Chapter 20

ZERO-SUM TWO-PERSON GAMES

T.E.S. RAGHAVAN*

University of Illinois at Chicago

Contents

References 757
Appendix, by T. Radzik and T.E.S. Raghavan. Duels 761

*The author would like to thank Ms Evangelista Fe, Swaminathan Sankaran and the anonymous referees for the many detailed comments that improved the presentation of this article.

Handbook of Game Theory, Volume 2, Edited by R.J. Aumann and S. Hart

Many parlor games like the game of "Le Her" or "Morra" [Dresher (1963)], involve just two players and finitely many moves for the players. When the number of actions available to a player is finite, we could in theory, reduce the problem to a game with exactly one move for each player, a move where actions for the players are various master plans for the entire game chosen from among the finitely many possible master plans. These master plans are called *pure strategies*. The original game is often analyzed using such a reduced form called games in normal form (also called strategic form). Certain information about the original game can be lost in the process [Kuhn (1953), Aumann and Maschler (1972)]. The reduction however helps to focus our understanding of the strategic behavior of intelligent players keen on achieving certain guaranteed goals against all odds.

Given a two-person game with just one move for each player, the players independently and simultaneously select one among finitely many actions resulting in a payoff for each player. If i, j are their independent choices in a play, then the game is defined by a pair of real matrices $A = (a_{ij})$, $B = (b_{ij})$ where a_{ij} is the payoff to player I and b_{ij} is the *payoff* to player II. The game is called zero sum if $a_{ij} + b_{ij} \equiv 0$. Thus in zero-sum games, what one player gains, the opponent loses. In such games A suffices to determine the payoff.

We can use the following example [Dresher (1963)] to illustrate what we have said so far.

Example. From a deck of three cards numbered 1, 2, 3 player I picks a card at will. Player II tries to guess the card. After each guess player I signals either High or Low or Correct, depending on the guess of the opponent. The game is over as soon as the card is correctly guessed by player II. Player II pays player I an amount equal to the number of trials he made.

There are three pure strategies for player I. They are:

α: Choose 1,

β: Choose 2,

γ: Choose 3.

For player II the following summarizes the possible pure strategies, excluding obviously "bad ones".

(a) Guess 1 at first. If the opponent says Low, guess 2 in the next round. If the opponent still says Low, guess 3 in the next round.

(b) Guess 1 at first. If the opponent says Low, guess 3 in the next round. If the opponent says High, guess 2 in the next round.

(c) Guess 2 at first. If the opponent says Low, guess 3; if the opponent says High, guess 1.

(d) Guess 3 at first. If the opponent says High, guess 1 in the next round. If the opponent says Low, guess 2 in the next round.

(e) Guess 3 at first. If the opponent says High, guess 2 in the next round. If the opponent still says High, guess 1 in the next round.

Thus the payoff matrix is given by

$$A = \begin{array}{c} \\ \alpha \\ \beta \\ \gamma \end{array} \begin{array}{c} \begin{array}{ccccc} (a) & (b) & (c) & (d) & (e) \end{array} \\ \begin{pmatrix} 1 & 1 & 2 & 2 & 3 \\ 2 & 3 & 1 & 3 & 2 \\ 3 & 2 & 2 & 1 & 1 \end{pmatrix} \end{array}.$$

A pair of pure strategies (i^*, j^*) for a payoff matrix $A = (a_{ij})$ is called a *saddle point* if $a_{ij^*} \leqslant a_{i^*j^*} \leqslant a_{i^*j} \quad \forall_i, \forall_j$. Thus $a_{i^*j^*}$ is the guaranteed gain for player I against any "j" player II chooses. It is also the maximum loss to player II against any "i" player I chooses. In general such a saddle point may not exist. When it does we have

$$a_{i^*j^*} = \max_i \min_j a_{ij} = \min_j \max_i a_{ij}.$$

The above game has no saddle point. In fact $\max_i \min_j a_{ij} = 1$ and $\min_j \max_i a_{ij} = 2$. When players face intelligent opponents, the saddle point strategies are the best course of actions for both players.

When a game has no saddle point, without violating the rules of the game, it might be possible for players to enlarge the available set of masterplans and look for a saddle point in the enlarged game. For example a master plan for a player in this new game could be based on the outcome of a suitably chosen random device from an available set of random devices, where the possible outcomes are themselves masterplans of the original game. If each player selects his masterplan based on the outcome of his independently chosen random device we enter into a new game called *a mixed extension* of the game. Here the set of available random devices are called mixed strategies for the original game and they will be the pure strategies for the mixed extension game. The expected payoff $K(x, y)$ is the outcome for the mixed extension game when player I uses the random device x and player II uses the random device y.

In our example above player I could rely on a random device $x^* = (x_1^*, x_2^*, x_3^*) = (\frac{2}{5}, \frac{1}{5}, \frac{2}{5})$ where say x_1^* is the probability of choosing pure strategy α in the original game. Player I can guarantee an expected gain of $\frac{9}{5}$ no matter which pure strategy player II uses. Similarly, the random device $y^* = (y_a^*, y_b^*, y_c^*, y_d^*, y_e^*) = (0, \frac{1}{5}, \frac{3}{5}, \frac{1}{5}, 0)$ where say $y_c^* = \frac{3}{5}$ is the probability of selecting the pure strategy c by the random device y^*. Player II by using y^* can guarantee an expected loss no more than $\frac{9}{5}$ whatever player I does. Thus, if x, y are any arbitrary probability vectors for players I and II, then the expected payoff $K(x, y)$ to player I satisfies

$$K(x, y^*) \leqslant K(x^*, y^*) \leqslant K(x^*, y) \quad \forall x, \text{ and } \forall y.$$

Thus for the mixed extension game, (x^*, y^*) is a saddle point. Here x^*, y^* are called *optimal mixed strategies* for players I and II respectively and $\frac{9}{5}$ is called the *value* of the game.

While Borel [von Neumann (1953)] suspected that games may not in general possess value even in the mixed extension, von Neumann (1928) laid the foundations of Game Theory by proving the decisive *minimax* theorem. This celebrated theorem remained dormant for a while. It was the classic monograph of [von Neumann and Morgenstern (1944)] that arrested the attention of mathematicians and social scientists and triggered research activity in Game theory.

Of the many existence proofs for the minimax theorem [von Neumann (1928), Kakutani (1941), Ville (1938), Loomis (1946), Nash (1950), Owen (1970) and so on] Kakutani's proof and Nash's proof are amenable for extensions to more general situations. Nash's proof is based on the Brouwer's fixed point theorem. Brouwer's theorem asserts that any continuous self map ϕ of a compact convex set X in Euclidean n-space R^n admits a fixed point. That is $\phi(x) = x$ for some $x \in X$. We will now prove the minimax theorem using the Brouwer's fixed point theorem.

Minimax theorem. (von Neumann). *Let $A = (a_{ij})$ be any $m \times n$ real matrix. Then these exists a pair of probability vectors $x^* = (x_1^*, x_2^*, \ldots, x_m^*)$* and *$y^* = (y_1^*, y_2^*, \ldots, y_n^*)$ such that for a unique constant v*

$$\sum_i a_{ij} x_i^* \geqslant v, \quad j = 1, 2, \ldots, n, \tag{1}$$

$$\sum_j a_{ij} y_j^* \leqslant v, \quad i = 1, 2, \ldots, m. \tag{2}$$

Equivalently if $K(x, y) = \sum_i \sum_j a_{ij} x_i y_j$ then (x^*, y^*) is a saddle point for $K(x, y)$. That is

$$\min_y \max_x K(x, y) = \max_x \min_y K(x, y),$$

where min *and* max *are taken respectively over the set of all probability vectors x for player I and probability vectors y for player* II.

Proof. Given any probability vector x for I and y for II if $p_i = p_i(x, y) = \sum_j a_{ij} y_j - \sum_i \sum_j a_{ij} x_i y_j$, $q_j = q_j(x, y) = \sum_i \sum_j a_{ij} x_i y_j - \sum_i a_{ij} x_i$, we are looking for a pair (x, y) with $p_i \leqslant 0$ and $q_j \leqslant 0$ for all i,j. Let $\varphi: (x, y) \to (\xi, \eta)$ be a map where ξ is a probability m-vector and η is a probability n-vector defined by coordinates

$$\xi_i = \frac{x_i + \max(p_i, 0)}{1 + \sum_1^m \max(p_k, 0)}, \quad \eta_j = \frac{y_j + \max(q_j, 0)}{1 + \sum_1^n \max(q_k, 0)}; \quad i = 1, 2, \ldots, m,$$
$$j = 1, 2, \ldots, n.$$

It can be easily checked that φ is a continuous self-map of the set of probability vector pairs. By the Brouwer's fixed point theorem, there exists an (x^*, y^*) such that $\varphi(x^*, y^*) = (x^*, y^*)$.

