Appendix 1: Utility Theory

Much of the theory presented is based on utility theory at a fundamental level. This theory gives a justification for our assumptions (1) that the payoff functions are numerical valued and (2) that a randomized payoff may be replaced by its expectation. There are many expostions on this subject at various levels of sophistication. The basic theory was developed in the book of Von Neumann and Morgenstern (1947). Further developments are given in Savage (1954), Blackwell and Girshick (1954) and Luce and Raiffa (1957). More recent descriptions may be found in Owen (1982) and Shubik (1984), and a more complete exposition of the theory may be found in Fishburn (1988). Here is a brief description of the basics of linear utility theory.

The method a 'rational' person uses in choosing between two alternative actions, a_1 and a_2, is quite complex. In general situations, the payoff for choosing an action is not necessarily numerical, but may instead represent complex entities such as "you receive a ticket to a ball game tomorrow when there is a good chance of rain and your raincoat is torn" or "you lose five dollars on a bet to someone you dislike and the chances are that he is going to rub it in". Such entities we refer to as *payoffs* or *prizes*. The 'rational' person in choosing between two actions evaluates the value of the various payoffs and balances it with the probabilities with which he thinks the payoffs will occur. He may do, and usually does, such an evaluation subconsciously. We give here a mathematical model by which such choices among actions are made. This model is based on the notion that a 'rational' person can express his preferences among payoffs in a method consistent with certain axioms. The basic conclusion is that the 'value' to him of a payoff may be expressed as a numerical function, called a *utility*, defined on the set of payoffs, and that the preference between lotteries giving him a probability distribution over the payoffs is based only on the expected value of the utility of the lottery.

Let $\mathcal{P}$ denote the set of payoffs of the game. We use P, P_1, P_2, and so on to denote payoffs (that is, elements of $\mathcal{P}$).

Definition. *A* **preference relation on** $\mathcal{P}$, *or simply preference on* $\mathcal{P}$, *is a (weak) linear ordering,* $\preceq$, *on* $\mathcal{P}$; *that is,*

(a) (linearity) if P_1 *and* P_2 *are in* $\mathcal{P}$, *then either* $P_1 \preceq P_2$ *or* $P_2 \preceq P_1$ *(or both), and*

(b) (transitivity) if P_1, P_2 *and* P_3 *are in* $\mathcal{P}$, *and if* $P_1 \preceq P_2$ *and* $P_2 \preceq P_3$, *then* $P_1 \preceq P_3$.

If $P_1 \preceq P_2$ *and* $P_2 \preceq P_1$, *then we say* P_1 *and* P_2 *are* **equivalent** *and write* $P_1 \simeq P_2$.

We assume that our 'rational' being can express his preferences over the set $\mathcal{P}$ in a way that is consistent with some preference relation. If $P_1 \preceq P_2$ and $P_1 \not\simeq P_2$, we say that our rational person *prefers* P_2 to P_1 and write $P_1 \prec P_2$. If $P_1 \simeq P_2$, we say that he is *indifferent* between P_1 and P_2. The statement $P_1 \preceq P_2$ means either he either prefers P_2 to P_1 or he is indifferent between them.

Unfortunately, just knowing that a person prefers P_2 to P_1, gives us no indication of how much more he prefers P_2 to P_1. In fact, the question does not make sense until a third point of comparison is introduced. We could, for example, ask him to compare P_2 with the payoff of P_1 plus \$100 in order to get some comparison of how much more he prefers P_2 to P_1 in terms of money. We would like to go farther and express all his preferences in some numerical form. To do this however requires that we ask him to express his preferences on the space of all lotteries over the payoffs.

Definition. *A* **lottery** *is a finite probability distribution over the set $\mathcal{P}$ of payoffs. We denote the set of lotteries by $\mathcal{P}^*$.*

(A finite probability distribution is one that gives positive probability to only a finite number of points.)

