merely to locate the box containing the ball, the box for which (3.72) *is a maximum is examined first or is never examined.*

More interesting and difficult problems arise in situations in which the testing disturbs the probability distribution. For a two-box model, this leads to functional equations of the form

$$f(p_1, p_2) = \text{Min} \begin{Bmatrix} p_1 + (1 - p_1)[1 + f(a_{12}, a_{22})] \\ p_2 + (1 - p_2)[1 + f(a_{11}, a_{21})] \end{Bmatrix}, \tag{3.73}$$

which is easily resoluble. However, for three boxes we obtain

$$f(p_1, p_2, p_3) = \text{Min}_i \{p_1 + (1 - p_i)[f(p_{1i}^*, p_{2i}^*, p_{3i}^*)]\}, \tag{3.74}$$

where

$$\begin{aligned} p_{11}^* &= \frac{a_{12}p_2 + a_{13}p_3}{1 - p_1}, \\ p_{21}^* &= \frac{a_{22}p_2 + a_{23}p_3}{1 - p_1}, \\ p_{31}^* &= \frac{a_{32}p_2 + a_{33}p_3}{1 - p_1}, \end{aligned} \tag{3.75}$$

and so on.

Functional equations of this type occur frequently in the theory of sequential analysis, in connection with problems in which the distribution is unknown and each observation yields additional information concerning it.

3.10. A Simple Three-choice Equation

Equation (3.70) may be written

$$f(x, y) = \text{Max}\begin{Bmatrix} A\colon & p_1[r_1x + f(s_1x, y)] \\ B\colon & p_2[r_1y + f(x, s_1y)] \\ C\colon & p_3[sx + sy + tf(x, y)] \end{Bmatrix} \tag{3.76}$$

by virtue of the homogeneity of $f(x, y)$, and, as a consequence, in the simpler form

$$f(x, y) = \text{Max}\begin{bmatrix} A\colon & p_1[r_1x + f(s_1x, y)] \\ B\colon & p_2[r_1y + f(x, s_1y)] \\ C\colon & \dfrac{p_3s[x + y]}{1 - p_3t} \end{bmatrix}. \tag{3.77}$$

For this equation we can prove

THEOREM 3.5. *If* $0 < r_1, s, t, p_1, p_2, p_3 < 1$, $r_1 + s_1 = 1$, *there are at most three decision regions.*

The proof, which we shall merely sketch, is more interesting than the result and is applicable to more general situations. The basic idea is to employ a continuity method, using an appropriate parameter, which in this case is s. For $s = 0$, there are actually two regions, as we know from the previous results. It is now not difficult to show that as s varies between 0 and 1, the number of regions does not exceed three.

3.11. The Equation $f(x, y) = \text{Max}\,[x + f(ax, by), y + f(cy, dx)]$

As another example of an equation in the case of which special techniques are applicable, let us consider

$$f(x, y) = \text{Max}\begin{bmatrix} A\colon & x + f(ax, by) \\ B\colon & y + f(cy, dx) \end{bmatrix}, \tag{3.78}$$

where we shall assume that $0 \leq a, b, c, d < 1$. Under these conditions we know that there is a unique solution. Actually, these conditions are too strong, since $0 \leq cd < 1$ is sufficient to ensure existence and uniqueness.

The principal result we shall obtain is

THEOREM 3.6. *All optimal strategies are periodic from some point on.*

Let us note that an A-choice sends (x, y) into (ax, by), and that a B-choice sends (x, y) into (cy, dx). We observe that the motion induced by B^2 sends (x, y) into (cdx, cdy), or, more precisely, if the optimal policy is $B^2 0$ (read B^2 optimal), then

$$f(x, y) = (y + dx) + cdf(x, y). \tag{3.79}$$

From this we conclude that if $B^2 0$ is an optimal policy, then $B^2 0 = \bar{B}$ ($\bar{C}$ denotes the fact that, for any C, the sequence of moves represented by C is repeated periodically); and in fact for a point (x, y) where this takes place,

$$f(x, y) = f_B(x, y) = \frac{y + dx}{1 - cd}. \tag{3.80}$$

Next, suppose that for a point (x, y) the optimal strategy has the form A^kBA^kB0. Then

$$\begin{aligned} f(x, y) &= x(1 + a + a^2 + \cdots + a^{k-1}) + b^k y + f(cb^k y, da^k x) \\ &= x\frac{1-a^k}{1-a} + b^k y + cb^k y\frac{1-a^k}{1-a} \\ &\qquad + b^k da^k x + f(a^k b^k cdx, a^k b^k cdy) \\ &= P(x, y) + a^k b^k cdf(x, y). \end{aligned} \tag{3.81}$$

From this we conclude that if A^kBA^kB0 is optimal, then the A^kB pattern repeats periodically, i.e.,

$$A^kBA^kB0 = \overline{A^kB}. \tag{3.82}$$

Similarly, if BA^kBA^k0 is optimal, we obtain

$$BA^kBA^k0 = \overline{BA^k}. \tag{3.83}$$

Also, in this case,

$$f_{\overline{BA^k}}(x, y) = \frac{(y + b^k dx)\left(1 - c\dfrac{1-a^k}{1-a}\right)}{a - cda^k b^k}. \tag{3.84}$$

We are now in a position to classify completely the optimal strategies. First, broader classifications are obtained, and from these obvious eliminations are made to achieve the final list. As a first crude classification for the optimal strategies beginning with A,

$$A0 = \begin{cases} (1) & [\overline{A}], \\ (2) & [A^l\overline{B}], \\ (3) & [A^lB0] \end{cases} \tag{3.85}$$

(we put brackets around a strategy when no further subclassification will be made using this form). Considering those strategies of (3), above, which are not in (2), we have, since $B^20 = \overline{B}$,

$$A0 = A^lB0 = A^lBA0;$$

then,

$$A^lBA0 = \begin{cases} (3') & [A^lB\overline{A}], \\ (3'') & A^lBA^kB0. \end{cases} \tag{3.86}$$

Next, in case (3″) we have two cases according as $l \geq k$ or $l < k$:

Case 3″ ($l \geq k$)

In this case we have, if $A0$ is optimal, $A0 = A^lBA^kB0 = A^{l-k}(A^kBA^kB0)$; and since at the state reached after $l - k$ applications of A, A^kBA^kB0 would be optimal, the above, together with (3.82), implies that

$$A0 = A^{l-k}\overline{A^kB} = [A^lBA^k, l \geq k]. \tag{3.87}$$

CASE 3″ ($l < k$)
This case leads to three subcases:

$$A0 = A^l BA^k B0 = \begin{cases} (3''_1) & A^l BA^k B\overline{A}, \\ (3''_2) & A^l BA^k BA^r B0, \\ (3''_3) & [A^l BA^k \overline{B}]. \end{cases} \tag{3.88}$$

Subcase $3''_1$
This subcase implies that, since $l < k$,

$$A0 = A^l(BA^k BA^k 0) = [A^l \overline{BA^k}], \tag{3.89}$$

via (3.82).
Subcase $3''_2$
In this subcase we again have two cases, according as $k > r$ or $k \leq r$.
If $k > r$, we have $A^l BA^k BA^r B0 = A^l BA^{k-r}(A^r BA^r B0) = A^l BA^{k-r}\overline{A^r B} = [A^l BA^r \overline{BA^r}, k > \text{Max}\,(l, r)]$. On the other hand, if $k \leq r$, $A^l BA^k BA^r B0 = A^l(BA^k BA^k 0) = [A^l \overline{BA^k}]$.
Collating the classification carried out above, we see that a list which includes all optimal strategies beginning with A is

$$A0 = \begin{cases} (a_1) & \overline{A}, \\ (a_2) & A^l \overline{B}, \quad l = 1, 2, \cdots, \\ (a_3) & A^l B\overline{A}, \quad l = 1, 2, \cdots, \\ (a_4) & A^l \overline{BA^k}, \quad l = 1, 2, \cdots; \quad k = 1, 2, \cdots, \\ (a_5) & A^l BA^k \overline{B}, \quad 1 \leq l < k, \\ (a_6) & A^l BA^k \overline{BA^r}, \quad k > \text{Max}\,(l, r) \geq 1. \end{cases} \tag{3.90}$$

Next we consider optimal strategies beginning with B. It is quite clear that either $B0 = \overline{B}$ or $B0 = BA0$. Thus, we see immediately from (3.89) that a list of possible optimal strategies beginning with B is

$$\begin{array}{lll} (b_1) & \overline{B}, & \\ (b_2) & BA^l \overline{B}, & l = 1, 2, \cdots, \\ (b_3) & \overline{BA^l}, & l = 1, 2, \cdots, \\ (b_4) & BA^l \overline{BA^k}, & k < l, \\ (b_5) & B\overline{A}. & \end{array} \tag{3.91}$$

Although it is now possible to obtain the decision regions explicitly by computing the results of the allowable optimal strategies, the amount of effort required is so great that another technique is employed. In place of this approach, a combination of the geometric treatment discussed in the next section, together with the analytic approach already employed, yields the information concerning the number of connected decision regions. Since the proof requires a detailed investigation, we shall omit it here.

3.12. An Illustration of the Geometric Approach

Let us consider the equation

$$f(x, y) = \text{Max}\begin{bmatrix} A\colon & ax + rf(bx, y) \\ B\colon & cy + sf(x, dy) \end{bmatrix} \tag{3.92}$$

and obtain its solution by using the geometric techniques discussed in Section 2.8. As we know, the solution will be to employ B whenever

$$\frac{y}{x} \geq \frac{a(1-s)}{c(1-r)} \tag{3.93}$$

and to employ A whenever the reverse inequality holds, assuming, as we shall, that $0 < a, b, c, d, r, s < 1$.

To prove this result, we consider first the set of all strategies of the form ABT and construct their subenvelope from above, $E_{AB} = \overline{\text{Env}}_T L(ABT)$. Similarly, we form $E_{BA} = \overline{\text{Env}}_R L(BAR)$. Then let $\bar{E} = \overline{\text{Env}}\,\{E_{AB}, E_{BA}\}$. For a given strategy T,

$$\begin{aligned} f_{ABT}(x, y) &= ax + rcy + rsf(bx, dy), \\ f_{BAT}(x, y) &= cy + sax + rsf(bx, dy), \end{aligned} \tag{3.94}$$

so that $L(ABT)$ and $L(BAT)$ intersect at the normalized point $\bar{y}_1$ corresponding to $y/x = a(1-s)/c(1-r)$, $x + y = 1$. (More precisely, $\bar{y}_1 = a(1-s)/[c(1-r) + a(1-s)]$.) Note in particular that $\bar{y}$ is independent of the choice of strategies. Furthermore, for $\bar{y} > \bar{y}_1$, $L(BAT)$ lies below $L(ABT)$. Thus, $\bar{E}$ consists of E_{AB} for $\bar{y} \leq \bar{y}_1$ and of E_{BA} for $\bar{y} \geq \bar{y}_1$. Hence, with respect to the strategies included in the subenvelope $\bar{E}$, A is an optimal initial choice to the left of $\bar{y}_1$, and B is optimal to the right of $\bar{y}_1$. To complete the proof of the theorem, we need only show that this property is preserved after we pass to the full envelope by taking the envelope of $\bar{E}$ and the lines of the strategies not yet considered. These lines are of the form $L(A^kBT)$, $L(B^kAR)$, $k > 1$; $L(A^\infty)$ and $L(B^\infty)$.

If a line $L(A^kBT)$, k > 1 touches the envelope E at a point $\bar{y}_0$, $L(ABT)$ also touches the envelope and to the right of $\bar{y}_0$, since A transforms the decision at (x, y) into one at (bx, y), and the normalized (bx, y) is larger than the normalized (x, y). Thus, if $\bar{y}_0 \geq \bar{y}_1$, $L(ABT)$ would touch E to the right of $\bar{y}_1$, which is impossible, since $L(ABT)$ lies below the subenvelope $\bar{E}$ for $\bar{y} \geq \bar{y}_1$. A symmetric argument disposes of the $L(B^kAR)$, $k > 1$. As for the two lines $L(A^\infty)$ and $L(\bar{B}^\infty)$, these are limits of the $L(A^kBT)$, $L(B^kAR)$, respectively, as $k \to \infty$, and they can in no way affect the ultimate envelope E. In fact, clearly A^∞ is optimal only at $\bar{y} = 0$, and B^∞ at $\bar{y} = 1$. This completes the proof of the theorem.

The problem considered above has a finite analogue, as discussed in Section 3.5, whose solution constitutes a more precise form of the above result. Namely, letting $N = 1, 2, \cdots$, we consider

$$f(N, x, y) = \text{Max}\begin{bmatrix} A:\ ax + rf(N-1, bx, y) \\ B:\ cy + sf(N-1, x, dy) \end{bmatrix}, \tag{3.95}$$

where $f(0, x, y) \equiv 0$. Determining $f(N, x, y)$ for each N reduces to considering strategies s_N of N steps, each either A or B, and calculating $f_{s_N}(N, x, y)$ via (3.95) and then determining which of these is largest. Thus, for each point (x, y), $f(N, x, y)$ is the maximum of 2^N numbers. Again we wish to characterize the optimal initial choice, which now, of course, depends on N.

THEOREM 3.7. *There exists an $N_0 = N_0(b, d, r, s)$ such that for $N \geq N_0$, the optimal initial choice for (N, x, y) is A if $\bar{y} \leq \bar{y}_1$ and B if $\bar{y} \geq \bar{y}_1$. More precisely, we may take*

$$N_0 = \text{Max}\left[2 + \frac{\log e}{\log \frac{1}{b}}, 2 + \frac{\log e'}{\log \frac{1}{d}}, 2\right], \tag{3.96}$$

where

$$e = \frac{(1 - r + sd)(1 - s)}{1 - r},$$
$$e' = \frac{(1 - s + rb)(1 - r)}{1 - s}.$$

Furthermore, this is best possible in the sense that given a, b, c, d, r, there always exist values of s such that (1) N_0 is as large as we please; and (2) for all $N, 2 \leq N < N_0$ there are points to the right of $\bar{y}_1 = \bar{y}_1(s)$ at which A is an optimal initial choice for an N-step strategy.

PROOF. For $N = 1$ it is clear that one always chooses A at (x, y) if $ax \geq cy$, and chooses B otherwise. For $N = 2$ we see that $L(AB)$ and $L(BA)$ intersect at $\bar{y}_1$; and thus, for $N = 2$ the optimal initial choice would be A to the left of $\bar{y}_1^*$ and B to the right, except possibly for interference from the strategies A^2 and B^2. That is, it may be that A^2 appears in this "2-envelope" to the right of $\bar{y}_1$, as in Fig. 3.10. Since $L(A^2)$ is above $L(AB)$ at $\bar{y} = 0$, this would mean that $L(AB)$ is completely dominated.

Also, the intersection of $L(A^2)$ and $L(BA)$ that occurs at $\bar{y}_2$ is to the right of $\bar{y}_1$. This is numerically equivalent to

$$\frac{c}{a}(1 - s + rb) \geq \frac{a(1 - s)}{c(1 - r)},$$
$$b \geq \frac{1 - s}{1 - r}. \tag{3.97}$$

The symmetric situation for B^2, i.e., $L(B^2)$ intersecting $L(AB)$ to the left of $\bar{y}_1$, entails

$$d \geq \frac{1 - r}{1 - s}.$$

Since b and d lie between 0 and 1, both of these phenomena cannot occur simultaneously. For convenience we shall suppose henceforth that if one of them occurs, it is for A^2, i.e., $b \geq (1 - s)/(1 - r)$, so that $L(A^2)$ dominates $L(AB)$ as in Fig. 3.10.