Thus $x_i^* \sum_k \max(p_k, 0) = \max\ (p_i, 0)$ for all i. For some $i, x_i^* > 0, p_i \leqslant 0$. Thus $\max(p_k, 0) = 0$ for all k. A similar argument shows that $\max(q_l, 0) = 0$ for all l. These are precisely the equalities we wanted for $v = \sum_i \sum_j a_{ij} x_i^* y_j^*$. The second assertion that (x^*, y^*) is a saddle point for the mixed extension $K(x, y)$ easily follows from multiplying respectively the inequalities (1) and (2) by y_j and x_i and taking the sum over j and i respectively. □

The existence theorem does not tell us how to compute a pair of optimal strategies. An algebraic proof was given by Weyl (1950). Thus if the entries of the payoff are rational, by the algebraic proof one can show that the value is rational and a pair of optimal strategies have rational entries. Brown and Von Neumann (1950) suggested a solution by differential equations that roughly mimics the above tatonnement process $(x, y) \to (\xi, \eta)$. It turned out that the minimax theorem can be proved via linear programming in a constructive way leading to an efficient computational algorithm *a la* the simplex method [Dantzig (1951)]. Interestingly the minimax theorem can also be used to prove a version of the duality theorem of linear programming.

Equivalence of the minimax theorem and the duality theorem

Given an $m \times n$ real matrix $A = (a_{ij})$ and given column m-vector b and column n-vector c we have the following two problems called dual linear programs in standard form.

$$\begin{array}{ll} \text{Primal:} & \max c \cdot x \\ \text{subject to} & Ax \leqslant b, \quad x \geqslant 0, \\ \text{Dual:} & \min b \cdot y \\ \text{subject to} & A^T y \geqslant c, \quad y \geqslant 0. \end{array}$$

Any x satisfying the constraints of the primal is called a *feasible solution to the primal*. Feasible solutions to the dual are similarly defined. Here $c \cdot x$ and $b \cdot y$ are called the *objective functions* for the primal and the dual.

The following is a version of the fundamental theorem of linear programming due to von Neumann [Dantzig (1951)].

Duality Theorem. *If the primal and the dual problems have at least one feasible solution, then the two problems have optimal solutions; further, at any optimal solution, the value of the two objective functions coincide.*

Indeed the duality theorem and the minimax theorem are equivalent. We will see that either one implies the other.

Duality theorem ⇒ minimax theorem

Given a payoff matrix $A=(a_{ij})$, by adding a constant c to all the entries, we can assume $a_{ij}>0$ for all i,j and $v>0$. Our inequalities (1) and (2) can be rewritten as

$$\sum_i a_{ij}\left(\frac{x_i}{v}\right)\geqslant 1, \quad \text{for all } j,$$

$$\sum_j a_{ij}\left(\frac{y_j}{v}\right)\leqslant 1, \quad \text{for all } i,$$

where say, $p=(p_1,\dots,p_m)$, $q=(q_1,\dots,q_n)$ with

$$p_i=\frac{x_i}{v}\geqslant 0, \quad \sum_i p_i=\frac{1}{v}, \tag{3}$$

$$q_j=\frac{y_j}{v}\geqslant 0, \quad \sum_j q_j=\frac{1}{v}. \tag{4}$$

Given a matrix $A=a_{ij}$ with all entries positive consider the dual linear programs

Problem A. $\min \sum_i p_i$ subject to $\sum_i a_{ij}p_i\geqslant 1$ for all j and $p_i\geqslant 0$ for all i.

Problem B. $\max \sum_j q_j$ subject to $\sum_j a_{ij}q_j\leqslant 1$ for all i and $q_j\geqslant 0$ for all j.

Here $q=0$ is feasible for Problem B. For large N the vector $p=(N,\dots,N)$ is feasible for Problem A. By the duality theorem we have an optimal solution p^*,q^* to the two problems. Further $\sum_i p_i^*=\sum_j q_j^*$ at any optimal pairs. By normalizing p^* and q^* we have optimal x^* and y^* satisfying (1), (2) for the payoff A with value $v=1/\sum_j q_j^*$. Thus the minimax theorem is equivalent to solving the above dual linear programs.

Before reducing the dual linear programs to a single game problem we need the following theorems.

Theorem on skew symmetric payoffs. *Let $A=-A^{\mathrm{T}}$ be a payoff matrix. Then the value of the game is zero and both players have the same set of optimal strategies.*

Proof. Let if possible $v<0$. Let y be optimal for player II. Thus $Ay=-A^{\mathrm{T}}y<0$ and hence $A^{\mathrm{T}}y>0$. This contradicts $v<0$. Similarly we can show that $v>0$ is not possible. Thus $v=0$ and optimal strategies of one player are also optimal for the other player. □

Equalizer theorem. *Let $A=(a_{ij})$ be a payoff with value v. Let the expected payoff to player* I *when he uses any optimal strategy x and when player* II *uses a fixed* *column be v. Then player* II *has an optimal strategy which chooses this column with positive probability.*

Proof. We can as well assume $v = 0$ and the particular column is the nth column. It is enough to show that the system of inequalities

$$\sum_j a_{ij} u_j \leqslant 0$$
$$u_n = 1$$

has a solution $u \geqslant 0$. We could normalize this and get the desired strategy. This is equivalent to finding a solution $u, w \geqslant 0$ to the matrix equation

$$\begin{bmatrix} A & I \\ e_n & 0 \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} = \mathbf{p} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

(Here e_n is nth unit vector). It is enough to show that the point $\mathbf{p}$ is in the closed cone K generated by the columns of the matrix

$$B = \begin{bmatrix} A & I \\ e_n & 0 \end{bmatrix}.$$

Suppose not. Then by the strong separation theorem [Parthasarathy and Raghavan (1971)] we can find a strictly separating hyperplane ($\mathbf{f}$, α) between the closed cone K and the point $\mathbf{p}$. This shows that the vector $\mathbf{f} \geqslant 0$ and the scalar $\alpha < 0$. Further $\mathbf{f} \cdot a_j \geqslant 0$ for all columns a_j and $\mathbf{f} \cdot a_n + \alpha \geqslant 0$. Normalizing $\mathbf{f}$ we get an optimal strategy for player I which gives positive expectation when player II selects column n. This contradicts the assumption about column n. □

Minimax theorem ⇒ Duality theorem

Consider the following $m + n + 1 \times m + n + 1$ skew symmetric payoff matrix

$$B = \begin{bmatrix} 0 & A & -b \\ -A^{\mathrm{T}} & 0 & c \\ b^{\mathrm{T}} & -c^{\mathrm{T}} & 0 \end{bmatrix}.$$

Since B is a skew symmetric payoff the value of B is 0. Any optimal mixed strategy (η^*, ξ^*, θ) for player II is also optimal for player I. If $\theta > 0$, then the vectors $y^* = (1/\theta) \cdot \eta^*, x^* = (1/\theta) \cdot \xi^*$ will be feasible for the two linear programming problems. Further when player II uses the optimal strategy and I uses the last row, the expected income to player I is $b \cdot y^* - c \cdot x^* \leqslant 0$. Since the inequality $b \cdot y^* - c \cdot x^* \geqslant 0$ is always true for any feasible solutions of the two linear programs, the two objective functions have the same value and it is easy to see that the problems have y^*, x^* as their optimal solutions. We need to show that for some optimal (η^*, ξ^*, θ),

$\theta > 0$. Suppose not. Then by the equalizer theorem player II pays a *penalty* if he chooses column n with positive probability, namely $-b\cdot\eta^* + c\cdot\xi^* > 0$. Thus either $b\cdot\eta^* < 0$ or $c\cdot\xi^* > 0$. Say $c\cdot\xi^* > 0$. For any feasible $\bar{x}$ for the primal and for the feasible y° of the dual we have $c\cdot\bar{x} - b\cdot\bar{y} \leqslant 0$. For large N, $\bar{x} = x^\circ + N\xi^*$ is feasible for the primal and $c\cdot(x^\circ + N\xi^*) - b\cdot y^\circ > 0$ a contradiction. A similar argument can be given when $b\cdot\eta^* < 0$.

Extreme optimal strategies. In general optimal strategies are not unique. Since the optimal strategies for a player are determined by linear inequalities the set of optimal strategies for each player is a closed bounded convex set. Further the sets have only *finitely many extreme points* and one can effectively enumerate them by the following characterization due to Shapley and Snow (1950).

Theorem. *Let A be an $m \times n$ matrix game with $v \neq 0$. Optimal mixed strategies x^* for player* I *and y^* for player* II *are extreme points of the convex set of optimal strategies for the two players if and only if there is a square submatrix $B = (a_{ij})_{i\in I, j\in J}$, such that*

i. *B nonsingular.*
ii. $\sum_{i\in I} a_{ij}x_i^* = v \quad j\in J \subset \{1,2,\ldots,n\}$.
iii. $\sum_{j\in J} a_{ij}y_j^* = v \quad i\in I \subset \{1,2,\ldots,m\}$.
iv. $x_i^* = 0$ *if* $i \notin I$.
v. $y_j^* = 0$ *if* $j \notin J$.

