The Theory of Games and Linear Programming

S. VAJDA

LONDON: METHUEN & CO. LTD
NEW YORK: JOHN WILEY & SONS, INC

First published in 1956
Reprinted, with minor corrections, 1957

CATALOGUE NO. 4060/U (METHUEN)

PRINTED IN GREAT BRITAIN
BY W. S. COWELL LTD, AT THE BUTTER MARKET, IPSWICH

CONTENTS

CHAPTER I

AN OUTLINE OF THE THEORY OF GAMES

The germ of the theory of games, as we know it today, is contained in notes by E. Borel (Ref. 5). He conjectured a special case of the so-called Fundamental Theorem, but was held up by his failure to prove it. In 1928, John von Neumann gave a talk to the Mathematical Society in Göttingen (Ref. 22) and gave a proof of that theorem. Very little was, however, heard of the theory until 1944, when the monumental work by J. von Neumann and O. Morgenstern (Ref. 24) appeared. The mathematical contents of this book included some further developments, while the second-named author's reputation as a leading economist gave a clue to a particular application for the theory. It was soon noticed – which is not surprising – that a theory of competitive games could throw some light not only on features of relatively peaceful economic competition, but also on those of combat and warfare. From then on, practical and theoretical aspects were developed side by side, and by now the theory of games has given rise to an extended activity which comprises research even in such abstract subjects as set theory, and topology.

In this book, only some knowledge of – fairly elementary – algebra and of linear analytic geometry will be assumed.

The theory of Linear Programming also drew its stimulus from problems in economics, but has reached the status of a theory in its own right. It deals with a maximizing (or minimizing) problem that cannot be solved by the methods of calculus. It is an interesting fact that problems of the theory of games can be shown to form a special case of those of Linear Programming.

In most applications of either the theory of games, or of Linear Programming, the computations called for are so extensive that it would seem impossible to carry them out in the lifetime of those who are interested in the answer. A great deal of research has therefore gone into computing methods and the advent of automatic computers has made it possible to hope for answers within a reasonable time. This book will deal with the computational aspect rather thoroughly, but without specific attention to problems posed by automatic computing. Use will be made of geometric representations where it might facilitate the understanding by those readers who find such methods attractive. However, those who prefer their mathematics to be formal will find that the treatment is precise.

2. We now introduce the problems which will have our attention by a simple, though typical, example: the game of Matching Pennies. Its rules are these: each of two players, A and B, puts down a penny, head or tail up, without showing it to his opponent. The pennies are then uncovered and A pockets both if they show the same side; otherwise B gets them both.

This perhaps not very exciting game is convenient to introduce some jargon which is already well-established in the literature. We have here two players, hence we speak of a *two-person* game. One player wins what the other loses or, in other words, the sum of their gains is zero. Hence the clumsy, but generally accepted name: *zero-sum* game. A *game*, in this context, is a collection of rules which determine what the players may do, and who wins eventually what, dependent on what the players choose to do. A game consists of *moves*; that of Matching Pennies provides only for one move for each player. Any particular instance of a game is called a *play*.

The fundamental feature of games considered by this theory is the fact that the gain or loss of either player depends not only on his own actions, but also on those of his opponent. It is clear that this feature appears also in some problems of economics, for instance at an auction, or at the Stock Exchange, and also in warfare.

One type of game which takes its inspiration from formalized situations of warfare is that of the so-called Blotto games, of which the following is a very elementary example from Ref. 26: General A commands four and Colonel B(lotto) five companies. The General can reach a town on two different roads and can send each company on whatever road he wishes. The Colonel can order any of his companies to defend any road. A wins if, on any road, he has more companies than Blotto. Several more complicated versions exist. We refer the reader to Ref. 20, Ref. 3, and the very amusing Ref. 28.

We continue to give examples of games and to discuss their formal features.

In games with more than one move for a player it is natural to imagine that each of them decides on his choice as the play develops. However, it is often convenient to imagine that they decide in advance what they would do under all possible future circumstances. A collection of choices for each possible situation is called a *strategy*. In games with only one move a strategy is identical with a move.

As an illustration we consider a very simple game* with 'bluffing'. By this term we mean that a player bids high on a weak hand, in order to induce his opponent to believe that he has a strong hand, and therefore to abandon a policy which would lead him to success. This is not a technical definition of bluffing, which would in any case be very difficult to construct. But the reader will understand the term if he thinks of such popular games as Poker or Bridge.

The game is as follows: A gets one of two possible cards, either the 'high' card or the 'low' card, with equal probability. If he receives the high card, then he must bid £2. If he receives the low card, then he may pay £1, or he may bid £2. B may, but need not, challenge a bid. If he does not, then he pays £1. If he challenges, he wins £2 if A had the low card, but loses £2 if A had the high card.

This game also exhibits what is called a *chance move*, i.e. a move whose outcome depends on chance. In this case it is provided by the deal of the first card to A. We shall now construct the strategies for this game, and analyse their effects.

Strategies of A:

If the card is high, the bid is compulsory, but if the card is low, then he may either pay £1 (strategy A_1), or bid £2 (he bluffs, strategy A_2).

* This is a simplification of the game mentioned on p. 91 of the volume which contains Ref. 27. For a more complicated game see Ref. 12a.

Strategies of B:

If A pays, B can only accept £1. But if A bids, then B may either challenge (strategy B_1), or not challenge (strategy B_2).

The outcomes are computed as follows:

(A_1, B_1). If the card was low, A pays £1. If the card was high, A bids £2, and when challenged, wins £2. The probability of each of these events is 1/2, hence the average payment will be 10/- to A.

(A_1, B_2). If the card is low, A pays £1. If the card is high, A bids £2, and B pays £1. The average payment is 0.

(A_2, B_1). If the card is low, A bids £2 and, when challenged, loses £2. If the card is high, A bids £2 and, when challenged, wins £2. The average is thus again 0.

(A_2, B_2). If the card is low, A bids £2 and B pays £1. If the card is high, A bids £2 and B pays £1. Hence the payment is £1 to A.

The reduction of a game's structure to strategies is called *normalizing*. There exists an elaborate theory referring to games in their *extensive* form, i.e. taking account of the succession of moves and, in particular, of the pattern of *information*, i.e. of what each player knows about his own and his opponent's earlier moves. We shall not enter into this here and refer the reader to Refs. 24 and 20.

Not all games are zero-sum games. In many purchases it is reasonable to assume that the value which the buyer attaches to the goods which he acquires is higher than the price paid, while the latter is at least the value attached to the goods by the seller. Again, the loss of a particular ship might mean much more for the successful prosecution of a war than the enemy realizes. Formally, however, a *non-zero-sum* game with two players can be thought of as a zero-sum three-person game by assuming that any balance between the two payments is made up, or received by, a third player who has no influence on the progress of the game, and whose only purpose is that of making or receiving payments.

Generalizing further, one can think of *n-person* games, in which n opponents play. Such games can again be zero-sum or non-zero-sum games, depending on the balance of all payments at the end of the play. The theory contains, so far, fewer results as one proceeds from simpler to more complicated generalizations. In this book we shall only deal with two-person games, but we mention in passing that there are two types of n-person games, which are distinguished by the names *cooperative* and *non-cooperative*. The former are those dealt with in Ref. 24 and their distinguishing feature is the possibility that the players form *coalitions* outside the game proper, by coordinating the strategies which they will apply when the game starts. It has been tried, not with complete success, to consider such arrangements as parts of the game itself, and then to exclude outside combinations by its rules. Games in which each player looks exclusively after himself are called non-cooperative (Ref. 21) and two-person zero-sum games belong, of course, always to this category.

3. We must now make it quite clear that the theory of games approaches the question of what is the best procedure for the players by a well-defined philosophy. This you may accept or reject. For a mathematician, at any rate, there is great attraction in the fact that it leads to mathematically interesting results. The approach will gradually emerge as we proceed; it is not that of a gambler who plays for his excitement. Our player analyzes the situation in an entirely unemotional way.

Let us consider a more sophisticated game than Matching Pennies with the following rules:

A and B choose, independently, one of the values −1, 0 or 1. Let A's choice be denoted by s and that of B by t. Then B pays A the amount $s(t-s)+t(t+s)$.

It is again clear that whatever A chooses, he cannot be certain of the outcome, because it depends also on the choice of B. To obtain an overall picture of all possibilities, we tabulate the average potential outcomes in what is called a *pay-off table*, or pay-off matrix, and which looks in the present case as follows:

		B chooses $t=$		
		−1	0	1
A chooses $s=$	−1	2	−1	−2
	0	1	0	1
	1	−2	−1	2

We have entered the amounts which A wins, and hence B loses, given the combinations of the respective choices which determine the cells of the table. We shall adhere to the convention that the gains of player A are indicated. We shall also call him the *First*, or *Maximizing*, Player and his opponent the *Second*, or *Minimizing*, Player. The latter will try to lead to as small an amount in the table as possible, since his gains are the negatives of the entries. If any entry is negative, then the First Player loses, and the Second Player wins, the absolute value of the indicated amount. With these conventions we may think of a game as defined by its pay-off table.

We see at a glance how the amount that A wins depends on the choice of B as well. Of course, neither player knows (without cheating) what the other will do. The theory of games now assumes that both argue as follows:

For every choice that I can make, I must fear that my opponent makes that choice which makes my gain (or average gain, if there is a chance move) the smallest possible under the circumstances. Hence, if I make that choice which makes this smallest gain as large as possible, then I am as safe as I can ever reasonably expect to be.

If both argue in this cautious (or perhaps defeatist) way, then in the present game A observes that he gains at least -2, 0 or -2, dependent on whether he makes $s=-1$, 0 or 1. He will therefore choose $s=0$, since this gives him the most amongst those safe expectations. B, arguing in a similar fashion, will decide to choose $t=0$. The actual gain of both players will then be zero.

This type of argument is fundamental to our analysis. We ask now what A and B would do if they knew what the other had decided to do or rather if they assumed that he would argue as explained above and would act accordingly.

If A assumed that B, following this argument (or for any other reason) has chosen 0, then he would choose 0 himself, because this gives him the largest entry in the second column, which corresponds to B's choice of 0. Hence he would do precisely what he has already decided to do on the basis of his previous argument. Similarly, B would not be affected in his decision even if he knew in advance that A had chosen 0, as can be seen from the table. We see that the 'second order' approach, i.e. the assumption that the opponent acts in accordance with the philosophy of the theory of games, will not induce either player to change his 'first order' decision, taken on the basis of that philosophy.

This feature of our present game depends on the fact that there is an entry in the pay-off table which is the smallest in its row and, at the same time, the largest in its column. The position of that entry is often described, by an obvious geometrical analogy, as a *saddle point*, and although this expression ignores the lack of continuity in the positions and is therefore not entirely appropriate, we shall use it because it is already well established. However, we shall avoid von Neumann and Morgenstern's expression 'specially strictly determined game' (p. 150 of Ref. 24), which is the same as a game with saddle point.

Of course, not every zero-sum two-person game has a saddle point: Matching Pennies is a trivial counter example. As a further illustration, consider the following game:

The two players choose, independently, a number out of 1, 2 or 3. If both choose the same number, A pays to B the amount of the chosen number. Otherwise A receives the amount of his own number from B. The pay-off table of this Eluding Game is as follows:

		B 1	B 2	B 3
A	1	−1	1	1
A	2	2	−2	2
A	3	3	3	−3

The minimum of the column maxima (viz. 2) is different from the maximum of the row minima (viz. −1). This means, as is obvious but will be proved analytically in Chapter III, that there is no saddle point in this table. We must find a method for 'solving' such games and must, of course, first agree on what we mean by a 'solution'.

In such a simple example as Matching Pennies it is easy to see what one should do. If the game is played but once, then no choice is better than another. But if the game is repeated, then it would certainly be a mistake for a player to show a preference for either Head or Tail, because his opponent could easily take advantage of such a tendency. One should therefore use Head and Tail with equal frequencies in the long run, and in such a manner that it is impossible to guess which comes next. Such random choice at equal long-term frequencies could be achieved, for instance, by spinning the coin.

This argument applies to both players and we assume that both act accordingly. Then, in the long run, both will win 1 in half the games and lose 1 in the other half. On the average they will thus win 0.

To this result we apply an analysis which is similar to that applied to games with a saddle point. Let us assume that A knows, or thinks he knows, that B will use his two possible strategies with equal frequencies. Then he has no reason to change his own first order choice of equal frequencies. As a matter of fact, he could now choose any frequencies he likes without changing the average pay-off. However, it would not be wise for him to make a change, because B could punish him if he did. The same argument applies to B.

The feature of stability which we demonstrated for games with saddle point has now been shown to exist also in other games, provided we consider also combinations of strategies with given frequencies. Such a combination is called a *mixed strategy*, and single strategies, which appear as labels in the pay-off table, will be called *pure*. They are, of course, a special case of mixed strategies, where the relative frequencies are zero for all pure strategies but one, which latter is used all the time. Therefore, whenever we speak of strategies, we mean mixed ones, unless otherwise stated.

A pair of strategies with that feature of stability which we have described is called a *solution* (see III.3 for a rigorous definition). A strategy which appears in a solution is called *optimal*.

To illustrate these concepts by another game, let us take the game with bluffing introduced earlier. Its pay-off table is

		B	
		B_1	B_2
A	A_1	1/2	0
	A_2	0	1

The solution of this game is for A to take the first pure strategy twice as often as the second, and for B to do the same. We write this down by simply quoting the relative frequencies, thus: (2/3, 1/3) for A as well as for B. This is indeed a solution, because no player can do better if the other sticks to his strategy, but could gain less or lose more than 1/3 (which is what A gains now) if he departed from the optimal solution and his opponent took advantage of it. Incidentally, it is interesting to note that bluffing, which is a feature of A's second pure strategy, is actually applied once out of three times in his optimal strategy.

It is possible for a game to have more than one solution. Thus, for instance, the game with pay-off table

$$\begin{array}{cc} & \text{B} \\ \text{A} & \begin{pmatrix} 1 & 2 & 4 \\ 4 & 2 & 1 \end{pmatrix} \end{array}$$

has the solution: (1/3, 2/3) for A and (0, 1, 0) for B, and also the solution: (2/3, 1/3) for A and again (0, 1, 0) for B. It follows that the strategies $[\frac{1}{3}t + \frac{2}{3}(1-t),\ \frac{2}{3}t + \frac{1}{3}(1-t)]$ for A and (0, 1, 0) for B are also solutions for any value of t between 0 and 1. Now it would hardly be justified to speak of an optimal strategy without any qualification, if it could only make up a solution with some particular optimal strategy of the opponent. However, we prove in III.3 that if s_A, s_B are strategies of A and B respectively which form a solution, and if this is also true of t_A and t_B, then (s_A, t_B) and (t_A, s_B) are also solutions and all these pairs produce the same pay-off. The pay-off resulting from a solution is called the *value* of the game.

We have seen that not every game has a saddle point. However, it is a fundamental fact in the theory of games that every zero-sum two-person game with a finite number of strategies has a solution, possibly in mixed strategies for one or for both of the players. The proof will be given in Chapter III.

A game in which both players use all their strategies in their optimal solutions is called *completely mixed* (Ref. 13). Matching Pennies and the game with bluffing are such games, but the following is not:

$$\begin{array}{cc} & \text{B} \\ \text{A} & \begin{pmatrix} 4 & 1 & 3 \\ 2 & 3 & 4 \end{pmatrix} \end{array}$$

The reader will easily verify that its solution is : (1/4, 3/4) for A and (1/2, 1/2, 0) for B. This game also supplies an illustration of the concept of *dominance*. We say that a pure strategy dominates another, of the same player, if for any choice of the opponent the first is at least as favourable as the other, and for at least one choice is definitely better. The pay-off matrix above shows that B would be foolish ever to use his third strategy, which is dominated by his second. It is therefore not surprising that the third strategy does not occur in B's optimal strategy, and it could have been ignored from the beginning.

It must be understood, though, that even a dominated strategy may be optimal. For instance, in the game

$$\begin{array}{cc} & \text{B} \\ \text{A} & \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \end{array}$$

there are two solutions: (1, 0) for A and (1, 0) for B, and also (0, 1) for A and again (1, 0) for B. Thus both pure strategies of A are optimal, although the second dominates the first.

We have not yet said anything about how to find the solutions of a game, but only how to test whether a pair of strategies is a solution or not. We now add the remark that if the pay-off table is skew-symmetric, so that it remains unaltered if the roles of the two players are interchanged, the value of the game is clearly zero. General methods of solving games are given later, in particular in Chapter IX, but the impatient reader will be glad to know that the solutions of all games which we mention are recorded in the Directory of Games at the end of the book.

CHAPTER II

GRAPHICAL REPRESENTATION

It will be helpful to readers who like to think in geometrical terms if we now introduce a graphical representation of concepts in the theory of games. For a presentation in the two-dimensional plane, we consider games where one of the two players has only two pure strategies at his disposal and we again choose

$$
\begin{array}{cc} & B \\ A & \begin{pmatrix} 4 & 1 & 3 \\ 2 & 3 & 4 \end{pmatrix} \end{array}
$$

as an example. We shall exhibit geometrical models for A as well as for B, the former being the player with two pure strategies. The results can easily be adapted to the case when it is the minimizing player who has only two pure strategies.

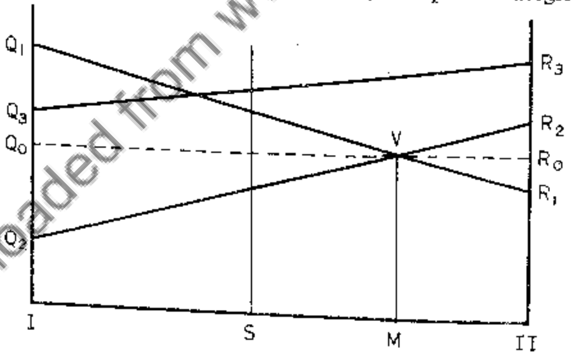

Fig. II.1

A can choose either of his two pure strategies, or a mixture of both. His choice may be marked on a straight line of length 1, as shown in figure II.1, the points I and II representing respectively the two strategies. Any intermediate point represents a mixture, so that for instance the strategy (1/4, 3/4) will be given by a point whose distances from II and I are in proportion 1 : 3. Points outside the stretch I–II need not concern us here, because they could not be arrived at by two non-negative frequencies.

Perpendicularly above the line I–II we mark heights corresponding to the amounts which A wins when B uses one of his strategies, say the first. In this case, when A uses the pure strategy referred to as I, he gets 4, and if he uses II, he gets 2. Moreover, if he chooses any combination (m, n) with $m+n=1$, he gains $4m+2n$. All these points with their heights above the appropriate points on I–II, lie on the straight line Q_1R_1 in the figure. The lines Q_2R_2 and Q_3R_3, corresponding to B's other strategies, are also shown.

Now assume that A has chosen the mixed strategy represented by S in the diagram. Whatever B does, A will then obtain at least the amount represented by the height of the lowest intersection of the vertical through S with a line corresponding to one of B's strategies (here B's second strategy). Such a height can be found for any point in I–II, and according to the theory of games, A should choose a point to maximize this height. He should therefore choose the strategy represented by M, which is at a distance of 3/4 from I and 1/4 from II. His strategy is thus (1/4, 3/4), as we have already seen in I.3. It gives him a gain of 2·5, which is the height of V above M and the value of the game.

The diagram also shows that B should never use his third pure strategy, because this line is everywhere above that of his second. Whether the first or the second pure strategy is better for B depends on A's choice; if A chooses M, then any combination of B's first two pure strategies leads to the same pay-off.

The reader will appreciate that the pattern of strategy lines can be very different from that of our present example and we mention, in particular, two cases: (*a*) when there is an optimal pure strategy and (*b*) when there is more than one optimal strategy for A. These are illustrated in Figure II.2.

Now we turn to B. If there is a saddle point, the answer is trivial. If there is none, then it is clear from figure II.1 that B must use his first two pure strategies only and it will, in fact, be proved later (in IX.2) that if A has only two pure

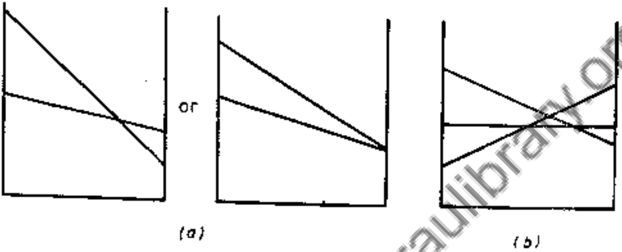

Fig. II.2

strategies, B need never use a combination of more than two himself. B's optimal mixture is characterized by the fact that whatever A does, the pay-off is the same. Hence we can find B's mixture by drawing Q_0R_0 parallel to I–II through V and taking the ratio Q_0Q_1/Q_0Q_2, or also R_0R_1/R_0R_2 (see figure II.1). These ratios are equal by virtue of simple properties of similar triangles and also equal to the ratio of the frequencies with which the pure strategies are used in the optimal mixture.

2. Another possible procedure for finding B's optimal strategies would consist of a straightforward generalization of the diagram to three dimensions. However, such a generalization would be impossible if B had more than three pure strategies. We therefore introduce another method, which has the advantage that it can give an overall picture referring simultaneously to both players (still assuming that one of them has only two pure strategies to choose from).

In this new method any strategy of B is represented by a point, the coordinates of which are the two pay-offs against the two pure strategies of A. In our illustrative example they will be the points $P_1=(4, 2)$, $P_2=(1, 3)$, and $P_3=(3, 4)$.

If B uses any mixture $pP_1+qP_2+rP_3$ (the notation will be found self-explanatory), then the pay-off to A is the weighted average of the pay-offs against each pure strategy, the weights being p, q and r, adding up to 1. Hence B's strategy (p, q, r) will be represented by a point whose abscissa and ordinate

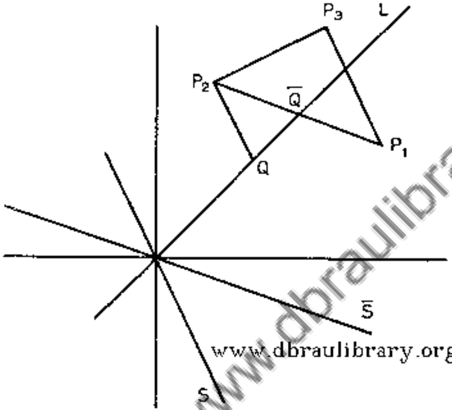

Fig. II.3

are the appropriately weighted averages of the abscissae and ordinates of the points of B's pure strategies. We therefore draw the 'strategy polygon' of B. This is a polygon which includes all points representing pure strategies and also any point that lies on a straight line between any other two points of the polygon. Thus it is a convex region and is known as the *convex hull* of the given points. All its vertices correspond to pure strategies, but some pure strategies may also correspond to points inside the polygon.