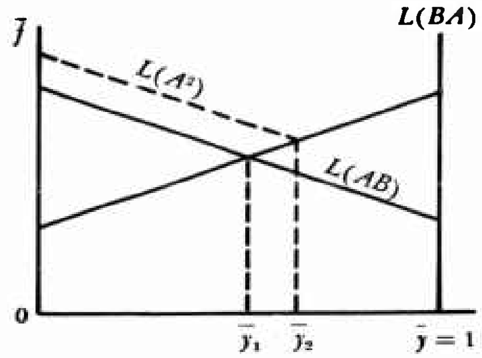

Fig. 3.10

We now proceed to consider any fixed $N \geq 3$. The subenvelope of the lines of the N-strategies of the forms

$$A^k BS_{N-k-1},\ B^k AR_{N-k-1}, \qquad k \geq 1, \tag{3.98}$$

is again separated into two parts by $\bar{y}_1$ such that, insofar as these strategies are concerned, A is the optimal initial decision for $\bar{y} \leq \bar{y}_1$, and B for $\bar{y} \geq \bar{y}_1$. There remains then to account for the two strategies A^N and B^N. If B^N appears in this N-envelope to the left of $\bar{y}_1$, B^2 will appear in the 2-envelope to the left of $\bar{y}_1$, which is not the case (according to arrangements made above). Thus, B^N does not alter the character of the initial decisions as determined by $\bar{y}_1$. The last possibility to consider is, Does A^N appear in the envelope at a point to the right of $\bar{y}_1$? If it does, then A^{N-2} moves this point to one farther to the right, at which A^2 appears in the 2-envelope. This last point, of course, must then be to the left of $\bar{y}_2$. Translating these statements into numerical terms yields

$$\frac{1}{b^{N-2}} \frac{a(1 - s)}{c(1 - r)} \leq \frac{a}{c} (1 - s + rb), \tag{3.99}$$

or

$$\left(\frac{1}{b}\right)^{N-2} \leq \frac{(1 - s + rb)(1 - r)}{1 - s} = e. \tag{3.100}$$

If $e < 1$, then this is impossible, and we may take the N_0 of the theorem equal to 2. On the other hand, if $e \geq 1$, we have $N \leq 2 + (\log e)/[\log (1/b)]$. Thus, in general, removing all asymmetries, we may take

$$N_0 = \text{Max} \left[2, 2 + \frac{\log e}{\log \frac{1}{b}}, 2 + \frac{\log e'}{\log \frac{1}{d}} \right], \tag{3.101}$$

where

$$e' = \frac{(1 - r + sd)(1 - s)}{1 - r}.$$

To demonstrate the "best possible" part of the theorem, we observe that given a, b, c, d, r and taking s close to 1, e, and consequently N_0, can be made as large as we please. We may ensure in the following manner that there are points to the right of $\bar{y}_1$ at which A is an optimal initial choice for $N < N_0$.

We first arrange for $L(B^2)$ to intersect $L(BA)$ to the right of $\bar{y}_2$. This is equivalent to $a/c(1 - s + rb) < a/cd$ or $1/d > 1 - s + rb$. This last, of course, is satisfied for s sufficiently close to 1. Then, for $N = 2$, A^2 dominates all BAT to the left of y_2, and, moreover, A^2 dominates all BS to the left of y_2, where $\bar{y}_2 > \bar{y}_1$. Now assume $k - 1 < N_0 - 1$, so that $A^{-(k-1)}y_2 = y_k > y_1$, and assume also that $(1)_{k-1}$: A^{k-1} dominates all ABT_{k-3} to the left of $\bar{y}_k$, and that $(2)_{k-1}$: A^{k-1} dominates all BT_{k-2} to the left of $\bar{y}_k$. We observe that under the conditions provided for earlier, for $k \geq 3$, $(1)_{k-1}$ implies $(2)_{k-1}$. We now wish to complete the induction by deriving $(1)_k$. That is, we must verify that A^k dominates all ABT_{k-2} to the left of $y_{k+1} = A^{-1}(y_k)$. This, however, is equivalent to A^{k-1} dominating all BT_{k-2} to the left of $A\bar{y}_{k+1} = \bar{y}_k$, which is precisely $(1)_{k-1}$. Since $k < N_0$, $\bar{y}_{k+1} > \bar{y}_1$, and the induction is completed.

CHAPTER 4

THE FUNCTIONAL EQUATION $f(x) = \underset{0 \leq y \leq x}{\text{Max}} \{g(y) + h(x - y) + f[ay + b(x - y)]\}$ AND RELATED TOPICS

4.1. Introduction

In this chapter we shall study a number of equations possessing a common structure. Since these equations are intractable analytically unless we make some simplifying assumptions, we shall devote our efforts to showing that certain simple hypotheses yield a number of interesting and important results.

We shall begin with a discussion of the equation

$$f(x) = \underset{0 \leq y \leq x}{\text{Max}} \{g(y) + h(x - y) + f[ay + b(x - y)]\}, \tag{4.1}$$

whose origin is described in Section 1.6, Problem 1.8, and continue with

$$f_k(x, y) = \underset{0 \leq z \leq x}{\text{Max}} [p_k\phi(y + c_k z) + (1 - p_k)f_{k+1}(x - z, y + c_k z)], \tag{4.2}$$

devoting some time to presenting a simple dynamic programming problem that gives rise to the above functional equation.

After this we shall turn our attention to the equation

$$f(x) = \underset{0 \leq y \leq x}{\text{Max}} \left[g(y) + h(x - y) + \int_0^y f(y - s)k(s)\, ds\right], \tag{4.3}$$

and to the equation of optimal inventory

$$u(x) = \underset{y \geq x}{\text{Min}} \left\{g(y - x) + a[M + u(0)][1 - F(y)] + a\int_0^y u(y - s)F'(s)\, ds\right\}, \tag{4.4}$$

discussed in Section 1.8.6.

4.2. The Equation $f(x) = \underset{0 \leq y \leq x}{\text{Max}} \{g(y) + h(x - y) + f[ay + b(x - y)]\}$

Let us, in this section, consider the functional equation

$$f(x) = \underset{0 \leq y \leq x}{\text{Max}} \{g(y) + h(x - y) + f[ay + b(x - y)]\}, \qquad f(0) = 0. \tag{4.5}$$

We shall begin our discussion of (4.5) by proving

THEOREM 4.1. *If*

$$\begin{aligned}
&\text{(a)} \qquad g(0) = h(0) = 0, \\
&\text{(b)} \qquad g'(x), h'(x) > 0, \qquad g''(x), h''(x) \geq 0, \qquad \textit{for } x > 0, \\
&\text{(c)} \qquad \sum_{n=0}^{\infty} g(c^n x) < \infty, \qquad \sum_{n=0}^{\infty} h(c^n x) < \infty,
\end{aligned} \tag{4.6}$$

where $c = \text{Max}\,[a, b]$, *then the optimal policy consists in choosing* $y = 0$ *or* x *at each stage.*

PROOF. Let $f(x)$ be defined as above, and define, in addition,

$f_N(x)$ = return obtained using an optimal policy when only N stages are allowed. (4.7)

We have

$$f_1(x) = \underset{0 \leq y \leq x}{\text{Max}}\,[g(y) + h(x - y)], \tag{4.8}$$

and, generally,

$$f_{N+1}(x) = \underset{0 \leq y \leq x}{\text{Max}}\,\{g(y) + h(x - y) + f_N[ay + b(x - y)]\}, \tag{4.9}$$

for $N \geq 1$.

As we know from Section 2.3 of the chapter on existence and uniqueness theorems, the limit of $f_N(x)$ as $N \to \infty$ is $f(x)$, the unique solution to (4.5). Let us now demonstrate that the hypotheses of (4.6) yield the result that $f_N(x)$ is monotone increasing and convex.

To establish the result for $N = 1$, we observe that the convexity of g and h yields the convexity of $g(y) + h(x - y)$ as a function of y in $[0, x]$. Hence, the maximum is attained at an end point and

$$f_1(x) = \text{Max}\,[g(x), h(x)], \tag{4.10}$$

which is monotone increasing and convex. Let us now argue inductively. If the result has been established for N, it follows that $g(y) + h(x - y) + f_N[ay + b(x - y)]$ is convex and thus that its maximum occurs at $y = 0$ or x. Hence,

$$f_{N+1}(x) = \text{Max}\,[g(x) + f_N(ax), h(x) + f_N(bx)], \tag{4.11}$$

which shows that $f_{N+1}(x)$ is monotone increasing and convex.

Letting $N \to \infty$, we see that

$$f(x) = \text{Max}\,[g(x) + f(ax), h(x) + f(bx)], \tag{4.12}$$

which shows that the optimal policy is to choose $y = 0$ or x. Alternatively, we could use the convexity of $f(x)$, obtained as a limit of convex functions $f_N(x)$, to establish (4.12) directly from (4.5).

Prior to a further study of (4.12), we shall establish a similar result for the equation in (4.6).

4.3. A Result Concerning Equation (4.2)

Consider the following problem: We are given initially x dollars and a quantity y of a serum, together with the prerogative of purchasing additional amounts of the serum at specified times $t_1 < t_2 < \cdots$. At the kth purchasing opportunity, t_k, a quantity $c_k z$ of serum may be purchased for z dollars, where c_k is a monotone-increasing function of k. Given the probability that an epidemic occurs between t_k and t_{k+1}, and the condition that if an epidemic occurs we may only use the amount of serum on hand, the problem is to determine the purchasing policy that maximizes the over-all probability of successfully combating an epidemic, given the probability of success with a quantity w of serum available.

The condition $c_k > c_{k-1}$ is imposed to indicate the cheaper cost of serum at a later date because of technological improvement. Let

p_k = probability that the epidemic occurs between t_k and t_{k+1}, assuming that it has not occurred previously,

$\phi(w)$ = probability of combating the epidemic successfully with a quantity w of serum,

$f_k(x, y)$ = over-all probability of success using an optimal purchasing policy from t_k on, given x dollars and a quantity y of serum on hand. (4.13)

Invoking the principle that an optimal policy must possess an optimal continuation after any initial action, we obtain in the usual manner the functional equation

$$f_k(x, y) = \operatorname*{Max}_{0 \le z \le x} [p_k \phi(y + c_k z) + (1 - p_k) f_{k+1}(x - z, y + c_k z)]. \qquad (4.14)$$

In order to state the following result in simple form, let us assume that $p_k = p$. We have then

THEOREM 4.2. *If*

(a) $\phi(0) = 0$,

(b) $\phi(w)$ *is monotone increasing and convex, for all values of w that occur,* (4.15)

then the optimal policy consists in purchasing no serum at $t_1, t_2, \cdots, t_{k-1}$ and in using all available money at t_k, where k is chosen so as to maximize

$$[1 - (1 - p)^{k-1}]\phi(y) + (1 - p)^{k-1}\phi(y + c_k x). \qquad (4.16)$$

The proof is obtained in very much the same manner as above, employing the function

$f_{k,n}(x, y)$ = over-all probability of success using an optimal purchasing policy from t_k on, given x dollars and a quantity y of serum on hand and exactly n subsequent purchasing times, (4.17)

which satisfies the functional equation

$$f_{k,n}(x, y) = \underset{0 \le z \le x}{\text{Max}} \; [p\phi(y + c_k z) + (1 - p) f_{k+1,n-1}(x - z, y + c_k z)]. \tag{4.18}$$

The case in which this convexity property fails is much more difficult to resolve and will usually require a purchase of a certain quantity of serum at each purchasing point. A simpler problem of this type is discussed in more detail in Section 4.5.

4.4. The Equation $f(x) = \text{Max}\,[g(x) + f(ax), h(x) + f(bx)]$

Let us now turn to a discussion of the equation

$$f(x) = \text{Max}\,[g(x) + f(ax), h(x) + f(bx)]. \tag{4.19}$$

It is difficult to obtain any analytic representation for the function $f(x)$ or any description of the optimal policy, unless one makes some further assumptions concerning g and h. We shall pursue the analysis to the point where these assumptions are required and then illustrate the general method of attack by proving the following result:

THEOREM 4.3. *The solution of*

$$f(x) = \text{Max}\,[cx^d + f(ax), ex^f + f(bx)], \qquad f(0) = 0, \tag{4.20}$$

subject to

$$\begin{aligned} &\text{(a)} \qquad 0 < a, b < 1, \qquad c, d > 0, \\ &\text{(b)} \qquad 0 < d < f, \end{aligned} \tag{4.21}$$

is given by

$$\begin{aligned} f(x) &= cx^d + f(ax), \qquad 0 \le x \le x_0 \\ &= ex^f + f(bx), \qquad x_0 \le x, \end{aligned} \tag{4.22}$$

where

$$x_0 = \left[\frac{\dfrac{c}{1 - a^f}}{\dfrac{e}{1 - b^d}}\right]^{1/(f-d)}. \tag{4.23}$$

We shall represent by A the operation of choosing $g(x) + f(ax)$ and by B the operation of choosing $h(x) + f(bx)$. A solution corresponding to an optimal sequence of choices may be represented symbolically by

$$S = A^{a_1}B^{b_1}A^{a_2}B^{b_2}\cdots, \tag{4.24}$$

where the a_i and b_i are positive integers or zero.

We suspect from our previous work that the x values where

$$AB + \text{optimal continuation} = BA + \text{optimal continuation} \tag{4.25}$$

will play an important role in determining the solution. If A and then B is used, we obtain

$$f(x) = g(x) + h(ax) + f(abx), \tag{4.26}$$

while B and then A yields

$$f(x) = h(x) + g(bx) + f(abx). \tag{4.27}$$

The equation corresponding to (4.21) is then

$$g(x) + h(ax) = h(x) + g(bx). \tag{4.28}$$

Let us now make the simplifying assumption that this equation has exactly one non-zero solution, $\bar{x}$, and that $AB > BA$ for $0 \leq x < \bar{x}$ and that $AB < BA$ for $x \geq \bar{x}$. Without some condition of this type it seems very difficult to obtain a general solution.

Let us now show that either A^∞ or B^∞—that is to say, A or B repeated indefinitely—is the optimal sequence in $[0, \bar{x}]$. Let

$$S_1 = B^{b_1}A^{a_2}B^{b_2}\cdots, \tag{4.29}$$

be an optimal sequence for some x in $[0, \bar{x}]$. This may be written

$$B^{b_1-1}(BA)A^{a_2-1}B^{b_2}\cdots. \tag{4.30}$$

Since the result of applying B is to decrease x, after $(b_1 - 1)$ applications of B, the point x will still be in the interval $[0, x]$. In this interval, BA plus optimal continuation is inferior to AB plus optimal continuation. It follows, therefore, that S_1 is majorized by

$$S_2 = B^{b_1-1}(AB)A^{a_2-1}B^{b_2}\cdots. \tag{4.31}$$

If $b_1 - 1 \neq 0$, we may continue in this way until we arrive at an optimal sequence for which A is a first move, provided that b_1 is not ∞, which is equivalent to saying: provided that A is used at all.

We see then that A is either used first or not at all. It follows that it is only necessary to compare B^∞ and A^kB^∞, of which a special case is A^∞, in $[0, \bar{x}]$. The return corresponding to B^∞ is

$$H(x) = h(x) + h(bx) + \cdots, \tag{4.32}$$

whereas A^kB^∞ yields

$$G_k(x) = g(x) + \cdots + g(a^{k-1}x) + H(a^kx). \tag{4.33}$$

If for $0 \leq x \leq x_1$ we have $H(x) > g(x) + H(ax) = G_1(x)$, then clearly $H(x) > g(x) + [g(ax) + H(a^2x)] = G_2(x)$.

In order to continue, we must now make an assumption concerning the solutions of $H(x) = g(x) + H(ax)$ and similarly of the equation $G(x) = h(x) + G(bx)$. Imposing the condition that there are unique non-zero solutions, and proceeding by a systematic enumeration of cases, we may obtain the solution to (4.19). In place of a detailed account of the results in the general case, let us consider the simpler equation represented by (4.20). The equation $AB = BA$ takes the simple form

$$\frac{cx^d}{1-a^f} = \frac{ex^f}{1-b^d}, \tag{4.34}$$

whose unique non-trivial solution is $x = x_0$, as given by (4.23). The equations $H(x) = g(x) + H(ax)$, $G(x) = h(x) + G(bx)$ turn out to be equivalent to (4.34).

Since we have assumed $f > d$, it is clear that A will be the best move for small x, where $0 \leq x \leq x_2$. If $x_2 < x < x_0$, and close enough to x_2 so that A and B both transform it into the interval $[0, x_0]$, it is clear that B at x implies that BA^∞ will be the optimal sequence, whereas A implies A^∞. Since $BA^\infty < A^\infty$ in $[0, x_0]$, it follows that B is not chosen at x. Continuing this procedure, we see that A is used throughout $[0, x_0]$. In exactly the same way we see that B is chosen for $x > x_0$.

A result of this type is useful for approximation purposes, since an increasing function of reasonably smooth growth can be approximated to some degree of accuracy by cx^d. Approximation of $g(x)$ by cx^d is equivalent to approximation of $\log g(x)$ by $\log c + d \log x$, and, finally, to $\log g(e^u)$ by a straight line $c_1 + d_1 u$.

Let us point out, finally, that the change of variable

$$x = e^u, \qquad f(e^u) = \phi(u), \tag{4.35}$$

converts (4.19) into the form

$$\begin{aligned} \phi(u) = \text{Max}\,[g_1(u) + \phi(u - a_1),\ h_1(u) + \phi(u - b_1)],& \\ \phi(-\infty) = 0, \qquad u > -\infty,& \end{aligned} \tag{4.36}$$

which is also an equation of an interesting type.

It would be of some interest to determine the simplest conditions upon g and h which would ensure that an optimal policy always has the simple form shown above.