Proof. (Necessary.) After renumbering the columns and the rows, we can assume for an extreme optimal pair (x^*, y^*), $x_i^* > 0$ $i = 1,2,\ldots,p$; $y_j^* > 0$ $j = 1,2,\ldots,q$. If row i is actively used (i.e. $x_i^* > 0$,) then $\sum_{j=1}^n a_{ij}y_j^* = v$. Thus we can assume $\tilde{A} = (a_{ij})_{i\in\bar{I}, j\in\bar{J}}$ such that

$$\sum_{i=1}^{p} a_i{}^j x_i^* \begin{cases} = v, & j\in\bar{J} \supset \{1,2,\ldots,q\}, \\ > v, & \text{for } j\notin\bar{J}, \end{cases}$$

$$\sum_{j=1}^{q} a_{ij} y_j^* \begin{cases} = v, & i\in\bar{I} \supset \{1,2,\ldots,p\}, \\ < v, & \text{for } i\notin\bar{I}. \end{cases}$$

We claim that the p rows and the q columns of $\tilde{A} = (a_{ij})_{i\in\bar{I}, j\in\bar{J}}$ are independent. Suppose not. For some $(\pi_1, \ldots, \pi_p) \neq 0$

$$\sum_{i=1}^{p} a_{ij}\pi_i = 0, \quad j\in\bar{J},$$

$$\sum_{i=1}^{p} a_{ij}x_i^* = v, \quad j\in\bar{J}.$$

Thus $\sum_{i=1}^p \pi_i \sum_{j=1}^q a_{ij}y_j^* = v\sum_{i=1}^p \pi_i = 0$. As $v \neq 0$, $\sum_{i=1}^p \pi_i = 0$. Since $\sum_i a_{ij}x_i^* > v$ for

$j \notin \bar{J}$ we can find $\varepsilon > 0$ sufficiently small such that $x'_i = x^*_i - \varepsilon\pi_i \geqslant 0$, $x''_i = x^*_i + \varepsilon\pi_i \geqslant 0$, $x'_i = x''_i = \pi_i \equiv 0$, $i > p$ and

$$\sum_{i=1}^{m} a_{ij}x'_i = \sum_{i=1}^{m} a_{ij}x^*_i - \varepsilon \sum_{i=1}^{m} a_{ij}\pi_i \geqslant v, \quad \text{for all } j,$$

$$\sum_{i=1}^{m} a_{ij}x''_i = \sum_{i=1}^{m} a_{ij}x^*_i + \varepsilon \sum_{i=1}^{m} a_{ij}\pi_i \geqslant v, \quad \text{for all } j.$$

Thus x', x'' are optimal and $x^* = (x' + x'')/2$ is not extreme optimal, a contradiction. Similarly the first q columns of $\tilde{A}$ are independent. Hence a nonsingular submatrix B containing the first p rows and the first q columns of $\tilde{A} = (a_{ij})_{i\in\bar{I}, j\in\bar{J}}$ and satisfying conditions (i)–(iv) exists.

Conversely, given such a matrix B satisfying (i)–(iv), the strategy x^* is extreme optimal, otherwise $x' \neq x''$, $x^* = (x' + x'')/2$ and we have $\sum_i a_{ij}x'_i \geqslant v$, $\sum_i a_{ij}x''_i \geqslant v$ for $j \in J$ with $\sum_i a_{ij}x^*_i = v$ for $j \in J$. Thus $\sum_{i\in I} a_{ij}x'_i = \sum_{i\in I} a_{ij}x''_i = v$ for $j = J$ and the matrix B is singular. □

Since there are only finitely many square submatrices to a payoff matrix there could be only finitely many extreme points and as solutions of linear equations, they are in the same ordered subfield as the data field. The problem of efficiently loacting an extreme optimal strategy can be handled by solving the linear programming problem mentioned above. Among various algorithms to solve a linear programming problem, the simplex algorithm is practically the most efficient. Linear inequalities were first investigated by Fourier (1890) and remained dormant for more than half a century. Linear modeling of problems in industrial production and planning necessitated active research and the pioneering contributions of Kantorovich (1939), Koopmans (1951) and Dantzig (1951) brought them to the frontiers of modern applied mathematics.

Simplex algorithm. Consider the canonical linear programming problem

$$\max \sum_j b_j y_j$$

subject to

$$\sum_{j=1}^{n} a_{ij}y_j = d_i, \quad i = 1, 2, \ldots, m,$$

$$y_j \qquad \geqslant 0, \quad j = 1, 2, \ldots, n.$$

Any solution $y = (y_1, \ldots, y_n)$ to the above system of inequalities is called a feasible solution. We could also write the system as

$$y_1C^1 + y_2C^2 + \cdots + y_nC^n = d,$$

$$y_1, y_2, \ldots, y_n \geqslant 0,$$

where $C^1, C^2, \ldots, C^n$ are the columns of the matrix A and d is the column vector with coordinates $d_i, i = 1, 2, \ldots, m$. It is not hard to check that any extreme point $y = (y_1, y_2, \ldots, y_n)$ of the convex polyhedra of feasible solutions can be identified with a set of linearly independent columns $C^{i_1}, C^{i_2}, \ldots, C^{i_k}$ such that the y_j are zero for coordinates other *than* $i_1, i_2, \ldots, i_k$. By slightly perturbing the entries we could even assume that every extreme point of feasible solutions has exactly m coordinates positive.

We call two extreme points adjacent if the line segment joining them is a one dimensional face of the feasible set. Algebraically, two adjacent extreme points can be identified with two bases which differ by exactly one basis vector. The new basis is chosen by bringing a column from outside into the current basis which in turn determines the removal of an appropriate column from the current basis. An iteration consists of searching for an improved value of the objective function at an adjacent extreme point. The algorithm terminates if no improvements in the value of the objective function with adjacent extreme points is possible.

Fictitious play. An intuitively easy to implement algorithm to find the approximate value uses the notion of fictitious play [Brown (1951)].

Assume that a given matrix game, $A = (a_{ij})_{m \times n}$ has been played for t rounds with $(i_1, j_1), (i_2, j_2), \ldots, (i_t, j_t)$ as the actual choices of rows and columns by the two players. Let rows $1, 2, \ldots, m$ appear $k_1, k_2, \ldots, k_m$ times in the t rounds. For player II one way of learning from player I's actions is to pretend that the proportions $(k_1/t, \ldots, k_m/t)$ are the true mixed strategy choices of player I. With such a belief, the best choice for player II in round $t+1$ is to choose any column j_{t+1} which minimizes the fictitious expected payoff $(1/t)\sum_{i=1}^{m} k_i a_{ij}$.

Suppose in the above data, columns $1, 2, \ldots, n$ appear $l_1, l_2, \ldots, l_n$ times in the first t rounds. Player I can also pretend that the true strategy of player II is $(l_1/t, l_2/t, \ldots, l_n/t)$. With such a belief the best choice for player I in round $t+1$ is to choose any row i_{t+1} which maximizes the fictitious expected income $1/t \sum_{j=1}^{n} l_j a_{ij}$. The remarkable fact is that this naive procedure can be used to approximate the value of the game. We have the following

Theorem. *Let (x^t, y^t) be the strategies $(k_1/t, \ldots, k_m/t)$, $(l_1/t, \ldots, l_n/t)$ $t = 1, 2, \ldots$ where (x^1, y^1) is arbitrary and (x^t, y^t) for $t \geqslant 2$ is determined by the above fictitious play. Then*

$$v = \lim_{t \to \infty} \min_j \frac{1}{t} \sum_{i=1}^{m} k_i a_{ij} = \lim_{t \to \infty} \max_i \frac{1}{t} \sum_{j=1}^{n} l_j a_{ij}.$$

The above procedure is only of theoretical interest. It is impractical and the convergence to the value is known to be very slow. Even though $v(t) = \min_j (1/t) \sum_{i=1}^{m} k_i a_{ij} \to v$ as $t \to \infty$, the mixed strategies $\xi(t) = (k_1/t, k_2/t, \ldots, k_m/t)$ and $\eta(t) = (l_1/t, \ldots, l_n/t)$ may not converge.

A proof can be given [Robinson (1951)] by showing that for any skew symmetric payoff A,

$$\lim_t \min_j \sum_i k_i a_{ij} = 0.$$

In general the sequence of strategies $\{(x^t, y^t)\}$ oscillates around optimal strategies.

Completely mixed games. A mixed strategy x for player I is *called completely mixed* iff $x > 0$ (i.e. all rows are essentially used). Suppose x, y are completely mixed optimal strategies for the two players. The inequalities $Ay \leqslant v \cdot 1$ are actually equalities, for otherwise $v = (x, Ay) < (x, v \cdot 1) = v$, a contradiction. In case $v = 0$, the matrix A is singular. We call a matrix game A *completely mixed* if every optimal strategy is completely mixef for both players.

Theorem. *If a matrix game A is completely mixed, then A is a square matrix and the optimal strategies are unique.*

Proof. Without loss of generality $v \neq 0$. In case $y' \neq y''$ are two extreme optimal strategies for player II, then $Ay' = v \cdot 1, Ay'' = v \cdot 1$ and $A(y' - y'') = 0$. Thus rank $A < n$. Since the extreme $y' > 0$, by the Shapley – Snow theorem, A has an $n \times n$ submatrix which is nonsingular. This contradicts rank $A < n$. Thus rank $A = n$ and the extreme optimal strategy is unique for player II. A similar argument applies for player I and shows that rank $A = m$, and the extreme optimal strategy is unique for player I. □

We have a formula to compute the value v for completely mixed games, and it is given by solving $Ay = v \cdot 1$. The unique solution y is optimal for player II. Since y is a probability vector $y = vA^{-1} \cdot 1$ gives $v = \det A/(\sum_i \sum_j A_{ij})$ where det A is the determinant of A and A_{ij} are the cofactors of A.

In case the payoff is a square matrix it can be shown that when one player has on optimal strategy which is not completely mixed then his opponent also possesses an optimal strategy that is not completely mixed [Kaplansky (1945)].