If P_1, P_2 and P_3 are payoffs, the probability distribution, p, that chooses P_1 with probability 1/2, P_2 with probability 1/4, and P_3 with probability 1/4 is a lottery. We use lower case letters, p, p_1, p_2 to denote elements of $\mathcal{P}^*$. Note that the lottery p that gives probability 1 to a fixed payoff P may be identified with P, since receiving payoff P is the same as receiving payoff P with probability 1. With this identification, we may consider $\mathcal{P}$ to be a subset of $\mathcal{P}^*$.

We note that if p_1 and p_2 are lotteries and $0 \le \lambda \le 1$, then $\lambda p_1 + (1-\lambda)p_2$ is also a lottery. It is that lottery that first tosses a coin with probability λ of heads; if heads comes up, then it uses p_1 to choose an element of $\mathcal{P}$ and if tails comes up, it uses p_2. Thus $\lambda p_1 + (1-\lambda)p_2$ is an element of $\mathcal{P}^*$. Mathematically, a lottery of lotteries is just another lottery.

We assume now that our 'rational' person has a preference relation not only over $\mathcal{P}$ but over $\mathcal{P}^*$ as well. One very simple way of setting up a preference over $\mathcal{P}^*$ is through a utility function.

Definition. *A* **utility function** *is a real-valued function defined over $\mathcal{P}$.*

Given a utility function, $u(P)$, we extend the domain of u to the set $\mathcal{P}^*$ of all lotteries by defining $u(p)$ for $p \in \mathcal{P}^*$ to be the expected utility: i.e. if $p \in \mathcal{P}^*$ is the lottery that chooses $P_1, P_2, \ldots, P_k$ with respective probabilities $\lambda_1, \lambda_2, \ldots, \lambda_k$, where $\lambda_i \ge 0$ and $\sum \lambda_i = 1$, then

$$u(p) = \sum_{i=1}^{k} \lambda_i u(P_i) \tag{1}$$

is the expected utility of the payoff for lottery p. Thus given a utility u, a simple preference over $\mathcal{P}^*$ is given by

$$p_1 \preceq p_2 \qquad \text{if and only if} \qquad u(p_1) \le u(p_2), \tag{2}$$

i.e. that lottery with the higher expected utility is preferred.

The basic question is, can we go the other way around? Given an arbitrary preference, $\preceq$ on $\mathcal{P}^*$, does there exist a utility u defined on $\mathcal{P}$ such that (2) holds? The answer is no in general, but under the following two axioms on the preference relation, the answer is yes!

A1. *If p_1, p_2 and q are in $\mathcal{P}^*$, and $0 < \lambda \le 1$, then*

$$p_1 \preceq p_2 \qquad \textit{if, and only if} \qquad \lambda p_1 + (1-\lambda)q \preceq \lambda p_2 + (1-\lambda)q. \tag{3}$$

A2. *For arbitrary p_1, p_2 and q in $\mathcal{P}^*$,*

$$p_1 \prec p_2 \qquad \textit{implies} \qquad \textit{there exists a } \lambda > 0 \textit{ such that } p_1 \prec \lambda q + (1-\lambda)p_2 \tag{4}$$

and similarly,

$$p_1 \prec p_2 \qquad \textit{implies} \qquad \textit{there exists a } \lambda > 0 \textit{ such that } \lambda q + (1-\lambda)p_1 \prec p_2 \tag{5}$$

Axiom A1 is easy to justify. Consider a coin with probability $\lambda > 0$ of coming up heads. If the coin comes up tails you receive q. If it comes up heads you are asked to choose between p_1 and p_2. If you prefer p_2, you would naturally choose p_2. This axiom states that if you had to decide between p_1 and p_2 before learning the outcome of the toss, you would make the same decision. A minor objection to this axiom is that we might be indifferent between $\lambda p_1 + (1-\lambda)q$ and $\lambda p_2 + (1-\lambda)q$ if λ is sufficiently small, say $\lambda = 10^{-100}$, even though we prefer p_1 to p_2. Another objection comes from the person who dislikes gambles with random payoffs. He might prefer a p_2 that gives him \$2 outright to a gamble, p_1, giving him \$1 with probability 1/2 and \$3.10 with probability 1/2. But if q is \$5 for sure and $\lambda = 1/2$, he might prefer $\lambda p_1 + (1-\lambda)q$ to $\lambda p_2 + (1-\lambda)q$ on the basis of larger expected monetary reward, because the payoff is random in either case.