4.5. The Functions g and h Both Concave

Let us now return to the equation

$$f(x) = \underset{0 \leq y \leq x}{\text{Max}} \{g(y) + h(x - y) + f[ay + b(x - y)]\}, \qquad f(0) = 0 \tag{4.37}$$

and assume that g and h are both concave increasing functions of x. The problem is now much more complex, and, in general, the optimal y will not be at end point.

We shall prove

THEOREM 4.4. *Let*

$$\begin{aligned} &\text{(a)} \qquad g(0) = h(0) = 0, \\ &\text{(b)} \qquad g'(x), h'(x) \geq 0 \quad \textit{for } x \geq 0, \\ &\text{(c)} \qquad g''(x), h''(x) \leq 0 \quad \textit{for } x \geq 0, \end{aligned} \tag{4.38}$$

and consider the sequence of approximations to f defined by

$$\begin{aligned} f_0(x) &= \underset{0 \leq y \leq x}{\text{Max}} [g(y) + h(x - y)] \\ f_{n+1}(x) &= \underset{0 \leq y \leq x}{\text{Max}} \{g(y) + h(x - y) + f_n[ay + b(x - y)]\}, \\ &\qquad n = 0, 1, 2, \cdots. \end{aligned} \tag{4.39}$$

For each n, there is a unique $y_n = y_n(x)$ that yields the maximum. If $b \leq a$, we have $y_1 \leq y_2 \leq y_3 \cdots$, and the reverse inequalities for $b \geq a$. In particular, if $y_n(x) = x$ for some n in the case $b \leq a$, then $y_m(x) = x$ for $m \geq n$, and the solution of the original equation in (4.37) *will be furnished by $y = x$.*

This result is important in connection with determining approximate solutions, since it is quite simple to determine numerically y_1, y_2, and even y_3.

We shall begin by assuming that all the maxima occur within the interval $[0, x]$ and shall then consider the case in which one $y_n(x) = x$. Considering the function $f_1(x)$, we see that its maximum, y, is determined by the equation

$$g'(y) = h'(x - y). \tag{4.40}$$

Since the left-hand side is monotone increasing and the right-hand side is monotone decreasing, there is at most one solution. If we assume $h'(x) > g'(0)$, $g'(x) > h'(0)$, there will be exactly one solution of (4.40), which we call $y_1 = y_1(x)$. Differentiating (4.40), we obtain

$$y_1' g''(y_1) = (1 - y_1')h''(x - y_1), \tag{4.41}$$

which yields

$$y_1' = \frac{h''(x - y_1)}{g''(y_1) + h''(x - y_1)} > 0, \tag{4.42}$$

and

$$1 - y_1' > 0. \tag{4.43}$$

Turning to the expression for f, we have

$$f_1(x) = g(y_1) + h(x - y_1), \tag{4.44}$$

whence

$$f_1'(x) = g'(y_1)y_1' + (1 - y_1')h'(x - y_1) = h'(x - y_1), \tag{4.45}$$

using (4.40). Thus, $f_1'(x) > 0$ and $f_1''(x) = (1 - y_1')h''(x - y_1) < 0$, which means that $f_1(x)$ is concave.

Let us now turn to the function $f_2(x)$,

$$f_2(x) = \operatorname*{Max}_{0 \leq y \leq x} \{g(y) + h(x - y) + f_1[ay + b(x - y)]\}. \tag{4.46}$$

Assuming that there is a maximum inside the interval, we obtain

$$g'(y) - h'(x - y) + (a - b)f_1'[ay + b(x - y)] = 0, \tag{4.47}$$

which we write

$$g'(y) + (a - b)f_1'[ay + b(x - y)] = h'(x - y). \tag{4.48}$$

The left-hand side is again strictly decreasing and the right-hand side strictly increasing, so that there is at most one solution which we call $y_2 = y_2(x)$, if it exists. Note that if there is no solution of (4.48), then

$$f_2(x) = g(x) + f_1(ax). \tag{4.49}$$

Let us, however, assume that there is a solution. Then,

$$f_2(x) = g(y_2) + h(x - y_2) + f_1[ay_2 + b(x - y_2)], \tag{4.50}$$

whence, as above, using (4.48),

$$f_2'(x) = h'(x - y_2) + bf_1'[ay_2 + b(x - y_2)]. \tag{4.51}$$

Using (4.48) again, this may be written

$$f_2' = \frac{ah'(x - y_2) - bg'(y_2)}{a - b}. \tag{4.52}$$

This procedure is perfectly general, and we obtain, under our assumption concerning the existence of an internal maximum,

$$f_n' = \frac{ah'(x - y_n) - bg'(y_n)}{a - b}, \qquad n = 1, 2, 3, \cdots. \tag{4.53}$$

We now wish to show that if $b < a$, then $y_1 \leq y_2 \leq \cdots$, and, conversely, if $a < b$, that $y_1 \geq y_2 \geq \cdots$. The two cases are really one, since we may interchange the roles of y and $x - y$ if we so wish. Since $f_1' > 0$, we see, on comparing (4.48) and (4.40), that $y_1 < y_2$.

The equation for y_3 is

$$g'(y) + (a - b)f_2'[ay + b(x - y)] = h'(x - y). \tag{4.54}$$

If we can show that $f_2'(x) > f_1'(x)$, the same argument as that for y_1, y_2 shows that $y_3 > y_2$. Comparing (4.45) and (4.51), we see that $f_2' > f_1'$, since $h'(x - y_2) > h'(x - y_1)$.

To obtain the result for general n, always assuming that the maxima occur at inner points, we use (4.53). We know that $f_n'(x) > f_{n-1}'(x)$ implies that $y_{n+1} > y_n$. Since the function

$$r(y) = \frac{ah'(x - y) - bg'(y)}{a - b} \tag{4.55}$$

is monotone increasing in y and $y_n > y_{n-1}$, via an inductive hypothesis, it follows that $f_n' > f_{n-1}'$ and thus that $y_{n+1} > y_n$.

Let us now consider the situation in which some $y_n(x) = x$. If $n = 1$, it is easy to see that $y_n(x) = x$, $n \geq 1$, since $y_1(x) = x$ means that $g'(y) \geq h'(x - y)$ for $0 \leq y \leq x$. Since

$$\begin{aligned} &\frac{\partial}{\partial y}\{g(y) + h(x - y) + f_1[ay + b(x - y)]\} \\ &= g'(y) - h'(x - y) + (a - b)f_1'[ay + b(x - y)] \end{aligned} \tag{4.56}$$

and $a \geq b$, we see that this expression is positive if $g'(y) \geq h'(x - y)$ for $0 \leq y \leq x$.

Hence, $y_2(x) = x$, and, similarly, $y_n(x) = x$.

Let us now take the case in which $y_2(x) = x$, $y_1(x) \neq x$. Since $y_2(x) = x$ implies that $g'(y) - h'(x-y) + (a-b)f_1'[ay + b(x-y)] \geq 0$ for all $0 \leq y \leq x$, we have, in particular, $g'(x) - h'(0) + (a-b)f_1'(ax) \geq 0$. Since

$$\begin{aligned} f_2'(x) &= g'(x) + af_1'(ax) \\ &= g'(x) - h'(0) + (a-b)f_1'(ax) + h'(0) + bf_1'(ax) \\ &\geq h'(0) \end{aligned} \tag{4.57}$$

and $f_1'(x) = h'(x - y_1) \leq h'(0)$, we see that $f_2'(x) \geq f_1'(x)$. This, as above, implies that $y_3 \geq y_2 = x$, and the process continues.

Let us note, finally, that if $g'(y) \geq h'(x-y)$ for all y in $[0, x]$, then $g'(y) \geq h'(z-y)$ for y in $[0, z]$ for all $z \leq x$.

In closing this discussion of the functional equation, let us observe that if an interior maximum exists, we must have

$$g'(y) - h'(x-y) + (a-b)f'[ay + b(x-y)] = 0, \tag{4.58}$$

and

$$f'(x) = h'(x-y) + bf'[ay + b(x-y)]. \tag{4.59}$$

These equations may be solved explicitly for y and $f(x)$ if g and h are quadratic. This particular solution also furnishes a useful approximation to the solution of the general case.

4.6. The Equation $f(x) = \max_{0 \leq y \leq x} [g(y) + h(x-y) + \int_0^y f(y-s)k(s)\,ds]$

As another application of the techniques we have developed, let us now consider the functional equation

$$f(x) = \max_{0 \leq y \leq x} \left[g(y) + h(x-y) + \int_0^y f(y-s)k(s)\,ds \right], \tag{4.60}$$

where we shall assume

$$\begin{array}{ll} \text{(a)} & g(0) = h(0) = 0, \\ \text{(b)} & g'(y) > 0, \quad h'(y) > 0, \quad g'(0) < h'(0), \\ \text{(c)} & k(s) \geq 0, \\ \text{(d)} & g''(y) > 0, \quad h''(y) > 0, \\ \text{(e)} & h(y) - g(y) \text{ is monotone increasing in } y. \end{array} \tag{4.61}$$

We shall use the successive approximations defined by

$$\begin{aligned} f_0(x) &= \max_{0 \leq y \leq x} [g(y) + h(x-y)], \\ f_{n+1}(x) &= \max_{0 \leq y \leq x} \left[g(y) + h(x-y) + \int_0^y f_n(y-s)k(s)\,ds \right]. \end{aligned} \tag{4.62}$$

Let us consider $f_0(x)$. For x small the maximum is attained at $y = 0$, since $g'(0) < h'(0)$. Furthermore, we see that $g(y) + h(x - y)$ is monotone decreasing for small x, as a function of y. Since $g'(x)$ surpasses $h'(0)$ for x large, $g(y) + h(x - y)$ will eventually possess a turning point that is a minimum (see Fig. 4.1). The maximum will

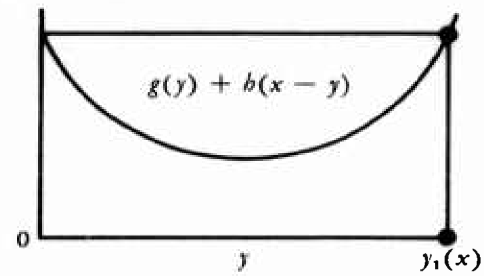

Fig. 4.1

stay at $y = 0$ until the point x_1, where $g(x) = h(x)$, at which point $y = 0$ and $y = x_1$ yield the same value for $g(y) + h(x - y)$. There is only one turning point, since $g'(y) - h'(x - y) = 0$ can have only one solution for $0 \leq y \leq x$. This is a consequence of our assumption that $g''(y) > 0$, $h''(y) > 0$, which means that $g'(y)$ is monotone increasing, whereas $h'(x - y)$ is monotone decreasing. It follows that

$$\begin{aligned} f_0 &= h(x), \qquad 0 \leq x \leq x_1 \\ &= g(x), \qquad x_1 < x \\ &= \operatorname{Max}\,[g(x), h(x)]. \end{aligned} \tag{4.63}$$

Consider the function f_1. We have

$$f_1 = \operatorname*{Max}_{0 \leq y \leq x} \left[g(y) + h(x - y) + \int_0^y f_0(y - s)k(s)\,ds \right]. \tag{4.64}$$

The function $\Psi_2(y) = g(y) + h(x - y) + \int_0^y f_0(y - s)k(s)\,ds$ is monotone decreasing for small y and possesses turning points for the y values satisfying

$$g'(y) + \int_0^y f_0'(y - s)k(s)\,ds = h'(x - y). \tag{4.65}$$

Since $f_0(x)$ is again a convex function, we see that the left-hand side of (4.65) is monotone increasing for $0 \leq y \leq x$, whereas the right-hand side is monotone decreasing. Hence, there is again one solution at most. Let x_2 be the first value of x for which

$$h(x) = g(x) + \int_0^x f_0(x - s)k(s)\,ds. \tag{4.66}$$

Since $f_0 > 0$, $k(s) \geq 0$, we have $x_2 < x_1$.

Furthermore, since $g(x) - h(x) + \int_0^x f_0(x - s)k(s)\,ds$ is monotone increasing, the solution takes the form

$$\begin{aligned} f_1(x) &= \operatorname{Max}\left[h(x), g(x) + \int_0^x f_0(x-s)k(s)\,ds\right] \\ &= h(x) \qquad \text{for } 0 \le x < x_2 \\ &= g(x) + \int_0^x f_0(x-s)k(s)\,ds \qquad \text{for } x > x_2. \end{aligned} \tag{4.67}$$

From this it is clear that $f_1(x)$ is again convex. In exactly the same fashion we obtain

$$\begin{aligned} f_n &= h(x), \qquad 0 \le x < x_n \\ &= g(x) + \int_0^x f_{n-1}(x-s)k(s)\,ds, \qquad x > x_n. \end{aligned} \tag{4.68}$$

Since $f_1 > f_0$, we obtain $f_{n+1} > f_n$ and $x_{n+1} < x_n < \cdots x_0$. The numbers x_n are monotone increasing and approach a limit $\bar{x}$. Since f_n converges to $f(x)$, the solution of (4.60), we obtain

$$\begin{aligned} f &= h(x), \qquad 0 \le x < \bar{x} \\ &= g(x) + \int_0^x f(x-s)k(s)\,ds, \qquad x \ge \bar{x}. \end{aligned} \tag{4.69}$$

This proves that $\bar{x}$ does not equal zero, since $f(x) = h(x)$ for a small positive interval about 0, as we see on comparing

$$g(x) + \int_0^x f(x-s)k(s)\,ds = g(x) + 0(x^2) \tag{4.70}$$

with $h(x)$ for x small.

The number $\bar{x}$ is determined as the non-zero root of

$$h(\bar{x}) = g(\bar{x}) + \int_0^{\bar{x}} h(\bar{x}-s)k(s)\,ds. \tag{4.71}$$

4.7. The Optimal Inventory Equation

Consider the functional equation

$$u(x) = \operatorname*{Min}_{y \ge x}\left[g(y-x) + a\left\{[M + u(0)]e^{-by} + b\int_0^y e^{-bv}u(y-v)\,dv\right\}\right], \tag{4.72}$$

where we assume

$$\begin{aligned} &\text{(a)} \qquad g(0) = 0, \quad g'(y) \ge 0, \quad g''(y) \ge 0, \\ &\text{(b)} \qquad b, M > 0, \quad 0 < a < 1. \end{aligned} \tag{4.73}$$

We shall approximate to u by means of the sequence

$$u_0(x) = a[M + u_0(0)]e^{-bx} + b \int_0^x e^{-bv} u_0(x - v)\, dv,$$

$$\begin{aligned} u_{n+1}(x) = \operatorname*{Min}_{y \geq x} \Big[g(y - x) + a \Big\{ [M + u_n(0)] e^{-by} \\ + b \int_0^y e^{-bv} u_n(y - v)\, dv \Big\} \Big]. \end{aligned} \tag{4.74}$$

The function $u_0(x)$ obtained by setting $y = x$ for all x corresponds to a policy of never ordering.

Let us now determine some of the important properties of $u_0(x)$. Using (4.74) and setting $x = 0$, we obtain

$$u_0(0) = \frac{aM}{1 - a}. \tag{4.75}$$

Thus, the equation for u_0 takes the form

$$u_0(x) = \frac{aMe^{-bx}}{1 - a} + b \int_0^x e^{-bv} u_0(x - v)\, dv, \tag{4.76}$$

which is a simple representative of a renewal equation and may be solved explicitly. For our purposes, however, there is no need of this, since the properties we require may be obtained directly from (4.76). Since the solution may be obtained quite easily, we note that it is given by

$$u_0 = \frac{M}{b(1 - a)} e^{-bx}(e^{abx} - 1) + \frac{aM}{1 - a} e^{-bx}. \tag{4.77}$$

Referring to (4.76), we have

$$u_0'(x) = -abMe^{-bx} + ab \int_0^x e^{-bv} u_0'(x - v)\, dv, \tag{4.78}$$

which shows inductively or by direct solution via iteration that $u_0'(x) < 0$. Furthermore, $u'(0) = -abM$. Differentiating again, we obtain

$$\begin{aligned} u_0''(x) &= ab^2Me^{-bx} + abe^{-bx}u_0'(0) + ab \int_0^x e^{-bv} u_0''(x - v)\, dv \\ &= ab^2M(1 - a)e^{-bx} + ab \int_0^x e^{-bv} u_0''(x - v)\, dv, \end{aligned} \tag{4.79}$$

which shows that $u_0''(x) > 0$, again directly by iteration or inductively.