For Z-matrices (square matrices with off-diagonal entries nonpositive) if the value is positive, then the maximizer cannot omit any row (this results in a submatrix with a nonpositive column which the minimizer will choose even in the original game). Thus the game is completely mixed. One can infer many properties of such matrices by noting the game is completely mixed. It is easy to check that since $v > 0$ the matrix is non-singular and its inverse is nonnegative: For completely mixed games $A = (a_{ij})$ with value zero, the cofactor matrix (A_{ij}) has all $A_{ij} > 0$ or all $A_{ij} < 0$. This can be utilized to show that any Z-matrix with positive value has all principal minors positive [Raghavan (1978), (1979)].

The reduction of matrix games to linear programming is possible even when the strategy spaces are restricted to certain polyhedral subsets. This is useful in solving for stationary optimal strategies in some dynamic games [Raghavan and Filar (1991)].

Polyhedral constraints and matrix games. Consider a matrix game A where players I(II) can use only mixed strategies constrained to lie in polyhedra $X(Y)$. Say $X = \{x: B^T x \leqslant c,\ x \geqslant 0\}$ where $X \subset$ set of mixed strategies for player I. Let

$$Y = \{y: Ey \geqslant f, \quad y \geqslant 0\}.$$

We know that $\max_{x\in X} \min_{y\in Y} x^{\mathrm{T}} Ay = \min_{y\in Y} \max_{x\in X} x^{\mathrm{T}} Ay$. The linear program $\min_{y\in Y} x^{\mathrm{T}} Ay$ has a dual

$$\max_{z\in T} f\cdot z,$$

where $T = \{z: E^{\mathrm{T}} z \leqslant A^{\mathrm{T}} x, z \geqslant 0\}$. Thus player I's problem is

$$\max_{(z,x)\in K} f\cdot z,$$

where $K = \{(z, x): E^{\mathrm{T}} z \leqslant x,\ B^{\mathrm{T}} x \leqslant c,\ z \geqslant 0, x \geqslant 0\}$. This is easily solved by the simplex algorithm.

Dimension relations. Given a matrix game $A = (a_{ij})_{m\times n}$ let X, Y be the convex sets of optimal strategies for players I and II. Let the value $v = 0$. Let $J_1 = \{j: y_j > 0$ for some $y \in Y\}$. Let $J_2 = \{j: \sum_i a_{ij} x_i = 0$ for any $x \in X\}$. It is easy to show that $J_1 \subset J_2$. From the equalizer theorem we proved earlier that $J_2 \subset J_1$. Thus $J_1 = J_2 = J$, say. Similarly we have $I_1 = I_2 = I$ for player I. Since the rows outside I_2 and the columns outside J_2 are never used in optimal plays, we can as well restrict to the game $\tilde{A}$ with rows in I_2 and columns in J_2 and with value 0. Let $\tilde{X}, \tilde{Y}$ be the vector spaces generated by X, Y, respectively. The following theorem independently due to [Bohnenblust, Karlin and Shapley (1950)] and [Gale and Sherman (1950)] characterizes the intrinsic dimension relation between optimal strategy sets and essential strategies of the two players.

Dimension theorem. $|I| - \dim \tilde{X} = |J| - \dim \tilde{Y}$.

Proof. Consider the submatrix $\tilde{A} = (a_{ij})_{i\in I, j\in J}$ For any $y \in Y$ let $\bar{y}$ be the restriction of y to the coordinates $j \in J$. We have $\tilde{A}y = 0$. Let $\tilde{A}\pi = 0$ for some π in $R^{|J|}$. Since we can always find an optimal strategy y^* with $y_j^* > 0$ for all $j \in J$, we have $z = y^* - \varepsilon\pi \geqslant 0$ for small ε and $\tilde{A}z = 0$. Clearly $z \in \tilde{Y}$. Further since the linear span of y^* and z yield π the vector space $\{u: \tilde{A}u = 0\}$ coincides with $\tilde{Y}$. Thus dim $\tilde{Y} = |J| - \operatorname{rank} \tilde{A}$. A Similar argument shows that dim $\tilde{X} = |I| - \operatorname{rank} \tilde{A}$. Hence rank $\tilde{A} = |I| - \dim \tilde{X} = |J| - \dim \tilde{Y}$. □

Semi-infinite games and intersection theorems

Since a matrix payoff $A = (a_{ij})$ can be thought of as a function on $I \times J$, where $i \in I = \{1, 2, \ldots, m\}$ and $j \in J = \{1, 2, \ldots, n\}$, a straightforward extension is to prove the existence of value when I or J is not finite. If exactly one of the two sets I or J is assumed infinite, the games are called *semi-infinite games* [Tijs (1974)]. Among them the so-called S-games of Blackwell and Girshick (1954) are relevant in statistical decision theory when one assumes the set of states to be finite of nature.

Let Y be an arbitrary set of pure strategies for player II. Let $I = \{1, 2, \ldots, m\}$ be the set of pure strategies for player I. The bounded kernel $K(i, y)$ is the payoff to player I when "i" is player I's choice and $y \in Y$ is the choice of player II. Let $S = \{(s_1, s_2, \ldots, s_m): s_i = K(i, y), i = 1, 2, \ldots, \mathrm{m};\ y \in Y\}$. The game can also be played as follows. Player II selects an $s \in S$. Simultaneously player I selects a coordinate i. The outcome is the payoff s_i to player I by player II.

Theorem. *If S is any bounded set then the S-game has a value and player* I *has an optimal mixed strategy. If* con S *(convex hull of S) is closed, player* II *has an optimal mixed strategy which is a mixture of at most m pure strategies. If S is closed convex, then player* II *has an optimal pure strategy.*

Proof. Let $t^* \in T = \operatorname{con} \bar{S}$ be such that $\min_{t \in T} \max_i t_i = \max_i t_i^* = v$. Let $(\xi, x) = c$ be a separating hyperplane between T and the open box $G = \{x: \max_i x_i < v\}$. For any $\varepsilon > 0$ and for any $i, t_i^* - \varepsilon < v$. Thus $\xi_i \geqslant 0$ and we can as well assume that ξ is a mixed strategy for player I. By the Caratheodory theorem [Parthasarathy and Raghavan (1971)] the boundary point t^* is a convex combination of at most m points of $\bar{S}$. It is easy to check that $c = v$ and ξ is optimal for player I. When S is closed the convex combination used in representing t^* is optimal for player II; otherwise t^* is approximated by t in S which is in the ε neighborhood of t^*. $\square$

The sharper assertions are possible because the set S is a subset of R^m. Many intersection theorems are direct consequences of this theorem [Raghavan (1973)]. We will prove Berge's intersection theorem and Helly's theorem which are needed in the sequel. We will also state a geometric theorem on spheres that follows from the above theorem.

Berge's intersection theorem. *Let $S_1, S_2, \ldots, S_k$ be compact convex sets in $\mathbf{R}^m$. Let $\bigcap_{i \neq j} S_i \neq \phi$ for $j = 1, 2, \ldots, m$. If $S = \bigcup_{i=1}^k S_i$* is convex then $\bigcap_{i=1}^k S_i \neq \phi$.

Proof. Let players I and II play the S-game where I chooses one of the indices $i = 1, 2, \ldots, k$ and II chooses an $x \in S$. Let the payoff be $f_i(x) =$ distance between x and S_i. The functions f_i are continuous convex and nonnegative. For any optimal

$\mu_i I_{x_1} + \mu_2 I_{x_2} + \cdots + \mu_p I_{x_p}$ of player II, we have

$$v \geqslant \sum_j \mu_j f_i(x_j) \geqslant f_i\left(\sum_j \mu_j x_j\right).$$

Thus player II has an optimal pure strategy $x^\circ = \sum_j \mu_j x_j$. If an optimal strategy $\lambda = (\lambda_1, \ldots, \lambda_k)$ of player I skips an index say 1, then player II can always choose any $y \in \bigcap_{i \neq 1} S_i$ and then $0 = \sum_i \lambda_i f_i(y) \geqslant v$. Thus $v = 0$ and $x^\circ \in \bigcap_{i=1}^k S_i$. If $\lambda_i > 0 \forall i$, then $f_i(x^\circ) \equiv v$. Since $x^\circ \in S_i$, for some i we have $v = 0$ and $x^\circ \in \bigcap_{i=1}^k S_i$. □

Helly's theorem. *Let $S_1, S_2, \ldots, S_k$ be convex sets in $\mathbf{R}^m$. Let $\bigcap_{i \in I} S_i \neq \phi$ if $1 \leqslant |I| \leqslant m+1$. Then $\bigcap_{i=1}^k S_i \neq \phi$.*

Proof. By induction we can assume that for any $|I| = r \geqslant (m+1)$, $\bigcap_{i \in I} S_i \neq \phi$ and prove it for $|I| = r+1$. Say, $I = \{1, 2, \ldots, r+1\}$. Let $a_j \in S_i$ if $i \neq j$. Let $C = \operatorname{con}\{a_1, \ldots, a_{(r+1)}\}$. Since $r > n$ by the Caratheodary theorem $C = \bigcup_{i \in I} C_i$ where $C_i = \operatorname{con}\{a_1, \ldots, a_{i-1}, a_{i+1}, a_{r+1}\}$. (Here we define $a_0 = a_{r+1}$ and $a_{r+2} = a_1$). Further $C_i \subset S_i$. By Berge's theorem $\bigcap_{i \in I} C_i \neq \phi$. □

The following geometric theorem also follows from the above arguments.