Axiom A2 is more debatable. It is called the continuity axiom. Condition (4) says that if $p_1 \prec \lambda q + (1-\lambda)p_2$ when $\lambda = 0$, then it holds for λ sufficiently close to 0. It might not be true if q is some really horrible event like death. It is safe to assume that for most people, $p_2 =$ \$100 is strictly preferred to $p_1 =$ \$1, which is strictly preferred to $q =$ death. Yet, would you ever prefer a gamble giving you death with probability λ and \$100 with probability $1-\lambda$, for some positive λ, to receiving \$1 outright? If not, then condition (4) is violated. However, people do not behave as if avoiding death is an overriding concern. They will drive on the freeway to get to the theater or the ballpark for entertainment, even though they have increased the probability of death (by a very small amount) by doing so. At any rate, Axiom A2 implies that there is no payoff infinitely less desirable or infinitely more desirable than any other payoff.

Theorem 1. *If a preference relation, $\preceq$, on $\mathcal{P}^*$ satisfies A1 and A2, then there exists a utility, u, defined on $\mathcal{P}$ that satisfies (2). Furthermore, u is uniquely determined up to change of location and scale.*

If a utility $u(P)$ satisfies (2), then for arbitrary real numbers a and $b > 0$, the utility $\hat{u}(P) = a + bu(P)$ also satisfies (2). Thus the uniqueness of u up to change of location and scale the strongest uniqueness that can be obtained.

For a proof see Blackwell and Girshick and for extensions see Fishburn. The proof is constructive. We may choose p and q in $\mathcal{P}^*$ arbitrarily, say $p \prec q$, and define $u(p) = 0$ and

$u(q) = 1$. This merely fixes the location and scale. Then for any p' such that $p \prec p' \prec q$, we may define $u(p') = \text{glb}\{\lambda : p' \prec \lambda q + (1-\lambda)p\}$. For p' not preferenced between p and q, $u(p')$ may be defined by extrapolation. For example, if $p \prec q \prec p'$, we first find $\tau = \text{glb}\{\lambda : q \prec \lambda p' + (1-\lambda)p\}$, and then define $u(p') = 1/\tau$. The resulting function, u, satisfies (2).

One may conclude from Theorem 1 that if a person has a preference relation on $\mathcal{P}^*$ that satisfies Axioms 1 and 2, then that person is *acting* as if his preferences were based on a utility defined on $\mathcal{P}$ and that of two lotteries in $\mathcal{P}^*$ he prefers the one with the larger expected utility. Ordinarily, the person does not really think in terms of a utility function and is unaware of its existence. However, a utility function giving rise to his preferences may be approximated by eliciting his preferences from a series of questions.

Countable Lotteries and Bounded Utility. It is sometimes desirable to extend the notion of a lottery to countable probability distributions, i.e. distributions giving all their weight to a countable number of points. If this is done, it must generally be assumed that utility is bounded. The reason for this is that if equation (1) is to be satisfied for countable lotteries and if u is unbounded, then there will exist lotteries p such that $u(p) = \pm\infty$. This may be seen as follows.