Consider the function

$$\begin{aligned} u_1(x) &= \operatorname*{Min}_{y \geq x} \Big[g(y - x) + a \Big\{ [M + u_0(0)]e^{-by} + b \int_0^y e^{-bv} u_0(y - v)\, dv \Big\} \Big] \\ &= \operatorname*{Min}_{y \geq x} [g(y - x) + u_0(y)]. \end{aligned} \tag{4.80}$$

Since $g(0) = 0$, it is clear that $u_1(x) \leq u_0(x)$ for all $x \geq 0$. Now consider the function $\Psi_0(y, x) = g(y - x) + u_0(y)$. Setting $\partial\Psi_0/\partial y = 0$, we obtain

$$g'(y - x) = -u_0'(y), \tag{4.81}$$

and we are interested only in solutions $y \geq x$. By hypothesis, the function $g'(y - x)$ is monotone increasing in y, whereas (4.79) shows that $-u_0'(y)$ is monotone decreasing.

Let us assume that $g'(0) > 0$. Let x_1 be the value of x such that $g'(0) = -u_0'(x)$ (see Fig. 4.2). If there is no solution, which is to say

$$g'(0) > abM, \tag{4.82}$$

then $\partial\Psi_0/\partial y$ is positive for all $y \geq x$ and all x and the minimum occurs at $y = x$. This

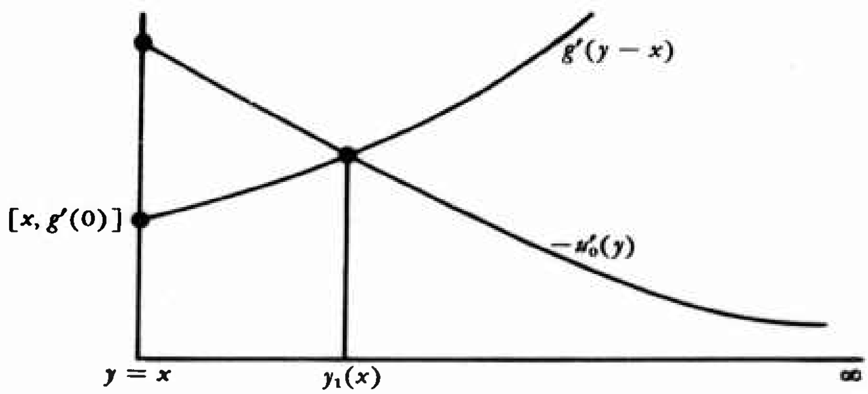

Fig. 4.2

means that the solution of (4.72) is $u(x) \equiv u_0(x)$. Let us assume, then, that $abM > g'(0)$. For $x \geq x_1$, $\partial\Psi_0/\partial y \geq 0$, and the optimal $y = x$. For $x < x_1$ there is a unique x_1 determined by (4.81), which is equivalent to

$$g'(y - x) = a\left[Mbe^{-by} - b\int_0^y e^{-br}u_0'(y - v)\,dv\right]. \tag{4.83}$$

We have then

$$\begin{aligned} u_1(x) &= u_0(x), \qquad x \geq x_1 \\ &= g(y_1 - x) + a\Big\{e^{-by_1}[M + u_0(0)] \\ &\qquad\qquad + b\int_0^{y_1} u_0(y_1 - v)e^{-br}\,dv\Big\} \\ &= 0 \leq x < x_1. \end{aligned} \tag{4.84}$$

For $x \geq x_1$, we have $u_1' = u_0'$, while for $0 < x < x_1$ we obtain

$$\begin{aligned} u_1'(x) &= -g'(y_1 - x) + y_1'[g'(y_1 - x) + u_0'(y_1)] \\ &= -g'(y_1 - x) \\ &= u_0'(y_1) \end{aligned} \tag{4.85}$$

on referring to (4.81). Since $y_1 > x$, and $-u_0'(x)$ is increasing, we obtain

$$-u_1'(x) = -u_0'(y_1) < -u_0'(x). \tag{4.86}$$

From (4.85) we obtain $u_1''(x) = u_0''(y_1)y_1'$. Since $u_0''(y_1) > 0$, the sign of u_1'' depends on that of y_1'. From (4.81) we obtain $g''(y_1 - x)(y_1' - 1) = -u_0''(y)y_1'$, which shows that $0 \leq y_1' \leq 1$. Hence, $u_1'' \geq 0$.

Consider now the equation for $u_2(x)$:

$$\begin{aligned} u_2(x) = \operatorname*{Min}_{y \geq x} \Big[G(y - x) + a \Big\{ &[M + u_1(0)]e^{-by} \\ &+ b \int_0^y e^{-bv} u_1(y - v)\, dv \Big\} \Big]. \end{aligned} \tag{4.87}$$

Taking the partial derivative of the expression within the brackets with respect to y, we obtain

$$\frac{\partial \Psi_1}{\partial y} = g'(y - x) + ab \int_0^y e^{-bv} u_1'(y - v)\, dv - abMe^{-by}. \tag{4.88}$$

Setting this equal to zero, we obtain

$$g'(y - x) = abMe^{-by} - ab \int_0^y e^{-bv} u_1'(y - v)\, dv. \tag{4.89}$$

Let us consider the function

$$a\left[Mbe^{-by} - b \int_0^y e^{-bv} u_1'(y - v)\, dv \right] = \phi_2(y). \tag{4.90}$$

We have

$$\begin{aligned} \phi_2'(y) &= -ab^2Me^{-by} - abe^{-by}u_1'(0) - ab \int_0^y e^{-nv} u_1''(y - v)\, dv \\ &= e^{-by}ab[-bM - u_1'(0)] - ab \int_0^y e^{-b} u_1''(y - v)\, dv. \end{aligned} \tag{4.91}$$

Since $-u_1'(0) \leq -u_0'(0) = abM$, the quantity $-bM - u_1'(0)$ is negative. Hence, $\phi_2'(y) < 0$. Thus, there is one solution at most of the equation in (4.89).

Since $-u_1'(x) \leq -u_0'(x)$ for all x, the curve $\phi_2(y)$ lies below the curve $\phi_1(y) = abMe^{-by} - ab \int_0^y e^{-bv} u_0'(y - v)\, dv$ for all y. Therefore, the intersection of ϕ_2 with $g'(y - x)$ always lies to the left of the intersection of $g'(y - x)$ with $\phi_1(y)$, as shown in Fig. 4.3.

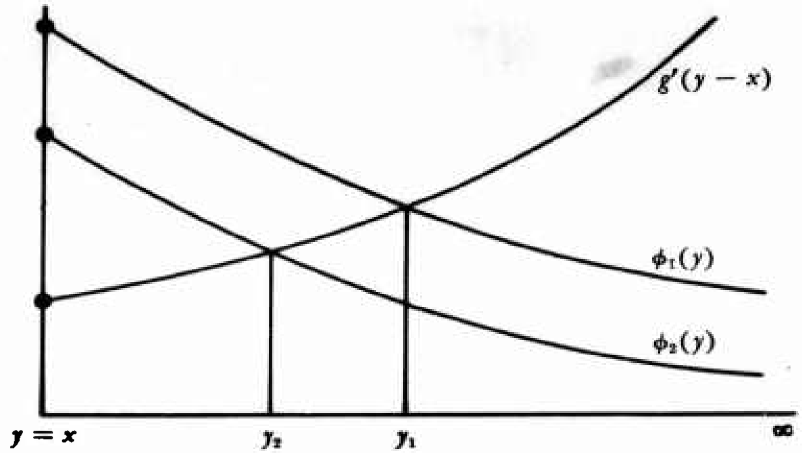

Fig. 4.3

Let x_2 be the solution of $g'(0) = \phi_2(x)$. Then $x_2 < x_1$. Therefore,

$$\begin{aligned} u_2 &= a\left\{e^{-bx}[M + u_0(0)] + b\int_0^x u_1(x-v)e^{-bv}\,dv\right\}, \qquad x \geq x_2 \\ &= g(y_2 - x) + \Psi_2(y_2, x), \qquad 0 \leq x < x_2. \end{aligned} \tag{4.92}$$

From this we conclude that

$$\begin{aligned} u_2'(x) &= -abMe^{-bx} + ab\int_0^x u_1'(x-v)e^{-bv}\,dv, \qquad x > x_2 \\ &= -g'(y_2 - x), \qquad 0 \leq x \leq x_2. \end{aligned} \tag{4.93}$$

Comparing the expressions for u_1' and u_2' for $x > x_1$ and for $0 \leq x \leq x_2$, we readily conclude that $-u_2'(x) \leq -u_1'(x)$. For $x_1 \leq x \leq x_2$, we note that

$$g'(0) \geq abMe^{-bx} - ab\int_0^x u_1'(x-v)e^{-bv}\,dv. \tag{4.94}$$

Since $g'(y_1 - x) > g'(0)$, we see that $-u_1'(x) \geq u_2'(x)$ for all x.

Let us now examine the convexity of $u_2(x)$. The difficult region is $0 \leq x \leq x_2$. Here we have, using (4.93),

$$u_2'' = -g''(y_2 - x)(y_2' - 1). \tag{4.95}$$

We see that the sign of u_2'' depends on that of $y_2' - 1$. Referring to (4.89), the equation which defines y_2, we have, differentiating with respect to x, with $y \equiv y_2$,

$$\begin{aligned} g''(y-x)(y'-1) = y'\Big[&-ab^2Me^{-by} - abe^{-by}u_1'(0) \\ &-ab\int_0^y e^{-bv}u_1''(y-v)\,dv\Big]. \end{aligned} \tag{4.96}$$

Since $u_1'' \geq 0$ and $-u_1'(0) \leq -u_0'(0) = abM$, we see that the coefficient of y' is negative. Since $g'' \geq 0$, this shows that $y' \geq 0$. Referring to (4.96), we see that this implies that $y' - 1 \leq 0$. Hence, $u_2'' \geq 0$.

We now have all the results required for an inductive proof. The expression for u_n is

$$u_n = zg(y_n - x) + a\left\{e^{-by_n}[M + u_{n-1}(0)] + b\int_0^{y_n} u_{n-1}(y_n - v)e^{-bv}\,dv\right\} \tag{4.97}$$

for $0 \leq x \leq x_n$, and

$$u_n = a\left\{e^{-bx}[M + u_{n-1}(0)] + b\int_0^{x} u_{n-1}(x - v)e^{-bv}\,dv\right\} \tag{4.98}$$

for $x \geq x_n$, where

$$\begin{aligned} 0 &\leq x_n \cdots \leq x_2 \leq x_1, \\ 0 &\leq y_n \cdots \leq y_2 \leq y_1. \end{aligned} \tag{4.99}$$

Using the monotone properties and letting $n \to \infty$, we obtain for $u(x)$ a representation

$$u = g[y(x) - x] + a\left\{e^{-by(x)}[M + u(0)] + \int_0^{y(x)} u[y(x) - v]e^{-bv}\,dv\right\} \tag{4.100}$$

for $0 \leq x \leq x_\infty$, and

$$u = a\left\{e^{-bx}[M + u(0)] + \int_0^{x} u(x - v)e^{-bv}\,dv\right\},$$

for $x \geq x_\infty$.

We now wish to show that $x_\infty \neq 0$. For small x and y we have

$$\begin{aligned} u(x) &= \operatorname*{Min}_{y \geq x}\left[g'(0)(y - x) + a\left\{[M + u(0)]e^{-by} + b\int_0^{y} e^{-bv}[u(0) + 0(y - v)]\right\}\right] \\ &= \operatorname*{Min}_{y \geq x}\,[g'(0)(y - x) + aM - abMy + 0(y^2)]. \end{aligned} \tag{4.101}$$

This shows that for small x the minimum is not at $y = x$.

4.8. The Solution for Linear Cost

Let us consider the simple case in which $g(y) = cy$, where c is a positive constant. The equation is now

$$u(x) = \underset{y\geq x}{\text{Min}}\left[c(y - x) + a\left\{[M + u(0)]e^{-by} + b \int_0^y e^{-bv}u(y - v)\, dv\right\}\right]. \quad (4.102)$$

For $0 \leq x < x_\infty$, we choose $y = y(x)$ where

$$c = abMe^{-by} + ab \int_0^y e^{-bv}u'(y - v)\, dv. \quad (4.103)$$

Since the equation is independent of x, the solution, which we know to be unique, is $y = x_\infty$.

The optimal policy is then to choose $y = x_\infty$ if $x < x_\infty$ and to choose $y = x$ if $x \geq x_\infty$. The function $u(x)$ satisfies the equation

$$\begin{aligned} u(x) &= a\left\{[M + u(0)]e^{-bv} + b \int_0^x e^{-bv}u(x - v)\, dv\right\}, \qquad x \geq x_\infty \\ &= c(x_\infty - x) + a\left\{[M + u(0)]e^{-bx_\infty} + b \int_0^{x_\infty} e^{-bv}u(x_\infty - v)\, dv\right\}, \\ &\qquad 0 \leq x \leq x_\infty. \end{aligned} \quad (4.104)$$

For $0 \leq x \leq x_\infty$, we have $u'(x) = -c$, whence $u(x) = u(0) - cx$ in $[0, x_\infty]$. Since x_∞ is determined by (4.103), we may use the second equation in (4.104) to find $u(0)$. Having determined the solution in $[0, x_\infty]$, the solution for larger x is found by solving the first equation in (4.104), a simple renewal equation.

CHAPTER 5

THE EQUATION $u(n) = \underset{1 \le i \le M}{\text{Max}} \left[\sum_{j=1}^{R} a_{ij} u(n-j) + c_i \right]$ AND RELATED TOPICS

5.1. Introduction

In this chapter we consider a number of functional equations that are more or less loosely connected. The first equation is

$$u(n) = \underset{1 \le i \le M}{\text{Max}} \left[\sum_{j=1}^{R} a_{ij} u(n-j) + c_i \right], \tag{5.1}$$

the homogeneous form of which was encountered in Section 1.6, Problem 1.11. After discussing the asymptotic behavior of the solutions of (5.1) for the case in which the a_{ij} are all non-negative, we shall discuss a problem in production planning that gives rise to the functional equation

$$f_N(x) = \text{Max}\,[f_{N-1}(Ax), f_{N-1}(Bx)], \qquad N = 1, 2, \cdots, \tag{5.2}$$

where x is a two-dimensional vector, A and B are 2×2 positive matrices, and $f_0(x) = c_1x_1 + c_2x_2$. This problem seems extremely difficult, and we are able only to contribute some partial results, which are, however, of interest in themselves.

We shall close with a solution of the simple testing equation

$$f(x) = \text{Min}\begin{bmatrix} 1 + xf(1) \\ 1 + f(ax) \end{bmatrix}, \qquad x > 0,$$
$$f(0) = 0. \tag{5.3}$$

5.2. The Equation $u(n) = \underset{1 \le i \le K}{\text{Max}}\ [\sum_{j=1}^{R} a_{ij} u(n-j)]$

We shall begin our discussion with the homogeneous equation

$$u(n) = \underset{1 \le i \le K}{\text{Max}} \left[\sum_{j=1}^{R} a_{ij} u(n-j) \right], \qquad n \ge R, \tag{5.4}$$

where $u(l)$ is a given non-negative quantity for $0 \le l \le R-1$.

Our first result is

THEOREM 5.1. *Consider equation* (5.4), *in which we assume that*

(a) $\quad a_{ij} \ge 0;$

(b) *there is one equation,* $r^R = \sum_{j=1}^R a_{kj} r^{R-j}$, *whose largest positive root is greater than the corresponding roots of the other equations of this type;*

(c) *for this index* k, $a_{k1} \neq 0$. (5.5)

Under these circumstances, the solution of (5.4) *is given by*

$$u(n) = \sum_{j=1}^{R} a_{kj} u(n-j), \tag{5.6}$$

for n *sufficiently large.*

PROOF. For the sake of simplicity, consider the third-order case with $k = 2$:

$$u(n+3) = \text{Max} \begin{Bmatrix} A\colon\ a_1 u(n+2) + b_1 u(n+1) + c_1 u(n) \\ B\colon\ a_2 u(n+2) + b_2 u(n+1) + c_2 u(n) \end{Bmatrix}, \tag{5.7}$$

where $u(0)$, $u(1)$, $u(2)$ are preassigned positive quantities. Let us assume that of the two equations

$$\begin{aligned} r^3 &= a_1 r^2 + b_1 r + c_1, \\ r^3 &= a_2 r^2 + b_2 r + c_2, \end{aligned} \tag{5.8}$$

it is the first that has the largest positive root, and let ρ be this root.

Let us first show inductively that

$$e\rho^n \leq u(n) \leq f\rho^n \tag{5.9}$$

for two positive constants e and f. Consider the lower inequality first. Let e be chosen so that the inequality is valid for $n = 0, 1, 2$. Then, since

$$u(n+3) \geq a_1 u(n+2) + b_1 u(n+1) + c_1 u(n), \tag{5.10}$$

we obtain

$$u(3) \geq e(a_1\rho^2 + b_1\rho + c) = e\rho^3, \tag{5.11}$$

and clearly an inductive argument yields the inequality for all n.