We are only concerned with points inside and on the edges and vertices of the strategy polygon, because points outside it could only occur if at least one weight were allowed to be negative. Now given the point of the polygon, or the strategy chosen by B, A should choose from his own strategies the one giving the highest pay-off. If he chooses his first pure strategy he gains (the amount of) the abscissa of B's strategy point, and if he chooses his second pure strategy he gains the ordinate of B's strategy point. Therefore if this point is below the line L which bisects the first and the third quadrant, then A should choose his first pure strategy, and if it is above, his second. If the point is on L, then A gets the same amount whichever strategy he chooses, whether pure or mixed. B should select that point of the polygon which gives A least, i.e., whose larger coordinate is smallest.

To find this point geometrically, imagine a right-angled wedge with its vertex on L and its sides parallel to the axes of coordinates and extending in their negative directions, starting from the left of and below* the polygon and moving upwards until it reaches the polygon. The point of contact, or one of them, indicates an optimal strategy of B.

It will be seen that if L intersects the polygon at all, and if the intersection with the smallest abscissa, or ordinate, is on an edge of negative slope (i.e. one that slopes downwards to the right), then that point of intersection gives B's optimal strategy ($\bar{Q}$ in figure II.3). However, if these conditions are not both fulfilled, then it will be one of the vertices which

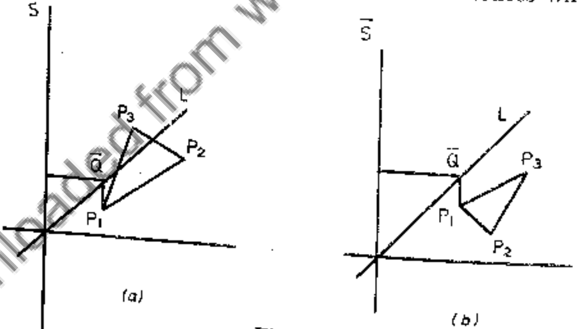

Fig. II.4

indicates B's strategy, and this is then a pure strategy, as will be seen in figure II.4. We can express the construction also by saying that we draw a lower horizontal and a left-hand

* We imagine that the axis of abscissae is horizontal, that of ordinates vertical, and that the directions of the axes are positive and negative as in the diagram.

vertical support to the polygon and if the bisector meets first one of these, then the solution contains a pure strategy of B, given by the vertex from which that horizontal or vertical emanates. It should be noted that this can happen even if, on its further course, the bisector L meets a side of the polygon (as in figure II.4(a)). In (a) neither pure strategy of A dominates the other, while in (b) the first pure strategy dominates. (The two parts of figure II.4 refer to the games defined by

$$(a)\quad \begin{pmatrix} 2 & 5 & 3 \\ 1 & 3 & 4 \end{pmatrix} \text{ and } (b)\quad \begin{pmatrix} 3 & 4 & 5 \\ 2 & 1 & 3 \end{pmatrix}$$

respectively.)

Clearly, if there is a point P_a in the polygon above and to the right of a point P_b, also in the polygon, then P_a need not be considered because P_b has both coordinates smaller than P_a. This explains the significance of the slopes of the edges of B's strategy polygon. In figure II.3 the first intersection of L with the polygon is with a line of negative slope and the horizontal and vertical supports are therefore irrelevant in this case.

3. We now look for a representation of A's strategies which can be superimposed on the same diagram as that showing B's and thus give a complete picture for both players at one glance. We must therefore find a way of showing the gains of A and how they depend on his choice of strategy.

Let us denote A's two pure strategies by S_1 and S_2. Let him choose the mixture (m, n) with $m+n = 1$. If B chooses the strategy point $P = (x_1, y_1)$, then the pay-off to A is mx_1+ny_1, and we want to find a convenient geometrical meaning for this expression.

We draw a line through the origin whose equation is $mx+ny = 0$ and call it A's strategy line. Since both m and n are non-negative, this line will either coincide with one of the axes, or it will have a negative slope. Notice also that if $m=1$, and hence $n=0$, the strategy line will be the axis of ordinates and mx_1+ny_1 is then simply the distance of P from this line, i.e. the abscissa of P. A similar remark applies to $m=0$ and $n=1$, and the ordinate of P. However, the distance of P from a line $mx+ny=0$ is, in general, not simply the average of abscissa and ordinate with weights m and n respectively. A different representation is then required.

Draw through P a parallel to $mx+ny=0$, i.e. the line whose equation is $mx+ny = mx_1+ny_1$; intersect this line with $x=y$, i.e. with L (see figure II.3). The point of intersection, Q, say, has an abscissa, and ordinate, equal to $(mx_1+ny_1)/(m+n) = mx_1+ny_1$, which is the expression for the pay-off.

Now consider again A's strategy line $mx+ny=0$. B should choose that strategy point available to him through which a parallel to A's strategy line intersects L in a point with as small an abscissa (or ordinate) as possible. In general, this will be the vertex which is first met if a line parallel to A's strategy line sweeps upwards from the left bottom to the right top. In special cases, viz. when the first side met of the polygon is parallel to the strategy line, all points of it are equally good for B and, in particular, any of the two vertices on this side, which correspond to pure strategies.

A will, of course, choose the strategy line, say $\bar{S}$, which makes the abscissa of that point on L, which B obtains by the argument just described, as large as possible (see the point $\bar{Q}$ in figures II.3 and II.4). If B's polygon intersects L first (i.e. nearest to the left-hand bottom) in a side of negative slope, then A cannot do better than to choose his line parallel to that side (as in figure II.3). But if there is no intersection at all, or first an intersection with an edge of positive slope (as in figure II.4), then A should choose that of the two axes of coordinates which is farther away from B's best point. (Remember that he cannot choose a line with positive slope!)

In all these cases it is seen that B's optimal strategy is the correct answer to A's optimal choice, and vice versa. In other words, a solution always exists, at least in terms of mixed strategies. The optimal strategies can be found from the diagrams. Thus, in figure II.3, B's optimal strategy is given by (1/2, 1/2, 0), while A's is given by the coefficients of the equation of $\bar{S}$. The numerical values of these coefficients are easily derived from the fact that the line $\bar{S}$ is parallel to P_1P_2, i.e. to

$$(a_{22}-a_{21})x+(a_{11}-a_{12})y = 0$$

and we find that the coefficients have the ratio 1/4 : 3/4. In figure II.4 the optimal strategies form a saddle point and are even easier to find. The value of the game is the abscissa of $\bar{Q}$ in each case.

These considerations contain implicitly a geometrical proof of the fundamental theorem, when one of the players has only two pure strategies. The next chapter contains an algebraic proof which holds generally.

ALGEBRA OF THE THEORY OF GAMES

We shall now introduce algebraic notations and algebraic treatment, in order to obtain general results.

Let A have n strategies and B have m strategies. The pay-off for the combination of A's i-th and B's j-th strategy will be denoted by a_{ij}. If A uses his pure strategies in proportions $x_1 : \ldots : x_n$ where $x_1 + \ldots + x_n = 1$, then we denote this mixed strategy by $(x) = (x_1, \ldots, x_n)$. Similarly, a mixed strategy of B is denoted by $(y) = (y_1, \ldots, y_m)$, where $y_1 + \ldots + y_m = 1$. Of course, $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ are all non-negative. This will always be understood whenever we use the notation (x) or (y).

If A and B choose, respectively, strategies (x) and (y) then the average pay-off is $\sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j a_{ij}$.

Now consider how the strategies should be chosen. If A chooses the strategy $(\bar{x})$, say, then he must fear that B will choose that (y) which makes $\sum_i \sum_j \bar{x}_i y_j a_{ij}$ as small as possible. Now if B can do this by a mixed strategy, then he can also do it by at least one pure strategy, because this pay-off is a weighted average of amounts $\sum_i \bar{x}_i a_{ij}$ with positive weights y_j adding up to unity, and can therefore not be smaller than the smallest of the $\sum_i \bar{x}_i a_{ij}$. Hence A need only scrutinize B's pure strategies and note which gives the smallest pay-off for any of his (A's) own strategies. Given $(\bar{x})$, we write for that smallest outcome

$$\min_j \sum_i a_{ij}\bar{x}_i = \sum_i a_{ij(\bar{x})}\bar{x}_i \quad . \quad . \quad . \quad (1)$$

where $j\,(\bar{x})$ describes that strategy for which the minimum is obtained, when $(\bar{x})$ is given.

A wishes to make this minimum as large as he can. He will therefore choose his (x) accordingly, i.e. so that he obtains

$$\max_{(x)} \min \sum_i a_{ij} x_i = \max_{(x)} \sum a_{ij(x)} x_i = v_1, \text{ say} \quad . \quad (2)$$

We describe v_1 as the *maximin* and note that the 'pure' maximin, viz. $\max_i \min_j a_{ij}$, is clearly smaller than or equal to v_1.

We may repeat this argument from B's point of view. We start by assuming that he chooses $(\bar{y})$ and then observes that A would do best by choosing that strategy which produces

$$\max_i \sum_j a_{ij}\bar{y}_j = \sum_j a_{i(\bar{y})j}\bar{y}_j \quad . \quad . \quad . \quad (3)$$

B will therefore choose his own strategy such that the pay-off becomes

$$\min_{(y)} \max_i \sum_j a_{ij}y_j = \min_{(y)} \sum_j a_{i(y)j}y_j = v_2, \text{ say} \quad . \quad (4)$$

Again it is clear that the 'pure' *minimax*, viz. $\min_j \max_i a_{ij}$, is larger than or equal to v_2.

It should be understood that when we speak of a maximum or a minimum, there may be more than one strategy which produces it and when we mention *the* optimal strategy, we may be referring to just one of them.

The Fundamental Theorem of the theory of games, or Minimax Theorem (M.M.T.) states that always $v_1 = v_2$. It is easy to show that $v_1 \leqslant v_2$, i.e. that the maximin is never larger than the minimax. Indeed, define $(\bar{x})$ by

$$v_1 = \min_j \sum_i \bar{x}_i a_{ij}$$

and $(\bar{y})$ by

$$v_2 = \max_i \sum_j \bar{y}_j a_{ij},$$

then $v_1 \leqslant \sum_i \sum_j \bar{x}_i \bar{y}_j a_{ij}$ and $v_2 \geqslant \sum_i \sum_j \bar{x}_i \bar{y}_j a_{ij}$.

The right-hand sides of the inequalities are equal, so $v_1 \leqslant v_2$. The M.M.T. will therefore be proved if it is also shown that $v_1 \geqslant v_2$.

It also follows that if the pure maximin equals the pure minimax (i.e. if there is a saddle point: the equivalence of these two statements will be shown in III.3), then they both equal $v_1 = v_2$.

We shall prove the M.M.T. in the next section. Assuming its validity, we see from (1) and (3) that

$$\sum_i \sum_j a_{ij}\bar{x}_i\bar{y}_j = \sum_i a_{ij(\bar{x})}\bar{x}_i = \sum_j a_{i(\bar{y})j}\bar{y}_j.$$

These equations tell us that the two strategies $(\bar{x})$ and $(\bar{y})$ which were chosen to give, respectively, the best protection against the pure strategies indicated by $j(\bar{x})$ and $i(\bar{y})$, also give the best protection against one another. Consequently, if A chooses $(\bar{x})$, B need not choose $j(\bar{x})$ to incur the least possible loss, but might equally well choose (y), which will generally be a mixed strategy. This could not be true unless every pure strategy which forms part of $(\bar{y})$ gave, together with $(\bar{x})$, the same pay-off as does the pure strategy $j(\bar{x})$.

Nevertheless, it is not true to say that B can use either $(\bar{y})$, or $j(\bar{x})$, and view A's choice with equanimity. On the contrary, he can do this only if he chooses $(\bar{y})$, because otherwise A might be able to deviate from his optimal strategy and get more than the minimax v_2.

2. We come now to the proof of the M.M.T. Many proofs are now known. Historically the first was that given by von Neumann in Ref. 22. It is by no means elementary, nor is that given by the same author in Ref. 23. J. Ville produced an elementary proof in 1938 (Ref. 27), of which more will be said later, and Ref. 24 contains, on pp. 153 ff. the proof reproduced here. A further proof, actually for a more general theorem, will be given when we deal with Linear Programming, in VIII.3.

It is, then, our aim to prove that $v_1 = v_2$, and we know already that it is sufficient to show that $v_1 \geqslant v_2$.

The proof follows from two lemmas which we must first demonstrate.

Lemma 1: The theorem of the supporting hyperplanes. (T.S.H.)

Let the values a_{ij} $(i=1, \ldots, n;\ j=1, \ldots, m)$ be given. We speak of a set $(a_{1j}, \ldots, a_{nj})$ as a point A_j in n-dimensional space and say that a further point $A=(a_1, \ldots, a_n)$ belongs to the convex hull of $A_1, \ldots, A_m$ if we can find non-negative values $t_1, \ldots, t_m$ adding up to unity such that

$$a_i = t_1 a_{i1} + \ldots + t_m a_{im}$$

for all

$$i=1, \ldots, n.$$

This hull is indeed convex, i.e. if two points belong to it, then all points on a straight line between them also belong to it. The reader will have no difficulty in deriving this from the definition. The T.S.H. (in the form in which we need it) states that, if the point $0=(0, \ldots, 0)$ does not belong to the convex hull of $A_1, \ldots, A_m$, then one can find values $s_1, \ldots, s_n$ such that for any point A belonging to the convex hull we have

$$s_1a_1+\ldots s_na_n>0.$$

For $n=2$ or 3 this is intuitively clear (see figure III.1). These

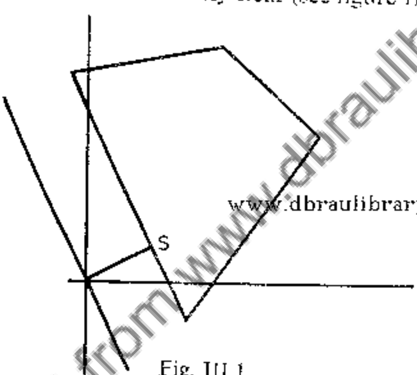

Fig. III.1

special cases also explain the name of the theorem, a hyperplane being the generalization of line and plane to higher dimensions.

We now give the proof of the lemma for general n.

If 0 does not belong to the convex hull C, then there is a point in C, different from 0, for which the square of the 'distance from 0', i.e. the sum of the squares of C's coordinates, is smallest. Let it be the point $S=(s_1, \ldots, s_n)$. Consider now an arbitrary point of C, say $A=(a_1, \ldots, a_n)$. For every t between 0 and 1, the point with coordinates $ta_i+(1-t)s_i$ will also belong to C. Moreover, it will be at least as distant from 0 as S is, i.e.

$$\sum [ta_i+(1-t)s_i]^2 \geq \sum_i s_i^2.$$

It follows by elementary algebra that, if $t>0$,

$$2\sum_i s_i(a_i-s_i)+\sum_i (a_i-s_i)^2 t \geq 0.$$

If t tends to zero, the inequality tends to

$$\sum_i s_i(a_i-s_i)\geq 0, \text{ i.e. } \sum_i s_i a_i \geq \sum_i s_i^2 > 0$$

(because S is different from 0). This proves the theorem.

Lemma 2: The theorem of the alternative for matrices. (T.A.M.)

Let the values a_{ij} be given and consider the convex hull C of the points $A_1, \ldots, A_m$ (defined as before) and of the points $(1, 0, \ldots, 0)$, $(0, 1, 0, \ldots, 0)$, $\ldots$ $(0, 0, \ldots, 0, 1)$. The point 0 either belongs to C, or does not belong to it. The T.A.M. states that in the former case there exists some (y) such that

$$a_{i1}y_1+\ldots+a_{im}y_m \leq 0 \text{ for all } i, \qquad \text{(i)}$$

and that in the latter there exists some (x) such that

$$a_{1j}x_1+\ldots+a_{nj}x_n>0 \text{ for all } j. \qquad \text{(ii)}$$

Case 1. The point 0 belongs to C, i.e. there exist values $t_1, \ldots, t_{m+n}$, all non-negative and adding up to unity, such that

$t_1a_{i1}+\ldots+t_ma_{im}+t_{m+i}=0$, or $t_1a_{i1}+\ldots+t_ma_{im}=-t_{m+i}\leq 0$.

Now $t_1+\ldots+t_m$ must be positive, since otherwise all t_i $(i=1, \ldots m+n)$ would be zero and could not add up to 1. It follows that $y_j=t_j/(t_1+\ldots+t_m)$ satisfy (i).

Case 2. (figure III.1) The point 0 does not belong to C. Then, by the T.S.H., we can find $(s_1, \ldots, s_n)$ such that

$$s_1a_{1j}+\ldots+s_na_{nj}>0 \qquad \text{for all } j\ (1, \ldots, m)$$

and

$$s_1>0, \ldots s_n>0$$

(because the points $(1, 0, \ldots, 0)$ etc. belong also to C.) Consequently the values $x_i=s_i/(s_1+\ldots+s_n)$ satisfy (ii). This proves Lemma 2, the T.A.M.

This lemma will now be used to prove the Fundamental

Theorem (M.M.T.). In case 1 above we have (y) such that, for all i, $\sum_j a_{ij}y_j \leqslant 0$, and hence also $\max_i \sum_j a_{ij}y_j \leqslant 0$.

Therefore

$$v_2 = \min_{(y)} \max_i \sum_i \sum_j a_{ij}y_j \leqslant 0.$$

In the second case, there exists (x) such that, for all j, $\sum_i a_{ij}x_i > 0$, and hence also $\min_j \sum_i a_{ij}x_i > 0$. Therefore

$$v_1 = \max_{(x)} \min_j \sum_i \sum_j a_{ij}x_i > 0.$$

Thus we have shown it is not possible that simultaneously $v_1 < 0$ and $v_2 > 0$.

Now consider the pay-off table with entries $a_{ij} - k$, where k is some arbitrary constant, positive or negative. In exactly the same way as before we can show that it is not possible that $v_1 < k < v_2$ should hold, whatever k. Therefore v_1 can never be smaller than v_2, which proves the M.M.T.

Because of the great mathematical interest of this theorem, we outline J. Ville's proof mentioned above. It follows from (4) that

$$\max_i \sum_j a_{ij}y_j \geqslant v_2$$

for all (y). In other words, there is for every (y) at least one i such that

$$\sum_j a_{ij}y_j \geqslant v_2.$$

It is shown in Ref. 27 that it is then possible to find an $(\bar{x})$ such that

$$\sum_j a_{ij}\bar{x}_i y_j \geqslant v_2$$

for every (y). It follows that

$$v_2 \leqslant \min_{(y)} \sum_i a_{ij}\bar{x}_i y_j \leqslant \max_{(x)} \min_{(y)} \sum_i \sum_j a_{ij}x_i y_j$$
$$= \max_{(x)} \min_j \sum_i \sum_j a_{ij}x_i y_j = v_1.$$

3. In I.3 we introduced the concept of a *solution* of a game.

We now give a precise definition of this on rather different lines, which we later show to be equivalent to that in 1.3.

Any pair $(\bar{x})$ and $(\bar{y})$ is a solution of the game (a_{ij}) if

$$\sum_i \sum_j a_{ij}\bar{x}_i\bar{y}_j = \max_{(x)} \min_{(y)} \sum_i \sum_j a_{ij}x_iy_j = \min_{(y)} \max_{(x)} \sum_i \sum_j a_{ij}x_iy_j.$$

Nothing that we have said leads to the conclusion that there can only be one solution to a game. But since there can only be one single maximin or minimax, it is clear that all solutions lead to the same value $v = v_1 = v_2$, viz. the value of the game. Now let both the pair $(\bar{x})$, $(\bar{y})$ and $(\bar{\bar{x}})$, $(\bar{\bar{y}})$ be solutions. Then the pair $(\bar{x})$, $(\bar{\bar{y}})$ is also a solution, because, by virtue of the definition of a solution,

$$\sum_i \sum_j a_{ij}\bar{x}_i\bar{\bar{y}}_j \geqslant \sum_i \sum_j a_{ij}\bar{x}_i\bar{y}_j \geqslant \sum_i \sum_j a_{ij}\bar{\bar{x}}_i\bar{y}_j$$

and also

$$\sum_i \sum_j a_{ij}\bar{\bar{x}}_i\bar{y}_j \geqslant \sum_i \sum_j a_{ij}\bar{x}_i\bar{y}_j \geqslant \sum_i \sum_j a_{ij}\bar{x}_i\bar{\bar{y}}_j$$

It follows that all terms in these inequalities must have equal value.

We now consider the following question which is clearly of practical importance. Let a pair of strategies (x') and (y') be given. How can we test whether this pair is a solution of a given game?

If (x'), (y') is a solution, then necessarily

(a) $\sum_i \sum_j a_{ij}x'_iy'_j \leqslant \sum_i \sum_j a_{ij}x'_iy_j$ for all (y)

and

(b) $\sum_i \sum_j a_{ij}x'_iy'_j \geqslant \sum_i \sum_j a_{ij}x_iy'_j$ for all (x).

But these two inequalities are also sufficient for (x'), (y') to be a solution, as will now be shown. In words, we prove that if both players have chosen their strategies and none would be better off by changing his, as long as the other keeps his unchanged, then the pair is a solution. The proof is as follows:

From (a) and (b) we have

(c) $\max_{(x)} \sum_i \sum_j a_{ij}x_iy'_j = \sum \sum a_{ij}x'_iy'_j = \min_{(y)} \sum_i \sum_j a_{ij}x'_iy_j$

But clearly we also have

$$\text{(d)}\quad \min_{(y)} \max_{(x)} \sum_i \sum_j a_{ij}x_iy_j \leqslant \max_{(x)} \sum_i \sum_j a_{ij}x_iy'_j$$

and

$$\text{(e)}\quad \max_{(x)} \min_{(y)} \sum_i \sum_j a_{ij}x_iy_j \geqslant \min_{(y)} \sum_i \sum_j a_{ij}x'_iy_j$$

Because the left-hand sides of (d) and (e) are equal, they must be equal to the middle term of (c) as well. This completes the proof. It also shows that the definition of a solution given here is equivalent to that of I.3.

If the game has a saddle point, i.e. if we can find pure strategies (x_0) and (y_0) satisfying (a) and (b), then it follows from (c), (d) and (e) that the value of the game equals the maximum of the row minima and also the minimum of the column maxima. For a large pay-off matrix this provides a much more convenient criterion for the existence of a saddle point than a search for it throughout the whole matrix.