Suppose we have a utility function $u(P)$ on $\mathcal{P}$, and suppose that for a lottery, p, that chooses $P_1, P_2, \ldots$ with respective probabilities $\lambda_1, \lambda_2, \ldots$ such that $\sum_{n=1}^{\infty} \lambda_n = 1$, the extension of u to countable lotteries satisfies

$$u(p) = \sum_{n=1}^{\infty} \lambda_n u(P_n). \tag{1'}$$

If $u(P)$ is unbounded, say unbounded above, then we could find a sequence, $P_1, P_2, \ldots$, such that $u(P_n) \ge 2^n$. Then if we consider the lottery, p, that chooses P_n with probability 2^{-n} for $n = 1, 2, \ldots$, we would have

$$u(p) = \sum_{n=1}^{\infty} 2^{-n} u(P_n) \ge \sum_{n=1}^{\infty} 2^{-n} 2^n = \infty.$$

Then p would be a lottery that is infinitely more desirable than P_1, say, contradicting Assumption A2.

Since the extension of utility to countable lotteries seems innocuous, it is generally considered that utility indeed should be considered to be bounded.

Exercises. 1. Does every preference given by a utility as in (1) satisfy A1 and A2?

2. Take $\mathcal{P} = \{P_1, P_2\}$, and give an example of a preference on $\mathcal{P}^*$ satisfying A2 but not A1.

3. Take $\mathcal{P} = \{P_1, P_2, P_3\}$, and give an example of a preference on $\mathcal{P}^*$ satisfying A1 but not A2.

Appendix 2: Owen's Proof of the Minimax Theorem

We now set out to prove the minimax theorem for finite games, that every matrix game has a value. The following non-constructive proof is due essentially to Guillermo Owen.

Theorem. *Every finite matrix game has a value.*

Proof. We will show that if a matrix does not have a value, then there is a submatrix (of smaller size) that does not have a value. Then by repeated application of this, we can reduce the size of the matrix without a value down to one row or one column, where we know the game has a value.

Let A be an $m \times n$ matrix, and let $\underline{V}(A)$ and $\overline{V}(A)$ denote the lower and upper values of the game with matrix A. Let $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$ denote minimax strategies for players I and II respectively, in the game with matrix A. Then

$$\sum_{j=1}^{n} a_{ij}\hat{q}_j \leq \overline{V}(A) \qquad \text{for } i = 1, \ldots, m \tag{1}$$

$$\sum_{i=1}^{m} \hat{p}_i a_{ij} \geq \underline{V}(A) \qquad \text{for } j = 1, \ldots, n \tag{2}$$

If equality holds for all the inequalities in (1) and (2), then

$$\underline{V}(A) = \sum_{j=1}^{n} \underline{V}(A)\hat{q}_j = \sum_j \sum_i \hat{p}_i a_{ij} \hat{q}_j = \sum_i \hat{p}_i \overline{V}(A) = \overline{V}(A) \tag{3}$$

so that the game A has a value.

Assume A does not have a value. Then at least one of the inequalities in (1) and (2) is strict. Without loss of generality, assume this occurs in (2) when $j = n$:

$$\sum_{i=1}^{m} \hat{p}_i a_{in} > \underline{V}(A). \tag{3}$$

This implies that $n \geq 2$. Define A' to be the submatrix of A obtained by deleting the nth column. We will show that the following two conditions are satisfied.

$$\underline{V}(A') \leq \underline{V}(A) \tag{4}$$

$$\overline{V}(A') \geq \overline{V}(A) \tag{5}$$

From this we may conclude that A' does not have a value because

$$\underline{V}(A') \leq \underline{V}(A) < \overline{V}(A) \leq \overline{V}(A').$$

To see (5), let $\hat{\mathbf{p}}'$ and $\hat{\mathbf{q}}'$ be the players' minimax strategies for A', and let $\mathbf{q}^0$ denote the vector $\mathbf{q}^0 = (\hat{q}'_1, \ldots, \hat{q}'_{n-1}, 0)$. Then,

$$\overline{V}(A') = \max_i \sum_{j=1}^{n-1} a_{ij}\hat{q}'_j = \max_i \sum_{j=1}^{n-1} a_{ij}q^0_j \geq \min_q \max_i \sum_{j=1}^{n} a_{ij}q_j = \overline{V}(A) \tag{6}$$

(Intuitively, it cannot help Player II when playing A to be restricted to strategies that do not use column n.)