To obtain the upper inequality, we proceed similarly. The constant f may be chosen so that the upper inequality is valid for $n = 0, 1, 2$. Then

$$u(3) \leq \text{Max} \begin{bmatrix} f(a_1\rho^2 + b_2\rho + c_1) = f\rho^3 \\ f(a_2\rho^2 + b_2\rho + c_2) < f\rho^3 \end{bmatrix}, \tag{5.12}$$

where the last inequality is a consequence of the maximal property of ρ. It is again clear that an inductive argument yields the upper inequality.

To prove Theorem 5.1, we show that the assumption that B is employed infinitely often leads to an eventual contradiction of the lower inequality. If B is used for $m = n + 3$, $n \geq 0$, we have

$$\begin{aligned} u(n+3) &= a_2 u(n+2) + b_2 u(n+1) + c_2 u(n) \\ &\leq f\rho^n[a_2\rho^2 + b_2\rho + c_2]. \end{aligned} \tag{5.13}$$

Using the maximal property of ρ, we have $a_2\rho^2 + b_2\rho + c_2 \leq c_3\rho^3$, where $0 < c_3 < 1$. Hence, we obtain $u(n + 3) \leq c_3 f\rho^{n+3}$.

Now consider $u(n + 4)$. We obtain

$$
\begin{aligned}
u(n+4) &= \mathrm{Max}\begin{bmatrix} a_1u(n+3) + b_1u(n+2) + c_1u(n+1) \\ a_2u(n+3) + b_2u(n+2) + c_2u(n+1) \end{bmatrix} \\
&\leq \mathrm{Max}\begin{bmatrix} f\rho^{n+1}(a_1c_3\rho^2 + b_1\rho + c_1) \\ f\rho^{n+1}(a_2c_3\rho^2 + b_2\rho + c_2) \end{bmatrix} \\
&\leq fc_4\rho^{n+4},
\end{aligned} \tag{5.14}
$$

where $0 < c_4 < 1$, and the constant fc_4 is again independent of n. Observe that the condition $a_1 \neq 0$ is essential for our proof. In exactly the same fashion we find that $u(n + 5) \leq fc_5\rho^{n+5}$, $0 < c_5 < 1$. Having established the relation $u(m) \leq c_6 f\rho^m$ for $m = n + 3, n + 4, n + 5$, three consecutive values of m, where $c_6 = \mathrm{Max}\,(c_3, c_4, c_5)$, it follows from the recurrence relation (5.4), that this inequality is valid for all subsequent m.

We see then that the effect of employing B once is to reduce the constant f. It follows that a choice of B infinitely often will eventually lead to a contradiction of the lower bound $u(n) \geq e\rho^n$. Consequently, there is a number n_0 dependent on the coefficients and initial values such that for $n \geq n_0$, B is not employed. The proof given above enables one to obtain an upper bound for the number of times that B is employed. Combining this fact with the easily demonstrated fact that a choice of A for any three consecutive values of n implies that it is chosen for all larger values of n, we may obtain a number n_0 with the property that for $n \geq n_0$, A is always used.

The condition that $a_1 \neq 0$ is necessary for the truth of the result in general. It is not difficult to verify that if

$$
\begin{aligned}
u(n+2) &= \mathrm{Max}\begin{bmatrix} A\colon\ bu(n) \\ B\colon\ cu(n+1) + \epsilon\, u(n) \end{bmatrix}, \\
u(0) &= 1, \\
u(1) &= c + \epsilon,
\end{aligned} \tag{5.15}
$$

where $c_2 < b < c$ and ϵ is sufficiently small and positive, then the optimal pattern is B for odd n, A for even n.

A finer analysis will show in the general case that the optimal pattern is always eventually periodic.

The case in which at least two characteristic equations have the same maximum root is more difficult to handle. It is easily seen that for large n only those choices corresponding to largest roots will be used, and it is not difficult to show by consideration of the quantities

$$
\begin{aligned}
\phi(n) &= \mathrm{Min}\,[V(n), V(n+1), V(n+2)] \\
\Psi(n) &= \mathrm{Max}\,[V(n), V(n+1), V(n+2)],
\end{aligned} \tag{5.16}
$$

where $V(n) = u(n)\rho^{-n}$, the first of which is monotone increasing and the second monotone decreasing, that $u(n)\rho^{-n}$ approaches a limit as $n \to \infty$. However, any further information concerning asymptotic behavior seems to be difficult to obtain.

5.3. The Inhomogeneous Equation

Let us now consider the equation

$$u(n) = \operatorname*{Max}_{1\le i\le M}\left[\sum_{j=1}^{R} a_{ij}u(n-j) + g_i\right], \tag{5.17}$$

for the case in which $a_{ij} \geq 0$ and $g_i \geq 0$, where again the quantities $u(l)$, $0 \leq l \leq R-1$ are given positive constants. The most interesting case is that in which

$$\sum_{j=1}^{R} a_{ij} = 1 \qquad \text{for } i = 1, 2, \cdots, M.$$

Since each equation has largest characteristic root equal to 1, it is the forcing term that dominates the situation for large n.

From the theory of linear difference equations, it is known that the solution of any recurrence relation of the form

$$u(n) = \sum_{j=1}^{R} a_{ij}u(n-j) + g_i, \qquad n \geq R,$$
$$u(l) = c_l, \qquad 0 \leq l \leq R-1, \tag{5.18}$$

where $\sum_{j=1}^{R} a_{ij} = 1$, $a_{ij} \geq 0$ has the form

$$u(n) = \frac{ng_i}{\sum_{j=1}^{R} ja_{ij}} + d_i + 0(\alpha_i^n), \qquad 0 < \alpha_i < 1, \tag{5.19}$$

for large n, where d_i is a constant dependent on the initial conditions.

We should suspect, then, that the solution of (5.17) would be determined, for large n, by the index i for which $g_i/\sum_j ja_{ij}$ assumes its maximum. This is indeed true. We shall prove

THEOREM 5.2. *Let*

$$c = \operatorname*{Max}_{i} \frac{g_i}{\sum_{j=1} ja_{ij}} \tag{5.20}$$

be attained for the single value $i = s$. If $a_{s1} > 0$, the solution of (5.17) is given by

$$u(n) = \sum_{j=1}^{R} a_{sj}u(n-j) + g_s \tag{5.21}$$

for $n \geq n_0$, where n_0 is an integer dependent on the initial conditions and coefficients.

PROOF. Let us establish first the inequalities

$$nc - K \leq u(n) \leq nc + K \tag{5.22}$$

for all n, with a suitable choice of K. For $0 \leq n \leq R$, we may choose K so that the inequality is valid. Let us now establish it inductively for $n \geq R + 1$. We have

$$\begin{aligned} u(n) &\geq \sum_{j=1}^{R} a_{sj} u(n-j) + g_s \\ &\geq \sum_{j=1}^{R} a_{sj}[(n-j)c - K] + g_s \\ &= nc - K, \end{aligned} \tag{5.23}$$

using the value for c given in (5.20).

To establish the upper bound, we use the fact that if the ith choice is made at n, we have

$$\begin{aligned} u(n) &= \sum_{j=1}^{R} a_{ij} u(n-j) + g_i \\ &\leq \sum_{j=1}^{R} a_{ij}[(n-j)c + K] + g_i \\ &\leq nc + K + g_i - \sum_{j=1}^{R} j a_{ij} \\ &\leq nc + K, \end{aligned} \tag{5.24}$$

using the optimal property of the index s.

Following the same reasoning as that above, let us show that if any other choice than the sth is made at n, the upper bound will be decreased. As before, this will show that the index s must be selected for all large n.

Referring to (5.24), we see that if $i \neq s$, we have

$$u(n) \leq nc + K + g_i - \sum_{j=1}^{R} j a_{ij} < nc + k - d_1, \tag{5.25}$$

where $d_1 > 0$. Consider now the situation at $n + 1$. We obtain, for some i,

$$\begin{aligned} u(n+1) &= \sum_{j=a}^{R} a_{ij} u(n+1-j) + g_i \\ &\leq a_{i1}[nc + K - d_1] + \sum_{j=2}^{R} a_{ij}(1 + n - j)c + K + g_i \\ &\leq (n+1)c + K - a_{i1}d_1 + g_i - c\sum_{j=1}^{R} j a_{ij} \\ &\leq (n+1)c + K - d_2, \end{aligned} \tag{5.26}$$

where $d_2 > 0$, since if $i = s$, we have $a_{s1} > 0$, and if $i \neq s$, we have

$$g_i - c \sum_{j=1}^{R} j a_{ij} < 0.$$

Continuing in this fashion, we see that for $n, n+1, \cdots, n+R$ we obtain a positive constant d such that $f(n) \leq nc + K - d$, if $i \neq s$ is selected at n. Having established this upper bound for R consecutive values, it follows via induction that the inequality persists for all larger values. We observe then that a repeated application of choices different from s cannot yield an optimal policy, since eventually we shall obtain a contradiction to the lower inequality.

If there are several choices yielding the same c, the above argument shows that we may restrict ourselves to considering only these choices. The asymptotic behavior will be the same as above, namely, $u(n) \sim nc$, and the result of varying choices will be negligible. Nevertheless, it is an interesting open problem to determine the asymptotic form of the solution in this case.

5.4. A Class of Problems Arising in Production Planning

Let us consider the following simplified problem. We are given an initial stock x and y of two quantities A and B, and means of producing more of A and B using the initial amounts. Specifically, we may divide x into two parts, u_1 to be used to produce more A, and u_2 to be used to produce more B, and y into two corresponding parts, v_1 and v_2. The new amount of A will be $f(u_1, v_1)$ and that of B will be $g(u_2, v_2)$, where f and g are given functions. This operation is now to be repeated N times, and the general problem is that of maximizing $h(x_N, y_N)$, where h is a given function.

Problems of this type arise in planning production schedules where different techniques are applicable at each stage of production.

Since the mathematical problem in its above generality seems to be hopelessly beyond our reach, let us consider the simpler situation, where

$$\begin{aligned} f(u, v) &= a_1 u + a_2 v, \\ g(u, v) &= b_1 u + b_2 v, && (5.27) \\ h(x, y) &= c_1 x + c_2 y, && (5.28) \end{aligned}$$

and all the coefficients involved are non-negative.

Another criterion function of interest, which applies to "bottleneck" situations, is

$$h(x, y) = \text{Min}\,(x, y). \qquad (5.29)$$

We shall not, however, discuss any of these very interesting and important problems here.

If we define

$$\begin{aligned} W_0(x, y) &= c_1 x + c_2 y = h(x, y), \\ W_N(x, y) &= h(x_N, y_N), && (5.30) \end{aligned}$$

we obtain for W_N the functional equation

$$W_N(x, y) = \text{Max}\,[W_{N-1}(a_1x + a_2y, 0),\ W_{N-1}(0, b_1x + b_2y),$$
$$W_{N-1}(a_1x, b_2y),\ W_{N-1}(a_2y, b_1x)] \qquad (5.31)$$

for $N \geq 1$, since the linearity of $h(x, y)$ will force u_1 to be either 0 or x and, similarly, v_1 to be either 0 or y.

We see, therefore, that we are given four matrices,

$$A_1 = \begin{pmatrix} a_1 & a_2 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \\ b_1 & b_2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} a_1 & 0 \\ 0 & b_2 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & a_2 \\ b_1 & 0 \end{pmatrix}, \qquad (5.32)$$

and the problem is that of forming a vector,

$$\begin{pmatrix} x_N \\ y_N \end{pmatrix} = C_1C_2 \cdots C_N \begin{pmatrix} x \\ y \end{pmatrix}, \qquad (5.33)$$

where each C_i is an A_j, $j = 1, 2, 3, 4$, in such a way as to maximize the inner product of (x_N, y_N) with a given vector.

In this form the problem may readily be generalized. However, even in its simplest forms it seems extremely difficult. If we seek to determine not the actual optimal policy, in the general case, but merely the order of magnitude of $W_N(x, y)$, the problem is still difficult. In the succeeding section we shall present a preliminary result in this direction.

5.5. The Problems of Largest Characteristic Root

In this section we shall present a preliminary result for the following problem:

> Given a finite set $\{A_i\}$ of non-negative square matrices, determine for each N the matrix $C_N = B_1B_2 \cdots B_N$, where each B_i is an A_j, which possesses the largest characteristic root.

The problem has not been resolved even in the simplest non-trivial case of 2×2 matrices, where the set consists of two non-commuting matrices A and B.

Let us introduce the notation

$$\phi(A) = \text{characteristic root of } A \text{ of largest absolute value}. \qquad (5.34)$$

It is a classical result of Perron that $\phi(A)$ is positive if A is non-negative, unless all the characteristic roots of A are zero. To simplify the presentation we shall assume that the A_i are actually positive. We now prove

THEOREM 5.3. *$\lambda = \text{Lim}_{N\to\infty}\, \phi(C_N)^{1/N}$ exists.*

PROOF. Since C_N^2 is a possible candidate for C_{2N}, it follows that

$$\phi(C_N)^2 = \phi(C_N^2) \leq \phi(C_{2N}). \qquad (5.35)$$

Letting N run through the values $\{2^k\}$, $k = 0, 1, 2, \cdots$, we see that

$$\lambda(k) = \phi(C_{2^k})^{1/2^k} \qquad (5.36)$$

is a monotone-increasing function of k.

To show that it is bounded, let us introduce the majorant, M, of the matrices $\{A_i\}$, defined by the property that the klth element in M is the maximum of the set of klth elements occurring in the $\{A_i\}$. We shall employ the notation

$$A << B \tag{5.37}$$

to indicate that $a_{ij} \leq b_{ij}$ for all i and j. It is known that if A and B are positive matrices, then $B >> A$ implies $\phi(B) \geq \phi(A)$. The converse is, however, not true.

Since $M >> A_i$, we have $M^N >> C_N$, whence $\phi(C_N) \leq \phi(M^N) = \phi(M)^N$. Consequently, the sequence $\phi(C_N)^{1/N}$ is uniformly bounded for all N, and thus $\lambda(k)$ is uniformly bounded. Since $\lambda(k)$ is monotone increasing, $\text{Lim}\,\lambda(k)$ exists as $k \to \infty$, and we set

$$\lambda = \operatorname*{Lim}_{k \to \infty} \lambda(k). \tag{5.38}$$

It remains to show that $\phi(C_N)^{1/N}$ has a limit. Let k be a fixed large number and write, for $N > 2^k$, $N = 2^k q + r$, where $0 \leq r \leq 2^k - 1$, q and r being integers.

Since $(C_2k)^q C_r$ is a possible choice for C_N, it follows that $\phi(C_N) \geq \phi(C_2k)^q C_r$. Since $C_r >> aI$, where I is the identity matrix, for some $a > 0$, we have $(C_2k)^2 C_r >> a(C_2k)^q$. Hence,

$$\phi(C_N) \geq a\phi(C_2k)^q, \tag{5.39}$$

or

$$\phi(C_N)^{1/N} \geq a^{1/N}\phi(C_2k)^{q/(2^k q+r)} \geq a^{1/N}[\phi(C_2k)^{1/2^k}]^{q/q+1}. \tag{5.40}$$

Letting $N \to \infty$, we have

$$\underline{\text{Lim}}\,\phi(C_N)^{1/N} \geq \phi(C_2k)^{1/2^k}. \tag{5.41}$$

Since this holds for every k, we obtain, finally, $\underline{\text{Lim}}\,\phi(C_N)^{1/N} \geq \lambda$.

In the above proof we have used positivity only in the statement $C_r >> aI$. A finer analysis based on the asymptotic form of $(C_2k)^q$ for large q will show that non-negativity is sufficient.

To obtain the inequality $\overline{\text{Lim}}\,\phi(C_N)^{1/N} \leq \lambda$, we write $2^k = qN_m + r$, where $\{N_m\}$ is a sequence on which $\overline{\text{Lim}}$ is obtained. Then consider the matrix $C^q_{N_m} C_r$. We have, as before, $\phi(C_2k) \geq a\phi(C_{N_m})^q$, whence

$$\phi(C_2k)^{1/2^k} \geq a^{1/2^k}[\phi(C_{N_m})^{1/N_m}]^{q/q+1}. \tag{5.42}$$

Letting $k \to \infty$, we obtain

$$\lambda \geq \phi(C_{N_m})^{1/N_m} \tag{5.43}$$

and thus $\lambda \geq \overline{\text{Lim}}$. Combining the two inequalities we obtain equality.

In very much the same fashion, we may prove

THEOREM 5.4. *Let M_N denote the smallest majorant of all the products $B_1B_2 \cdots B_N$ where each B_i is an A_j. Then*

$$\lim_{N\to\infty} \phi(M_N)^{1/N} = \mu. \tag{5.44}$$

It is immediate that $\mu \geq \lambda$, and it is conjectured on the basis of no evidence pro or con that $\mu = \lambda$.

5.6. A Testing Problem

In this section we shall present the solution to Problem 1.3, posed in Section 1.2.