4. If a game has no saddle point, then the determination of a solution or even of its value is no easy matter. We refer this question to a later part of the book (Chapter IX in particular) because it will be answered by a computing routine which solves a more general problem. This latter will now be introduced.

Let a pay-off matrix be given. If A chooses (x), then he can be certain of obtaining at least $\min_j \sum_i a_{ij}x_i = v$. We have therefore

$$a_{1j}x_1 + \ldots + a_{nj}x_n \geqslant v \qquad \text{for } j=1, \ldots, m.$$

$$x_1 + \ldots\ldots + x_n = 1$$

and

$$x_1, \ldots, x_n \geqslant 0.$$

A wants to make v as large as possible. (It will then be the value of the game.) This value is not necessarily positive, but will be so if we add a constant to all a_{ij} so as to make them all positive. This increases the value of the game by the same constant, but does not change the solution. We may therefore assume that v is positive and introduce new variables $x_i' = x_i/v$ which will then also assume non-negative values only. Dividing the inequalities by v, we obtain the system

$$\sum_{i=1}^{n} a_{ij}x_i' \geqslant 1 \qquad \text{for } j=1, \ldots, m$$

and

$$\sum_{i=1}^{n} x_i' = 1/v.$$

The right-hand side of the last equation must be minimized.

Repeating this argument for B, we obtain the set

$$\sum_{j=1}^{m} a_{ij}y_j' \leqslant 1 \qquad \text{for } i=1, \ldots, n$$

and

$$\sum_{j=1}^{m} y_j' \qquad \text{is to be maximized.}$$

We have thus found that the games problem can be reduced to a special case of a more general problem, which can be expressed as follows:

Let a_{ij}, b_j and c_i be given constants, for $i=1, \ldots, n$ and $j=1, \ldots, m$. Find non-negative values x_i and y_j such that

$$\sum_i a_{ij}x_i = b_j, \text{ making } \sum_i c_i x_i \text{ as small as possible}$$

and

$$\sum_j a_{ij}y_j = c_i, \text{ making } \sum_j b_j y_j \text{ as large as possible.}$$

The two problems, which are respectively generalizations of those which A and B have to solve, are in a relation to one another, which we call *dual*. It will be shown (in VIII.3) that the maximum of $\sum_j b_j y_j$ equals the minimum of $\sum_i c_i x_i$. This is, in the special case of the theory of games, equivalent to the M.M.T.

Yet another way of reducing the games problem to one of L.P. will be shown in IX.1.

5. As long as we consider only pure strategies, the pay-offs

form a discrete set. With the introduction of mixed strategies this is, of course, no longer true. We can now state that A has at his disposal a set of strategies $\sum_i a_{ij}x_i = X(x_1, \ldots, x_n)$, say, and B the set $\sum_j a_{ij}y_j = Y(y_1, \ldots, y_m)$. The pay-off resulting from any particular pair of strategies is $\sum_i \sum_j a_{ij}x_i y_j = P(x_1, \ldots, x_n; y_1, \ldots, y_m)$, say.

If we consider a pay-off function $A_{00} + A_{01}x + A_{10}y + A_{11}xy$ where x and y take values between 0 and 1, then this game is equivalent to a finite discrete game in which the players have two strategies each. But we can further generalize by assuming for X and Y some other, not necessarily linear, functions and for P some other, not necessarily bilinear, function. In general cases, e.g. when the pay-off function is not continuous, our proof of the M.M.T. loses its validity. We cannot enter here into more detail than to mention that J. Ville (Ref. 27) has proved that an analogous theorem holds if the pay-off is a continuous function of two variables in the closed unit square (see also Ref. 20, p. 186), while more complicated cases are tackled, for instance, in Refs. 16 and 17.

Finally, we mention that a very thorough account of the theory is given in the first two chapters of Ref. 4 which is, however, not a book for beginners.

CHAPTER IV

AN OUTLINE OF LINEAR PROGRAMMING

The term Linear Programming (L.P.) describes the solution of the following problem:

Let a set of equations, or *constraints*, be given as follows:

$$a_{11}x_1 + \ldots + a_{n1}x_n = b_1$$
$$\qquad \ldots \qquad (1)$$
$$a_{1m}x_1 + \ldots + a_{nm}x_n = b_m$$

It is required to find non-negative values of the variables $x_1, \ldots, x_n$, which satisfy the constraints and make the value of a given linear form (L.F.)

$$C = c_1x_1 + \ldots + c_nx_n \quad . \quad . \quad . \quad (2)$$

as small as possible.

We can assume that the m equations are not linearly dependent, because otherwise we could drop one or more of them.

The same problem, but without the condition that the variables must be non-negative, can be transformed into the above by substituting $x_i = x_i' - x_i''$. Moreover the above form is also appropriate to represent the case when some or all equations are replaced by inequalities, since

$$a_{1j}x_1 + \ldots + a_{nj}x_n \leqslant b_j \qquad \text{or} \qquad a_{1j}x_1 + \ldots + a_{nj}x_n \geqslant b_j$$

is equivalent to

$$a_{1j}x_1 + \ldots + a_{nj}x_n + x_{n+j} = b_j \quad \text{or} \quad a_{1j}x_1 + \ldots + a_{nj}x_n - x_{n+j} = b_j$$

with the proviso that x_{n+j} is also non-negative. We call these x_{n+j} *additional* (or *slack*) variables.

If it were our aim to maximize a given L.F., then we could reduce this problem to the previous one by requiring that the negative of the L.F. should be minimized. In this case we must, of course, change the sign of the smallest value of the minimized L.F., in order to obtain the largest value of the original L.F. which was to be maximized. Alternatively, we could consider the maximizing problem as one in its own right. We shall therefore speak of 'optimizing', whenever we mean either maximizing or minimizing, as the case may be.

We saw in the last chapter that an algebraic formulation of the games problem leads to L.P. We now give further examples which show how ability to deal with an L.P. problem may prove useful in a great variety of situations.

Perhaps the earliest problem of this type to appear in print is due to F. L. Hitchcock (Ref. 12, see also Ref. 14 and Chapter XXIII of Ref. 15 by G. B. Dantzig). Assume that there are a_s ships in the ports P_s and that we wish to move b_t of them to ports of destination Q_t. We assume that it would not matter which ship goes where, if it were not for the fact that the cost of moving a ship from one port to the other is different for different pairs of ports.

We denote the number of ships to be moved eventually from P_s to Q_t by y_{st} and assume that the cost of thus moving one ship is known and is c_{st}. It is our task to make the total removal process as cheap as possible, i.e. to minimize $\sum_s \sum_t c_{st} y_{st}$, subject to $\sum_s y_{st} = b_t$ for all t, and to $\sum_t y_{st} = a_s$ for all s, and remembering, of course, that the y_{st} must be non-negative. If we have also information on the cost of transportation between ports of the same set, and perhaps also to and from ports which are neither P nor Q, then we can also consider the possibility of trans-shipment, i.e. calling at intermediate ports. Moreover, the cost of storage and docking could also be taken into consideration. (For a simple method of solution see Ref. 7a.)

This transportation problem, at least in its simplest form, is formally identical with that of allocating n persons to n jobs in such a way that their total value is maximized, when the value of any person in any job is known and measurable.

A problem from another field is this: You are told that you need at least b_1 units of fats, b_2 of proteins and b_3 of carbohydrates per week to remain strong and healthy and you know also that the units of various food items, say F_i, contain respectively a_{ij} $(j=1, 2, 3)$ units of these ingredients. You wish to buy food at least total expense, but so that you meet the nutritional requirements. If, then, c_i is the price of one unit of F_i and x_i the amount which you buy of it, then you want to minimize $\sum_i c_i x_i$ subject to $\sum_i \sum_j a_{ij} x_i \geq b_j$ for $j=1$, 2 and 3. We have introduced the $\geq$ sign, because it can happen that it is cheaper to buy some item in excess of the minimum requirement (Cf. Ref. 25).

Variations of this problem may be more realistic than this simple form of it. For instance, you might not like to have too much starch in your diet and may be prepared to pay for limiting the amount of certain food items. In a related problem, that of provisioning at sea, it can be argued that what matters is economy not primarily in cost but rather in weight or bulk, and that one must therefore minimize the over-supply which leads to the difference between the two sides of the inequalities above.

L.P. models are appropriate in many more situations. A warehouse owner wants to buy when prices are lowest and to sell when they are highest, but his storing capacity is limited, and this restriction affects different goods in different ways (Ref. 6); there are different ways in which blending of petrol can satisfy various requirements and the cost, and the selling prices for the different products are also different (see Ref. 10); one might wish to start an advertising campaign and select various channels whose efficiency is expressible in quantitative terms, so as to minimize the cost or, conversely, so as to maximize the effect for a given total cost. Similar examples will be found in Ref. 15.

Most if not all these illustrations had a flavour of economics and we mention a final example from the theory of the firm (see Ref. 11). We assume that a firm can use various production processes, P_j and that their intensity can be measured. We also assume that the intensity is such that when it is multiplied by a constant, the requirements for all materials R_i used in the process are multiplied by the same factor. Denote the amount of R_i needed in P_j when its intensity is unity by a_{ij}. Then, if P_j is used at the (non-negative) intensity level y_j, the requirement which arises for

R_i is $\sum_j a_{ij}y_j$. If the profit of the product of process P_j at unit level is p_j, we may imagine that the firm tries to use its various production methods at such levels that the total profit $\sum_j p_j y_j$ is maximized, subject to the fact that not more than r_i is available of R_i, i.e. subject to $\sum_j a_{ij}y_j \leqslant r_i$.

We have here assumed that the income from sales is proportional to the amount sold and that the amount of raw material necessary for any process is proportional to its level. Hence the name linear (i.e. proportional) programming. In more recent literature this term has been replaced by Activity Analysis of Production and Allocation (the title of Ref. 15), but we retain the older and shorter term as a convenient name for a well-defined problem. (For further examples consult also Ref. 19.*) Economists would, no doubt, have many more remarks to make as to the validity of the models, but we do not wish to intrude into their domain.

One might perhaps think, for a moment, that such problems could be solved using differential calculus. However we shall soon see that the solutions we are seeking will be found, not in general at points where some derivative is zero, but always at points on the boundaries to the region of possible values of the variables introduced by the requirement that these should be non-negative. It is therefore clear that special techniques are required to handle these problems efficiently.

2. Before showing any of the methods which have been devised for L.P., it is of interest to answer a question which may be asked in connection with practical problems, such as that of the advertiser which we have mentioned. Assume, first, that m constraints are given and that it is required to minimize an L.F., C. We may think of this as the cost of a process, and of the right-hand side of the first constraint, say b_1, as of its product whose amount has been fixed, while the remaining $m-1$ constraints represent side-conditions. It may then be asked whether, given the minimum of C as the expense

* See also Refs. 29, 30 and 31.

which can be allowed, it would be possible to produce, in fact, more than b_1. (See Ref. 29, App. to first paper.)

Imagine that only the second to m-th constraints exist and consider all values b_1 and C which can be obtained by those x_i which satisfy these constraints. If values b_1' and C', and also values b_1'' and C'' can be produced respectively by making $x_i = x_i'$ or $x_i = x_i''$, then the pair $tb_1' + (1-t)b_1''$ and $tC' + (1-t)C''$ can also be produced, by the analogous combination of x_i' and x_i''. Hence if we consider, in a Cartesian plane, points (b_1, C) which are obtainable from variables satisfying the last $m-1$ constraints, these points will form a convex area.

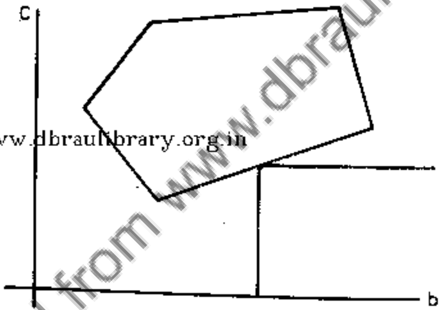

Fig. IV.1

For any given b_1 the smallest value of C will be the ordinate of the lower intersection of a vertical with abscissa b_1 and the convex area. It follows that, if the value of b_1 lies between the abscissa of the point of the smallest C (under constraints 2 to m, i.e. of the lowest point of the area) and the maximum b_1 obtainable (again under constraints 2 to m), then no larger b_1 can be obtained for the corresponding C. The necessary qualification of this statement when the area has a horizontal lower edge is obvious.

3. The various computational routines for solving the L.P. problem differ in the way in which the goal is reached. In

this chapter we proceed by first finding a set of non-negative values $x_1, \ldots, x_n$ that satisfy the constraints; only then shall we be concerned with finding out whether other non-negative values of the variables would produce a smaller value of C. In Chapters VIII and XI we introduce methods in which we start from values that are not necessarily non-negative and then seek values that are all non-negative.

The number of equations, m, will generally be smaller than the number of variables, n; but even so we cannot be certain that a solution in non-negative variables exists. This is illustrated by the constraints $x_1+x_2=1$, $x_1-x_3=2$. By subtraction we obtain $x_2+x_3=-1$, which is clearly impossible with non-negative values.

To make the subsequent discussion simpler, we introduce the following terms:

A *solution* is any set x_i which satisfies all constraints, irrespective of the signs of the values. (It does not take any account of the L.F. and is thus not necessarily a solution of the L.P. problem as a whole.)

A *feasible* solution (f.s.) is a solution with non-negative x_i. A *basis* is a set of m variables such that the matrix of their coefficients in the m constraints is not singular. These m variables are *basic variables* (b.v.'s) – in relation to that basis – and the others are *non-basic variables* (n.b.v.'s). The *basic solution* (b.s.) associated with a basis is obtained by putting the n.b.v.'s equal to zero and solving the constraints for the b.v.'s.

A *basic feasible solution* (b.f.s.) is, of course, a basic solution that is also feasible.

It would be more in keeping with usage in algebra to call the aggregate of the sets

$$(a_{u_1j}, \ldots, a_{u_mj})$$

for $j=1, \ldots, m$ a basis (as is done, for instance, in Ref. 9), but we find it convenient to use this term for the set of corresponding variables.

4. The first procedure to be explained is the so-called Simplex Method (S.M.), due to G. B. Dantzig (Ref. 15,

Chapter XXI). The name derives from the accident that one of the first examples to be tackled by this method contained the constraint $\sum_i x_i = 1$, which is the equation of a simplex in n-dimensional geometry. The name is now used for the procedure whatever the form of the (linear) constraints.

In this method our steps are as follows:

To begin with, we find a b.f.s. Then we determine whether the optimum has already been reached. If not, then we delete one of the b.v.'s from the basis and introduce another variable instead. The new basis is treated in the same way and the steps are repeated, if necessary. It will be shown that this process must eventually terminate by either producing the optimum or by showing that the constraints are contradictory, or by showing that the values of the L.F. which can be produced by b.f.s.'s are unbounded (from below if the L.F. is to be minimized, or from above in the opposite case).

A simple illustration of the last-mentioned case is given by $x_1 - x_2 = 0$, maximize $x_1 + x_2$ for non-negative x_1 and x_2.

We assume now that a b.f.s. has been found, by inspection or by the physical nature of the problem. Otherwise complications arise which will be dealt with in Chapter VII.

An example will help to make the principle of the method clear. Let the following problem be given:

$$-2x_1 + x_2 + x_3 = 2$$
$$x_1 - 2x_2 + x_4 = 2$$
$$x_1 + x_2 + x_5 = 5$$

Minimize $x_2 - x_1 = C$, say, for non-negative x_i.

The constraints may, for instance, be thought of as having arisen from inequalities, because each of x_3, x_4 and x_5 occurs only in one of the equations.

In this case it is easy to see, by inspection, that a b.f.s. is given by $x_1 = 0$, $x_2 = 0$, $x_3 = 2$, $x_4 = 2$, $x_5 = 5$. We can express the constraints as equations for the b.v.'s in terms of the n.b.v.'s, as follows:

$$x_3 = 2 + 2x_1 - x_2$$
$$x_4 = 2 - x_1 + 2x_2$$
$$x_5 = 5 - x_1 - x_2$$

We must also express C in terms of the n.b.v.'s. As it happens, this is already done.

We consider next whether the minimum of C has already been reached. Now the coefficient of x_1 in the expression for C is negative. Consequently an increase of x_1 will decrease C further. But if we increase x_1, then x_3, x_4 and x_5 will change and we must be careful not to make any of these variables negative.

For x_3 this danger does not exist at this stage, because an increase of x_1 would increase x_3 as well. Regarding the other variables, we find that x_1 may be increased until it is 2, which makes $x_4=0$, $x_3=6$ and $x_5=3$. This result is quite acceptable, because the number of positive variables is then again three, as before.

The new basis consists now of x_1, x_3 and x_5. To start with the next stage, we express these variables, and C, in terms of the n.b.v.'s x_2 and x_4. This is easily done by solving the second equation for the new b.v., viz. x_1, thus:

$$x_1 = 2+2x_2-x_4$$

and substituting into the other equations, giving

$$\begin{aligned} x_3 &= 6+3x_2-2x_4 \\ x_5 &= 3-3x_2+x_4 \\ C &= -2-x_2+x_4. \end{aligned}$$

We can decrease C further by increasing x_2. We notice also that only x_5 is in danger of becoming negative through an excessive increase of x_2, and that the latter variable can be increased to 1 without causing any harm, but not further. Substitution into the other equations gives $x_1=4$, and $x_3=9$.

Once more we express the b.v.'s, and C, by the n.b.v.'s:

$$\begin{aligned} x_1 &= 4-x_4/3-2x_5/3 \\ x_2 &= 1+x_4/3-x_5/3 \\ x_3 &= 9-x_4-x_5 \\ C &= -3+2x_4/3+x_5/3. \end{aligned}$$

No increase of either x_4 or x_5 would decrease C any further and we have therefore reached the final solution. The lowest obtainable value for C is -3, and it is produced by $x_1=4$, $x_2=1$ and $x_3=9$.

5. It is clear that in carrying out the computations it is unnecessary to record everything that we have written down for the sake of clarity. The process which we have performed can be expressed equally well in a series of tableaux, as follows:

			x_1	x_2
			-1	1
	x_3	2	-2	1
	x_4	2	1*	-2
	x_5	5	1	1
		0	1	-1

			x_4	x_2
				1
	x_3	6	2	-3
-1	x_1	2	1	-2
	x_5	3	-1	3*
		-2	-1	1

			x_4	x_5
	x_3	9	1	1
-1	x_1	4	1/3	2/3
1	x_2	1	$-1/3$	1/3
		-3	$-2/3$	$-1/3$

It is convenient to imagine the equations re-written, so that all variables are on the left-hand side and only the constants on the right-hand side. The rows (apart from the last) refer to the b.v.'s and the columns to the n.b.v.'s, while the last row gives the expression of C. The entries on the left of the names of the variables are the values of their coefficients c_i, say, in the original expression for C, and immediately underneath the names of the n.b.v.'s we have entered their coefficients in that expression; zeros are omitted. The first entry in each row on the right of the name of the variable is its value, i.e. the right-hand side of the equation to which the row refers. The other entries are the coefficients of the n.b.v.'s. The recording of the original coefficients in C allows a useful check because, as will be seen in the next chapter, we obtain the bottom value of each column by multiplying each entry of that column by the coefficient c_i in the same row, adding, and then subtracting from the sum the value c_j at the top of the column. (For instance, in the last tableau, last column, $-2/3+1/3=-1/3$, or in the last column of the second tableau $(-1)(-2)-1=1$.)

At each step we decide to bring one of those n.b.v.'s into the basis, for which the entry at the bottom of the table has positive sign (or negative sign, if we wished to maximize the L.F.), and to remove from the basis that variable (or one of those, in the case of a tie) which shows the smallest ratio of the constant and the coefficient in the column thus chosen. (For instance, in the first tableau, x_1 is to be made basic and since 2/1 is smaller than 5/1, x_4 is to be made non-basic.) The value in the column of the variable to be introduced into, and the row of the variable to be removed from, the basis is called the *pivot* and has been marked by an asterisk.

The transformation of one tableau into the next can also be done by a simple routine which will be described in Chapter VI.

6. If one deals with specific problems which can be solved by L.P., then it is useful to be aware of special features which apply to that particular case. Thus it is easy to show that if the transportation problem is solved by this method, the coefficients of the n.b.v.'s will in all equations and at all stages be 0, 1 or -1. Hence the pivot is always 1 and the numbers of the ships to be moved on specified routes will always be integers. (Ref. 14, Chapter XXIII by G. B. Dantzig.)

Of course, problems which could be solved by L.P. often have special features which allow altogether different methods to be used and we do certainly not wish to suggest that the general methods which we describe are the best in all circumstances. The transportation problem, for one, has such a simple structure that special methods have been proposed for its solution which may lead more quickly to the required result.

CHAPTER V

GRAPHICAL REPRESENTATION OF L.P. (1)

In order to obtain an overall view of the L.P. problem, and also in order to appreciate the nature of the complications that can arise, the reader will welcome a geometrical representation in two dimensions, which is applicable to inequalities in two variables or, equivalently, to equations whose number is two less than that of the variables.

Consider the example of IV.4, to which we shall, in this chapter, refer as example 1. The most natural geometrical picture is that in Cartesian coordinates x_1 and x_2. Since they must not assume negative values, only the first quadrant of the plane is of interest. Having solved the constraints for the b.v.'s, i.e. in the example for x_3, x_4 and x_5, we consider the equations $x_3=0$ etc. as the equations of straight lines. All variables must be non-negative and therefore only one side of any of these lines will concern us. In figure V.1 the lines are drawn and shading indicates areas which are forbidden. By an easily understood analogy, we shall speak of feasible, basic, etc., points.

We observe that only the polygon a b c d e (including edges and vertices) is feasible and we wish to find that point of the polygon whose coordinates minimize $x_2-x_1=C$.

All lines $C=x_2-x_1=$constant are parallel and one of them has been drawn with an arrow indicating the direction of decreasing C. Thus the feasible point which minimizes C is 'c'. It is the intersection of $x_4=0$, i.e. $x_1-2x_2=2$ and $x_5=0$, i.e. $x_1+x_2=5$, so that its coordinates are $x_1=4$, and $x_2=1$. The resulting value of C is $1-4=-3$ and no feasible point produces a smaller value. We have already obtained this result in the last chapter and we observe that the succession of tableaux is reflected in our graph by the succession of points $x_1=x_2=0$ (i.e. 'a'), $x_2=x_4=0$ (i.e. 'b') and, finally, $x_4=x_5=0$ (i.e. 'c').

Our graph also shows that it happens only in exceptional circumstances that the point which leads to the optimum of the L.F. lies on more than two lines, or that there is more than one such point.