Theorem. *Let $S_1, S_2, \ldots, S_m$ be compact convex sets in a Hilbert space. Let $\bigcap_{i \neq j} S_i \neq \phi$ for $j = 1, 2, \ldots, m$, but $\bigcap_{i=1}^m S_i = \phi$. Then there exists a unique $v > 0$ and a point x_0 such that the closed sphere $S(x_0, v)$ with center x_0 and radius v has nonnull intersection with each set S_i while spheres with center x_0 and with radius $< v$ are disjoint with at least one S_i. In fact no sphere of radius $< v$ around any other point in the space has nonempty intersection with all the sets S_i.*

When both pure strategy spaces are infinite, the existence of value in mixed strategies fails to hold even for very simple games. For example, if $X = Y =$ the set of positive integers and $K(x, y) = \operatorname{sgn}(x - y)$, then no mixed strategy $p = (p_1, p_2, \ldots)$ on X can hedge against all possible y in guaranteeing an expected income other then the worst income -1. In a sense if p is revealed, player II can select a sufficiently large number y such that the chance that a number larger than y is chosen according to the mixed strategy p is negligible. Thus

$$\sup_p \inf_y K^*(p, y) = -1 \quad \text{where } K^*(p, y) = \sum_x p(x) K(x, y).$$

A similar argument with the obvious definition of $K^*(p, q)$ shows that

$$\sup_p \inf_q K^*(p, q) = -1 < \inf_q \sup_q K^*(p, q) = 1.$$

The failure stems partly from the noncompactness of the space P of probability measures on X.

Fixed point theorems for set valued maps

With the intention of simplifying the original proof of von Neumann, Kakutani (1941) extended the classical Brouwer's theorem to set valued maps and derived the minimax theorem as an easy corollary. Over the years this extension of Kakutani and its generalization [Glicksberg (1952)] to more general spaces have found many applications in the mathematical economics literature. While Brouwer's theorem is marginally easier to prove, and is at the center of differential and algebraic topology, the set valued maps that are more natural objects in many applications are somewhat alien to mainstream topologists.

Definition. For any set Y let 2^Y denote a collection of nonempty subsets of Y. Any function $\phi: X \to 2^Y$ is called a *correspondence* from X to Y. When X, Y are topological spaces, the correspondence ϕ is *upper hemicontinuous at* x iff given an open set G in Y and given $G \supset \phi(x)$, there exists a neighborhood $N \ni x$, such that the set $\phi(N) = \bigcup_{y \in N} \phi(y) \subset G$. The correspondence ϕ is *upper hemicontinuous on* X iff it is upper hemicontinuous at all $x \in X$.

Kakutani's fixed point theorem. *Let X be compact convex in $\mathbf{R}^n$. Let 2^X be the collection of nonempty compact convex subsets of X. Let $\phi: X \to 2^X$ be an upper hemi continuous correspondence from X to X. Then $x \in \phi(x)$ for some x.*

In order to prove minimax theorems in greater generality, Kakutani's theorem was further extended to arbitrary locally convex topological vector spaces. These are real vector spaces with a Hausdorff topology admitting convex bases, where vector operations of addition and scalar multiplication are continuous. The following theorem generalizes Kakutani's theorem to locally convex topological vector spaces [Fan (1952), Glicksberg (1952)].

Fan–Glicksberg fixed point theorem. *Let X be compact convex in a real locally convex topological vector space E. Let Y be the collection of nonempty compact convex subsets of X. Let ϕ be an upper hemicontinuous correspondence from X to Y. Then $x \in \phi(x)$ for some x.*

The following minimax theorems of [Ville (1938) and Ky Fan (1952)] are easy corollaries of the above fixed point theorem.

Theorem. (Ville). *Let X, Y be compact metric spaces. Let $K(x, y)$ be continuous on $X \times Y$. Then*

$$\min_{\nu} \max_{\mu} \int\int K(x, y)\, \mathrm{d}\mu(x)\, \mathrm{d}\nu(y) = \max_{\mu} \min_{\nu} \int\int K(x, y)\, \mathrm{d}\nu(x)\, \mathrm{d}\mu(y)$$

where μ, ν may range over all probability measures on X, Y, respectively.

Ky Fan's minimax theorem. *Let X, Y be compact convex subsets of locally convex topological vector spaces. Let $K:X \times Y \to \mathbf{R}$ be continuous. For every $\bar{x} \in X$, $\bar{y} \in Y$, let $K(\bar{x}, y): Y \to \mathbf{R}$ be a convex function and $K(x, \bar{y}): X \to \mathbf{R}$ be a concave function. Then*

$$\max_{x \in X} \min_{y \in Y} K(x, y) = \min_{y \in Y} \max_{x \in X} K(x, y).$$

Proof. Given any $\bar{x} \in X$, $\bar{y} \in Y$ let

$$\Lambda(\bar{x}) = \{v : v \in Y, \quad \min_{y \in Y} K(\bar{x}, y) = K(\bar{x}, v)\}$$

and

$$\Gamma(\bar{y}) = \{u : u \in X, \quad \min_{x \in X} K(x, \bar{y}) = K(u, \bar{y})\}.$$

The sets $\Lambda(\bar{x})$ and $\Gamma(\bar{y})$ are compact convex and the function $\phi:(x, y) \to \Gamma(\bar{y}) \times \Lambda(\bar{x})$ is an upper hemicontinuous map from $X \times Y$ to all nonempty compact convex subsets of $X \times Y$. Applying Fan–Glicksberg fixed point theorem we have an $(x^\circ, y^\circ) \in \phi(x^\circ, y^\circ)$. This shows that (x°, y°) is a saddle point for $K(x, y)$. $\square$

In Ky Fan's minimax theorem, the condition that the function K is jointly continuous on $X \times Y$ is somewhat stringent. For example let $X = Y =$ the unit sphere S of the Hilbert space l_2 endowed with the weak topology. Let $K(x, y)$ be simply the inner product $\langle x, y \rangle$. Then the point to set maps $y \to \Lambda(y)$, $x \to \Gamma(x)$ are not upper hemicontinuous. [Nikaido (1954)]. Hence Ky Fan's theorem is inapplicable to this case. Yet $(0, 0)$ is a saddle point!

However in Ville's theorem, the function $K^*(\mu, \nu) = \int\int K(x, y)\, \mathrm{d}\mu(x)\, \mathrm{d}\nu(y)$ is *bilinear* on $P \times Q$ where $P(Q)$ are the set of probability measures on $X(Y)$. Further K^* is *jointly continuous* on $P \times Q$ where P, Q are compact convex metrizable spaces viewed as subsets of $C^*(X)(C^*(Y))$ in their weak topologies. Ville's theorem now follows from Fan's theorem.

Definition. Let X be a topological space. A function $f: X \to R$ is called upper semicontinuous iff $\{x : f(x) < c\}$ is open for each c. If g is upper semicontinuous, $-g$ is called lower semicontinuous.

In terms of mixed extensions of K on $X \times Y$, the following is a strengthening of Ville's theorem.

Theorem. *Let X, Y be compact Hausdorff. Let $K:X \times Y \to \mathbf{R}$ be such that $K(x, \cdot)$ and $K(\cdot, y)$ are upper semicontinuous and bounded above for each $x \in X$, $y \in Y$. Then*

$$\inf_{\nu \in Q} \sup_{\mu \in p} K^*(\mu, \nu) = \sup_{\mu \in P} \inf_{\nu \in Q} K^*(\mu, \nu),$$

where $P(Q)$ are the set of probability measures with finite support on $X(Y)$.

Proof. Let X^*, Y^* denote regular Borel probability measures on X and Y respectively. When X or Y is finite the assertion follows from Blackwell's assertion for S-games. Let $\tilde{Y}$ be any finite subset of Y. For $\varepsilon > 0$, the set of ε-optimal strategies μ of the maximizer in the mixed extension K^* on $X^* \times \tilde{Y}^*$ is weakly compact. This decreasing net of ε-optimals have a nonempty intersection with

$$\inf_{\nu \in Q} \max_{\mu \in X^*} K^*(\mu, \nu) = \max_{\mu \in X^*} \inf_{\nu \in Q} K^*(\mu, \nu). \tag{7}$$

For Y metrizable, K^* is well-defined on $X^* \times Y^*$. However by (7)

$$\inf_{\nu \in Q} \max_{\mu \in X^*} K^*(\mu, \nu) \geqslant \inf_{\nu \in Q} \sup_{\mu \in P} K^*(\mu, \nu)$$

$$\max_{\mu \in X^*} \inf_{\nu \in Q} K^*(\mu, \nu) \leqslant \sup_{\mu \in P} \inf_{\nu \in Q} K^*(\mu, \nu).$$

Thus the assertion follows when Y is metrizable. The general case can be handled as follows. Associate a family G of continuous functions ϕ on Y with $K(x, y) \geqslant \phi(y)$ for some x. We can assume G to be countable with

$$\inf_{\nu \in Q} \sup_{\phi \in G} \int \phi(y)\, \mathrm{d}\nu(y) = \inf_{\nu \in Q} \int \sup_{\mu \in P} K(\mu, \nu).$$

Essentially G can be used to view Y as a metrizable case. □

Other extensions to non-Hausdorff spaces are also possible [Mertens (1986)].