Let $f(x)$ equal the expected time consumed using an optimal procedure. Then

$$f(x) = \text{Min} \begin{Bmatrix} L\colon\ 1 + xf(1) \\ A\colon\ 1 + f(ax) \end{Bmatrix}, \qquad 1 \geq x > 0, \qquad 0 \leq a < 1. \tag{5.45}$$

For x close to zero it is clear that $f(x) > xf(1)$, since $f(x) > 1$ for all x. Therefore, in some interval $[0, x_0]$ we have

$$f(x) = 1 + xf(1), \tag{5.46}$$

where $f(1)$ is some, as yet undetermined, constant.

In $[x_0, x_0/a]$ we obtain

$$f(x) = \text{Min} \begin{Bmatrix} 1 + xf(1) \\ 1 + [1 + axf(1)] \end{Bmatrix}. \tag{5.47}$$

Hence, we must compare $xf(1)$ with $1 + axf(1)$. If we assume that $1 + axf(1) \leq xf(1)$ for $x_0 \leq x \leq x_0/a$, we turn to the next interval $[x_0/a, x_0/a^2]$, and so on. Since, eventually, $x_0/a^k > 1$ if $x_0 \neq 0$, we must in this way either cover the interval $[0, 1]$ or obtain a point x_1 where $1 + xf(1) > 1 + f(ax)$. This certainly is true at $x = 1$, since

$$f(1) = \text{Min} \begin{Bmatrix} 1 + f(1) \\ 1 + f(a) \end{Bmatrix} = 1 + f(a). \tag{5.48}$$

Let us show that A is used to the right of of x_2, where x_2 is the first point at which

$$x_2 f(1) = f(ax_2). \tag{5.49}$$

If L is employed for $x_2/a > x > x_2$, we have

$$f_L(x) = 1 + xf(1) = 1 + x[1 + f(a)]. \tag{5.50}$$

If A is used, we have

$$f_A(x) = 1 + f(ax) = 1 + [1 + axf(1)], \tag{5.51}$$

since $ax < x_2$, which means that L is used there. At $x = x_2$, the two straight lines $y = 1 + x[1 + f(a)]$, $y = 2 + axf(1)$ intersect. Hence, for $x > x_2$, one is above the other. At $x = 0$, one intercept is 1, the other 2; hence,

$$f_{AL} = 2 + axf(1) < 1 + x[1 + f(1)] = f_{LA} \tag{5.52}$$

for $x_2 < x < x_2/a$. Similarly, we show that $f_{AL} < f_{LA}$ in $x_2/a < x \leq x_2/a^2$, and so on. Only a finite number of such steps are required.

It remains to compute $f(1)$ and x_2. Let $a^k < x_2 < a^{k-1}$. Then

$$f(1) = k + f(a^k) = k + 1 + a^k f(1), \tag{5.53}$$

whence

$$f(1) = \frac{k+1}{1-a^k}. \tag{5.54}$$

Since this is a convex function of k, the minimum occurs at either a unique k or at two adjacent k's. Having determined k, we have $f(1)$, and then

$$x_0 = \frac{1}{(1-a)f(1)} = \frac{1-a^k}{(1-a)(k+1)}. \tag{5.55}$$

CHAPTER 6
GAMES OF SURVIVAL

6.1. Introduction

In this chapter we shall present some results concerning a class of games, which we call "games of survival," in which two players with finite fortunes, f_1 and f_2, respectively, in *chips* play a normalized finite zero-sum two-person game. The game is continued until the fortune of one of the players is reduced to zero, or ad infinitum if this never occurs. The payoff in *money* is $(1, 0)$ if player two is ruined before player one, and $(0, 1)$ if the reverse holds.

Another way of viewing this is that each player is playing so as to maximize the probability that he will survive his opponent.

We shall first consider a simple game using the functional-equation approach of the previous chapters, and then present a more powerful technique that utilizes more of the actual structure of the process.

6.2. The 2 × 2 Game

Let us consider the situation in which two players, A and B, possessing fortunes x and y, respectively, play the zero-sum game defined by the matrix

$$\Gamma = \begin{pmatrix} -1 & a \\ c & -b \end{pmatrix}, \tag{6.1}$$

where a, b, and c are positive integers, with the purpose in mind of ruining the opponent.

Since the game is zero-sum, we shall set $x + y = d$ and specify the state of the fortunes of the players by x, the quantity held by A. Let us define, for $0 < x < d$, x integral,

$$f(x) = \text{probability that } B \text{ is ruined before } A \text{ when } A \text{ has } x \text{ and both players use optimal play,} \tag{6.2}$$

setting

$$\begin{aligned} f(x) &= 0, \qquad x \leq 0 \\ &= 1, \qquad x \geq d. \end{aligned} \tag{6.3}$$

If this function exists, it satisfies the equation

$$\begin{aligned} f(x) &= \operatorname*{Min}_q \operatorname*{Max}_p \left[p_1 q_1 f(x-1) + p_1 q_2 f(x+a) \right. \\ &\qquad \left. + p_2 q_1 f(x+c) + p_2 q_2 f(x-b)\right] \\ &= \operatorname*{Max}_p \operatorname*{Min}_q [\cdots], \end{aligned} \tag{6.4}$$

for $x = 1, 2, \cdots, d - 1$.

To simplify the formulas which occur, we shall set $V[f(x)]$ as the value of the game whose matrix is

$$\begin{pmatrix} f(x-1) & f(x+a) \\ f(x+c) & f(x-b) \end{pmatrix}. \tag{6.5}$$

We shall use the notation $V(M)$ to denote the value of the game whose matrix is M.

The functional equation of (6.4) therefore has the form

$$\begin{aligned} f(x) &= V[f(x)], \qquad && x = 1, 2, \cdots, d-1, \\ f(x) &= 0, && x \leq 0 \\ &= 1, && x \geq d. \end{aligned} \tag{6.6}$$

Although it is not immediately seen that $f(x)$ exists, there is no difficulty in defining

$$\begin{aligned} f_n(x) = {} & \text{probability that } B \text{ is ruined before } A \text{ when } n \text{ rounds of the game are} \\ & \text{played with both sides using optimal play and } A \text{ possessing } x. \end{aligned} \tag{6.7}$$

This function satisfies the equations

$$\begin{aligned} f_0(x) &= 1, && x \geq d \\ &= 0, && x \leq d-1, \\ f_{n+1}(x) &= V[f_n(x)], \qquad && n = 0, 1, \cdots, \qquad x = 1, 2, \cdots, d-1, \\ f_{n+1}(x) &= 1, && x \geq d \\ &= 0, && x \leq 0, \end{aligned} \tag{6.8}$$

assuming that in the n-stage process A plays to maximize this probability, and B plays to minimize it. The situation is unsymmetrical, since there is always in the n-stage process a non-zero probability that both sides survive. As $n \to \infty$, this probability approaches zero, and the situation becomes symmetrical.

It is clear that $f_1(x) \geq f_0(x)$ for all x, and hence, inductively, that $f_{n+1}(x) \geq f_n(x)$. It follows from the trivial observation that $0 \leq f_n(x) \leq 1$ for all x and n and that $f_n(x)$ converges, as $n \to \infty$, for all x to a function that we call $f(x)$. That $f(x)$ satisfies (6.6) is a consequence of the fact that the value of a game is a continuous function of the game matrix.

Since $f_0(x)$, and consequently each $f_n(x)$, is a monotone-increasing function of x, it follows that $f(x)$ is monotone. Let us now demonstrate the important result that it is actually strictly monotone.

We have

$$f(1) = V\begin{pmatrix} 0 & f(a) \\ f(c) & 0 \end{pmatrix}. \tag{6.9}$$

If $f(a)$ and $f(c)$ are positive, then $f(1) > 0$. Let us assume, to the contrary, that $f(x) = 0, 1, 2, \cdots, k < d$, but $f(k+1) \neq 0$. That a k with this property exists is clear.

Then

$$f(k) = V\begin{pmatrix} f(k-1) & f(k+a) \\ f(k+c) & f(k-b) \end{pmatrix} = V\begin{pmatrix} 0 & f(k+a) \\ f(k+c) & 0 \end{pmatrix}. \qquad (6.10)$$

Since $f(k+a) \geq f(k+b) > 0$, $f(k+c) \geq f(k+1) > 0$, it follows that $f(k) > 0$, which is a contradiction, unless $k = 0$. Thus, $f(1) > 0$.

Now,

$$f(2) = V\begin{pmatrix} f(1) & f(a+2) \\ f(c+2) & f(2-b) \end{pmatrix}. \qquad (6.11)$$

Since $f(1) > 0$, $f(a+2) \geq f(a+1)$, $f(c+2) \geq f(c+1)$, $f(2-b) \geq 0$, we must have $f(2) > f(1)$, unless $f(2-b) = 0$ and the solution of the game is $p_2 = q_2 = 1$. This is clearly not so, since $q_2 = 1$ is a better response to $p_2 = 1$. Similarly, we prove, using induction, that

$$0 = f(0) < f(1) < f(2) < \cdots < f(d) = 1, \qquad (6.12)$$

with strict inequality at every step.

With these preliminaries disposed of, we now turn to the question of uniqueness. Let us set

$$T(p, q, f) = p_1q_1f(x-1) + p_1q_2f(x+a) + p_2q_1f(x+c) + p_2q_2f(x-b). \qquad (6.13)$$

Let f and g be solutions of

$$\begin{aligned} f(x) &= \underset{q}{\text{Min}}\ \underset{p}{\text{Max}}\ T(p, q, f) = \underset{p}{\text{Max}}\ \underset{q}{\text{Min}}\ T(p, q, f), \\ g(x) &= \underset{q}{\text{Min}}\ \underset{p}{\text{Max}}\ T(p, q, f) = \underset{p}{\text{Max}}\ \underset{q}{\text{Min}}\ T(p, q, f), \end{aligned} \qquad (6.14)$$

satisfying the boundary conditions

$$\begin{aligned} f(x) = g(x) &= 0, \qquad x \leq 0 \\ &= 1, \qquad x \geq d, \end{aligned} \qquad (6.15)$$

with the further assumption that $g(x)$ is uniformly bounded.

Under the assumption that $f(x) \not\equiv g(x)$, let

$$\Delta = \text{Max}\,|f(x) - g(x)|, \qquad (6.16)$$

and let y be the largest integer in $[0, d]$ for which this maximum, assumed to be not equal to zero, is attained.

If we set $p_i = p_i(y)$, $q_i = q_i(y)$, $\bar{p}_i = \bar{p}_i(y)$, $\bar{q}_i = \bar{q}_i(y)$ to be sets of values for which the min-max is assumed, we have

$$\begin{aligned} f(y) &= T(p, q, f), \\ g(y) &= T(\bar{p}, \bar{q}, g). \end{aligned} \qquad (6.17)$$

From the properties of min-max, we have

$$\begin{aligned} &\text{(a)} \qquad f(y) = T(p, q, f) \geq T(\bar{p}, q, f) \\ &\text{(b)} \qquad \qquad \qquad \leq T(p, \bar{q}, f), \end{aligned} \tag{6.18}$$

and

$$\begin{aligned} &\text{(a)} \qquad g(y) = T(\bar{p}, \bar{q}, g) \leq T(\bar{p}, q, g) \\ &\text{(b)} \qquad \qquad \qquad \geq T(p, \bar{q}, g). \end{aligned} \tag{6.19}$$

Combining (6.18a) with (6.19a), we obtain

$$f(y) - g(y) \geq T(\bar{p}, q, f) - T(\bar{p}, q, g) = T(\bar{p}, q, f - g), \tag{6.20}$$

while (6.18b) and (6.19b) yield

$$f(y) - g(y) \leq T(p, \bar{q}, f) - T(p, \bar{q}, g) = T(p, \bar{q}, f - g). \tag{6.21}$$

From these two inequalities, we conclude that

$$\Delta = |f(y) - g(y)| \leq \text{Max}\,[|T(\bar{p}, q, f - g)|, |T(p, \bar{q}, f - g)|]. \tag{6.22}$$

Since

$$\begin{aligned} |T(\bar{p}, q, f - g)| &\leq T(\bar{p}, q, \Delta) = \Delta, \\ |T(p, \bar{q}, f - g)| &\leq T(p, \bar{q}, \Delta) = \Delta, \end{aligned} \tag{6.23}$$

we conclude that (6.22) is actually an equality, which means that the inequalities in (6.20) and (6.21) must also be equalities.

Consider the relation

$$f(b) - g(y) = \sum_{i,j} \bar{p}_i q_j [f(y + a_{ij}) - g(y + a_{ij})], \tag{6.24}$$

where we set

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} -1 & a \\ c & -b \end{pmatrix}. \tag{6.25}$$

Since $\sum_{i,j} \bar{p}_i q_j = 1$, if $|f(y + a_{ij}) - q(y + a_{ij})| < D$, $\bar{p}_i q_j$ must be zero. By assumption, y was the largest integer in $[0, d]$ for which $|f(x) - g(x)| = \Delta$. Hence, $\bar{p}_i q_i = 0$ whenever $a_{ij} > 0$.

It follows that $\bar{p}_1 q_2 = 0$, $\bar{p}_2 q_1 = 0$. Since $\bar{p}_1 + \bar{p}_2 = 1$, both $\bar{p}_1$ and $\bar{p}_2$ cannot be zero, which means that q_1 or $q_2 = 0$. Coming back to the game matrix

$$\begin{pmatrix} f(x - 1) & f(x + a) \\ f(x + c) & f(x - b) \end{pmatrix}, \tag{6.26}$$

we see that the strict monotonicity of $f(x)$ makes it impossible for $q_1 = 0$ or $q_2 = 0$ to be optimal play for B for $x = y$.

We have thus obtained the desired contradiction.

The method we have employed is quite general and can be used to treat many particular types of $m \times n$ games. The general case, however, in which one only assumes that the entries in the matrix are positive or negative integers still presents difficulties.

6.3. More General Results

Let us now consider the game $\Omega(f_1, f_2)$ characterized by the payoff matrix, (Γ_{ij}), in which the elements are non-zero integers, and the finite fortunes f_1 and f_2 of each player. We shall show that Ω is inessential and has some easily described optimal strategies.* We shall also show that if $\mathrm{Max}_{i,j} |\Gamma_{ij}|$ is small enough compared with the combined fortunes, then to play at the nth play a δ^n-optimal strategy for Γ is an ε-optimal strategy for Ω, if δ is sufficiently small. (δ^n is the nth power of δ.)

We assume that every column of Γ has a positive entry and that every row has a negative entry. Otherwise, there would be a negative column or a positive row. In the first case, player 2 can always force player 1's fortune to become non-positive by repeatedly playing the negative column. In the second case, player 1 can force player 2's fortune to become non-positive by repeatedly playing the positive row.