There is, of course, no reason why just x_1 and x_2 should be

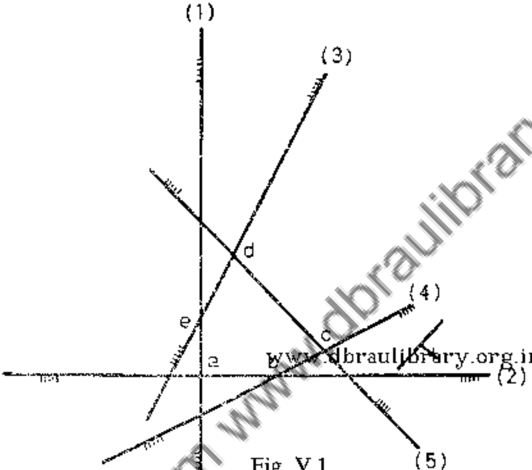

Fig. V.1

the coordinates of the system and any other pair of variables would do equally well. It is also clear that this representation in two dimensions can be used whenever there are just two more variables than constraints.

If we apply this method to the game defined by

$$\begin{pmatrix} 4 & 1 & 3 \\ 2 & 3 & 4 \end{pmatrix}$$

which we have already considered, then we find that the L.P. problem of player A can be written as follows:

$$\begin{aligned} x_1 + x_2 \qquad\qquad &= 1 \\ 4x_1 + 2x_2 - x_3 - \nu &= 0 \\ x_1 + 3x_2 - x_4 - \nu &= 0 \\ 3x_1 + 4x_2 - x_5 - \nu &= 0 \end{aligned}$$

Maximize ν for non-negative $x_1, \ldots, x_5$.

(ν may be positive, negative, or zero.)

If we choose now x_1 and v as the Cartesian variables and express all other variables in terms of these, then we obtain again figure II.1. Note, however, that the line $v=0$ is not a boundary of the area of feasible points.

We now turn to examples intended to introduce the reader to some of the complications that can arise. These examples will deviate only slightly from example 1, but the deviations will be sufficient to alter essential aspects of the problem.

One of the special features of example 1 was the fact that the origin of the system, viz. $x_1=x_2=0$, was a feasible point. We now turn to a case where this is not true and where there may be some difficulty in finding one simply by inspecting the constraints.

Example 2.

$$\begin{aligned} -2x_1+\ x_2+x_3&=-2\\ x_1-2x_2+x_4&=\ \ 2\\ x_1+\ x_2+x_5&=\ \ 5 \end{aligned}$$

Minimize x_2-x_1.

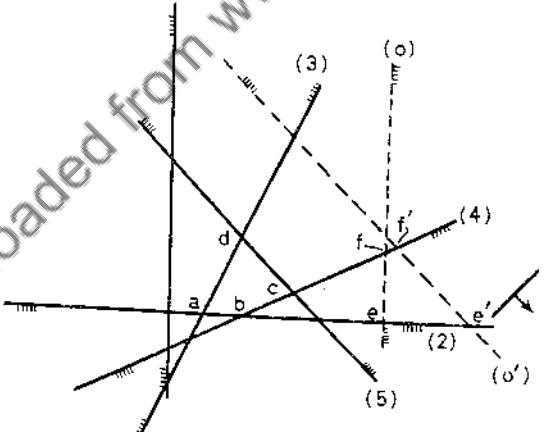

Fig. V.2

(Figure V.2 contains two dotted lines, marked (o) and (o′) which do not concern us at this stage.)

Example 3.

$$-2x_1+x_2+x_3=2$$
$$-x_1+2x_2-x_4=8$$
$$x_1+x_2+x_5=5$$

Minimize x_2-x_1.

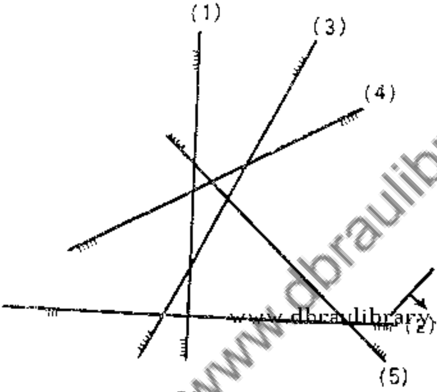

Fig. V.3

The constraints are contradictory for non-negative variables. This can be demonstrated by multiplying them, respectively, by 1, -1 and 1, and adding. We obtain $x_3+x_4+x_5=-1$, which is clearly impossible. Studying the shading in figure V.2 we find indeed that there is no feasible point anywhere in the plane.

The next example illustrates a case with a feasible point that lies on three lines. It is clear that if this point is taken as basic, then one of the b.v.'s vanishes, because there are not enough lines from which the point has a positive distance.

Example 4.

$$2x_1-x_2+x_3=4$$
$$x_1-2x_2+x_4=2$$
$$x_1+x_2+x_5=5$$

Minimize x_2-x_1.

Finally, consider the following example:

Example 5.

$$-2x_1+x_2+x_3=2$$
$$x_1-x_2+x_4=2$$
$$x_1+x_2+x_5=5$$
$$\text{Minimize } x_2-x_1.$$

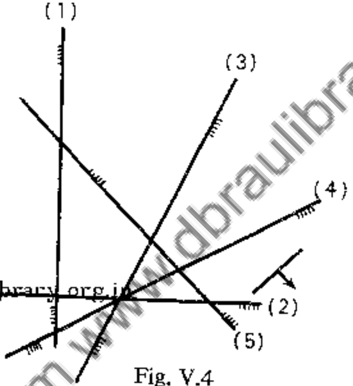

Fig. V.4

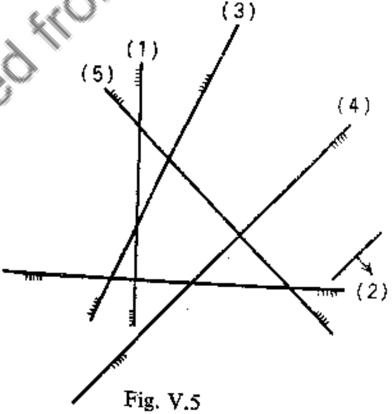

Fig. V.5

In this case the line $x_4=0$ is parallel to the line $x_2-x_1=0$ and hence all feasible points of $x_4=0$ produce the same minimal value of the L.F.

The algebraic consequences of such complications or peculiarities will be considered in Chapter VII, while the following chapter deals with such straightforward cases as example 1.

CHAPTER VI

ALGEBRA OF THE SIMPLEX METHOD

Chapter IV has served as an introduction to the S.M. and in the present chapter we derive its fundamental features by algebraic means.

Systems of linear inequalities, without any optimizing requirement, have been studied and the reader will find relevant titles in the bibliography of Ref. 15. We shall not enter here into any detail, because these papers have not led to practical numerical methods.

We consider the problem defined by (1) and (2) of IV.1 and assume that a b.f.s. has been found. Let the b.v.'s be denoted by $x_{u_1}, \ldots, x_{u_m}$ and the n.b.v.'s (which have the value zero) by $x_{u_{m+1}}, \ldots, x_{u_n}$. The m constraints may then be written

$$a_{u_1j}x_{u_1}+\ldots+a_{u_mj}x_{u_m}=b_j-(a_{u_{m+1}j}x_{u_{m+1}}+\ldots+a_{u_nj}x_{u_n}) \quad (1)$$

for $j=1, \ldots, m$. Solving this set for the b.v.'s we obtain

$$x_{u_s}=\sum_{t=1}^{m} d_{ts}[b_t-\sum_{k=1}^{n-m} a_{u_{m+k}t}x_{u_{m+k}}] \quad . \quad (2)$$

for $s=1, \ldots, m$.

The d_{ts} are the elements of the matrix which is inverse to that given by the left-hand sides of the constraints, as written in (1). We call the latter matrix the 'matrix of coefficients' of the b.v.'s. The d_{ts} have the fundamental property that $\sum_s a_{u_sj}d_{ts}=1$ for $j=t$ and $=0$ otherwise, and also that $\sum_t a_{u_st}d_{tr}=1$ for $r=s$ and $=0$ otherwise. In (2) there also appear expressions

$$\sum_{t=1}^{m} a_{u_{m+k}t}d_{ts}$$

which we denote by $z_{u_su_{m+k}}$. For uniformity, we also write $\sum_{t=1}^{m} d_{ts}b_t=z_{u_s0}$ and have then, by (2)

$$x_{u_s}=z_{u_s0}-\sum_{k=1}^{n-m} z_{u_su_{m+k}}x_{u_{m+k}}. \quad . \quad . \quad (2a)$$

It can be verified by using the fundamental properties of d_{ts}, that $z_{u_s u_{m+k}}$ can be thought of as satisfying the equations

$$a_{u_{m+k}j}=a_{u_1j}z_{u_1u_{m+k}}+\ldots+a_{u_mj}z_{u_mu_{m+k}} \quad . \tag{3}$$

for $j=1, \ldots, m$; $k=1, \ldots, n-m$.

We also want to express the L.F. in terms of the n.b.v.'s. Therefore we substitute for the b.v.'s and obtain

$$C=\sum_{s=1}^{m} c_{u_s}z_{u_s0}+\sum_{k=1}^{n-m}\left[c_{u_{m+k}}-\sum_{s=1}^{m} c_{u_s}z_{u_su_{m+k}}\right]x_{u_{m+k}} \quad . \tag{4}$$

We can rewrite this as

$$C=z_{00}-\sum_{k=1}^{n-m} z_{0u_{m+k}}x_{u_{m+k}} \quad . \quad . \quad . \tag{4a}$$

where

$$z_{00}=\sum_{s=1}^{m} c_{u_s}z_{u_s0} \text{ and } z_{0u_{m+k}}=\sum_{s=1}^{m} c_{u_s}z_{u_su_{m+k}}-c_{u_{m+k}}.$$

Since the values of the $x_{u_{m+k}}$ are zero, those of the x_{u_s} are z_{u_s0}, and the value of C is z_{00}, which is of course the sum of these values, each multiplied by its coefficient c_{u_s}.

We assume in this chapter that the z_{u_s0} are positive (not zero).

It can be seen from (4a) that if the $z_{0u_{m+k}}$ are all positive or zero, then C cannot be further decreased by increasing any of the n.b.v.'s. On the other hand, if some $z_{0u_{m+k}}$ is negative, then an increase of the corresponding n.b.v. will decrease C. This situation is sometimes described by saying that we may interpret (3) to mean that an activity, to which $x_{u_{m+k}}$ refers, is equivalent to the combination with weights $z_{u_su_{m+k}}$ of the activities to which the x_{u_s} refer, and that C can be reduced, if the 'cost', i.e. the coefficient, of the single activity $x_{u_{m+k}}$ is smaller than the cost of the equivalent combined activity, made up of the b.v.'s of the stage which we have reached.

If there is more than one variable which could profitably be increased, then we are free to select any of them. No simple rule can be given which would tell us with certainty which choice leads to the final result in the smallest number of steps; but it has been found to be a good plan to take that variable with the smallest (most negative) coefficient.

We repeat that, if all coefficients of the n.b.v.'s in the expression for C are positive or zero, then the final solution has been reached.

2. When we increase one of the n.b.v.'s, say $x_{u_{m+h}}$, from zero to $x'_{u_{m+h}}$, the value of x_{u_s} becomes (see (2))

$$x'_{u_s} = z_{u_s0} - z_{u_su_{m+h}} x'_{u_{m+h}}.$$

Now if $z_{u_su_{m+h}}$ is negative, then an increase of $x_{u_{m+h}}$ increases x_{u_s} as well, and its value certainly remains positive. But if the coefficient is positive, then we must not increase the value of $x_{u_{m+h}}$ beyond $z_{u_s0}/z_{u_su_{m+h}}$. Therefore, since none of the b.v.'s can be allowed to become negative, we make $x'_{u_{m+h}}$ equal to the smallest of the ratios $z_{u_s0}/z_{u_su_{m+h}}$ $(s=1, \ldots, m)$ with a positive denominator. If the smallest of these ratios occurs for $s=r$, then it is the variable x_{u_r} which will be made zero and thereby dropped from the basis, by making $x_{u_{m+h}} = z_{u_r0}/z_{u_ru_{m+h}}$. The other b.v.'s will then have the values

$$z_{u_s0} - z_{u_su_{m+h}} \frac{z_{u_r0}}{z_{u_ru_{m+h}}}$$

and will thus still be positive (or zero in special cases). The new basis* is the same as before, except that x_{u_r} has been replaced by $x_{u_{m+h}}$. We express the new b.v.'s by the new n.b.v.'s, do the same for the L.F., and start the next stage.

3. At every stage we are faced with a set of constraints which, collecting all variables on the left-hand side, can be written as follows:

$$x_{u_s} + \sum_{k=1}^{n-m} z_{u_su_{m+k}} x_{u_{m+k}} = z_{u_s0} \quad (s=1, \ldots, m)$$

$$C + \sum_{k=1}^{n-m} z_{0u_{m+k}} x_{u_{m+k}} = z_{00}.$$

* It follows from the argument in Section 7 that it is, in fact, a basis in the sense of the definition of this term.

These equations can be written in tableau form, thus:

			$\underline{x_{u_{m+1}}}$	$\underline{x_{u_{m+h}}}$	$\underline{x_{u_n}}$
			$c_{u_{m+1}}$	$c_{u_{m+h}}$	c_{u_n}
c_{u_1}	$\underline{x_{u_1}}$	z_{u_10}	$z_{u_1u_{m+1}}$	$z_{u_1u_{m+h}}$	$z_{u_1u_n}$
c_{u_r}	$\underline{x_{u_r}}$	z_{u_r0}	$z_{u_ru_{m+1}}$	$z_{u_ru_{m+h}}$	$z_{u_ru_n}$
c_{u_m}	$\underline{x_{u_m}}$	z_{u_m0}	$z_{u_mu_{m+1}}$	$z_{u_mu_{m+h}}$	$z_{u_mu_n}$
	$\underline{C}$	z_{00}	$z_{0u_{m+1}}$	$z_{0u_{m+h}}$	z_{0u_n}

The underlined symbols are labels of columns and rows, while all other entries are numerical and the z_{u_s0}, in particular, are positive. The $c_{u_1}, \ldots, c_{u_n}$ are the coefficients in the first expression for C. If the b.v.'s do not appear in the original expression of the L.F., then in the first tableau $z_{00}=0$ and $z_{0u_{m+k}}=-c_{u_{m+k}}$. This was so in example 1 of Chapters IV and V.

If we want to minimize (maximize) the L.F., then a positive (negative) $z_{0u_{m+k}}$ indicates a variable $x_{u_{m+k}}$ which could be brought into the basis. (Remember that the entries in the tableau other than z_{00} and z_{u_s0} are the negatives of the coefficients in the equations (2) and (4).) Having decided for $x_{u_{m+h}}$, say, we divide z_{u_s0} by $z_{u_su_{m+h}}$ in those rows in which the latter value is positive and take that x_{u_s} out of the basis which produces the smallest quotient. In this chapter we assume that there is no tie between those quotients which we compare.

As we proceed, changing the set of variables at every stage, we could, in principle, solve the set of the constraints again and again for those variables which become basic, and construct our new tableau. Fortunately, this clumsy procedure is unnecessary, because we can give rules which tell us how to transform one tableau into the next quite simply.

4. We assume now that the variable to be dropped from the basis is denoted by x_{u_r}, and that to be introduced by $x_{u_{m+h}}$. Then $z_{u_r u_{m+h}} > 0$. From the equation

$$x_{u_r} + \sum_{k=1}^{n-m} z_{u_r u_{m+k}} x_{u_{m+k}} = z_{u_r 0}$$

we obtain, first,

$$x_{u_{m+h}} + \sum_{\substack{k=1 \\ \neq h}}^{n-m} \frac{z_{u_r u_{m+k}}}{z_{u_r u_{m+h}}} x_{u_{m+k}} + \frac{1}{z_{u_r u_{m+h}}} x_{u_r} = \frac{z_{u_r 0}}{z_{u_r u_{m+h}}}$$

and then, by substituting for $x_{u_{m+h}}$ into the other equations,

$$x_{u_s} + \sum_{\substack{k=1 \\ \neq h}}^{n-m} \left[z_{u_s u_{m+k}} - \frac{z_{u_r u_{m+k}}}{z_{u_r u_{m+h}}} z_{u_s u_{m+h}} \right] x_{u_{m+k}} - \frac{z_{u_s u_{m+h}}}{z_{u_r u_{m+h}}} x_{u_r}$$

$$= z_{u_s 0} - \frac{z_{u_r 0} z_{u_s u_{m+h}}}{z_{u_r u_{m+h}}}$$

for $s = 1, \ldots, m$, but $\neq r$, while the equation for C can be transformed into the same equation, replacing x_{u_s} by C and all u_s, whenever it appears as a suffix of z, by 0.

We write the new row for $x_{u_{m+h}}$ into the place which has just been vacated by x_{u_r} and the new column for x_{u_r} into the place where $x_{u_{m+h}}$ was.

Let us introduce the following abbreviations:

S_s for $z_{u_s u_{m+h}}/z_{u_r u_{m+h}}$ when $s \neq r$,
S_r for $-1/z_{u_r u_{m+h}}$, and
S_0 for $z_{0 u_{m+h}}/z_{u_r u_{m+h}}$.

Also, write

T_k for $z_{u_r u_{m+k}}/z_{u_r u_{m+h}}$ when $k \neq h$,
T_h for $1/z_{u_r u_{m+h}}$, and
T_0 for $z_{u_r 0}/z_{u_r u_{m+h}}$.

The value $z_{u_r u_{m+h}}$, which appears in all these expressions, is the pivot.

With these symbols, we can write the new tableau as follows:

				$x_{u_{m+k}}$ $(k \neq h)$	x_{u_r}
				$c_{u_{m+k}}$	c_{u_r}
$(s \neq r)$ c_{u_s}	x_{u_s}	$z_{u_s0} - S_s z_{u_r0}$		$z_{u_s u_{m+k}} - S_s z_{u_r u_{m+k}}$	$-S_s$
$c_{u_{m+h}}$	$x_{u_{m+h}}$	$-S_r z_{u_r0}$		$-S_r z_{u_r u_{m+k}}$	$-S_r$
	C	$z_{00} - S_0 z_{u_r0}$		$z_{0u_{m+k}} - S_0 z_{u_r u_{m+k}}$	$-S_0$

or also as follows:

$(s \neq r)$ c_{u_s}	x_{u_s}	$z_{u_s0} - T_0 z_{u_s u_{m+h}}$	$z_{u_s u_{m+k}} - T_k z_{u_s u_{m+h}}$	$-z_{u_s u_{m+h}} T_h$
$c_{u_{m+h}}$	$x_{u_{m+h}}$	T_0	T_k	T_h
	C	$z_{00} - T_0 z_{0u_{m+h}}$	$z_{0u_{m+k}} - T_k z_{0u_{m+h}}$	$-z_{0u_{m+h}} T_h$

Both forms are useful, and the reader is advised to express them in words, and to check his knowledge by the tableaux in Chapter IV, which illustrate the rules of transformation. These rules are greatly simplified if there is a zero either in the column or in the row of the pivot. In the former case the whole row remains unchanged, and in the latter case the whole column.

Having performed the transformation, we can check the new tableau by the formula which defined $z_{0u_{m+k}}$, where the u_s are now the subscripts of the new b.v.'s and the u_{m+k} those of the new n.b.v.'s.

5. The variable which we have introduced into the basis, viz. $x_{u_{m+h}}$, now has the value $z'_{u_{m+h}0} = z_{u_r0}/z_{u_r u_{m+h}}$, and the L.F. has been reduced by $z'_{u_{m+h}0} \cdot z_{0u_{m+h}}$. If the second factor is negative, then this 'reduction' will also be negative, i.e. the L.F. will be increased. This happens when we wish to maximize the L.F.

When the new value is not zero, then the L.F. changes from one tableau to the next, and always in the same sense. But since only a finite number of b.f.s.'s exists, these changes can not go on indefinitely, because we can never come back to an earlier value of the L.F. Hence the process must terminate.* (The case when the value of the new variable in the basis, or of any b.v., becomes zero will be studied in Chapter VII.)

6. We have now completed the description of the S.M., apart from some complications which will be resolved in the next chapter. But the detailed computing routine need not be precisely the one we have used and it may be of value to describe a variant of it. Detailed analysis of the S.M. shows that at every stage it is the set of $z_{0u_{m+k}}$ which decides the selection of the new b.v. and that, once this has been chosen as $x_{u_{m+k}}$, say, only the values z_{us0} and $z_{usu_{m+k}}$ matter. It is therefore reasonable to enquire whether these relevant values could be produced by some other method. We shall now exhibit such an alternative, which we call the Inverse Matrix Method (I.M.M.), for reasons which will presently emerge.

We repeat, for convenience, the following definitions:

$$z_{us0} = \sum_{t=1}^{m} d_{ts} b_t, \quad z_{usu_{m+k}} = \sum_{t=1}^{m} d_{ts} a_{u_{m+k}t}$$

$$z_{0u_{m+k}} = \sum_{s=1}^{m} c_{us} z_{usu_{m+k}} - c_{u_{m+k}}, \quad z_{00} = \sum_{s=1}^{m} c_{us} z_{us0}.$$

Consider the matrix

$$\begin{pmatrix} b_1 & a_{11} & & a_{n1} & 0 \\ & & & & \\ b_m & a_{1m} & & a_{nm} & 0 \\ 0 & -c_1 & & -c_n & 1 \end{pmatrix} = (M), \text{ say.}$$

We shall refer to its columns as the 0-th, 1st, ..., n-th, last respectively.

Let the b.v.'s be $x_{u_1}, \ldots, x_{u_m}$, and extract from (M) the matrix

$$\begin{pmatrix} a_{u_11} & & a_{u_m1} & 0 \\ & & & \\ a_{u_1m} & & a_{u_mm} & 0 \\ -c_{u_1} & & -c_{u_m} & 1 \end{pmatrix} = (A), \text{ say.}$$

* Some b.v.'s removed from a b.f.s. might reappear in a later b.f.s. though.

Also compute the inverse of (A), i.e.

$$\begin{pmatrix} d_{11} & & d_{m1} & 0 \\ & & & \\ d_{1m} & & d_{mm} & 0 \\ \sum_s c_{u_s} d_{1s} & & \sum_s c_{u_s} d_{ms} & 1 \end{pmatrix} = (A)^{-1}.$$

If we omit the last row and the last column from (A), then we have the matrix of the coefficients of the b.v.'s. Its inverse is the matrix obtained from $(A)^{-1}$ by omitting the last row and column.