To prove (4), let ϵ be a small positive number and consider the strategy

$$\mathbf{p}^\epsilon = (1-\epsilon)\hat{\mathbf{p}} + \epsilon\hat{\mathbf{p}}'.$$

Suppose (4) is not true, $\underline{V}(A') > \underline{V}(A)$. We will show that $\sum_{j=1}^{n} p_i^\epsilon a_{ij} > \underline{V}(A)$ for all j provided ϵ is sufficiently small, contradicting the definition of $\underline{V}(A)$ as the best Player I can guarantee. For $j = 1, \ldots, n-1$,

$$\sum_{i=1}^{m} p_i^\epsilon a_{ij} = (1-\epsilon)\sum_{i=1}^{m} \hat{p}_i a_{ij} + \epsilon \sum_{i=1}^{m} \hat{p}'_i a_{ij} \geq (1-\epsilon)\underline{V}(A) + \epsilon\underline{V}(A') > \underline{V}(A)$$

and for $j = n$,

$$\sum_{i=1}^{m} p_i^\epsilon a_{in} = (1-\epsilon)\sum_{i=1}^{m} \hat{p}_i a_{ij} + \epsilon \sum_{i=1}^{m} \hat{p}'_i a_{in} = \sum_{i=1}^{m} \hat{p}_i a_{in} + \epsilon \sum_{i=m}^{m} (\hat{p}'_i - \hat{p}_i) a_{in}.$$

Now since $\sum_{i=1}^{m} \hat{p}_i a_{in}$ is strictly greater than $\underline{V}(A)$, we can make $\sum_{i=1}^{m} p_i^\epsilon a_{in}$ also strictly greater than $\underline{V}(A)$ by choosing ϵ sufficiently small.

Appendix 3: Contraction Maps and Fixed Points

Definition 1. *A metric space, (X, d), consists of a nonempty set X of points together with a function d from $X \times X$ to the reals satisfying the four following properties for all x, y and z in X:*
(1) $d(x, x) = 0$ for all x.
(2) $d(x, y) > 0$ for all $x \neq y$.
(3) $d(x, y) = d(y, x)$ for all x and y.
(4) $d(x, y) \leq d(x, z) + d(z, y)$

Definition 2. *A metric space, (X, d), is said to be* **complete** *if every Cauchy sequence in X converges to a point in X. In other words, for any sequence $x_n \in X$ for which $\max_{n>m} d(x_m, x_n) \to 0$ as $m \to \infty$, there exists a point $x^* \in X$ such that $d(x_n, x^*) \to 0$ as $n \to \infty$.*

Examples. Euclidean space in N-dimensions, R^N, with points $\boldsymbol{x} = (x_1, \ldots, x_N)$, is an example of a complete metric space under any of the following metrics.
1. Euclidean distance, $\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}$.
2. L_1 distance: $\|\boldsymbol{x}\| = |x_1| + |x_2| + \cdots + |x_N|$.
3. Sup norm: $\|\boldsymbol{x}\| = \max\{|x_1|, |x_2|, \ldots, |x_N|\}$.

These metrics are equivalent in the sense that if for one of them $\|\boldsymbol{x}_n\| \to 0$ for a sequence $\boldsymbol{x}_n \in X$, then $\|\boldsymbol{x}_n\| \to 0$ for the others also. In what follows, we take a fixed metric space (X, d).

Definition 3. *A map $T : X \to X$ is a contraction map if there is a positive constant $c < 1$, called the contraction factor, such that*

$$d(Tx, Ty) \leq c\, d(x, y)$$

for all $x \in X$ and $y \in X$.