Let $\overline{\Omega}^{(n)}(f_1, f_2)$ be the game in which two players repeat Γ n times, or until one of the players has a non-positive fortune, if this occurs first. The payoff in money is $(0, 1)$ if player 1 ends with a non-positive fortune, and $(1, 0)$ otherwise. $\overline{\Omega}^{(n)}(f_1, f_2)$ is a constant-sum two-person game with value, say, $[\bar{v}^{(n)}(f_1, f_2), 1 - \bar{v}^{(n)}(f_1, f_2)]$. We observe that

1. Player 2 can always win as much money in $\overline{\Gamma}^{(n+1)}$ as in $\overline{\Gamma}^{(n)}$ by playing a $\overline{\Gamma}^{(n)}$-optimal strategy during the first n moves of $\overline{\Gamma}^{(n+1)}$ and by playing arbitrarily on the $(n+1)$th move. Hence,

$$\bar{v}^{(n)}(f_1, f_2) \geq \bar{v}^{(n+1)}(f_1, f_2).$$

2. Since each column has a positive entry, by repeatedly playing the strategy that assigns each pure strategy probability $1/i_0$, player 1 ensures that no matter what player 2 does, player 2's fortune will decrease each time with probability at least $1/i_0$. Player 1 thereby ensures, with a probability of at least $i_0^{-[f_2]-1}$, that player 2 will be bankrupted in at most $[f_2] + 1$ trials. ($[f_2]$ is the largest integer not larger than f_2.) Hence, if $n \geq [f_2] + 1$ and $(f_1, f_2) > (0, 0)$, then

$$\bar{v}^{(n)}(f_1, f_2) \in [\delta, 1],$$

where $\delta = i_0^{-[f_2]-1}$. By definition, we have also

$$\bar{v}^{(n)}(f_1, f_2) = 0 \qquad \text{if } f_1 \leq 0$$

and

$$\bar{v}^{(n)}(f_1, f_2) = 1 \qquad \text{if } f_2 \leq 0.$$

3. Let $G(\Delta)$ be the game value of Δ for each game Δ. If $(f_1, f_2) > 0$, after one

* In an inessential game, an optimal strategy for a player is one that secures for him the maximum amount he can ensure for himself. An ε-optimal strategy secures for him at least that amount less ε.

move of $\overline{\Gamma}^{(n+1)}(f_1, f_2)$, the players are playing $\overline{\Gamma}^{(n)}(f_1 + \Gamma_{ij}, f_2 - \Gamma_{ij})$. Hence,

$$\overline{v}^{(n+1)}(f_1, f_2) = G[\overline{v}^{(n)}(f_1 + \Gamma_{ij}, f_2 - \Gamma_{ij})].$$

4. Let $\varepsilon \geq 0$. Player 1 can always win as much in $\overline{\Gamma}^{(n)}(f_1 + \varepsilon, f_2 - \varepsilon)$ as in $\overline{\Gamma}^{(n)}(f_1, f_2)$. Hence,

$$\overline{v}^{(n)}(f_1 + \varepsilon, f_2 - \varepsilon) \geq \overline{v}^{(n)}(f_1, f_2).$$

We can now conclude:

(a) From (1) and (2),

$$\begin{aligned} \overline{v}^{(n)}(f_1, f_2) \to \overline{v}(f_1, f_2) &\in [\delta, 1] && \text{if } (f_1, f_2) > (0, 0) \\ &= 0 && \text{if } f_1 \leq 0 \\ &= 1 && \text{if } f_2 \leq 0. \end{aligned}$$

(b) From (3), if $(f_1, f_2) > (0, 0)$,

$$\overline{v}(f_1, f_2) = G[\overline{v}(f_1 + \Gamma_{ij}, f_2 - \Gamma_{ij})].$$

(c) From (4), for $\varepsilon \geq 0$,

$$\overline{v}(f_1 + \varepsilon, f_2 - \varepsilon) \geq \overline{v}(f_1, f_2).$$

DEFINITION. A strategy for player 1 is called *conditionally optimal* if the conditional distribution of his strategy at any play of Γ, given the course of the game up to that play, is an optimal strategy for the game $[\overline{v}(\phi_1 + \Gamma_{ij}, \phi_2 - \Gamma_{ij})]$, where (ϕ_1, ϕ_2) is the fortune distribution immediately before the play in question.

LEMMA 6.1. *If player 1's strategy is* conditionally optimal, *and if with probability 1 the fortune of one of the players (not necessarily always the same one) eventually becomes non-positive, then player 1 can expect at least $\overline{v}(f_1, f_2)$ in payoff.*

PROOF. It is sufficient to show that the probability that player 2's fortune becomes non-positive is at least $\overline{v}(f_1, f_2)$. Let $[(F_1^n, F_2^n) \mid n \geq 1]$ be the random variable of fortunes at play n, where, if the game ends at play N, $(F_1^{N+j}, F_2^{N+j}) \equiv (F_1^N, F_2^N)$ for $j \geq 1$. Then, since player 1's strategy is conditionally optimal, if $(F_1^n, F_2^n) > (0, 0)$,

$$\begin{aligned} E\overline{v}(F_1^{n+1}, F_2^{n+1}) &\geq EG[\overline{v}(F_1^n + \Gamma_{ij}, F_2^n - \Gamma_{ij})] \\ &= E\overline{v}(F_1^n, F_2^n), \end{aligned}$$

whereas, otherwise,

$$E\overline{v}(F_1^{n+1}, F_2^{n+1}) = E\overline{v}(F_1^n, F_2^n).$$

Hence, by induction,

$$E\overline{v}(F_1^n, F_2^n) \geq \overline{v}(f_1, f_2).$$

Let $[(P_1^n, P_2^n) \mid n \geq 1]$ be the random variable that is

$(0, 0)$ if neither player's fortune is non-positive by the end of the nth play,

(0, 1) if the first player's fortune is non-positive by the end of the nth play,
(1, 0) if the second player's fortune is non-positive by the end of the nth play.

Then

$$E\bar{v}(F_1^n, F_2^n) \leq EP_1^n + E(1 - P_2^n - P_1^n).$$

But, by assumption, the second term on the right tends to zero. Hence, where $\epsilon_n \to 0$,

$$EP_1^n + \epsilon_n \geq v(f_1, f_2),$$

which is the desired result.

LEMMA 6.2. *There is a conditionally optimal strategy for the first player which ensures that the probability that the game ends by the* nth *play tends uniformly to* 1 *as* n *tends to* ∞ *in the opponent's strategy.*

PROOF. First, we show that for each $(\phi_1, \phi_2) > (0, 0)$ there is an optimal strategy I for the first player for the game $[\bar{v}(\phi_1 + \Gamma_{ij}, \phi_1 - \Gamma_{ij})]$ such that for all J, $Pr\{\Gamma_{IJ} > 0\} > 0$. Suppose, on the contrary, that for some $(\phi_1, \phi_2) > (0, 0)$, for all optimal I, there is a J such that $Pr\{\Gamma_{IJ} > 0\} = 0$, or, since $\Gamma_{ij} \neq 0$, $Pr\{\Gamma_{IJ} < 0\} = 1$, which is the same thing. Then, since player 1 is playing optimally,

$$\bar{v}(\phi_1, \phi_2) \leq E\bar{v}(\phi_1 + \Gamma_{IJ}, \phi_2 - \Gamma_{IJ}).$$

From the monotonicity of $\bar{v}(\phi_1 + \epsilon, \phi_2 - \epsilon)$,

$$\bar{v}(\phi_1, \phi_2) \geq \bar{v}(\phi_1 + \Gamma_{IJ}, \phi_2 - \Gamma_{IJ}).$$

Combining,

$$\bar{v}(\phi_1, \phi_2) = \bar{v}(\phi_1 + \Gamma_{IJ}, \phi_2 - \Gamma_{IJ}),$$

or, weaker, from monotonicity again,

$$\bar{v}(\phi_1, \phi_2) = \bar{v}(\phi_1 - 1, \phi_2 + 1).$$

If $(\phi_1 - 1, \phi_2 + 1) > (0, 0)$, this implies that an optimal strategy I for the first player for the game $[\bar{v}(\phi_1 + \Gamma_{ij} - 1, \phi_2 - \Gamma_{ij} + 1)]$ is an optimal strategy for $[\bar{v}(\phi_1 + \Gamma_{ij}, \phi_2 - \Gamma_{ij})]$, since by using it against any J, the first player ensures for himself

$$\begin{aligned} E\bar{v}(\phi_1 + \Gamma_{IJ}, \phi_2 - \Gamma_{IJ}) &\geq E\bar{v}(\phi_1 + \Gamma_{IJ} - 1, \phi_2 - \Gamma_{IJ} + 1, \\ &\geq \bar{v}(\phi_1 - 1, \phi_2 + 1) \\ &= \bar{v}(\phi_1, \phi_2). \end{aligned}$$

Thus, for a fortune division $(\phi_1 - 1, \phi_2 + 1) > (0, 0)$, and by induction for a fortune division, $(\phi_1 - n, \phi_2 + n) > (0, 0)$, for all optimal strategies, I, there is a J such that $\Gamma_{IJ} < 0$. But eventually, perhaps for $n = 0$, whereas $\phi_1 \leq n + 1$. Therefore, for an optimal I and some J,

$$\begin{aligned} 0 < \delta \leq v(\phi_1 - n, \phi_2 + n) &\leq E\bar{v}(\phi_1 - n + \Gamma_{IJ}, \phi_2 + n - \Gamma_{IJ}) \\ &\leq \bar{v}(\phi_1 - n - 1, \phi_2 + n + 1) \\ &= 0, \end{aligned}$$

which is the contradiction for which we have been looking.

We have now proved that for $(\phi_1, \phi_2) > 0$, there is an optimal I such that for all J, $Pr\{\Gamma_{IJ} > 0\} > 0$. For each $(\phi_1, \phi_2) > (0, 0)$, fix such an I. Call it $I(\phi_1, \phi_2)$. From the compactness of the second player's set of strategies and the fact that $Pr\{\Gamma_{IJ} > 0\}$ is a continuous function of his strategy, $Pr\{\Gamma_{IJ} > 0\} \geq \rho(\phi_1, \phi_2) > 0$. Define $\sigma(\phi_1, \phi_2) = \text{Min}_k\, \rho(\phi_1 + k, \phi_2 - k) > 0$, where k is an arbitrary positive, zero, or a negative integer such that $(\phi_1 + k, \phi_2 - k) > (0, 0)$.

Now let player 1 use the conditionally optimal strategy that consists in playing $I(\phi_1, \phi_2)$ when the fortune distribution is (ϕ_1, ϕ_2). Let $Q^{(n)}$ be the probability that one player's fortune or the other's is exhausted on or before the nth play. Then, where $\sigma = \sigma(f_1, f_2)$,

$$Q^{([f_1+f_2]+1)} \geq 0,$$
$$Q^{(n+[f_1+f_2]+1)} \geq Q^{(n)} + (1 - Q^{(n)})\sigma^{[f_1+f_2]+1}.$$

By induction,

$$Q^{(N([f_1+f_2]+1))} \geq 1 - (1 - \sigma^{[f_1+f_2]+1})^{N-1}.$$

Hence, $Q^{(N)} \to 1$ as $N \to \infty$, which is the lemma.

Let $\underline{\Omega}^{(n)}(f_1, f_2)$ be the game in which the two players repeat Γ n times, or until one of the players has a non-positive fortune, if this occurs first, and the money payoff is $(1, 0)$ if player 2 ends with a non-positive fortune, or is $(0, 1)$, otherwise. $\underline{\Omega}^{(n)}(f_1, f_2)$ is a constant-sum two-person game with value $[\underline{v}^{(n)}(f_1, f_2), 1 - \underline{v}^{(n)}(f_1, f_2)]$. Obviously,

$$\underline{v}^{(n)}(f_1, f_2) \leq \bar{v}^{(n)}(f_1, f_2),$$

since any strategy for player 1 in $\underline{\Omega}^{(n)}(f_1, f_2)$ will ensure him as much money in $\bar{\Omega}^{(n)}(f_1, f_2)$. We therefore conclude, by the same reasoning as that stated earlier, that

$$\begin{array}{llll} \text{(a}'\text{)} & \underline{v}^{(n)}(f_1, f_2) \to \underline{v}(f_1, f_2) \in [0, 1 - \delta'] & & \text{if } (f_1, f_2) > (0, 0) \\ & \qquad\qquad\qquad = 0 & & \text{if } f_1 \leq 0 \\ & \qquad\qquad\qquad = 1 & & \text{if } f_2 \leq 0, \end{array}$$

where $\delta' > 0$;

$$\begin{array}{lll} \text{(b}'\text{)} & \underline{v}(f_1, f_2) = G[\underline{v}(f_1 + \Gamma_{ij}, f_2 - \Gamma_{ij})] & \text{if } (f_1, f_2) > (0, 0); \\ \text{(c}'\text{)} & \underline{v}(f_1 + \varepsilon, f_2 - \varepsilon) \geq \underline{v}(f_1, f_2) & \text{if } \varepsilon \geq 0. \end{array}$$

In addition,

$$\text{(d}'\text{)} \qquad \underline{v}(f_1, f_2) \leq \bar{v}(f_1, f_2).$$

DEFINITION. A strategy for player 2 is called *conditionally optimal* if the conditional distribution of his strategy at any play of Γ, given the course of the game up to that play, is an optimal strategy for the game $[\bar{v}(\phi_1 + \Gamma_{ij}, \phi_2 - \Gamma_{ij})]$, where (ϕ_1, ϕ_2) is the fortune distribution immediately before the play in question.

From Lemmas 6.1 and 6.2 and from the analogous Lemmas 6.1′ and 6.2′ that we do not write down, we conclude that each player has a conditionally optimal strategy which ensures that play ends by the nth play with probability tending uniformly to 1 as n tends to ∞ in the opponent's strategy. The first player's strategy ensures him $\underline{v}(f_1, f_2)$ on the average, and the second player's strategy ensures him $1 - \underline{v}(f_1, f_2) \geq 1 - \bar{v}(f_1, f_2)$ on the average. Since together the players can win no more than 1, we get

$$1 \geq \bar{v}(f_1, f_2) + [1 - \underline{v}(f_1, f_2)] \geq \bar{v}(f_1, f_2) + [1 - \bar{v}(f_1, f_2)] = 1.$$

This means that $\underline{v}(f_1, f_2) = \bar{v}(f_1, f_2) =$ (say) $v(f_1, f_2)$, and that $\Omega(\phi_1, \phi_2)$ is inessential with the solution $[v(f_1, f_2), 1 - v(f_1, f_2)]$.

v can be characterized as being the unique solution of

$$\begin{aligned} 0 \leq v(\phi_1, \phi_2) &= G[v(\phi_1 + \Gamma_{ij}, \phi_2 - \Gamma_{ij})] \leq 1 \qquad \text{if } (f_1, f_2) > (0, 0) \\ &= 0 \qquad \text{if } f_1 \leq 0 \\ &= 1 \qquad \text{if } f_2 \leq 0. \end{aligned}$$

For, if v^* is a solution,

$$\underline{v}^{(0)}(\phi_1, \phi_2) \leq v^*(\phi_1, \phi_2) \leq \bar{v}^{(0)}(\phi_1, \phi_2)$$

by definition, and so, by induction, using (a), (b), (a′), and (b′),

$$\underline{v}^{(n)}(\phi_1, \phi_2) \leq v^*(\phi_1, \phi_2) \leq \bar{v}^{(n)}(\phi_1, \phi_2).$$

Hence,

$$v(\phi_1, \phi_2) = \underline{v}(\phi_1, \phi_2) \leq v^*(\phi_1, \phi_2) \leq \bar{v}(\phi_1, \phi_2) = v(\phi_1, \phi_2),$$

giving

$$v(\phi_1, \phi_2) = v^*(\phi_1, \phi_2),$$

as was to be proved.

We thus have

THEOREM 6.1. *$\Omega(f_1, f_2)$ is inessential with the solution $[v(f_1, f_2), 1 - v(f_1, f_2)]$, where v is the unique solution in $\{(\phi_1, \phi_2) \mid \phi_1 > 0$ or $\phi_2 > 0\}$ of*

$$\begin{aligned} 0 \leq v(\phi_1, \phi_2) &= G[v(\phi_1 + \Gamma_{ij}, \phi_2 - \Gamma_{ij})] \leq 1 \qquad \textit{if } (\phi_1, \phi_2) > (0, 0) \\ &= 0 \qquad \textit{if } \phi_1 \leq 0 \\ &= 1 \qquad \textit{if } \phi_2 \leq 0. \end{aligned}$$

Each player has a conditionally optimal strategy that is optimal and which ensures that play ends by the nth *play with probability tending uniformly to* 1 *in the opponent's strategies.*

Let us turn now to the problem of the effective computation of an ε-optimal strategy for $\Omega(f_1, f_2)$. This is easy, if we are not interested in efficiency. That is, we need only find an n such that $\bar{v}^{(n)}(f_1, f_2) - \underline{v}^{(n)}(f_1, f_2) \leq \varepsilon - \delta$, where $\delta > 0$. Then a δ-optimal strategy for the first player for $\underline{\Omega}^{(n)}(f_1, f_2)$ provides an ε-optimal strategy for him for $\Omega(f_1, f_2)$. Thus, he can use the strategy on the first n moves of $\Omega(f_1, f_2)$ and act arbitrarily thereafter. Similarly, a δ-optimal strategy for the second player for $\overline{\Omega}^{(n)}(f_1, f_2)$ provides an ε-optimal strategy for him for $\Omega(f_1, f_2)$.

If $\text{Max}_{i,j} |\Gamma_{ij}|$ is small enough compared with f_1 and f_2, another class of interesting ε-optimal strategies exists. The repeated playing of an optimal strategy for Γ is an ε-optimal strategy for Ω. More precisely, let us remove the restriction that each Γ_{ij} be a non-zero integer. Let us require instead, say, that $G(\Gamma) \geq 0$ and that for some optimal strategy I, $Pr\{\Gamma_{Ij} > 0\} > 0$ for all j. If $G(\Gamma) = 0$, we require in addition that for some optimal J, $Pr\{\Gamma_{iJ} < 0\} > 0$ for all i. Define $\alpha = G(\Gamma)$, $\beta = \text{Min}_j\, Pr\{\Gamma_{Ij} > 0\}$, $\gamma = \text{Max}_{i,j} |\Gamma_{ij}|$.