The inner product* of the α-th row of $(A)^{-1}$ and the β-th column of (A) is, of course, zero when $\alpha \neq \beta$ and unity when $\alpha = \beta$. But if we form the inner product of the last row of $(A)^{-1}$ with the u_{m+k}-th column, which is not in (A), then we obtain

$$\sum_s \sum_t c_{u_s} d_{ts} a_{u_{m+k}t} + 1 . (-c_{u_{m+k}}) = \sum_s c_{u_s} z_{u_s u_{m+k}} - c_{u_{m+k}} = z_{0 u_{m+k}}.$$

If we compute all these inner products, then we can choose the variable $x_{u_{m+k}}$ to be introduced into the basis. We must then find x_{u_r}, the variable which should be dropped from it. To this end we have to find the $z_{u_s 0}$ and $z_{u_s u_{m+k}}$. But these are, respectively, the inner products of the successive rows of $(A)^{-1}$ multiplied by the 0-th and the u_{m+k}-th column of (M). Also note that the inner product of the last row of $(A)^{-1}$ and the 0-th column of (M) is

$$\sum_s \sum_t c_{u_s} d_{ts} b_t = \sum_s c_{u_s} z_{u_s 0} = z_{00}.$$

The matrices (M) and (A) can, to begin with, easily be written down from the constants in the constraints and the L.F. (A) has a last column consisting of 0's and a 1 at the bottom. But we want also to know how the inverse matrix of the next stage can be found by an easy routine. We assert that the transformation is the same as that for the tableaux in the S.M. Using the first form of this transformation, we can write the next inverse as a product of two matrices, viz.

* The inner product of a set $(a_1, .., a_n)$ and a set $(b_1, .., b_n)$ is defined as the sum $a_1 b_1 + \ldots + a_n b_n$.

$$\begin{pmatrix} 1 & 0 & -S_1 & 0 \\ 0 & 1 & -S_2 & 0 \\ & & & \\ 0 & 0 & -S_r & 0 \\ & & & \\ 0 & 0 & -S_m & 0 \\ 0 & 0 & -S_0 & 1 \end{pmatrix} \times (A)^{-1}$$

The proof is quite straightforward and need not be given here. It is carried out by computing the product of this new inverse with the new (A), which latter differs from the previous one merely in that the suffix u_r is everywhere replaced by the suffix u_{m+h}. If we remember the definitions of S_s etc., the result follows at once.

7. We have just seen that the inverse matrix is at every stage pre-multiplied by another matrix and we notice now that the value of the determinant of the latter is $-S_r$, i.e. the reciprocal of the pivot. It follows that at every transformation the determinant of (A) is multiplied by the pivot. The latter is always positive and hence, if we start with a non-singular matrix, none of the later matrices of coefficients will be singular. We cannot assert that the determinant will be positive, because the sign depends on the order in which we put the rows and columns.

8. As an illustration, we again choose example 1 of Chapter IV. We have

$$(M) = \begin{pmatrix} 2 & -2 & 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & -2 & 0 & 1 & 0 & 0 \\ 5 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$(A) = (A)^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} (x_3) \\ (x_4) \\ (x_5) \\ \end{matrix} \quad \vdots \quad \begin{matrix} z_{us0} & z_{us1} & S_s \\ 2 & -2 & -2 \\ 2 & 1 & -1 \\ 5 & 1 & 1 \\ 0 & 1 & 1 \end{matrix}$$

(The meaning of the entries on the right of the dotted line will become clear presently.)

	z_{us0}	z_{us1}	S_s
Inner product of the last row of $(A)^{-1}$ and the columns of (M) not in (A)	0	1	−1
		(x_1)	(x_2) .

Hence the new b.v. is x_1.

To continue, form the inner products of the rows of $(A)^{-1}$ with the 0-th and the 1st column of (M), to obtain z_{us0} and z_{us1}. The results are recorded in the appropriate columns above, to the right of the dotted line. The smallest ratio z_{us0}/z_{us1} with positive denominator is that for $u_s = 4$. We can now find the S_s and S_0, which have also been recorded, in the last column on the right.

The next tableau looks as follows:

					z_{us0}	z_{us2}		S_s
(x_3)	1	2	0	0	6	−3		−1
(x_1)	0	1	0	0	2	−2		−2/3
(x_5)	0	−1	1	0	3	3		−1/3
	0	−1	0	1	−2	1		1/3
					−2	1	−1	
						(x_2)	(x_4)	

Hence x_2 will be made a b.v. We compute z_{us2} (recorded above), find that the new n.b.v. is x_5, and compute S_s (also recorded). Similarly, we obtain

					z_{us0}		
(x_3)	1	1	1	0	9		
(x_1)	0	1/3	2/3	0	4		
(x_2)	0	−1/3	1/3	0	1		
	0	−2/3	−1/3	1	−3		
					−3	−2/3	−1/3
						(x_4)	(x_5)

This is the final tableau, because both z_{04} and z_{05} are negative, viz. −2/3 and −1/3.

9. It is worth noting that the inverse matrices $(A)^{-1}$ are already implicit in the Simplex tableaux if, as is often the case, we start off with variables whose matrix of coefficients is the

unit matrix, and which do not appear in the original L.F. To show this explicitly, remember formula (3) of this Chapter. We can also define values $z_{u_s u_t}$ for $t=1, \ldots, m$ satisfying similar equations. It is evident then that $z_{u_s u_t}=1$ when s equals t and $=0$ otherwise.

We now introduce the concept of an *extended Simplex tableau* (as distinct from the earlier *contracted* tableau), which has a column for each variable and looks as follows:

			x_1	x_n	
			c_1	c_n	
c_{u_1}	x_{u_1}	$z_{u_1 0}$	$z_{u_1 1}$	$z_{u_1 n}$	0
c_{u_m}	x_{u_m}	$z_{u_m 0}$	$z_{u_m 1}$	$z_{u_m n}$	0
	C	z_{00}	z_{01}	z_{0n}	1

Those columns which have the same label as one of the b.v.'s will, of course, be very simple indeed: they contain zeros everywhere except in the row of their own variable, where there is a 1. Note also that we have added a last column which has also zeros everywhere, except at the bottom, where there is again a 1.

If the original b.v.'s were $x_{v_1}, \ldots, x_{v_m}$, say, and their matrix of coefficients was the unit matrix, then $a_{v_i j}=0$ for $i \neq j$, and $=1$ for $i=j$. Consequently (compare (3)),

$$a_{v_j j}=\sum_s a_{u_s j} z_{u_s v_j}=1 \text{ and } a_{v_i j}=\sum_s a_{u_s j} z_{u_s v_i}=0 \text{ if } i \neq j.$$

This means that the matrix which has $a_{u_k t}$ in its t-th row and k-th column is the inverse of that which has $z_{u_t v_k}$ in its t-th row and k-th column. The latter is $(A)^{-1}$, apart from the last row and the last column. The entries z_{0v_k} in the last row equal $\sum_t c_{u_t} z_{u_t v_k} - c_{v_k}$ by construction of the Simplex tableau and hence are equal to the corresponding entries in the last row of $(A)^{-1}$, because $c_{v_k}=0$ by assumption. The last column of $(A)^{-1}$ consists of zeros and a final 1 at all stages.

It is a conspicuous feature of the I.M.M. that at every stage we refer back to the original coefficients in the constraints. It can therefore be assumed that the effect of rounding errors will on the whole be less in the I.M.M. than it might be in the S.M.

10. In our routines we have only considered b.f.s.'s. We justify this by showing that if an optimal f.s. exists which is not basic, then at least one optimal b.f.s. also exists. We show, moreover, that if the sets of variables that form optimal b.f.s.'s are bounded, then any optimal f.s. is a linear combination of optimal b.f.s.'s.

Suppose we have an optimal f.s. We may renumber the variables so that the positive ones are $x_1, \ldots, x_t$. Consider the matrix

$$\begin{pmatrix} a_{11} & a_{t1} \\ a_{1m} & a_{tm} \end{pmatrix}$$

If this matrix has rank r ($\leqslant t, m$), then we can assume, without loss of generality, that

$$\begin{vmatrix} a_{11} & a_{r1} \\ a_{1r} & a_{rr} \end{vmatrix} \neq 0.$$

We keep $x_{r+2}, \ldots, x_n$ constant and solve the first r constraints for $x_1, \ldots, x_r$, obtaining

$$x_s = v_s + w_s x_{r+1} \; (s = 1, \ldots, r)$$

$$C = v_0 + w_0 x_{r+1}.$$

If $x_{r+1} = 0$, then also $x_{r+2} = \ldots = x_n = 0$ and (by simple algebraic considerations) we can find a $m \times m$ matrix of coefficients of which the non-singular $r \times r$ matrix given above forms a part. Then $x_1, \ldots, x_n$ are already a b.f.s.

If x_{r+1} is positive, and also all x_s ($s = 1, \ldots, r$) are positiqe, then we can slightly change x_{r+1}, either by increasing it or by decreasing it, without making any of the positive variables negative. Therefore we must have $w_0 = 0$, otherwise one of these changes would decrease C, and yet C was assumed to be already minimal. All solutions obtained by changing x_{r+1} either way are, therefore, also optimal.

Now decrease x_{r+1} continuously until either itself, or at least one of the x_s becomes zero. We can alternatively increase x_{r+1} continuously and this will eventually either make one at least of the x_s vanish, or the optimal set turns out to be unbounded. We have thus found two optimal f.s.'s with fewer positive variables, of which the given optimal f.s. is a linear combination. Continuing in the same way, as far as possible, our statements are proved.

It is now easy to see that they remain true if everywhere we omit the word optimal. The form of the L.F. is then irrelevant to the proof and thus we may choose C identically equal to 0. Then every f.s. is optimal and the assertion has thus been proved.

DEGENERACY AND OTHER COMPLICATIONS

We have already hinted at possible complications and the time has come to face them. Difficulties can arise in the following ways:

(a) We cannot see how to start off with a b.f.s. In fact, it may be that none exists.

(b) The rules for carrying on fail, although we have not reached the final solution.

(c) The criterion for the choice of the variable to be dropped from the basis does not indicate a unique choice.

(d) Some of the b.v.'s are zero.

2. It is sometimes quite easy to find some b.f.s. to start off with. This is, for instance, the case when the system of equations originated from inequalities and the right-hand side has the same sign as the additional variable. To give a less trivial example, in the transportation problem we can choose any combination (i, j) and make y_{ij} equal to the smaller of a_i and b_j. Having thus disposed of either a_i or b_j completely, we obtain a reduced problem which can be attacked in the same way, and so on, until a first b.f.s. is reached. (For a first b.f.s. in the formulation of a games problem see IX.1.)

3. Now assume that no f.s. can be found by inspection. Let all b_j be non-negative, if necessary after multiplying equations by -1. Let there be $k \leqslant m$ variables such that each appears only in one equation and has a positive coefficient, a_{ij} say. (k may, of course, be zero.) Then put these x_i equal to b_j/a_{ij} and all other variables zero. Thus k equations will be satisfied, but the remaining $m-k$ equations will, in general, not hold. We therefore apply the following device: we add $m-k$ *artificial** non-negative variables $x_{n+k+1}, \ldots, x_{n+m}$, one on each of the left-hand sides of those equations which remain

* We reserve the expression 'additional variables' for those which are introduced to turn an inequality into an equation.

to be satisfied. Let us assume that these are the last $m-k$ of the set. They then become

$$a_{1j}x_1 + \ldots + a_{nj}x_n + x_{n+j} = b_j \qquad (j = k+1, \ldots, m).$$

Making $x_{n+k+1} = b_{k+1}, \ldots, x_{n+m} = b_m$, we satisfy these as well.

We do not want to have the artificial variables in the final solution. Therefore we introduce them into the L.F. as well, with large coefficients M (if the L.F. is to be minimized) or $-M$ (if it is to be maximized). Thus the L.F. becomes

$$\begin{aligned} &c_1x_1 + \ldots + c_nx_n \pm M(x_{n+k+1} + \ldots + x_{n+m}) \\ &= [c_1 \mp M(a_{1k+1} + \ldots + a_{1m})]\, x_1 + \ldots \\ &\quad + [c_n \mp M(a_{nk+1} + \ldots + a_{nm})]\, x_n \pm M(b_{k+1} + \ldots + b_m). \end{aligned}$$

M will be taken to be larger than any number with which it is compared during the process. Hence, if there is a f.s. to the system with the artificial variables equal to zero, then we shall reach it, because a L.F. containing M in the above form will not have its optimal value. If, then, we reach a final solution and it contains any of the artificial variables with non-zero values, we can conclude that no f.s. exists to the original system. On the other hand, once we remove an artificial variable from the basis, we need not consider it any further, because we have then reached a b.f.s. without the aid of that variable, which has become superfluous.

4. As an illustration, we take example 2 of Chapter V. We introduce an artificial variable x_0 into the first equation

$$-2x_1 + x_2 + x_3 - x_0 = -2,$$

and change the L.F. into $x_2 - x_1 + Mx_0$. We start with $x_4 = 2$, $x_5 = 5$, $x_0 = 2$, and express the L.F. in terms of x_1, x_2 and x_3, thus:

$$2M - (1+2M)x_1 + (1+M)x_2 + Mx_3.$$

			x_1	x_2	x_3
			-1	1	
	x_4	2	1	-2	0
	x_5	5	1	1	0
M	x_0	2	2*	-1	-1
		0	1	-1	0
	M	2	2	-1	-1

It is convenient to separate the coefficients of the L.F. into a portion which does not contain M and another which contains it. The last row of the tableau contains, in fact, numbers which we imagine multiplied by M. As long as there are non-zero values in this row, they are therefore overriding. In the present example it happens that we at once remove the artificial variable, and the M row need not be carried along any farther. The following tableaux are

			x_2	x_3
			1	
-1	x_1	1	$-1/2$	$-1/2$
	x_4	1	$-3/2$	$1/2^*$
	x_5	4	$3/2$	$1/2$
		-1	$-1/2$	$1/2$

			x_2	x_4
			1	
1	x_1	2	-2	1
	x_3	2	-3	2
	x_5	3	3^*	1
		-2	1	-1

			x_5	x_4
-1	x_1	4	$2/3$	$1/3$
	x_3	5	1	1
1	x_2	1	$1/3$	$-1/3$
		-3	$-1/3$	$-2/3$

The final solution does not contain the artificial variable. In figure V.2 the progress of the method is represented by the succession of points a, b, c.

5. We have seen that the entries in the M-row are

$$\sum b_t, \sum a_{1t}, \ldots, \sum a_{nt}$$

where the summation extends over those t which denote equations into which artificial variables were introduced. We can therefore interpret the M-method by saying that, to begin with, we try to minimize (or maximize) the artificial L.F.

$$\sum b_t - x_1 \sum a_{1t} - \ldots - x_n \sum a_{nt}.$$

If the optimum is zero, then we have found a f.s. and can proceed to optimize the L.F. which really concerns us. However, if it is positive, then no f.s. exists. The artificial L.F. can never be negative, because it equals the sum of non-negative variables $\sum_t x_{n+t}$.

6. A slightly different version of the M-method can be used when it is easy to solve the constraints for m variables. We may obtain, say

$$x_{u_s}+\sum_{k=1}^{n-m} z_{u_s u_{m+k}} x_{u_{m+k}}=z_{u_s 0} \qquad (s=1, \ldots, m).$$

Now let $z_{u_1 0}, \ldots, z_{u_t 0}$ be negative and the remaining $z_{u_s 0}$ (if there are any) non-negative. Let $z_{u_1 0}$ be the (algebraically) smallest, i.e. the most negative of the $z_{u_s 0}$. Then only one artificial variable, say x_{n+1}, need be introduced, subtracting it from the left-hand side of the 1-st to t-th equation. We also add $+Mx_{n+1}$ or $-Mx_{n+1}$ to the L.F. The following is then a b.f.s.:

$x_{u_1}=0, x_{u_2}=z_{u_2 0}-z_{u_1 0}, \ldots, x_{u_t}=z_{u_t 0}-z_{u_1 0}, x_{n+1}=-z_{u_1 0}$, $x_{u_s}=z_{u_s 0}$ for $s=t+1, \ldots, m$ as before.*

7. Since it is the object of the M-method to find, first, some b.f.s., the precise form of the L.F. is irrelevant at the first stages. It follows that we can also use this method to find the solution to a set of n linear equations in n variables, where intrinsically no question of optimization is involved. In such a case one would take, as the L.F. to be minimized, simply the sum of the artificial variables, multiplied by M. When they have been eliminated, the solution has been found. However, it must be remembered that in this elementary algebraic problem only the artificial variables are sign-restricted. One must therefore either replace each of the others by the difference of two sign-restricted variables, or proceed as in the M-method, but remember that only the artificial variables must at each stage retain their sign (or become zero). The reader will be able to find the correct procedure, if he wishes, but we do not pursue this matter further, because we mentioned this way of solving linear equations as a curiosity rather than as a method to be recommended. We could also invert a non-singular matrix, say (A), by a L.P. process. We make (A) the matrix of coefficients of a set of constraints, equal in number to that of the variables, and with such right-hand sides that

* Details of such a system were given independently by S. I. Gass in Ref. 31.

we know that the system has a solution. For instance, we can make each right-hand side equal to the sum of the coefficients of the same equation, so that the solution will be $x_i = 1$. We add a very simple L.F., for instance 0. No optimization is involved, because the system has only one single solution.

If we now use the M-method with m artificial variables, and use the I.M.M., we finally obtain $(A)^{-1}$, perhaps with some permutation of rows and columns, which can easily be traced.

8. Example 3 of Chapter V requires an artificial variable in its second equation. When the reader carries out the computation, he will find that x_0 will remain in the basis when he has reached the final solution. This need not surprise him, because he knows already that the constraints of this example are contradictory.

9. Our next example deals with a case where there are f.s.'s, but no finite minimum value:

$$-2x_1 + x_2 - x_3 = 2$$
$$x_1 - 2x_2 + x_4 = 2$$
$$x_1 + x_2 - x_5 = 5$$
$$\text{Minimize } x_2 - x_1.$$

This is example 1, except that the sign of x_5 has been changed. In geometrical terms (compare figure V.1) this means that the straight line $x_1 + x_2 = 5$ excludes now that side which contains zero. Thus that part of the plane which contains the feasible points is unbounded. We can reach the same final tableau as in Chapter IV with example 1, except that the signs of the entries in column x_5 are changed. This would suggest that x_5 ought to be brought into the basis; but none of the b.v.'s has a positive entry in that column and thus our rules fail to tell us which to remove from the basis.

We understand what this means when we remember why we considered only ratios with positive denominators at this stage. It was because we wanted to avoid making any of the variables negative through an excessive increase of the n.b.v.'s, while we did not mind increasing the b.v.'s further. In the present case, all of them increase when we increase x_5, and the L.F. would thereby be decreased. Consequently, the latter can be decreased without limit and no finite f.s. produces a minimum value.

We also notice that by giving x_5 larger and larger positive values, we can construct a sequence of f.s.'s with $m+1$ positive variables such that C tends to ∞. It is clear that a similar property holds whenever no finite optimal f.s. exists.

10. There is another way of discovering that the value of the L.F. is unbounded. We can introduce an 'artificial' constraint with one artificial variable, x_{n+1} say, thus:

$$x_1 + \ldots + x_n + x_{n+1} = \Omega$$

where Ω is a large but finite number. If there is a combination of finite $x_1, \ldots, x_n$ which produces the optimum L.F., then we obtain it and have $x_{n+1} = \Omega - \sum_1^n x_i$. On the other hand, if no such finite combination exists for the original problem, then (unless the constraints are contradictory) we obtain now a solution to the new problem, because the artificial equation cannot contradict any of the other constraints and ensures that the x_i are bounded. However, the expression for the final values of the x_i $(i=1, \ldots, n)$ will contain Ω, and this will be an indication that the original problem had no finite solution.

11. Next, we take example 4 of Chapter V. From the first tableau

		x_1	x_2
		-1	1
x_3	4	2	-1
x_4	2	1	-2
x_5	5	1	1
	0	1	-1

it would appear that we should introduce x_1 into the basis.

But which variable should be replaced by it? The ratios 2/1 and 4/2 are equal and we could choose either x_3 or x_4. Let us see what happens in either case. The second tableau will be one of the following two:

			x_2	x_3
			↓	
−1	x_1	2	−1/2	1/2
	x_4	0	−3/2	−1/2
	x_5	3	3/2	−1/2
		−2	−1/2	−1/2

			x_2	x_4
			↓	
	x_3	0	[illegible]	[illegible]
−1	x_1	2	[illegible]	[illegible]
	x_5	3	[illegible]	[illegible]
		[illegible]	[illegible]	[illegible]

In both cases one of the variables remaining in the basis has the value zero. This is a necessary consequence of the equality of the two ratios in the earlier tableau. In the tableau on the left this does not matter, because we have already reached the final stage. However, in the second tableau the pivot is in the row where the variable has zero value and hence the value of the variable now to be introduced into the basis, viz. x_2, will be zero in the next tableau as well. Also, the value of the L.F. will remain unchanged when we replace x_3 by x_2.

This latter fact must cause us some anxiety, since we have argued that the S.M. must terminate because the value of the L.F. changes at every stage, and now this change does not occur. We can therefore not be certain that we do not repeat some sequence of stages over and over again.

In the present case this would not happen. If we carry out the exchange of x_3 and x_2, we reach at once the final solution. The final values will be $x_2 = 0$, $x_1 = 2$ and $x_5 = 3$. We notice that those values which are not zero are the same as in the left-hand tableau above. This is as it must be, because there is only one point in figure V.5 which represents the smallest value of the L.F.

Thus we did not get into an ever-repeated loop, in spite of our misgivings. It is often said that the occurence of a cycle is very unlikely. This is equivalent to saying that it is difficult to construct an example where a repetition of a series of tableaux does in fact occur. One such example is due to

E. M. L. Beale (Ref. 2). An earlier example, privately circulated, is due to A. J. Hoffman.

12. These examples make us cautious and we investigate the position further. It is possible to avoid the recurrence of an earlier tableau by a device due to A. Charnes (Ref. 7, also 9).

In a diagram in two dimensions, equality of the critical ratios means, as we have seen, that there are more than two zero variables and hence more than two lines pass through the same point. An analogous interpretation can be given in more dimensions, but the simplest case already suggests a remedy. The difficulty would disappear if we dislocated the lines slightly and then put them together again when we have found the solution. This idea can be given analytical precision.