One can effectively use the intersection theorems on convex sets to prove more general minimax theorems.

General minimax theorems are concerned with the following problem: Given two arbitrary sets X, Y and a real function $K: X \times Y \to \mathbf{R}$, under what conditions on K, X, Y can one assert

$$\sup_{x \in X} \inf_{y \in Y} K(x, y) = \inf_{y \in Y} \sup_{x \in X} K(x, y).$$

A standard technique in proving general minimax theorems is to approximate the problem by the minimax theorem for matrix games. Such a reduction is often possible with some form of compactness of the space X or Y and a suitable continuity and convexity or quasi-convexity on the function K.

Definition. Let X be a convex subset of a topological vector space. Let $f: X \to \mathbf{R}$. For convex functions, $\{x: f(x) < c\}$ is convex for each c. Generalizing convex functions, a function $f: X \to \mathbf{R}$ is called *quasi-convex* if for each real c, $\{x: f(x) < c\}$ is convex. A function g is *quasi-concave* if $-g$ is quasi-convex.

Theorem. [Sion (1958)]. *Let X, Y be convex subsets of linear topological spaces with X compact. Let $K: X \times Y \to \mathbf{R}$ be upper semicontinuous in x (for each fixed y) and lower semicontinuous in y (for each x). Let $K(x, y)$ be quasi-concave in x and* *quasi-convex in y. Then*

$$\sup_{x\in X}\inf_{y\in Y} K(x,y) = \inf_{x\in Y}\sup_{x\in X} K(x,y).$$

Proof. Case (i). Both X, $Y$are compact and convex: Let if possible $\sup_x \inf_y K(x,y) < c < \inf_y \sup_x K(x,y)$. Let $A_x = \{y: K(x,y) > c\}$ and $B_y = \{x: K(x,y) < c\}$. Therefore we have finite subsets $A \subset X, B \subset Y$ such that for each $y \in Y$ and hence for each $y \in \mathrm{Con}\, B$, there is an $x \in A$ with $K(x,y) > c$ and for each $x \in X$ and hence for each $x \in \mathrm{Con}\, A$, there is a $y \in B$, with $K(x,y) < c$. Without loss of generality let A, B be with *minimal cardinality* satisfying the above conditions. We claim that there exists an $x_0 \in \mathrm{Con}\, A$ such that $K(x_0, y) < c$ for all $y \in B$ and hence for all $y \in \mathrm{Con}\, B$ [by quasi-convexity of $K(x_0, y)$]. Suppose not. Then if $B = \{y_0, y_1, \ldots, y_n\}$, for any $x \in X$ there exists a $y_j \in B$ such that $K(x, y_j) \geqslant c$. Let $C_i = \{x: K(x, y_i) \geqslant c\}$. The minimality of B means that $\bigcap_{i=0, i\neq j}^{n} C_i \neq \phi$. Further $\bigcap_{i=0}^{n} C_i = \phi$. By Helly's theorem the dimension of Con B is n. Thus it is an n-simplex. If G_i is the complement of C_i, then the G_i are open and since every open set G_i is an F_σ set the G_i's contain closed sets H_i that satisfy the conditions of Kuratowsky–Knaster–Mazurkievicz theorem [Parthasarathy and Raghavan (1971)]. Thus we have $\bigcap_{i=0}^{n} H_i \neq \phi$, and that $\bigcap_{i=0}^{n} G_i \neq \phi$. That is, for some $x_0 \in \mathrm{Con}\, A$, $K(x_0, y) < c$ for all y. Similarly there is a $y_0 \in \mathrm{Con}\, B$, such that $K(x, y_0) > c$ for all $x \in \mathrm{Con}\, A$. Hence $c < K(x_0, y_0) < c$, a contradiction.

Case (ii). Let X be compact convex: Let $\sup_x \inf_y K < c < \inf_y \sup_x K$. There exists $B \subset Y$, B finite, such that for any $x \in X$, there is a $y \in B$ with $K(x,y) < c$. The contradiction can be established for K on $X \times \mathrm{Con}\, B \subset X \times Y$. □

Often, using the payoff, a topology can be defined on the pure strategy sets, whose properties guarantee a saddle point. Wald first initiated this approach [Wald (1950)].

Given arbitrary sets X, Y and given a bounded payoff K on $X \times Y$, we can topologize the spaces X, Y with topologies $\mathscr{T}_X, \mathscr{T}_Y$ where a base for $\mathscr{T}_X$ consists of sets of the type

$$S(x_0, \varepsilon) = [x: K(x,y) - K(x_0, y) < \varepsilon \text{ for all } y\}, \quad \varepsilon > 0,\ x_0 \in X.$$

Definition. The space X is *conditionally compact* in the topology $\mathscr{T}_X$ iff for any given $\varepsilon > 0$, there exists a finite set $\{x_1, x_2, \ldots, x_{n(\varepsilon)}\}$ such that $\bigcup_{i=1}^{n(\varepsilon)} S(x_i, \varepsilon) = X$.

The following is a sample theorem in the spirit of Wald.

Theorem. *Let* $K: X \times Y \to \mathbf{R}$. *Let* X *be conditionally compact in the topology* $\mathscr{T}_X$. *For any* $\delta > 0$ *and finite sets* $A \subset X$, $B \subset Y$ *let there exist* $\bar{x} \in X, \bar{y} \in Y$ *such that* $K(x, \bar{y}) \leqslant K(\bar{x}, y) + \delta$ *for all* $x \in A$, $y \in B$. *Then* $\sup_X \inf_Y K(x,y) = \inf_Y \sup_X K(x,y)$.

Proof. Since X is conditionally compact in the topology $\mathcal{T}$, for $\varepsilon > 0$

$$\inf_Y \sup_X K \leqslant \inf_Y \max_A K + \varepsilon$$

for a finite set $A \subset X$. If B is any finite subset of Y, then by assumption $\inf_Y \max_A K \leqslant \sup_X \min_B K$. Thus

$$\inf_Y \sup_X K \leqslant \inf_{\mathcal{B}} \sup_X \min_B K + \varepsilon,$$

where $\mathcal{B}$ is the collection of finite subsets of Y.

Since X is $\mathcal{T}_X$ conditionally compact, the right side of the above inequality is finite or $-\infty$. We are through if $v = \inf_{\mathcal{B}} \sup_X \min_B K \leqslant \sup_X \inf_Y K + 2\varepsilon$. For otherwise if $\sup_X \inf_Y K < v - 2\varepsilon$, we have

$$\inf_Y K(x, y) < v - 2\varepsilon, \quad \text{for all } x$$

and thus for any finite set A

$$\inf_Y k(x_i, y) < v - 2\varepsilon, \quad \text{for all } x_i \in A.$$

That is for each x_i, and $\varepsilon > 0$, there exists a y_i such that $K(x_i, y_i) \leqslant K(x_i, y_i) + \varepsilon \leqslant v - \varepsilon$ for each i. Thus

$$\sup_X \min_B K(x, y_i) \leqslant v - \varepsilon, \quad \text{where } B = \{y_i, i \in I \text{ with } |I| \text{ finite}\}.$$

That is $v \leqslant v - \varepsilon$, a contradiction. □

Continuous payoffs. While the above general minimax theorems guarantee ε-optimals with finite steps, in mixed extensions, certain subclasses of games admit optimal mixed strategies *with finite steps*. These subclasses depend on the nature of the cones generated by $K(x, \cdot)$ and $K(\cdot, y)$. Let X, Y be compact metric. Let $K: X \times Y \to \mathbf{R}$ be continuous. Since K is bounded, we can assume $K > 0$. Let $C(X)$, $C(Y)$ be the Branch space of continuous functions on X, Y, respectively. The functions $h_\alpha(x) = K(x, \alpha)$ generate a cone whose closure we denote by C. Let $E = C - C$. Let P be the cone of nonnegative functions in $C(X)$. Any *positive linear* operator $A: E \to C(X)$ maps C into P.

Theorem. *Let C have non-null interior in $\bar{E}$ or let A be an isometry and the image cone $A(C)$ have non-null relative interior in its closed linear span. Then player* II *(the minimizer) has an optimal strategy with finite spectrum. If C has nonnull interior, both players have finite step optimals. Further, in this case $K(x, y)$ is separable. That is $K(x, y) = \sum_i \sum_j a_{ij} r_i(x) s_j(y)$.*

Proof. For any Borel probability measure v on Y the map $\tau: v \to K^*(v)$ is a continuous map of Y^* into $C(X)$. Further, $B = \tau(Y^*)$ is compact in $C(X)$. In fact when A is an isometry, $S = \operatorname{con}(A(B), 0)$ is compact and the range cone $A(\bar{C}) = \bigcup_{n=1}^{\infty} nS$. By the Baire Category Theorem the compact set S has nonnull interior and hence *finite dimensional.* The isometry A preserves extreme points. The function $K(\cdot,\alpha)$ is mapped into $M(\cdot,\alpha)$ under the isometry and any finite step optimal v_0 for $M(x,y)$ is optimal for K. Separability of K follows from the finite dimensional nature of C when C has interior. □

Extreme optimals. Extreme optimal strategies are hard to characterize for infinite games except for some special cases. Let X, Y be compact metric with a continuous payoff $K(x,y)$. Then the following theorem extends Shapley–Snow theorem to certain infinite games.