We assume that both f_1 and f_2 are positive and define $f = f_1 + f_2$. Define, for $\alpha = 0$,

$$\begin{aligned} p_0(\phi_1) &= \frac{1}{f+\gamma}\,\phi_1 && \text{if } 0 < \phi_1 < f \\ &= 0 && \text{if } \phi_1 \leq 0 \\ &= 1 && \text{if } \phi_1 \geq f, \end{aligned}$$

and for $\alpha > 0$,

$$\begin{aligned} p_\alpha(\phi_1) &= \frac{1 - \exp\left\{-\dfrac{\alpha}{\gamma^2}\,\phi_1\right\}}{1 - \exp\left\{-\dfrac{\alpha}{\gamma^2}\,(f+\gamma)\right\}} && \text{if } 0 < \phi_1 < f \\ &= 0 && \text{if } \phi_1 \leq 0 \\ &= 1 && \text{if } \phi_1 \geq f. \end{aligned}$$

Lemma 6.3. *If player* 1 *plays* I *repeatedly, then he can expect at least* $p_\alpha(f_1)$ *in payoff.* (I is any optimal strategy for Γ satisfying $Pr(\Gamma_{Ij} > 0) > 0$.)

Proof. Since $\beta > 0$, by the method of proof of Lemma 6.2, it follows that if player 1 plays I repeatedly, the probability that the game ends by the nth play tends to 1 as n tends to ∞. Hence, in order to prove Lemma 6.3, it is sufficient to show that for all N,

$$Ep_\alpha(F_1^N) \geq p_\alpha(f_1).$$

By induction, this would follow from

$$E\{p_\alpha(F_1^{N+1}) \mid F_1^N\} \geq p_\alpha(F_1^N).$$

We prove the latter.

Suppose that $\alpha = 0$. If $0 < F_1^N < f$, then for all (i, j), since $\Gamma_{ij} \leq \gamma$,

$$p_0(F_1^N + \Gamma_{ij}) \geq \frac{1}{f+\gamma}(F_1^N + \Gamma_{ij}).$$

Hence, if $0 < F_1^N < f$,

$$\begin{aligned} E\{p_0(F_1^{N+1}) \mid F_1^N\} &\geq \operatorname*{Min}_j E p_0(F_1^N + \Gamma_{Ij}) \\ &\geq \frac{1}{f+\gamma} \operatorname*{Min}_j E(F_1^N + \Gamma_{Ij}) \\ &\geq \frac{1}{f+\gamma} F_1^N \\ &= p_0(F_1^N). \end{aligned}$$

Since if $F_1^N \leq 0$ or $F_1^N \geq f$ our proposition is trivial, we have disposed of the case $\alpha = 0$.

Suppose now that $\alpha > 0$. Again, we need only consider $0 < F_1^N < f$. Then

$$p_\alpha(F_1^N + \Gamma_{ij}) \geq \frac{1 - \exp\left\{-\frac{\alpha}{\gamma^2}(F_1^N + \Gamma_{ij})\right\}}{1 - \exp\left\{-\frac{\alpha}{\gamma^2}(f+\gamma)\right\}}.$$

Hence,

$$\begin{aligned} E\{p_\alpha(F_1^{N+1}) \mid F_1^N\} &\geq \operatorname*{Min}_j E p_\alpha(F_1^N + \Gamma_{Ij}) \\ &\geq \operatorname*{Min}_j \frac{1 - E\exp\left\{-\frac{\alpha}{\gamma^2}(F_1^N + \Gamma_{Ij})\right\}}{1 - \exp\left\{-\frac{\alpha}{\gamma^2}(f+\gamma)\right\}} \\ &\geq \frac{1 - M\exp\left\{-\frac{\alpha}{\gamma^2} F_1^N\right\}}{1 - \exp\left\{-\frac{\alpha}{\gamma^2}(f+\gamma)\right\}}, \end{aligned}$$

where

$$\begin{aligned} M &= \operatorname*{Max}_j E\exp\left\{-\frac{\alpha}{\gamma^2}\Gamma_{Ij}\right\} \\ &\leq \operatorname*{Max}_j \left\{1 - \frac{\alpha}{\gamma^2} E\,\Gamma_{Ij} + (e-2)\left(\frac{\alpha}{\gamma}\right)^2\right\} \\ &\leq \left\{1 - \frac{\alpha^2}{\gamma^2} + (e-2)\frac{\alpha^2}{\gamma^2}\right\} \\ &< 1. \end{aligned}$$

Hence,

$$E\{p_\alpha(F_1^{N+1}) \mid F_1^N\} \geq \frac{1 - \exp\left\{-\frac{\alpha}{\gamma^2} F_1^N\right\}}{1 - \exp\left\{-\frac{\alpha}{\gamma^2}(f+\gamma)\right\}} = p_\alpha(F_1^N),$$

as was to be proved.

By symmetry, if $\alpha = 0$, we conclude

LEMMA 6.4. *If $\alpha = 0$, and if player 2 plays J repeatedly, then he can expect at least $p_0(f_2)$ in payoff.* (J is any optimal strategy for Γ satisfying $Pr(\Gamma_{iJ} < 0) > 0$.)

If $\alpha = 0$, Lemmas 6.3 and 6.4 give us, whenever $\Omega(f_1, f_2)$ is inessential with the solution $\{[v(f_1, f_2), 1 - v(f_1, f_2)]\}$,

$$p_0(f_1) \le v(f_1, f_2) \le 1 - p_0(f_2) = p_0(f_1) + \frac{\gamma}{f + \gamma}.$$

Thus, repeating I is $[\gamma/(f + \gamma)]$-optimal for player 1, and repeating J is $[\gamma/(f + \gamma)]$-optimal for player 2. If $\alpha > 0$, Lemma 6.3 gives us, whenever $\Omega(f_1, f_2)$ is inessential with the solution $\{[v(f_1, f_2), 1 - v(f_1, f_2)]\}$,

$$1 - \exp\left\{-\frac{\alpha}{\gamma^2} f_1\right\} \le p_\alpha(f_1) \le v(f_1, f_2) \le 1.$$

Thus, repeating I is $\exp\{-(\alpha/\gamma^2)f_1\}$-optimal for player 1, and any strategy is $\exp\{-(\alpha/\gamma^2)f_1\}$-optimal for player 2.

What if, instead of repeating I, player 1 repeats a δ-optimal I_δ, where δ is the smallest number for which I_δ is δ-optimal? If $\alpha > \delta$, no great harm is done, since it can be verified by precisely the proof given above that this is an $\exp\{-[(\alpha - \delta)/\gamma^2]f_1\}$-optimal strategy for player 1. If, however, $\alpha < \delta$, player 2 can expect at least $1 - \exp\{-[(\delta - \alpha)/\gamma^2]f_2\}$ in payoff. When $[(\delta - \alpha)/\gamma^2]f_2$ is large, this payoff is close to 1, so that I_δ is not a good strategy. Thus, if $\alpha = 0$, no matter how small γ is, it is not enough to repeat a δ-optimal strategy for sufficiently small δ. On the other hand, suppose that (I_n) is a sequence of strategies for player 1 whose nth member is δ^n-optimal for Γ and satisfies

$$\min_j Pr\{\Gamma_{I_n j} > 0\} \ge \beta' > 0,$$

where β' does not depend on n. Then

LEMMA 6.5. *If $\alpha = 0$ and player 1 plays I_n at the nth stage, then he can expect at least $p_\alpha(f_1) - [\delta/(1 - \delta)(f + \gamma)]$ in payoff.*

PROOF. The proof is almost identical with that of Lemma 6.3, where, instead of proving

$$Ep_\alpha(F_1^N) \ge p_\alpha(f_1),$$

one proves that

$$Ep_\alpha(F_1^N) \ge p_\alpha(f_1) - \frac{\delta + \cdots + \delta^{N-1}}{f + \gamma}.$$

It is now an easy step (left to the reader) to

THEOREM 6.2. *If $G(\Gamma) = \alpha > \delta$ and $\Omega(f_1, f_2)$ is inessential, repeating a strategy which is δ-optimal for Γ is $\exp\{-[(\alpha - \delta)/\gamma^2]f_2\}$-optimal for $\Omega(f_1, f_2)$. Let $G(\Gamma) = 0$, and let (I_n) be a sequence of strategies for player 1 whose nth member is δ^n-opti-*

mal for Γ *and satisfies*

$$\operatorname*{Min}_j Pr\{\Gamma_{I_n j} > 0\} \geq \beta' > 0,$$

where β' *does not depend on* n. *Then playing* I_n *at the* nth *stage is a* $\{[\gamma/(f+\gamma)] + [2\delta/(1-\delta)(f+\gamma)]\}$*-optimal strategy for player* 2.

The reader will observe that when each $\Gamma_{ij} \neq 0$, say, $|\Gamma_{ij}| \geq C$, we automatically have, for a δ^n-optimal I_n, when δ is sufficiently small.

$$\operatorname*{Min}_j Pr\{\Gamma_{I_n j} > 0\} \geq \frac{C - \delta^n}{C + \gamma} \geq \frac{C - \delta}{C + \gamma} > 0.$$

In closing, we wish to point out that the method of proof leading to Theorem 6.1 is trivially sufficient to handle the following generalized game of survival, in which the result of a play is a random state instead of a definite number. However, the method is apparently insufficient to handle more than a finite number of possible states or the possibility of "zeros." A finite set Σ with two distinguished points, σ_1 and σ_2, is given. Σ is partially ordered by $<$, which satisfies for some fixed n and all $\{x_i \mid 1 \leq i \leq n\}$,

$$x_1 < x_2 < \cdots < x_{n-1} < x_n \rightarrow x_1 = \sigma_2,\ x_n = \sigma_1.$$

For each $x \in \Sigma$, there is a set of random variables on Σ, $\{Y_{ij}(x) \mid 1 \leq i \leq i_0,\ 1 \leq j \leq j_0\}$, such that for all i and j, $Y_{ij}(\sigma_1) = \sigma_1$, $Y_{ij}(\sigma_2) = \sigma_2$; and for $x \neq \sigma_1, \sigma_2$,

$$Pr\{Y_{ij}(x) < x\} = 0 \rightarrow Pr\{x < Y_{ij}(x)\} = 1,$$
$$Pr\{x < Y_{ij}(x)\} = 0 \rightarrow Pr\{Y_{ij}(x) < x\} = 1.$$

In addition, for $x \neq \sigma_1, \sigma_2$ for each i, there is a j such that

$$Pr\{Y_{ij}(x) < x\} > 0;$$

and for each j, there is an i such that

$$Pr\{x < Y_{ij}(x)\} > 0.$$

Define $\prod_{n=1}^{N} Y^{(n)}_{i_n j_n}(x)$ by induction by

$$\prod_{n=1}^{M+1} Y^{(n)}_{i_n j_n}(x) = Y^{M+1}_{i_{M+1} j_{M+1}}\left[\prod_{n=1}^{M} Y^{(n)}_{i_n j_n}(x)\right],$$

where $\{[Y^{(n)}_{ij}(x) \mid 1 \leq i \leq i_0,\ 1 \leq j \leq j_0, x \in \Sigma]\}$ is a set of independent random variables, each distributed like $[Y_{ij}(x)]$. Then we finally require that $x < x'$ implies that for all N,

$$Pr\left\{\prod_{n=1}^{N} Y^{(n)}_{i_n j_n}(x') = \sigma_1\right\} \geq Pr\left\{\prod_{n=1}^{N} Y^{(n)}_{i_n j_n}(x) = \sigma_1\right\},$$

$$Pr\left\{\prod_{n=1}^{N} Y^{(n)}_{i_n j_n}(x') = \sigma_2\right\} \leq Pr\left\{\prod_{n=1}^{N} Y^{(n)}_{i_n j_n}(x) = \sigma_2\right\}.$$

All that we have said about $\{\Omega(f_1, f_2)\}$ up to Theorem 6.1, trivially modified, applies to the games $\{\Omega(x)\}$, in which two players repeatedly and simultaneously choose integers i_n and j_n at each time n, until $\prod_{n=1}^{N} Y_{i_n j_n}^{(n)}(x) = \sigma_1$ or σ_2, or ad infinitum, if this never occurs. The payoff is $(1, 0)$ if the game ends in the state σ_1, and is $(0, 1)$ if the game ends in the state σ_2. If the game goes on indefinitely, then the payoff is $[\alpha(C), \beta(C)]$ where $[\alpha(C), \beta(C)] \leq (1, 1)$ and $\alpha(C) + \beta(C) \leq 1$, and where $[\alpha(C), \beta(C)]$ can depend on the course of the game, C.

Similarly, Theorem 6.2 can be generalized by the use of expected values to the situation in which Σ is a set of reals satisfying, for $\sigma_1 > x > \sigma_2$,

$$\begin{aligned} Y_{ij}(x) &= x + a_{ij} && \text{if } \sigma_2 < x + a_{ij} < \sigma_1 \\ &= \sigma_1 && \text{if } \sigma_1 \leq x + a_{ij} \\ &= \sigma_2 && \text{if } \sigma_2 \geq x + a_{ij}, \end{aligned}$$

where a_{ij} is a real-valued random variable whose distribution depends on (i, j).

BIBLIOGRAPHY

1. ARROW, K. J., D. BLACKWELL, AND M. A. GIRSHICK, "Bayes and Minimax Solutions of Sequential Decision Problems," *Econometrica*, Vol. 17, Nos. 3–4, July–October, 1949, pp. 214–244.
2. ARROW, K. J., T. E. HARRIS, AND J. MARSCHAK, "Optimal Inventory Policy," Cowles Commission Paper No. 44, 1951.
3. BELLMAN, R., "On Games Involving Bluffing," *Rendiconti del Circolo Matematico di Palermo* (2), Vol. 1, 1952, pp. 1–18.
4. ———, "On the Theory of Dynamic Programming," *Proceedings of the National Academy of Sciences*, Vol. 38, No. 8, August, 1952, pp. 716–719.
5. ———, "On the Theory of Dynamic Programming," The RAND Corporation, Paper P-300 (rev.), August 23, 1952.
6. ———, AND D. BLACKWELL, "Some Two-person Games Involving Bluffing," *Proceedings of the National Academy of Sciences*, Vol. 35, No. 10, October, 1949, pp. 600–605.
7. BELLMAN, R., I. GLICKSBERG, AND O. GROSS, "On Some Variational Problems Occurring in the Theory of Dynamic Programming," *Proceedings of the National Academy of Sciences*, Vol. 39, No. 4, April, 1953, pp. 298–301.
8. BELLMAN, R., T. E. HARRIS, AND H. N. SHAPIRO, "Studies on Functional Equations Occurring in Decision Processes," The RAND Corporation, Paper P-382, August 18, 1952.
9. BUSH, R. R., AND C. F. MOSTELLER, "A Mathematical Model for Simple Learning," *Psychological Review*, Vol. 58, No. 5, September, 1951, pp. 313–325.
10. DVORETZKY, A. J., J. KIEFER, AND J. WOLFOWITZ, "The Inventory Problem—I: Case of Known Distributions of Demand" and "The Inventory Problem—II: Case of Unknown Distributions of Demand," *Econometrica*, Vol. 20, No. 2, April, 1952, pp. 187–222.
11. DVORETZKY, A. J., A. WALD, AND J. WOLFOWITZ, "Elimination of Randomization in Certain Statistical Decision Procedures and Zero-sum Two-person Games," *Annals of Mathematical Statistics*, Vol. 22, No. 1, March, 1951, pp. 1–21.
12. ESTES, W. K., "Toward a Statistical Theory of Learning," *Psychological Review*, Vol. 57, No. 2, March, 1950, pp. 94–107.
13. FLOOD, M. M., "On Stochastic Learning Theory," The RAND Corporation, Paper P-353, December 19, 1952.
14. ISAACS, R., "Games of Pursuit," The RAND Corporation, Paper P-257, November 17, 1951.
15. JOHNSON, S., AND S. KARLIN, "On Optimal Sampling Procedure for a Problem of Two Populations—I," The RAND Corporation, Paper P-328, October 1, 1952.
16. KARLIN, S., *A Mathematical Treatment of Learning Models—I*, The RAND Corporation, Research Memorandum RM-921, September 2, 1952.
17. ———, AND H. N. SHAPIRO, *Decision Processes and Functional Equations*, The RAND Corporation, Research Memorandum RM-933, September 3, 1952.
18. PEISAKOFF, M., *More on Games of Survival*, The RAND Corporation, Research Memorandum RM-884, June 20, 1952.
19. ROBBINS, H., "Some Aspects of the Sequential Design of Experiments," *Bulletin of the American Mathematical Society*, Vol. 58, No. 5, September, 1952, pp. 527–536.
20. VON NEUMANN, J., AND O. MORGENSTERN, *The Theory of Games and Economic Behavior*, 2d ed., Princeton University Press, Princeton, N.J., 1947.
21. WALD, A., *Statistical Decision Functions*, John Wiley & Sons, Inc., New York, 1950.
22. WOLFOWITZ, J., "On ϵ-Complete Classes of Decision Functions," *Annals of Mathematical Statistics*, Vol. 22, No. 3, September, 1951, pp. 461–465.

INDEX