Let a tableau be given and adjust it by adding certain polynomials P_{u_s} to the z_{u_s0} as follows:

		$x_{u_{m+k}}$
x_{u_s}	$z_{u_s0}+P_{u_s}$	$z_{u_su_{m+k}}$
L.F.	z_{00}	$z_{0u_{m+k}}$

The polynomials are defined by

$$P_{u_s}=\varepsilon^{u_s}+z_{u_su_{m+1}}\,\varepsilon^{u_{m+1}}+\ldots+z_{u_su_n}\,\varepsilon^{u_n}.$$

In words, the indices of the ε in P_{u_s} are the suffices of the n.b.v.'s and u_s. Their coefficients are the $z_{u_su_{m+k}}$, and $z_{u_su_s}$ (=1). These coefficients are given in the extended tableau (see Chapter VI.7). The symbol ε stands for a small positive number, i.e. one that is smaller than any other with which it is compared in the course of the calculations. We can ensure that $z_{u_s}+P_{u_s}>0$ for all s (even if $z_{u_s}=0$) by renumbering the variables so that the b.v.'s are $x_1, \ldots, x_m$; it is then the term ε^{u} that determines the sign of P_{u_s}.

We now show that this adjustment makes it impossible for any two ratios such as, for instance, $(z_{u_1 0}+P_{u_1})/z_{u_1 v_{m-1}}$ and $(z_{u_2 0}+P_{u_2})/z_{u_2 v_{m-1}}$ to be equal. If the first ratio is larger than the second without P_{u_1} and P_{u_2}, then the addition of the polynomials does not alter the relation, since the ε's are too small to have any effect. If the ratios (without the polynomials) are equal, however, then the relative magnitude of the adjusted ratios depends on the ratio associated with the lowest power of ε: the other powers will again be relatively ineffective. All ratios associated with the various powers of ε cannot be equal, because ε^{u_1} occurs only in the row for x_{u_1} and ε^{u_2} only in that of x_{u_2}. Thus the use of the adjusted tableau makes the possible choice of a b.v. to be dropped unique, and the value of the L.F. changes at every stage.

In our example the first two columns of the first and the second tableau are now, respectively:

1st tableau

x_3	$4+2\varepsilon-\varepsilon^2+\varepsilon^3$
x_4	$2+\varepsilon-2\varepsilon^2+\varepsilon^4$
x_5	$5+\varepsilon+\varepsilon^2+\varepsilon^5$
	0

2nd tableau

	x_3	$3\varepsilon^2+\varepsilon^3-2\varepsilon^4$
-1	x_1	$2+\varepsilon-2\varepsilon^2+\varepsilon^4$
	x_5	$3+3\varepsilon^2-\varepsilon^4+\varepsilon^5$
		$-(2+\varepsilon-2\varepsilon^2+\varepsilon^4)$

In the first tableau the decisive comparison is that of $-\varepsilon^2/2$ and $-2\varepsilon^2/1$. It leads to the replacement of x_4 by x_1. In the second tableau the value of x_3 is not zero anymore, but positive. The other columns are the same as before (for the second tableau take the alternative on the right-hand side). The third tableau will now be

			x_3	x_4
1	x_2	$\varepsilon^2+\varepsilon^3/3-2\varepsilon^4/3$	$1/3$	$-2/3$
-1	x_1	$2+\varepsilon+2\varepsilon^3/3-\varepsilon^4/3$	$2/3$	$-1/3$
	x_5	$3-\varepsilon^3+\varepsilon^4+\varepsilon^5$	-1	1
		$-2-\varepsilon+\varepsilon^2-1\varepsilon^3/3-1\varepsilon^4/3$	$-1/3$	$-1/3$

This is the final tableau and we can now put $\varepsilon=0$, which gives us the final b.f.s. as before.

It is not necessary to write down everything that we have exhibited for the sake of clarity. The coefficients of the powers of ε appear already in the extended tableau and our procedure can be described as follows:

Find first the column of that variable which we want to make basic, say $x_{u_{m+h}}$. Then find all positive $z_{us1}/z_{usu_{m+h}}$ and compare them. If there is a single minimum, then we have found the new n.b.v. But if there is more than one b.v. which produces the minimum, then all these b.v.'s qualify for further consideration. Take $z_{us1}/z_{usu_{m+h}}$ for all these b.v.'s (these ratios are not necessarily positive) and look for the algebraically smallest among them. If there are still ties, take the variables concerned in them and test $z_{us2}/z_{usu_{m+h}}$, and so on, until a single b.v. remains. We know that this must eventually happen.*

13. It may be of interest if we added a few remarks on the algebra of this case. The essential feature is that we obtain a b.f.s. $x_{u_1}, \ldots, x_{u_m}$ with some zero values among them. If for instance $x_{u_1}=0$, then

$$\begin{vmatrix} a_{u_1 1} & & a_{u_{m-1} 1} b_1 \\ & & \\ a_{u_1 m} & & a_{u_{m-1} m} b_m \end{vmatrix} = 0.$$

If a matrix constructed in such a way from the coefficients of $m-1$ variables and the right-hand sides of the constraints is singular, then we call the L.P. problem *degenerate*.

Some writers speak of a degenerate case if some submatrix of order m of

$$\begin{pmatrix} a_{11} & & a_{n1} & b_1 \\ & & & \cdot \\ a_{1m} & & a_{nm} & b_m \end{pmatrix}$$

is singular, even if that submatrix is composed of a_{ij} only.

* G. B. Dantzig has pointed out that the addition of $\varepsilon, \varepsilon^2, \ldots \varepsilon^m$ to the r.h.s. of the constraints, as suggested in his original paper, is computationally equivalent to Charnes' modification, provided the additional variables are placed ahead of the original ones.

Now if a matrix of the latter type, say

$$\begin{pmatrix} a_{u_1 1} & a_{u_m 1} \\ a_{u_1 m} & a_{u_m m} \end{pmatrix} = D_1$$

is singular, then we cannot solve for the x_{u_i} and cannot start the S.M. with them. In such a case we can show that, if there is a solution (not necessarily feasible), with $x_{u_{m+1}} = \ldots = x_{u_n} = 0$, then a matrix of type

$$\begin{pmatrix} a_{v_1 1} & a_{v_{m-1} 1} & b_1 \\ a_{v_1 m} & a_{v_{m-1} m} & b_m \end{pmatrix} = D_2$$

is also singular, so that the two definitions of degeneracy are, for our purpose, equivalent. The proof is as follows:

Let $$\sum_{i=1}^{m} a_{u_i j} x_{u_i} = b_j$$

hold for $j = 1, \ldots, m$. If $D_1 = 0$, then we can solve the system

$$\sum_{i=1}^{m} a_{u_i j} w_i = 0, \qquad j = 1, \ldots, m,$$

so that one of the w_i, say w_t, is not zero. Write $x_{u_t}/w_t = s$, say.

Then $$\sum_{i=1}^{m} a_{u_i j}(x_{u_i} - s w_i) = b_j$$

holds for all j. But $x_{u_t} - s w_t = 0$, so there are not more than $m-1$ non-zero terms on the left-hand side of the equations and hence a matrix of type D_2 is also singular.

14. Various other adjustments have been proposed which may be preferable in special cases. Thus for the transportation problem it has been found possible to introduce an ε of fixed magnitude which allows one to obtain the final values exactly by omitting the decimals which appear, due to the chosen value of ε. This method is very convenient in automatic

computation, where it is inconvenient to keep trace of the original order of the variables. (See Chapter XXIII, footnote on p. 366 of Ref. 15.)

15. We come now to example 5 of Chapter V. The tableaux are as follows:

		x_1	x_2
		−1	1
x_3	2	2	1
x_4	2	1*	−1
x_5	5	1	1
	0	1	−1

			x_1	x_2
				1
	x_3	6	2	−1
−1	x_1	2	1	1
	x_5	3	−1	2
		−2	−1	0

The second tableau exhibits the final solution, but we have here a feature which is new: z_{02} is zero. This does not induce us to carry the problem further, but the L.F. would not change if we introduced x_2 into the basis:

			x_4	x_5
	x_3	$7\frac{1}{2}$	3/2	1/2
−1	x_1	$3\frac{1}{2}$	1/2	1/2
1	x_2	$1\frac{1}{2}$	−1/2	1/2
		−2	−1	0

We could now again exchange x_5 and x_2, and so on. We see that a zero among the $z_{0u_{m+k}}$ indicates that the final b.f.s. is not the unique answer to the problem. By following up all such possibilities, we obtain all b.f.s.'s.

When we have two b.f.s.'s, then any linear combination of them is also a f.s., though not basic. Geometrically, in our example, all points on the line between $x_2=x_4=0$ and $x_4=x_5=0$ give the same minimal value to x_2-x_1. Algebraically, they are given by the values $x_1=2t+3\frac{1}{2}(1-t)$, $x_2=1\frac{1}{2}(1-t)$, $x_3=6t+7\frac{1}{2}(1-t)$, $x_4=0$, $x_5=3t$, where t is any value between 0 and 1 (inclusive).

16. It may be added that, from an algebraic point of view, the appearance of a $z_{0u_{m+k}}=0$ means that there is some u_{m+k} such that

$$c_{u_{m+k}} = \sum_{s=1}^{m} c_{u_s} z_{u_s u_{m+k}}.$$

But by definition we have also

$$a_{u_{m+k}j} = \sum_{s=1}^{m} a_{u_s j} z_{u_s u_{m-k}} \qquad \text{for } j=1, \ldots, m$$

(Chapter VI, formula (3))

and these relations, taken together, mean that the vectors

$$(c_i, a_{i1}, \ldots, a_{im})$$

with $i=u_1, \ldots u_m, u_{m+k}$ are not independent.

This can easily be understood from our graphical representation (Chapter V). It means that the L.F. is constant on one of the straight lines defined by the constraints. We see, also, that the condition is necessary, but not sufficient, for an infinity of solutions to exist.

DUALITY

Consider the following two schemes:

$$\sum_{i=1}^{N} a_{ij}x_i = b_j \qquad\qquad \sum_{j=1}^{m} a_{ij}y_{N+j} + y_i = c_i$$

$$(j=1, \ldots, m) \qquad\qquad (i=1, \ldots, N)$$

all x_i non-negative $\qquad$ $y_1, \ldots, y_N$ non-negative
no sign-restriction on the y_{N+j}

$$\text{Minimize } \sum_{1}^{N} c_i x_i = C \qquad\qquad \text{Maximize } \sum_{1}^{m} b_j y_{N+j} = B$$

On the left, we have the familiar L.P. problem, with m constraints on N variables. On the right, there are N equations in $N+m$ variables, of which N are sign-restricted, while the other m are not. We call such a pair *dual sets of equations.*

Now choose m variables x_{u_s} for which the equations can be solved. Also express C in terms of the remaining variables $x_{u_{m+k}}$ $(k=1, \ldots, N-m=n$, say). This leads to

$$x_{u_s} = z_{u_s 0} - \sum_{k=1}^{n} z_{u_s u_{m+k}} x_{u_{m+k}}$$

(X)

$$C = z_{00} - \sum_{k=1}^{n} z_{0 u_{m+k}} x_{u_{m+k}}.$$

Turning to the system in y, we eliminate the y_{N+j} by first solving the m equations with c_{u_s} on the right-hand side for them (which is possible by the choice of the u_s). We obtain

$$y_{N+j} = \sum_{s=1}^{m} (c_{u_s} - y_{u_s}) d_{js}.$$

Then we substitute into the remaining equations and have

$$y_{u_{m+k}} = c_{u_{m+k}} - \sum_{j=1}^{m} a_{u_{m+k} j} y_{N+j}$$

$$\text{(Y)} \qquad = -z_{0u_{m+k}} + \sum_{s=1}^{m} z_{u_s u_{m+k}} y_{u_s}$$

$$B = \sum_{s=1}^{m} (c_{u_s} - y_{u_s}) z_{u_s 0} = z_{00} - \sum_{s=1}^{m} z_{u_s 0} y_{u_s}.$$

All y_{u_s} ($s=1, \ldots, m$) and $y_{u_{m+k}}$ ($k=1, \ldots, n$) are, of course, sign-restricted.

2. With the following change of notation, viz.

old	new	old	new
k	i	s	j
$x_{u_{m+k}}$	x_i	y_{u_s}	y_{n+j}
$z_{u_s 0}$	$-b_j$	$C - z_{00}$	C
$z_{u_s u_{m+k}}$	$-a_{ij}$	$B - z_{00}$	B
$z_{0u_{m+k}}$	$-c_i$		

and remembering $x_{u_s} \geqslant 0$ and $y_{u_{m+k}} \geqslant 0$ we can write (X) and (Y) as follows:

$$\sum_{i=1}^{n} a_{ij} x_i \geqslant b_j \qquad\qquad \sum_{j=1}^{m} a_{ij} y_{n+j} \leqslant c_i$$

$$\text{Minimize } C = \sum_i c_i x_i \qquad\qquad \text{Maximize } B = \sum_j b_j y_{n+j}$$

Two sets of this form will be called *dual sets of inequalities.* They are equivalent to dual sets of equations, provided there is a b.f.s. to the latter for which the system can be solved; this is always possible by extending the system by the M-method, if necessary.

Comparing the two sets of inequalities we find that the number of variables in one scheme is equal to the number of inequalities in the other. The two matrices are 'transposes' of one another and the constants on the right-hand side of each scheme are those in the L.F. of the other. Also note that we minimize the L.F. of that scheme where the left-hand side is larger than the right-hand side, and maximize that where the opposite is true.*

* In this book we denote a L.F. to be minimized by C (for 'cost') and a L.F. to be maximized by B (for 'benefit').

The pair above is a slightly generalized form of that arrived at in the theory of games (in III.4), because there the constants on the right-hand side as well as the coefficients of the L.F. were 1. In any case, the results which hold for the dual sets will lead to results for the theory of games and we shall use this idea to derive yet another proof of the M.M.T.

3. To compare the results for the two sets we write them down in our familiar tableau form, again collecting all variables on one side of the equations (X) and (Y).

		$x_{u_{m+k}}$			y_{u_s}
x_{u_s}	z_{u_s0}	$z_{u_su_{m+k}}$	$y_{u_{m+k}}$	$-z_{0u_{m+k}}$	$-z_{u_su_{m+k}}$
C	z_{00}	$z_{0u_{m+k}}$	B	z_{00}	z_{u_s0}

The similarity of these two tableaux is conspicuous. The subscripts of the rows of x are those of the columns of y, and vice versa. The entries in the first column of the x-tableau are those in the last row of the y-tableau, while the remaining columns of the former contain the entries of the remaining rows of the latter, but with changed sign. The most important fact, however, is that, with the x- and y-variables chosen correspondingly, the values of B and C are equal.

As the set of the x_{u_s} changes, that of the $y_{u_{m+k}}$ changes simultaneously in such a way that the relation between the two tableaux remains unaltered and, in particular, the new values of B and C are again equal. The rules for transforming the y-tableau can be derived from those for its dual; alternatively, we can argue that the rules express merely the process of eliminating one variable from the basis and introducing another, and that therefore identical rules must apply to both tableaux. Both arguments produce the same result. This is due to the fact that each of them leads to one of the alternative descriptions of the new tableau in VI.4 which describe, in fact, the same relations.

Let us assume that we have found the optimal f.s. to the maximizing problem. The $-z_{0u_{m+k}}$ will then be positive or zero, and so will the z_{u_s0}. Now turn to the minimizing problem. Because the z_{u_s0} are non-negative, the x_{u_s} have the correct sign, and because the $z_{0u_{m+k}}$ are non-positive, we have again reached the final tableau. Consequently, the maximum of B equals the minimum of C.

We have thus proved the Fundamental Duality Theorem (F.D.T.) of L.P., which is as follows:

Given two dual sets, and provided that a finite minimum (maximum) value of C (B) exists, a maximum (minimum) value of B (C) exists as well, and the two optima are equal.

If now we have an optimal b.f.s., we can use the constraints to express the values of the corresponding b.v.'s as linear combinations of the b_j. We can then substitute these expressions for the x_{u_s} in the expression for C and regroup the result as a linear function of the b_j. It follows from the F.D.T. that the coefficients of the b_j in this function are the values of the basic variables in the corresponding optimal b.f.s. of the dual problem. This is called the regrouping principle in Refs. 8 and 10a.

Since the games problem is a special case of the L.P. problem, we have also given a new proof of the M.M.T.

4. Before going farther, we survey a few algebraic facts. From the two dual sets of inequalities we have

$$\sum_j b_j y_{n+j} \leqslant \sum_i \sum_j a_{ij} x_i y_{n+j} \leqslant \sum_i c_i x_i.$$

The F.D.T. tells us even more, viz that

$$\sum_j b_j y_{n+j} = \sum_i c_i x_i.$$

Hence the solutions of the two dual problems satisfy also

$$\sum_j b_j y_{n+j} \geqslant \sum_i c_i x_i,$$

which is a weaker form of the equation above. It follows, that the solution of the following set of inequalities, without optimizing requirement, is equivalent to the optimal f.s.'s of both sets of a dual pair:

$$
\begin{aligned}
&\sum_i a_{ij}x_i \geqslant b_j && (i=1, \ldots, n) \\
&\sum_j a_{ij}y_{n+j} \leqslant c_i && (j=1, \ldots, m) \qquad (I) \\
&\sum_j b_j y_{n+j} \geqslant \sum_i c_i x_i && \\
&x_i \geqslant 0,\ y_{n+j} \geqslant 0. &&
\end{aligned}
$$

The last result throws some light on the question of the conversion of a L.P. problem into a games problem. This is a reasonable question to ask, because we know that the opposite conversion is always possible.

Consider the game defined by the following pay-off matrix:

$$
\begin{array}{ccccc}
0 & 0 & -a_{11} & -a_{n1} & b_1 \\
0 & 0 & -a_{1m} & -a_{nm} & b_m \\
a_{11} & a_{1m} & 0 & 0 & -c_1 \\
a_{n1} & a_{nm} & 0 & 0 & -c_n \\
-b_1 & -b_m & c_1 & c_n & 0
\end{array}
$$

This is a skew-symmetric matrix and the value of the game is therefore zero (see end of Chapter I). Converting it into a L.P. problem, we have:

$$
\begin{aligned}
&a_{1j}x_1' + \ldots + a_{nj}x_n' - b_j z \geqslant v \quad \text{for all } j \\
&-a_{i1}y'_{n+1} - \ldots - a_{im}y'_{n+m} + c_i z \geqslant v \quad \text{for all } i \\
&\sum_j b_j y'_{n+j} - \sum_i c_i x_i' \geqslant v \\
&\sum_j y'_{n+j} + \sum_i x_i' + z = 1 \\
&x_i' \geqslant 0,\ y'_{n+j} \geqslant 0
\end{aligned}
$$

Maximize v.

Now the maximum of v is zero and we can thus replace v by zero on the right-hand side of the inequalities. We also know that a game has always a solution. If this particular game has a solution with $z \neq 0$, then we divide the inequalities by z, and $x_i'/z = x_i$ and $y'_{n+j}/z = y_{n+j}$ is a solution of (I) and hence of the dual sets. (Cf. Ref. 15, Chapter XX, by G. B. Dantzig.) Conversely, if x_i and y_{n+j} are finite optimizing values of the variables of a L.P. problem and of its dual, then

$$\left(\frac{x_i}{\sum_i x_i + \sum_j y_{n+j} + 1}, \frac{y_{n+j}}{\sum_i x_i + \sum_j y_{n+j} + 1}, \frac{1}{\sum_i x_i + \sum_j y_{n+j} + 1}\right)$$

is an optimal strategy of the game defined above. Consequently, if the L.P. problem (and hence its dual) has a finite optimal f.s., then the game has at least one solution such that $z \neq 0$. (As an example, where the game has a solution with $z = 0$ and others with $z \neq 0$, consider the L.P. problem $-x_2 \geqslant 0$ $x_1 + x_2 \geqslant 1$, minimize x_1.)

5. In the statement of the F.D.T. the existence of a finite optimal f.s. of one of the sets was assumed, the existence of a finite optimal f.s. of the other then followed. But we must also investigate the circumstances under which these two finite sets exist.

We would come to the conclusion that B is unbounded from above, if we had a negative z_{u_s0} for some u_s, and simultaneously all $-z_{u_su_{m+k}}$ for that u_s were negative or zero. At the same time, in the x-tableau we would find the value of x_{u_s} to be negative, but all other values in that row positive or zero. Then the information contained in the x_{u_s} row can be written

$$x_{u_s} + \sum_{k=1}^{n} z_{u_su_{m+k}} x_{u_{m+k}} = z_{u_s0}.$$

The left-hand side is positive or zero, and the right-hand side is negative. Hence, if the y-set has an unbounded L.F., then the dual x-set is contradictory; it can be shown in the same way that the roles of x and y in this statement can be reversed. (This does not imply that if one system is contradictory, its dual has an unbounded L.F.; it may also be contradictory.

Take, for instance, $x_1 - x_2 \geqslant 2$, $-x_1 + x_2 \geqslant -1$, $C = x_1 - 2x_2$ and its dual $y_1 - y_2 \leqslant 1$, $-y_1 + y_2 \leqslant -2$, $B = 2y_1 - y_2$.)

6. We imagine now that one out of a dual pair of sets is being tackled by the S.M. and that the other set is dealt with in parallel, as we have described for the proof of the F.D.T. If, then, we note the process for the second set, we obtain a new method for the solution of the L.P. problem, due to C. E. Lemke (Ref. 18). We call the new method the Dual Simplex Method (D.S.M.).

The rules for the D.S.M. may be described as follows (we speak, as heretofore, of an x-tableau when the L.F. is to be minimized, and of a y-tableau, when it is to be maximized):

Start with an $\begin{matrix} x\text{-} \\ y\text{-} \end{matrix}$ tableau where all $\begin{matrix} z_{0u_{m+k}} \\ z_{us0} \end{matrix}$ are $\begin{matrix} \text{non-positive.} \\ \text{non-negative.} \end{matrix}$

If the $\begin{pmatrix} z_{us0} \\ -z_{0u_{m+k}} \end{pmatrix}$ are all positive or zero, then we have finished. If not, we must find a b.v. and a n.b.v. to be exchanged. The choice of these can be derived from that in the dual tableau, which we imagine being treated by the S.M.

D.S.M.

Decide which variable to remove by choosing any of those with negative

$$\begin{pmatrix} z_{us0} \\ -z_{0u_{m+k}} \end{pmatrix}$$

Let the row thus chosen be that of $\begin{pmatrix} x_{ur} \\ y_{u_{m+h}} \end{pmatrix}$. Consider all $\begin{pmatrix} \text{negative } z_{uru_{m+k}} \\ \text{positive } z_{usu_{m+h}} \end{pmatrix}$ and

S.M.