Theorem. *Let $v \neq 0$ and let v_0 be optimal for player* II *with spectrum $\sigma(v_0) = Y$. Then an optimal strategy μ with spectrum $\sigma(\mu) = X_0$ is extreme iff $\{h_\alpha(x): K(x,\alpha), \alpha \in Y\}$ generates a dense linear manifold in $L_1(X_0, \mathscr{B}, \mu)$, where $\mathscr{B}$ is the class of Borel sets on X_0.*

The proof mimics Shapley–Snow arguments for finite games with the helpful hint that $L^\infty = L_1^*$.

Needless to say, the structure of optimal strategies can be quite complicated even for C^∞ payoffs K on the unit square. An example will settle what we want to convey [Glicksberg and Gross (1953)].

Example. Let μ_n, v_n be the nth moments of any arbitrary probability measures $\mu \neq v$ with infinite spectrum. Then the C^∞ kernel $K(x,y) = \sum_{n=0}^{\infty} (1/2^n)(x^n - \mu_n)(y^n - v_n)$ has μ, v as the unique optimal strategies.

Proof. The kernel K is analytic and for any optimal λ for player I we have $\sum_{n=0}^{\infty} (1/2^n)(\lambda_n - \mu_n)(y^n - v_n) \equiv 0$. Thus $\lambda_n = \mu_n$ for all $n \Rightarrow \lambda = \mu$. Similarly v is the unique optimal for the minimizer. □

More generally, given compact convex sets S, T of probability measures on the unit interval, one would like to know when they would be the precise set of optimal mixed strategies for a continuous game on the unit square. The following theorem is a partial answer to this question [Chin, Parthasarathy and Raghavan (1976)].

Theorem. *Let S, T be compact convex sets of probability measures on the unit interval with only finitely many extreme points given by $\{\mu^1, \mu^2, \ldots, \mu^p\}$ and $\{v^1, v^2, \ldots v^q\}$ respectively. Let the spectrum of at least one $\mu \in S$, $v \in T$ be the entire unit interval. Further for any $\varepsilon > 0$, let*

$$\max_{1 \le j \le q} \sup_{i \neq j} \frac{\mu^i(E_j)}{\mu^j(E_j)} < \varepsilon, \quad \max_{1 \le j \le q} \sup_{i \neq j} \frac{\mu^i(H_j)}{\mu^j(H_j)} < \varepsilon.$$

for some $\mu^r(E_r) > 0$, $v^s(H_s) > 0$, $r = 1, \ldots, p$, $s = 1, \ldots, q$. Then there exists a continuous payoff $K(x, y)$ on the unit square with S and T as the precise set of optimal strategies for the two players.

Proof. We will indicate the proof for the case $q = 1$. The general case is similar. One can find a set of indices $i_1, i_2, \ldots, i_{k-1}$, such that the matrix

$$\begin{bmatrix} 1 & \mu^1_{i_1} & \mu^1_{i_2} & \cdots & \mu^1_{i_{k-1}} \\ 1 & \mu^2_{i_1} & \mu^2_{i_2} & \cdots & \mu^2_{i_{k-1}} \\ 1 & \cdots & \cdots & \cdots & \cdots \\ 1 & \mu^k_{i_1} & \mu^k_{i_2} & \cdots & \mu^k_{i_{k-1}} \end{bmatrix}$$

is nonsingular where $\mu^j_n = \int_i^1 x^n \, d\mu^j$. Using this it can be shown that

$$K(x, y) = \sum_n \frac{[a_n + b_n x^{i_1} + \cdots h_n x^{i_{k-1}} - x^n][y^n - v^1_n]}{2^n \cdot M_n}$$

is a payoff precisely with optimal strategy sets S, T for some suitable constants M_n. □

Many infinite games of practical interest are often solved by intuitive guess works and ad hoc techniques that are special to the problem. For polynomial or separable games with payoff

$$K(x, y) = \sum_i \sum_j a_{ij} r_i(x) s_j(y),$$

$$K^*(\mu, v) = \sum_i \sum_j a_{ij} \left[\int r_i(x) \, d\mu(x) \right] \left[\int s_j(y) \, dv(y) \right] = \sum_i \sum_j a_{ij} u_i v_j,$$

where $u = (u_1, \ldots, u_m)$ and $v = (v_1, \ldots, v_n)$ are elements of the finite dimensional convex compact sets U, V which are the images of X^* and Y^* under the maps $\mu \to (\int r_1 \, d\mu, \int r_2 \, d\mu, \ldots, \int r_m \, d\mu)$; $v \to (\int s_1 \, dv, \ldots, \int s_n \, dv)$. Optimal μ^0, v^0 induce optimal points u^0, v^0 and the problem is reduced to looking for optimal u^0, v^0. The optimal u^0, v^0 are convex combinations of at most $\min(m, n)$ extreme points of U and V. Thus finite step optimals exist and can be further refined by knowing the dimensions of the sets U and V [Karlin (1959)]. Besides separable payoffs, certain other subclasses of continuous payoffs on the unit square admit finite step optimals. Notable among them are the convex and generalized convex payoffs and Polya-type payoffs.

Convex payoffs. Let X, Y be compact subsets of $\mathbf{R}^m$ and $\mathbf{R}^n$ respectively. Further, let Y be convex. A continuous payoff $K(x, y)$ is called *convex* if $K(\alpha, \cdot)$ is a convex function of y for each $\alpha \in x$. The following theorem of [Bohnenblust, Karlin and Shapley (1950)] is central to the study of such games.

Theorem. *Let $\{\phi_\alpha\}$ be a family of continuous convex functions on Y. If* $\sup_\alpha \phi_\alpha(y) > 0$ *for all $y \in Y$, then for some probability vector $\lambda = (\lambda_1, \lambda_2, \dots \lambda_{(n+1)})$, $\sum_{i=1}^{n+1} \lambda_i \phi_{\alpha_i}(y) > 0$ for all y, for some $n+1$ indices $\alpha_1, \alpha_2, \dots, \alpha_{n+1}$.*

Proof. The sets $K_\alpha = \{y : \phi_\alpha(y) \leqslant 0\}$ are compact convex and by assumption $\bigcap_\alpha K_\alpha = \phi$. By Helly's intersection theorem

$$\bigcap_{i=1}^{n+1} K_{\alpha_i} = \phi. \tag{10}$$

That is $\phi_{\alpha_i}(y) > 0$ for some $1 \leqslant i \leqslant n+1$ for each y. For any mixed strategy $p = (p_1, \dots, p_{n+1})$ on $\{\alpha_i, i = 1, \dots, n+1\}$ the kernel

$$\bar{K}(p, y) = \sum_{i=1}^{n+1} p_i \phi_{\alpha_i}(y) \tag{11}$$

admits a saddle point with value $\bar{v}$ by Fan's minimax theorem. Let (λ, y°) be such a saddle point. By (10) the value $\bar{v} > 0$. Thus

$$\sum_{i=1}^{n+1} \lambda_i \phi_{\alpha_i}(y) > 0, \quad \text{for all } y \tag{12}$$

□

As a consequence we have the following

Theorem. *For a continuous convex payoff $K(x, y)$ on $X \times Y$ as defined above, the minimizer has a pure optimal strategy. The maximizer has an optimal strategy using at most $n+1$ points in X.*

Proof. For any probability measure μ on X let $K^*(\mu, y) = \int_X K(x, y)\, \mathrm{d}\mu(x)$. By Ky Fan's minimax theorem K^* has a saddle point (μ°, y°) with value v. Given $\varepsilon > 0$, $\max_\alpha K(\alpha, y) - v + \varepsilon > 0$. From the above theorem $\sum_{i=1}^{n+1} \lambda_i^\varepsilon K(\alpha_i^\varepsilon, y) - v + \varepsilon > 0$ for all y. Since X is compact, by an elementary limiting argument an optimal λ with at most $n+1$ steps guarantee the value. Here y° is an optimal pure strategy for the minimizer. □

Weaker forms of convexity of payoffs still guarantee finite step optimals for games on the unit square $0 \leqslant x, y \leqslant 1$. The following theorem of Glicksberg (1953) is a sample (Karlin proved this theorem first for some special cases).