Decide which variable to introduce by choosing any $\begin{pmatrix} y_{us} \\ x_{u_{m+k}} \end{pmatrix}$ with a

$$\begin{pmatrix} \text{negative } z_{us0} \\ \text{positive } z_{0u_{m+k}} \end{pmatrix}.$$

Let the column thus chosen be that of $\begin{pmatrix} y_{ur} \\ x_{u_{m+h}} \end{pmatrix}$. Consider all positive $\begin{pmatrix} -z_{uru_{m+k}} \\ z_{usu_{m+h}} \end{pmatrix}$

D.S.M.—*cont.*

divide them into $\begin{pmatrix} z_{0u(m+k)} \\ z_{us0} \end{pmatrix}$. Take the smallest ratio (which will be positive), thus determining the $\begin{pmatrix} x_{um+k} \\ y_{ur} \end{pmatrix}$ to be introduced.

S.M.—*cont.*

and divide them into $\begin{pmatrix} -z_{0u(m+k)} \\ z_{us0} \end{pmatrix}$. Take the smallest ratio (which will be positive), thus determining the $\begin{pmatrix} y_{um+k} \\ x_{ur} \end{pmatrix}$ to be removed.

In the course of these steps, the value of the L.F. $\begin{pmatrix} C \\ B \end{pmatrix}$ will $\begin{pmatrix} \text{increase} \\ \text{decrease} \end{pmatrix}$, since it remains equal to the corresponding $\begin{pmatrix} B \\ C \end{pmatrix}$ in the dual tableau. This seems to be a move in the wrong direction, but is explained by the fact that a negative sign of a b.v. indicates that we have reached a value of the L.F. which has overshot the mark and is not available for a f.s.

When the variables to be exchanged have been determined, then we construct the next tableau by the familiar rules of transformation.

7. Of course, the D.S.M. has also its complications, dual to those resolved in Chapter VII. We can tackle all of them by taking our clue from the way in which we have overcome the difficulties in the S.M.

(a) How can we start off?

In the S.M. we used a version of the *M*-method with one artificial variable (VII.6) and we here develop the dual of that procedure.

Suppose we solve for *m* variables and express the L.F. also in terms of n.b.v.'s x_{um+k} ($k=1, \ldots, n$). If all coefficients in the L.F. have the correct sign, then we proceed by the D.S.M. without any difficulty. But if there are wrong signs in the L.F., associated with $x_{i_1}, \ldots, x_{i_g}$, say, then we introduce an artificial equation

$$x_{i_1} + \ldots + x_{i_g} + x_0 = M.$$

Of all coefficients with the wrong sign, choose that with the largest absolute value. Let this be d_{i_f}, say, viz. the coefficient of x_{i_f}. Then take x_{i_f} as a further b.v., solve the artificial equation for it and substitute in the constraints, and in the L.F. The variable x_{i_f} will thus disappear from the latter and all coefficients will have the desired sign.

As an illustration we take the case which is dual to example 2 of Chapter V (see also VII.4 and figure V.2).

Maximize $-2y_6 - 2y_7 + 5y_8$

subject to

$$\begin{aligned} -2y_6 + y_7 + y_8 + y_1 &= -1 \\ y_6 - 2y_7 + y_8 + y_2 &= 1 \\ y_6 + y_3 &= 0 \\ y_7 + y_4 &= 0 \\ y_8 + y_5 &= 0 \end{aligned}$$

for non-negative y_1, y_2, y_3, y_4 and y_5. Eliminating y_6, y_7 and y_8 we have

Maximize $2y_3 - 2y_4 - 5y_5$

subject to

$$\begin{aligned} 2y_3 - y_4 - y_5 + y_1 &= -1 \\ -y_3 + 2y_4 - y_5 + y_2 &= 1 \end{aligned}$$

with non-negative y_j.

We add the equation $y_3 + y_0 = M$, eliminate y_3, and obtain

Maximize $2M - 2y_0 - 2y_4 - 5y_5$,

subject to

$$\begin{aligned} -2y_0 - y_4 - y_5 + y_1 &= -1 - 2M \\ y_0 + 2y_4 - y_5 + y_2 &= 1 + M \\ y_0 + y_3 &= M \end{aligned}$$

for non-negative y_j.

This leads to the following succession of tableaux:

			M	y_0	y_4	y_5
				-2	-2	-5
	y_1	-1	-2	-2^*	-1	-1
	y_2	1	1	1	2	-1
	y_3		1	1	0	0
			2	2	2	5

We have here a M-column instead of the M-row. The artificial variable is being exchanged at once and we need not consider it any longer. The sequence of tableaux is then

			y_1	y_4	y_5
				-2	-3
	y_2	$1/2$	$1/2$	$3/2$	$-3/2$
2	y_3	$-1/2$	$1/2$	$-1/2^*$	$-1/2$
		-1	1	1	4

			y_1	y_3	y_5
				2	-5
	y_2	-1	2	3	-3^*
-2	y_4	1	-1	-2	1
		-2	2	2	3

			y_1	y_3	y_2
				2	
-5	y_5	$1/3$	$-2/3$	-1	$-1/3$
-2	y_4	$2/3$	$-1/3$	-1	$1/3$
		-3	4	5	1

Comparison of the two dual sets of tableaux will illustrate all features which we have mentioned and the reader is advised to work out the duals of the other examples of Chapter V as well.

8. In the last section we have solved the dual to example 2 of Chapter V by the D.S.M. We can, of course, also solve the original example 2 by the D.S.M. We add $x_1+x_0=M$. The vertical line marked (0) in figure V.2 is a line of equation $x_1=M$ where M is so large that the line lies outside all the intersections of the other lines. We have then the following tableaux which lead, naturally, to the same final result as the S.M. did in VII.4.

			M	x_2	x_0
				1	1
	x_3	−2	2	1	2
	x_4	2	−1	−2*	−1
	x_5	5	−1	1	−1
−1	x_1		1	0	1
			−1	−1	−1

			M	x_4	x_0
					1
	x_3	−1	3/2	1/2	3/2
1	x_2	−1	1/2	−1/2	1/2
	x_5	6	−3/2	1/2	−3/2*
−1	x_1		1	0	1
		−1	−1/2	−1/2	−1/2

From here on we need not consider x_0 any more.

			x_4	x_5
	x_3	5	1	1
1	x_2	1	−1/3	1/3
−1	x_1	4	1/3	2/3
		−3	−2/3	−1/3

We have started with $x_2=x_0=0$, represented by the point e in figure V.2. It is on the wrong side of $x_4=0$ and of $x_5=0$, hence the negative value of these variables in the first tableau (remember that M is overriding). On the other hand, the coefficients of the L.F. have already the right (negative) sign.

This means that if we increased any of the two n.b.v.'s, whose lines meet at the point e, i.e. if we moved on one of the two intersecting lines in a direction not forbidden by the shading of the other, then we move away from the direction in which the L.F. could be improved. This direction is, in fact, that in which we must move because, as we know, the value of the variable to be introduced into the basis changes from zero to a positive value. It remains to decide on which of the two lines we must move.

We want to move away from a shaded region to its boundary, i.e. in this case we want to eliminate either x_4 or x_5 from the basis. We have decided for x_4, hence we move towards the line $x_4 = 0$. Now if b, on this line, is an extreme point, then f, also on this line, cannot be, because the direction from b to f is the opposite of that from f to b. Thus the line and the direction in which we must move are uniquely defined. We move to f, and then to c. Here we are at last in a feasible point. The problem is solved.

9. (b) Does a bounded solution exist?

There is nothing new here that has not already been said in Chapter VII. We know that the existence of an unbounded solution is the dual to contradictory constraints and since the M-method supplies a criterion for the latter contingency, it is natural that the dual to the M-method, which we have just discussed, is a device for determining whether the solution is unbounded or not. This was discussed in VII.10.

10. (c) Which variable should be chosen for introduction into the basis?

The remedy, if there is a dilemma, is again the dual of that in Chapter VII for the corresponding case. We need not consider polynomials in ε, because we can derive the rule from that given in VII.12. The rule which emerges is, literally, the same as that for the S.M., but the values described by the same symbol are, of course, in different positions in the corresponding dual tableaux.

11. Finally we mention that we can, if we wish, use a combination of the S.M. and the D.S.M. to avoid the *M*-method. We first solve the constraints for *m* variables in terms of the others, and express the L.F. in terms of these others. If the constant terms in the constraints are then all non-negative, we can proceed by the S.M. If the coefficients in the L.F. are all of the right sign, we can proceed by the D.S.M. If neither of these things happens, then we can replace the L.F. by some other L.F. with coefficients of the right sign, and solve the modified problem by the D.S.M. This will lead to a b.f.s. of the original problem which can then be solved by the S.M.

CHAPTER IX

THE SOLUTION OF GAMES

In this chapter we apply our knowledge of **L.P.** to the solution of games. We know that any game can be written as a L.P. problem (see III.4), and the solution of the latter gives also that of the former. But the special features of the games problem allow some special treatment, as we shall see.

We have seen in Chapter VIII that each Simplex tableau refers simultaneously to two dual problems and since the problems of the two players in a zero-sum two-person game are dual in this sense, we can solve both problems simultaneously, and find the value of the game.

B's problem can be given the following L.P. form:

$$y_1 + \ldots + y_m = 1$$

$$a_{i1}y_1 + \ldots + a_{im}y_m + y_{m+i} = v \qquad (i=1, \ldots, n).$$

Minimize v for non-negative variables.

The quantity v is, of course, not a constant, but we can reduce the constraints to a more familiar form, which differs however from that in III.4.

We may assume without loss of generality that the constraints are ordered so that $a_{n1} \geqslant a_{i1}$ for all i. Subtract the second, ... etc. equation from the last and express the L.F. v by the last equation. This gives

$$y_1 + \ldots + y_m = 1$$

$$(a_{n1} - a_{i1})y_1 + \ldots + (a_{nm} - a_{im})y_m + y_{m+n} - y_{m+i} = 0$$

$$\text{for } i=1, \ldots, n-1.$$

Minimize $a_{n1}y_1 + \ldots + a_{nm}y_m + y_{m+n}$.

We have here n constraints on $n+m$ variables. No difficulty arises in selecting the first b.f.s. We choose $y_1 = 1$, $y_{m+n} = a_{n1} - a_{11}, \ldots, y_{m+n-1} = a_{n1} - a_{n-11}$. The value of the L.F. is then a_{n1}.

Using the I.M.M. we have

$$(A)=\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ a_{n1}-a_{11} & -1 & \cdots & 0 & 0 \\ \vdots & & & & \vdots \\ a_{n1}-a_{n-11} & 0 & \cdots & -1 & 0 \\ -a_{n1} & 0 & \cdots & 0 & 1 \end{pmatrix}$$

It is easily verified that the inverse matrix $(A)^{-1}$ is identical with (A), except that in the left-hand bottom corner there is a_{n1} instead of $-a_{n1}$. We notice that the entries in the first column of $(A)^{-1}$ are precisely the values of the basic variables, and that of the L.F. Since the 0-th column of (M) contains a 1 in the first row and only zeros in the other places, the first column of $(A)^{-1}$ will retain this meaning throughout.

We illustrate all this by the following game, which has been mentioned in I.3.

$$\begin{pmatrix} -1 & 1 & 1 \\ 2 & -2 & 2 \\ 3 & 3 & -3 \end{pmatrix}$$

The matrix (M) is as follows:

	y_1	y_2	y_3	y_4	y_5	y_6	
1	1	1	1	0	0	0	0
0	4	2	−4	−1	0	1	0
0	1	5	−5	0	−1	1	0
0	−3	−3	3	0	0	−1	1

We start with the basis y_1, y_4 and y_5 and obtain the following succession of tableaux (the column of z_{us0} is identical with the first column and need not be repeated):

						z_{us3}		S_3
(y_1)	1	0	0	0		1		1/6
(y_4)	4	−1	0	0		8		4/3
(y_5)	1	0	−1	0		6		−1/6
	3	0	0	1		6		1
					0	6	−1	
					(x_2)	(x_3)	(x_6)	

(y_1)	5/6	0	1/6	0	5/3			5/22
(y_4)	8/3	−1	4/3	0	22/3			−3/22
(y_3)	1/6	0	−1/6	0	−2/3			−1/11
	2	0	1	1	4			6/11
					4	−1	0	
					(x_2)	(x_5)	(x_6)	

(y_1)	5/22	5/22	−3/22	0			
(y_2)	4/11	−3/22	2/11	0			
(y_3)	9/22	−1/11	−1/22	0			
	6/11	6/11	3/11	1			
					−6/11	−3/11	−2/11
					(x_4)	(x_5)	(x_6)

The z_{0k} are all negative, hence we have reached the final stage.

We know from VIII.1 (equations (X) and (Y)) that $-z_{0u_{m+k}}$ is the value of y_{m+k} in the dual set and thus (6/11, 3/11, 2/11) is A's optimal strategy, while the value of the game is 6/11. It will be appreciated that, in the notation here used, only the final values of x_4, x_5 and x_6 and of y_1, y_2 and y_3 refer to strategies, while any of the additional variables, had they remained in the final solution, would have to be ignored from this point of view.

2. We now introduce a method of solving a game which is quite convenient when the pay-off matrix is small. Consider the dual sets of inequalities as they appear in a games problem (see III.4 and VIII.2) and choose the first bases as consisting of those additional variables which do not appear in the L.F. Whenever at a later stage a non-additional variable is introduced into the basis, this happens in both sets at the corresponding stages. Consequently, in the final solution the same number of non-additional variables will appear in both final tableaux.

This means, in other words, that the same number of additional variables will have eventually the value zero for both players. Let this number be r. Then there are in both sets r equations in r non-additional variables, which define the optimal strategies of both players. We may assume, without loss of generality, that the equations are the first r, and that the variables are, in the notation of III.4, $x_1', \ldots, x_r'$ and $y_1', \ldots, y_r'$. We have

$$a_{11}x_1' + \ldots + a_{r1}x_r' - v = 0$$

$$a_{1r}x_1' + \ldots + a_{rr}x_r' - v = 0$$

$$x_1' + \ldots + x_r' \qquad = 1$$

and similarly, with transposed a_{ij}, for $y_1, \ldots, y_r$.

The solution is

$$x_i' = \begin{pmatrix} a_{11} & a_{i-11} & 0 & a_{i+11} & a_{r1} & -1 \\ \vdots & & & & & \\ a_{1r} & a_{i-1r} & 0 & a_{i+1r} & a_{rr} & -1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \div \begin{pmatrix} a_{11} & a_{r1} & -1 \\ \vdots & & \\ a_{1r} & a_{rr} & -1 \\ 1 & 1 & 0 \end{pmatrix}$$

and similarly for y_j'.

By exchanging the last and the i-th column in the numerator we find that it equals $\sum_j A_{ji}$, where A_{ji} is the co-factor of a_{ij} in the determinant of a_{ij} $(i, j=1, \ldots, r)$. In view of the last constraint the denominator must be $\sum_i \sum_j A_{ji}$. We have also

$$v = \sum_i a_{ij}x_i' \text{ for any } j, \text{ and } v = \sum_j a_{ij}y_j' \text{ for any } i.$$

Now we do not know which submatrix of order r of the pay-off matrix leads to the final solution. We must therefore consider all of them, of orders 1, 2, $\ldots$, min (m, n) and solve the corresponding equations, as above. If some of the resulting x_i' or y_j' are negative, then the submatrix which led to this set is irrelevant to the solution of the game. But even when they are all non-negative, this is not yet sufficient. The other variables must also be non-negative, which means that inequalities of the types

$$a_{1r+t}x_1' + \ldots + a_{rr+t}x_r' \geqslant v = a_{11}x_1' + \ldots + a_{r1}x_r'$$

and

$$a_{r+s1}y_1' + \ldots + a_{r+sr}y_r' \leqslant v = a_{11}y_1' + \ldots + a_{1r}y_r'$$

must hold for all $t=1, \ldots, n-r$ and $s=1, \ldots, m-r$. If they hold, then we have reached a solution. (The notation is here again that which is appropriate when the x_i' and y_j' are the first r of their set.)

A games problem cannot have an unbounded solution, because the dual problem has always a solution. Hence, if we find all optimal b.f.s.'s of the corresponding L.P. problem, then it follows from the argument in VI.8 that we know all solutions of the games problem, since they are all linear combinations of those solutions, which correspond to optimal b.f.s.'s of the games problem.

3. We apply this idea to the game

$$\begin{pmatrix} 4 & 1 & 3 \\ 2 & 3 & 4 \end{pmatrix}.$$

For $r=1$ the search for a solution simply means looking for a saddle point. In this game there is none. Now take all sub-matrices of order 2, viz.

$$\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \quad \begin{pmatrix} 4 & 3 \\ 2 & 4 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 \\ 3 & 4 \end{pmatrix}$$

The third is irrelevant, because the second row dominates the first, and also the first column dominates the second; therefore the set of equations $y_1'+3y_2'=3y_1'+4y_2'$, or also the set $x_1'+3x_2'=3x_1'+4x_2'$ cannot possibly have non-negative values of all variables in their solutions. The first matrix gives $4x_1'+2x_2'=x_1'+3x_2'$, $x_1'+x_2'=1$, i.e. $x_1'=1/4$ and $x_2'=3/4$. Similarly we obtain $y_1'=y_2'=1/2$. The inequality $3\times\frac{1}{4}+4\times\frac{3}{4}\geqslant v=5/2$ holds and we have thus found a solution. If we treat the second matrix in the same way, we obtain positive values, but the necessary inequality does not hold. If A used nevertheless the values obtained from the second matrix, then B could spoil his game by choosing his own second strategy as a reply.

There is thus only one solution to this game.

CHAPTER X

GRAPHICAL REPRESENTATION OF L.P. (2)

We now introduce a geometrical model which generalizes the representation exhibited in II.2 and the following sections. We consider the set

$$a_{11}y_1+\ldots+a_{1m}y_m+y_{m+1}=c_1$$
$$a_{21}y_1+\ldots+a_{2m}y_m+y_{m+2}=c_2$$

Maximize $b_1y_1\ldots+b_my_m=B$, say, for non-negative variables.

We assume that none of the constants c_i or b_j is zero. Otherwise obvious modifications apply.

Dividing each equation by its right-hand side, introducing b_jy_j as a new variable and re-defining the coefficients suitably, the scheme can be brought into the following form: (the variables are, of course, different from the earlier ones)

$$a_{11}y_1+\ldots+a_{1m}y_m+(1/c_1)y_{m+1}=1$$
$$a_{21}y_1+\ldots+a_{2m}y_m+(1/c_2)y_{m+2}=1$$
$$\text{Maximize } y_1+\ldots+y_m=B, \text{ say.}$$

The new variables are not necessarily restricted to non-negative values: those y_j which stand for a variable whose coefficient b_j in the L.F. was negative, are now restricted to *non-positive* values. We assume without loss of generality that $y_1\ldots,y_f$ are non-negative and (if there are any others) $y_{f+1},\ldots,y_m$ are non-positive. The variables y_{m+1} and y_{m+2} are again non-negative, because they are the same as before.

To begin with, we assume that $B=1$. As in Chapter II, we consider in a Cartesian plane the points $P_j=(a_{1j}, a_{2j})$ for $j=1,\ldots,m+2$.

We must now show how to determine the region of all points whose abscissae and ordinates are found from those of the P_j by applying to them non-negative or non-positive weights y_j, as the case may be, while $B=1$.

If all the original b_j were positive (not zero) and $B=1$, then all points given by the left-hand sides of the constraints are those in the convex hull of the P_j. If only some b_j are positive, viz. $b_1, \ldots, b_f$, then we can certainly reach all points within the convex hull of $P_1, \ldots, P_f$. This area will be extended by the effect of $y_{f+1}, \ldots, y_{m+2}$ which we shall now analyse. Consider the variables y_{m+1} and y_{m+2} which have zero coefficients in the L.F. Whatever their value, the value of the L.F. will not be affected, but they will make an addition to the left-hand sides of the constraints, each of them affecting just one of the two left-hand sides. The two weights y_{m+1} and y_{m+2} will therefore extend the area which we have already obtained by moving it without limit in the directions parallel to the two directions from 0 to P_{m+1} and from 0 to P_{m+2}, i.e. in the direction of increasing ordinates, and of increasing abscissae.

Turning now to a P_s with negative weight attached to it, and choosing an arbitrary P_t in the area so far obtained, we may extend it in the direction parallel to that from P_s to P_t.

These constructions break down when all b_j are negative (which does not happen in a games problem). But then we change the sign of all b_j and require B to be minimized. We must remember, though, that the minimum obtained is then the negative of the maximum originally asked for.

When we have found the region of all points whose co-ordinates are defined by the two left-hand sides, while $B=1$, any area corresponding to other values of B can be found by similar contraction or dilatation from the origin. We want B to be as large as possible, and to have the point c=(1, 1) within it. We shall therefore choose a region as dilated as possible consistent with having c not outside it (which may of course involve some contraction from the case $B=1$). Generally we shall then have c placed on an edge of the region between two vertices, and then only two y_j will have non-zero values in the solution. In special cases c could coincide with one of the vertices of the area, or could lie on two lines, each connecting two points which originated from the P_j.

We take again the game

$$\begin{pmatrix} 4 & 1 & 3 \\ 2 & 3 & 4 \end{pmatrix}.$$

We have then the system

$$4y_1 + y_2 + 3y_3 + y_4 = 1$$
$$2y_1 + 3y_2 + 4y_3 + y_5 = 1.$$

In figure X.1 we have drawn the points P_j and the point c.

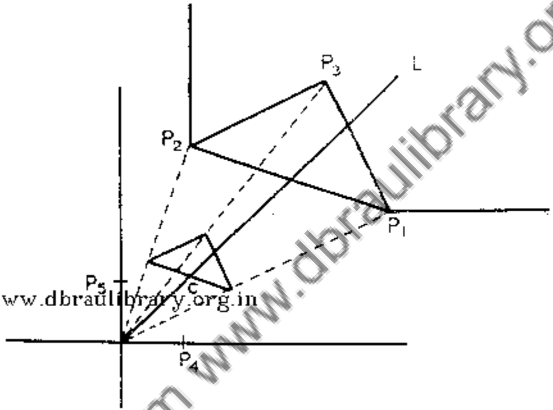

Fig. X.1

In games problems the points (1, 0) and (0, 1) will always be amongst the P_j.

The area which can be obtained by non-negative y_j, while $B=1$, contains certainly $P_1 P_2 P_3$. Because of P_4 and P_5 this area can be extended by drawing a vertical through P_2 upwards and a horizontal through P_1 to the right. The area contains all points to the right and above the broken line consisting of the vertical, of $P_2 P_1$, and of the horizontal. The point c can be reached by contracting the area towards the origin in the proportion $1 : 5/2$, whence $y_1 = y_2 = 1/5$ and $B = 2/5$. The equivalent L.P. problem has thus been solved. To go back to the game, we divide by 2/5 and obtain the mixed strategy (1/2, 1/2), while the value of the game is the reciprocal of 2/5, i.e. 5/2.

It is worth remarking that because of the presence of the points (1, 0) and (0, 1) and their effect on the area, the games problem always has a solution. However, in a L.P. problem it is possible that the line 0c does not intersect the area and that therefore no contraction or dilatation can bring c into it. This will happen when the constraints are contradictory. It can also happen that the area is unbounded in such a manner that no finite optimum of the L.F. exists.

We continue to be guided by the demonstration in Chapter II and turn now to A's problem. Its L.P. form can be written:

$$4x_1+2x_2 \geqslant 1$$
$$x_1+3x_2 \geqslant 1$$
$$3x_1+4x_2 \geqslant 1.$$

Minimize x_1+x_2 for non-negative variables.

This is a special case of inequalities in two variables, as follows:

$$a_{11}x_1+a_{21}x_2 \geqslant b_1$$
$$a_{1m}x_1+a_{2m}x_2 \geqslant b_2$$

Minimize $c_1x_1+c_2x_2=C$, say, for non-negative variables.

We assume again for the sake of simplicity that none of the c_i or b_j is zero. This system can be brought into the form

$$a_{11}x_1+a_{21}x_2 \geqslant 1 \qquad a_{1f+1}x_1+a_{2f+1}x_2 \leqslant 1$$
$$a_{1f}x_1+a_{2f}x_2 \geqslant 1 \qquad a_{1m}x_1+a_{2m}x_2 \leqslant 1$$

Minimize $x_1+x_2=C$, say.

This time either x_1 or x_2 (or both) will be non-positive, if either c_1 or c_2 (or both) were negative.

In the case where we dealt with two equations in $m+2$ variables the latter were interpreted as weights and the a_{1j} and a_{2j} as coordinates of points. Now, when we deal with two variables in m inequalities, we consider the variables as coefficients in equations of straight lines in a plane with coordinates a_1 and a_2.

Given any straight line S through the origin with equation $a_1x_1+a_2x_2=0$, draw a line parallel to S through P_j, to intersect the line L with equation $a_1=a_2$. (Remember that a_1 and a_2 are coordinates!) Denote the intersection by Q. Then Q has the abscissa $(x_1a_{1j}+x_2a_{2j})/(x_1+x_2)$ (compare II.3). In other words, $x_1a_{1j}+x_2a_{2j}$ is the abscissa of a point $\bar{Q}$ on L such that $0\bar{Q}/0Q=x_1+x_2$, where 0 is the origin. We shall refer to this statement as the Lemma.

If the value of the L.F. is 1, then $\bar{Q}=Q$. What restrictions does then an inequality $x_1a_{1j}+x_2a_{2j}\geq 1$ or ≤ 1 impose on x_1 and x_2, or rather on the line $x_1a_1+x_2a_2=0$ through the origin? By the Lemma any such line must satisfy the condition that a parallel to it through P_j intersects L in a point whose abscissa is at least (or at most) equal to 1. Moreover, if both x_1 and x_2 must be non-negative, or both non-positive, then only lines with negative slope are admitted; if they are of different type, then the lines must have positive slope.

As an illustration, consider the same game as before (see figure X.2). We have drawn the point b=(1, 1) and, through the origin, parallels to the lines from b to P_j ($j=1, 2, 3$). We

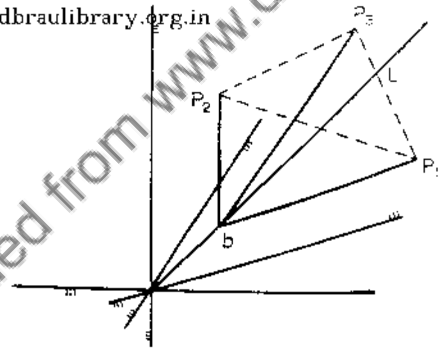

Fig. X.2

have also marked by shading those sectors into which the lines must not turn. It appears that in this case all lines with negative slope are admissible, as long as $C=1$. But we want to have C as small as possible. Therefore, taking any other value for C, the part played heretofore by b will be taken over by a point b(C) on L, such that the coordinates of b(C) equal $1/C$. As C increases, or decreases, b (C) moves nearer to (or away from) 0, and the lines through P_j and b (C) turn accordingly, while the admissible sectors change with them. Take now the situation when b (C) has moved into (5/2, 5/2) (see figure X.3). This point is the intersection of L and the

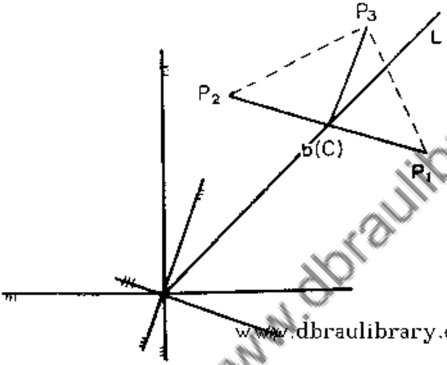

Fig. X.3

line P_1P_2 and it is clear that now the parallel to P_1b (C), which is also the parallel to P_2b (C), must not turn any farther on either side. It does not lie in any sector forbidden by the third parallel, and hence it is one, and the only one, line still left. Its equation is $a_1+3a_2=0$. C equals 2/5 and we have $x_1=1/10$ and $x_2=3/10$. It follows that the value of the game is 5/2 and the optimal mixture of strategies for A is (1/4, 3/4), as we already know.

CHAPTER XI

THE METHOD OF LEADING VARIABLES

A method which differs from those explained in the earlier chapters has been developed by E. M. L. Beale (Ref. 1) and we give here the gist of it, but must refer the reader for some more subtle detail to the original publication.

Imagine that a system of m equations in n variables is given and that it is required to minimize a L.F. for non-negative values of the variables. We solve the system for m of the variables, say $x_1, \ldots, x_m$ and obtain, say

$$x_j = z_{j0} + z_{jm+1}x_{m+1} + \ldots + z_{jn}x_n \qquad (j=1, \ldots, m)$$
$$C = z_{00} + z_{0m+1}x_{m+1} + \ldots + z_{0n}x_n$$

This transformation is possible, unless the system is contradictory and has no solution at all.

Of course, we can still not be certain that the system has a solution with non-negative x_j, and the values of C could be unbounded from below. But the latter possibility can be excluded by adding a further constraint with a further variable, viz.

$$1 = x_0 + \omega(x_{m+1} + \ldots + x_n),$$

where ω is finite and positive, and x_0 is also restricted to non-negative values.

If the original system has a finite solution then, provided ω is sufficiently small, the introduction of this new equation does not affect the final answer, because it does not add any further restrictions on the values $x_{m+1}, \ldots, x_n$. On the other hand, if the original set has only an infinite solution, then this will be shown by some of the variables being, in the final solution, of the order of magnitude $1/\omega$, and x_0 will then be zero. If the original system has no solution, the new one will not have any either.

We have now m equations with a variable on their left-hand side, and one equation with 1 on the left-hand side. This feature will be retained throughout the computations and we shall call the latter equation the *leading* equation and the variables appearing in it the *leading* variables. The method is accordingly called Method of Leading Variables (M.L.V.).

Our process can now be described as follows: we imagine, to begin with, that only the leading variables need be non-negative. Then it is easy to find the solution that minimizes C. For let the leading equation be

$$1 = \sum_{s} t_s x_s$$

and let

$$C = \sum_{s} T_s x_s$$

where s is summed over the $n-m+1$ leading variables, and where t_s may or may not contain ω.

If no t_s is positive, the leading equation cannot be satisfied with non-negative x_s and the whole system has no solution. Otherwise we can take some x_s, say x_p, such that t_p is positive, and we can use the leading equation to eliminate this variable from the expression for C. It becomes then

$$C = T_p/t_p + \sum_{s=0}^{n-m} (T_s - T_p t_s/t_p) x_s.$$

If all expressions in the brackets on the right-hand side are non-negative, then the set $x_p = 1/t_p$ and all other $x_s = 0$ produces the minimum of C for non-negative leading variables. Some of the t_s may be negative or zero, of course, but first we try to determine that x_p which makes all those brackets non-negative in which the t_s are positive. For this purpose we must choose p such that the ratio T_p/t_p is the smallest of all T_s/t_s for positive t_s. It can be shown that, at the very first stage, this choice of p ensures that all other brackets are also non-negative. This is so, because (i) at this stage C cannot be negatively infinite, by virtue of the introduction of the ω, and (ii) if one of the expressions in brackets were negative, then we could make C negatively infinite, which would thus lead to a contradiction. (For details see Ref. 1.) We call x_p the *principal* variable.

Until now we have ignored the existence of the non-leading equations. They contain constants, but we now multiply these by x_0, thereby changing our original problem. However, this need not disturb us, because it can be shown that the solution of the new problem can be translated into that of the original one without undue difficulty. We shall therefore show how to treat the new problem.

Consider the leading equation and the expression for C. If the principal variable is x_p, we make it $1/t_p$ and all the other leading variables zero. This minimizes C as long as only the leading variables are sign-restricted.

Next, we consider whether this choice of the principal variable leads to non-negative values for the non-leading variables. This depends on whether the coefficients z_{jp} $(j=1, \ldots, m)$ are non-negative. If they are, then our goal is reached, with leading variables as above and non-leading variables $x_j = z_{jp}/t_p$. But if some are negative, then a new set of leading variables must be chosen, by dropping x_p and choosing instead one of those x_j, say x_r, which spoilt the earlier result by being negative. We proceed by solving the equation with x_r for x_p and substituting into the other equations. In the leading equation, x_0 will have the coefficient 1, if it appears at all, because multiples of ω can be neglected, by comparison.

If we have not yet reached our goal when we determine the new principal variable, then the process of changing the leading equation must be repeated. It is shown in Ref. 1 that during the procedure the value of C will not decrease and that the process can be made to terminate. This depends on a criterion for the choice of the principal variable similar to the ϵ-method for breaking ties when exchanging variables in the S.M. or in the D.S.M.

Incidentally, since C is not unbounded from below at the first stage, and afterwards does not decrease, we can be certain, at every stage, that if we eliminated the principal variable from the leading equation and the expression for C, the latter would always appear with non-negative coefficients, otherwise we could produce an infinitely small value of C with non-negative leading variables, and this would contradict the non-decreasing tendency of C.

We must now refer to the question of how to translate the solution to our changed problem into that of the original one. Apart from multiples of ω the value of x_0 can, at the final stage, only be 0 or 1. If it is 1, then the two problems are identical. But if it is 0, then a further analysis is required, which cannot be given here. It may be said, however, that this difficulty does not arise if the solution to the original problem exists, and is finite.

It hardly needs mentioning that the M.L.V. can be adapted, by trivial changes, to the case of a L.F. which should be maximized.

2. The M.L.V. is independent of the D.S.M., but the two methods have certain similarities. We illustrate this by solving an example by the M.L.V. The geometrical interpretation, which will then follow, suggests that it is always possible to carry out the two processes in such a way that the set of basic variables in the D.S.M. is, at every stage, the same as the set of non-leading variables plus the principal variable in the M.L.V., and this can be proved formally. We take example 2 of Chapter V, which was again considered in VII.4 and in VIII.7.

The natural basic variables to begin with are x_3, x_4 and x_5. In the D.S.M. we also introduced x_1. In the present case, we add the leading equation containing x_0 and the variables x_1 and x_2, so that we start with

$$\begin{aligned} x_3 &= -2x_0+2x_1-x_2 \\ x_4 &= 2x_0-x_1+2x_2 \\ x_5 &= 5x_0-x_1-x_2 \\ 1 &= x_0+\omega(x_1+x_2) \\ C &= -x_1+x_2. \end{aligned}$$

The principal variable is x_1, and must be exchanged with x_4.

$$\begin{aligned} x_1 &= 2x_0-x_4+2x_2 \\ x_3 &= 2x_0-2x_4+3x_2 \\ x_5 &= 3x_0+x_4-3x_2 \\ 1 &= x_0-\omega(x_4+3x_2) \\ C &= -2x_0+x_4-x_2. \end{aligned}$$

The principal variable is x_2, and must be exchanged with x_5.

$$x_2 = x_0 + x_4/3 - x_5/3$$
$$x_1 = 4x_0 - x_4/3 - 2x_5/3$$
$$x_3 = 5x_0 - x_4 - x_5$$
$$1 = x_0 - \omega x_5$$
$$C = -3x_0 + 2x_4/3 + x_5/3$$

The principal variable is x_0, and we have reached the final set. It contains the solution, viz. $x_2=1$, $x_1=4$, $x_3=5$, $C=3$.

It is interesting to look again at figure V.2 and to inspect it to see what we have done. The line (0) which was drawn in connection with the D.S.M. is now irrelevant, and the line corresponding to the leading equation is marked (0′). We have here assumed that ω is small enough for the new line to be outside the intersections of the others. (It is easily seen how we could modify either the M.L.V. or the D.S.M. so as to make (0) and (0′) coincide.)

Our first leading variables were x_0, x_1 and x_2, and we minimized C subject to these being non-negative. This brought us to point e′ which, of all points in the triangle formed by the lines (0′), (1) and (2), produces the smallest value of C. We are here at a distance from (1), and hence x_1 is the principal variable; we are on the wrong side of (4), and hence x_4 must be made leading.

Our next position is f′ which, of all points in the 'triangle' (0′), (2), (4), makes C smallest. This triangle is bordered by the three lines mentioned, but lies on the unshaded sides of them, and is therefore unbounded. In f′ we are at a distance from (2), hence x_2 becomes principal variable; we must move towards the right side of (5), so that x_5 becomes leading. This brings us finally to c. This point lies on the triangle (0′), (4), (5) on the right side of all lines, and indicates x_0 as the principal variable.

REFERENCES

(1) E. M. L. Beale
An alternative method for linear programming.
Proc. Camb. Phil. Soc. **50**, 513–523; 1954.

(2) E. M. L. Beale
Cycling in the dual simplex algorithm.
Naval Research Logistics Quarterly, **2**, 269–275; 1955.

(3) D. W. Blackett
Some Blotto games.
Naval Research Logistics Quarterly, **1**, 55–60; 1954.

(4) David Blackwell and M. A. Girshik
Theory of games and statistical decisions.
New York and London, 1954.

(5) E. Borel
Sur les systèmes de formes linéaires à determinant symétrique gauche et la théorie générale du jeu.
C. R. de l'Acad. des Sciences, **184**, 52–54; 1927.
In English: Three short papers on games theory that involve chance and the skill of the players. Translated by L. J. Savage.
Econometrica, **20**, 97–117; 1953.

(6) A. S. Cahn
The warehouse problem. (Abstract.)
Bull. Am. Math. Soc., **54**, 1073; 1948.

(7) A. Charnes
Optimality and degeneracy in linear programming.
Econometrica, **20**, 160–170; 1952.

(7a) A. Charnes and W. W. Cooper
The stepping stone method of explaining linear programming calculations in transportation problems.
Management Science, **1**, 49–89; 1954.

(8) A. Charnes and W. W. Cooper
Scheduling of manufactured product. (Abstract.)
Econometrica, **21**, 480–481; 1953.

(9) A. Charnes, W. W. Cooper and A. Henderson
An introduction to linear programming.
New York and London; 1953.

(10) A. Charnes, W. W. Cooper and B. Mellon
Blending aviation gasolines – a study in programming interdependent activities in an integrated oil company.
Econometrica, **20**, 135–159; 1952.

(10a) A. Charnes, W. W. Cooper and B. Mellon
A model for programming and sensitivity analysis in an integrated oil company.
Econometrica, **22**, 193–217; 1954.

(11) R. Dorfman
Application of linear programming to the theory of the firm.
Berkeley, 1951.

(12) F. L. Hitchcock
The distribution of a product from several sources to numerous localities.
J. Math. and Phys., **20**, 224–230; 1941.

(12a) R. Isaacs
A card game with bluffing.
Amer. Mathem. Monthly, **62**, 99–108; 1955.

(13) I. Kaplansky
A contribution to von Neumann's theory of games.
Annals of Math., **46**, 474–479; 1945.

(14) T. C. Koopmans
Optimum utilization of the transportation system.
Econometrica, **17** (Suppl.) 136–145; 1949.

(15) T. C. Koopmans (edit.)
Activity analysis of production and allocation.
Cowles Comm. for Res. in Econ. Monograph No. 13.
New York and London, 1951.

(16) H. W. Kuhn and A. W. Tucker (edit.)
Contributions to the theory of games.
Ann. of Math. Studies No. 24. Princeton N.J., 1950.

(17) H. W. Kuhn and A. W. Tucker (edit.)
Contributions to the theory of games. II.
Ann. of Math. Studies No. 28. Princeton N.J., 1953.

(18) C. E. Lemke
The dual method of solving the linear programming problem.
Naval Research Logistics Quarterly, **1**, 36–47; 1954.

(19) G. Morton
Notes on linear programming.
Economica, **18**, 397–411; 1951.

(20) J. C. C. McKinsey
Introduction to the theory of games.
New York, Toronto and London, 1952.

(21) J. Nash
Non-cooperative games.
Annals of Math., **54**, 286–295; 1951.

(22) J. von Neumann
Zur Theorie der Gesellschaftsspiele.
Math. Annalen, **100**, 295–320; 1928.

(23) J. von Neumann
Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes.
Ergebnisse eines mathematischen Kolloquiums, **8**, 73–83; 1937.

(24) J. von Neumann and O. Morgenstern
Theory of games and economic behaviour.
Princeton N.J., 1944 (2nd ed. 1947).

(25) G. J. Stigler
The cost of subsistence.
J. of Farm Econ., **27**, 303–314; 1945.

(26) S. Vajda
Theory of games.
Science News, **28**, 29–40; 1953.

(27) J. Ville
Note sur la théorie générale des jeux où intervient l'habilité des joueurs.
Traité du calcul des probabilités et de ses applications, ed. É. Borel. *Tome IV, fasc. II.*, 105–113. Paris, 1938.

(28) J. D. Williams
The compleat strategyst.
New York, 1954.

(29) Conference on linear programming, May 1954. (Papers by S. Vajda, A. Land, G. Morton, J. A. C. Brown, E. M. L. Beale, and D. Prinz.)
London, 1955.

(30) Symposium on linear programming. (Papers by S. Vajda, E. M. L. Beale, and G. Morton and A. H. Land.)
J. Roy. Statist. Soc. B, **17**, 165–203; 1955.

(31) Proceedings of the Second Symposium in linear programming, Vols I, II.
Washington D.C. 1955.

A number of reports have been written for private circulation which are not available on the book market and have not been included in this list. Many of these reports are earlier versions of papers which have later been published.

The references are mentioned on the following pages:

REF. NO.	PAGE	REF. NO.	PAGE	REF. NO.	PAGE
1	96, 97, 98	11	32	21	5
2	66	12	31	22	1, 22
3	3	12a	3	23	22
4	29	13	10	24	1, 4, 5, 7, 22
5	1	14	31	25	32
6	32	15	31, 32, 33, 35, 46, 70, 77	26	2
7	66			27	22, 25, 29
7a	31	16	29	28	3
8	75	17	29	29	33, 34
9	35, 66	18	78	30	33
10	32	19	33	31	33
10a	75	20	3, 4, 29		

ABBREVIATIONS

		DEFINED ON PAGE
b.f.s.	basic feasible solution	35
b.s.	basic solution	35
b.v.	basic variable	35
D.S.M.	Dual Simplex Method	78
F.D.T.	Fundamental Duality Theorem	75
f.s.	feasible solution	35
I.M.M.	Inverse Matrix Method	52
L.F.	Linear Form	30
L.P.	Linear Programming	30
M.L.V.	Method of Leading Variables	97
M.M.T.	Minimax Theorem (Fundamental Theorem of the theory of games)	21
n.b.v.	non-basic variable	35
S.M.	Simplex Method	35
T.A.M.	Theorem of the Alternatives for Matrices	24
T.S.H.	Theorem of the Supporting Hyperplane	24

INDEX OF DEFINITIONS
(Terms with abbreviations see above.)

DIRECTORY OF GAMES

2 by 2 games	PAGE	VALUE	Solution for A	Solution for B
Matching Pennies $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$	2, 7, 8, 10	0	(1/2, 1/2)	(1/2, 1/2)
Bluffing $\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}$	3, 9, 10	1/3	(2/3, 1/3)	(2/3, 1/3)
$\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$	10	1	(0, 1), (1, 0)	(1, 0)
2 by 3 games				
$\begin{pmatrix} 1 & 2 & 4 \\ 4 & 2 & 1 \end{pmatrix}$	9	1	(2/3, 1/3), (1/3, 2/3)	(0, 1, 0)
$\begin{pmatrix} 4 & 1 & 3 \\ 2 & 3 & 4 \end{pmatrix}$	10, 12, 14, 41 89, 91, 94	5/2	(1/4, 3/4)	(1/2, 1/2, 0)
$\begin{pmatrix} 2 & 5 & 3 \\ 1 & 3 & 4 \end{pmatrix}$	17	2	(1, 0)	(1, 0, 0)
$\begin{pmatrix} 3 & 4 & 5 \\ 2 & 1 & 3 \end{pmatrix}$	17	3	(1, 0)	(1, 0, 0)
3 by 3 games				
$\begin{pmatrix} 2 & -1 & -2 \\ 1 & 0 & 1 \\ -2 & -1 & 2 \end{pmatrix}$	5	0	(0, 1, 0)	(0, 1, 0)
Eluding game $\begin{pmatrix} -1 & 1 & 1 \\ 2 & -2 & 2 \\ 3 & 3 & -3 \end{pmatrix}$	7, 86	6/11	(6/11, 3/11, 2/11)	(5/22, 4/11, 9/22)

DIRECTORY OF L.P. PROBLEMS

	PAGE
Example 1. $x_1=4$, $x_2=1$, $x_3=9$; $C=-3.$	36, 40, 54
Example 1 but with changed sign of x_5. Infinite solution.	63
Example 2 $x_1=4$, $x_2=1$, $x_3=5$; $C=-3.$	42, 60, 82, 99
Example 3 No solution	43, 63
Example 4 $x_1=2$, $x_5=3$; $C=-2.$	43, 64, 67
Example 5 $x_1=2$ or $3\frac{1}{2}$, $x_2=0$ or $1\frac{1}{2}$, $x_3=6$ or $7\frac{1}{2}$, $x_5=3$ or 0, $x_6=0$ or 0; $C=-2.$	44, 70
Example dual to example 2 $y_4=-2/3$, $y_5=1/3$; $B=-3.$	80

Various problems are mentioned on pp. 31ff. In particular, Transportation problem, mentioned on pp. 31, 39, 59, 69.