Theorem. *For a continuous payoff K on the unit square let $(\partial^l/\partial y^l) K(x, y) \geqslant 0$ for all x, for some power l. Then player* II *has an optimal strategy with at most $\frac{1}{2} l$ steps; (0, 1 are counted as $\frac{1}{2}$ steps). Player* I *has an optimal strategy with at most l steps.*

Proof. We can assume $v = 0$ and $(\partial^l/\partial y^l) K(x, y) > 0$. If I_y is the degenerate measure at y, we have $K^*(\mu, I_y) \geqslant 0$ for all optimal μ of the maximizer. By assumption $K^*(\mu, y) = 0$ has at most l roots counting multiplicities. Since interior roots have *even* multiplicities the spectrum of optimal strategies of the minimizer lie in this set. Hence the assertion for player II. We can construct a polynomial $p(y) \geqslant 0$ of degree $\leqslant l - 1$ with $K^*(\mu, y) - p(y)$ having exactly l roots. Let $y_1, y_2, \ldots, y_l$ be those roots. We can find for each $\bar{x}, K^*(\mu, y) - p(y) = K(\bar{x}, y) - p(\bar{x}, y)$, with the same roots. Next we can show $K(x, y) - p(x, y)$ has the same sign for all x, y. Indeed $p(x, y)$ is a polynomial game with value 0 whose optimal strategy for the maximizer serves as an optimal strategy for the original game $K(x, y)$. □

Bell.shaped kernels. Let $K: \mathbf{R}^2 \to \mathbf{R}$. The kernel K is called *regular bell shaped* if $K(x, y) = \varphi(x - y)$ where φ satisfies (i) φ is continuous on R; (ii) for $x_1 < x_2 < \cdots < x_n$; $y_1 < y_2 < \cdots < y_n$, det $\| \varphi(x_i - y_j) \|$ is nonnegative; (iii) For each $x_1 < \cdots < x_n$ we can find $y_1 < y_2 < \cdots < y_n$ such that det $\| \varphi(x_i - y_j) \| > 0$; (iv) $\int_R \varphi(u)\,\mathrm{d}u < \infty$.

Theorem. *Let K be bell shaped on the unit square and let φ be analytic. Then the value v is positive and both players have optimal strategies with finitely many steps.*

Proof. Let ν be optimal for the minimizer. If the spectrum $\sigma(\nu)$ is an infinite set then

$$\int_0^1 K(x - y)\,\mathrm{d}\nu(y) \equiv v$$

by the analyticity of the left-hand side. But for any φ satisfying (i), (ii), (iii) and (iv) $\int_0^1 K(x - y)\,\mathrm{d}\nu(y) \to 0$ as $x \to \infty$. This contradicts $v > 0$. □

Further refinements are possible giving bounds for the number of steps in such finite optimal strategies [Karlin (1959)].

References

Aumann, R.J. and M. Maschler. (1972) 'Some thoughts on the minimax principle' *Management Science*, **18**: 54–63.

Blackwell, D and G.A. Girshick. (1954) *Theory of Games and Statistical Decisions*. New York: John Wiley.

Bohnenblust, H., S. Karlin and L.S. Shapley. (1950) 'Solutions of discrete two person games'. In *Contributions to the theory of games*. Vol. I, eds. Kuhn, H.W. and A.W.Tucker. Annals of Mathematics Studies, **24**. Princeton, NJ: Princeton University Press, pp. 51–72.

Bohnenblust, H., S. Karlin and L.S. Shapley. (1950) 'Games with continuous convex payoff'. In *Contributions to the thoery of games*. Vol I, eds. Kuhn, H.W. and A.W. Tucker. Annals of Mathematics Studies, **24**. Princeton, NJ: Princeton University Press, pp. 181–192.

Brown, George W. (1951) 'Interative solutions of games by fictitious play'. In *Activity Analysis of Production and Allocation*, Cowles commission Monograph 13. Ed. T.C. Koopmans. New York: John Wiley. London: Chapman Hall, pp. 374–376.

Brown, George W. and J. von Neumann (1950) 'Solutions of games by differential equations'. In *Contributions to the theory of games*. Vol. I, eds. Kuhn, H.W. and A.W. Tucker. Annals of Mathematics Studies, **24**. Princeton, NJ: Princeton University Press, pp. 73–79.

Chin, H., T. Parthasarathy and T.E.S. Raghavan. (1976) 'Optimal strategy sets for continuous two person games', *Sankhya*, Ser. A **38**: 92–98.
Dantzig, George B. (1951) 'A proof of the equivalence of the programming problem and the game problem'. In *Activity analysis of production and allocation*, Cowles Commission Monograph 13. Ed. T.C. Koopmans. New York: John Wiley. London: Chapman and Hall, pp. 333–335.
Dresher, M. (1962) *Games of strategy*. Englewood Cliffs, NJ: Prentice Hall.
Fan, Ky. (1952) 'Fixed point and minimax theorems in locally convex topological linear spaces', *Proceedings of the National Academy of Sciences, U.S.A.*, **38**: 121–126.
Flourier, J.B. (1890) 'Second Extrait'. In: *Oeuvres*. Ed., G. Darboux, Paris: Gauthiers Villars, pp. 325–328 [(English translation D.A. Kohler (1973)].
Gale, D. and S. Sherman. (1950) 'Solutions of finite two person games'. In *Contributions to the theory of games*. Vol. I, eds. Kuhn, H.W. and A.W. Tucker. Annals of Mathematics Studies, **24**. Princeton, NJ: Princeton University Press, pp. 37–49.
Glicksberg, I. (1952) 'A further generalization of Kakutani's fixed point theorem', *Proceedings of the American Mathematical Society*, **3**: 170–174.
Glicksberg, I., and O. Gross. (1953) 'Notes on games over the square'. In *Contributions to the theory of games*. Vol. II, eds. Kuhn, H.W. and A.W. Tucker. Annals of Mathematics Studies, **28** Princeton, NJ: Princeton University Press, pp. 173–184.
Glicksberg, I. (1953) 'A derivative test for finite solutions of games', *Proceedings of the American Mathematical Society*, **4**: 595–597.
Kakutani, S. (1941) 'A generalization of Brouwer's fixed point theorem', *Duke Mathematical Journal*, **8**: 457–459.
Kantarovich, L.V. (1939) 'Mathematical methods in the organization and production'. [English translation: *Management science* (1960), **6**: 366–422.]
Kaplansky, I. (1945) 'A contribution to von Neumann's theory of games', *Annals of Mathematics*, **46**: 474–479.
Karlin, S. (1959) *Mathematical Methods and Theory in Games, Programming and Economics*. Volumes I and II. New York: Addison Wesley.
Koopmans, T.C. (1951) 'Analysis of production as an efficient combination of activities'. In *Activity analysis of production and allocation*, Cowles Commission Monograph 13. Ed. T.C. Koopmans. New York: John Wiley. London: Chapman and Hall, pp. 33–97.
Kohler, D.A. (1973) 'Translation of a report by Fourier on his work on linear inequalities', *Opsearch*, **10**: 38–42.
Kuhn, H.W. (1953) 'Extensive games and the problem of information'. In *Contributions to the theory of games*. Vol. II, eds. Kuhn, H.W. and A.W. Tucker. Annals of Mathematics Studies, **28** Princeton, NJ: Princeton University Press, pp. 193–216.
Loomis, I.H. (1946) 'On a theorem of von Neumann', *Proceedings of the National Academy of Sciences, U.S.A*, **32**: 213–215.
Mertens, J.W. (1986) 'The minimax theorem for U.S.C.–L.S.C. payoff functions'. *International Journal of Game Theory*, **15**: 237–265.
Nash, J.F. (1950) 'Equilibrium points in *n*-person games', *Proceedings of the National Academy of Sciences, U.S.A.*, **36**: 48–49.
Nikaido, H. (1954) 'On von Neumann's minimax theorem', *Pacific Journal of Mathematics*, **4**: 65–72.
Owen, G. (1967) 'An elementary proof of the minimax theorem', *Management Science*, **13**: 765.
Parthasarathy, T and T.E.S. Raghavan (1971) *Some Topic in Two Person Games*. New York: American Elsevier.
Raghavan, T.E.S. (1973) 'Some geometric consequences of a game theoretic results', *Journal of Mathematical Analysis and Applications*, **43**: 26–30.
Raghavan, T.E.S. (1978) 'Completely mixed games and M-matrices', *Journal of Linear Algebra and Applications*, **21**: 35–45.
Raghavan, T.E.S. (1979) 'Some remarks on matrix games and nonnegative matrices', *SIAM Journal of Applied Mathematics*, **36**(1): 83–85.
Raghavan, T.E.S. and J.A. Filar (1991) 'Algorithms for stochastic games-A survey', *Zeitschrift für Operations Research*, **35**: 437–472.
Robinson, J. (1951) 'An iterative method of solving a game', *Annals of Mathematics*, **54**: 296–301.
Shapley, L.S. and R.N. Snow (1950) 'Basic Solutions of Discrete Games'. In *Contributions to the theory of games*. Vol. I, eds. Kuhn, H.W. and A.W. Tucker. Annals of Mathematics Studies, **24**. Princeton, NJ: Princeton University Press, pp. 27–35.

Sion, M. (1958) 'On general minimax theorem', *Pacific Journal of Mathematics*, **8**: 171–176.

Ville, J. (1938) 'Note sur la théorie géneralé des jeux oú intervient l'habilité des joueurs'. In *Traité du calcul des probabilités et de ses applicatiońs.* Ed. by Emile Borel. Paris: Gauthier Villars, pp. 105–113.

von Neumann, J. (1928) 'Zur Theorie der Gesellschaftspiele', *Mathematische Annalen*, **100**: 295–320.

von Neumann, J. and O. Morgenstern. (1944) *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

von Neumann, J. (1953) 'Communication on the Borel notes', *Econometrica*, **21**: 124–127.

Wald, A. (1950) *Statistical Decision Functions*. New York: John Wiley.

Weyl, H. (1950) 'Elementary proof of a minimax theorem due to von Neumann'. In *Contributions to the theory of games*. Vol. I, eds. Kuhn, H.W. and A.W. Tucker. Annals of Mathematics Studies, **24**. Peinceton, NJ: Princeton University Press, pp. 19–25.