Definition 4. *A point, $x \in X$, is said to be a fixed point of a map $T : X \to X$, if $Tx = x$.*

Contraction Mapping Theorem. *A contraction map, T, on a complete metric space (X, d) has a unique fixed point, x_0. Moreover, for any $y \in X$, $d(T^n y, x_0) \to 0$ as $n \to \infty$.*

Proof. Let $c < 1$ denote the contraction factor of T. Then, $d(T^2 y, Ty) \leq c\, d(Ty, y)$, and inductively, $d(T^{n+1} y, T^n y) \leq c^n d(Ty, y)$ for every n. The inequality,

$$\begin{aligned} d(T^{n+1} y, y) &\leq d(T^{n+1} y, T^n y) + \cdots + d(Ty, y) \\ &\leq (c^n + \cdots + 1) d(Ty, y) \qquad (*) \\ &\leq d(Ty, y)/(1 - c), \end{aligned}$$

for all $n > 0$, implies that for all $n > m$,

$$d(T^{n+1}y, T^m y) \le \frac{d(T^{m+1}y, T^m y)}{1-c} \le \frac{c^m d(Ty, y)}{1-c}.$$

This shows that $\{T^n y\}$ is a Cauchy sequence. Hence, there is a point $x_0 \in R^d$ such that $T^n y \to x_0$ as $n \to \infty$. That x_0 is a fixed point follows from

$$\begin{aligned} d(Tx_0, x_0) &\le d(Tx_0, T^n y) + d(T^n y, x_0) \\ &\le c\, d(T^{n-1}y, x_0) + d(T^n y, x_0) \to 0, \end{aligned}$$

as $n \to \infty$, so that $d(Tx_0, x_0) = 0$. The fixed point must be unique because if z is another fixed point, then $d(x_0, z) = d(Tx_0, Tz) \le c\, d(x_0, z)$ shows that $d(x_0, z) = 0$. ∎

Corollary 1. *For any $y \in X$, $d(T^n y, x_0) \le c^n d(y, x_0)$, where x_0 is the fixed point of T.*

Proof. $d(T^n y, x_0) = d(T^n y, T^n x_0) \le c^n d(y, x_0)$. ∎

Corollary 2. *For any $y \in X$, $d(x_0, Ty) \le \frac{c}{1-c} d(Ty, y)$, where x_0 is the fixed point of T.*

Proof. From (*), letting $n \to \infty$, we have $d(x_0, y) \le d(Ty, y)/(1-c)$. Therefore, $d(x_0, Ty) \le c\, d(x_0, y) \le c\, d(Ty, y)/(1-c)$. ∎

Corollary 1 gives a bound on the rate of convergence of $T^n y$ to x_0. Corollary 2 gives an upper bound on the distance of Ty to x_0 based on the distance from y to Ty.

Proof of Theorem 1 of Section II.6. This proof is based on the Contraction Mapping Theorem and the following simple lemma.

Lemma. *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two matrices of the same dimensions, then*

$$|\mathrm{Val}(\boldsymbol{A}) - \mathrm{Val}(\boldsymbol{B})| \le \max_{i,j} |a_{ij} - b_{ij}|.$$

Proof. Let $z = \max_{i,j} |a_{ij} - b_{ij}|$. Then $\mathrm{Val}(\boldsymbol{A}) + z = \mathrm{Val}(\boldsymbol{A} + z) \ge \mathrm{Val}(\boldsymbol{B})$, because $b_{ij} \le a_{ij} + z$ for all i and j. Similarly, $\mathrm{Val}(\boldsymbol{B}) + z \ge \mathrm{Val}(\boldsymbol{A})$, completing the proof.

Proof of Theorem II.6.1. Let T be the map from R^N to R^N defined by $T\boldsymbol{x} = \boldsymbol{y}$ where $y_k = \mathrm{Val}\big(\boldsymbol{A}^{(k)}(\boldsymbol{x})\big)$ for $k = 1, \ldots, N$, where $\boldsymbol{A}^{(k)}(\boldsymbol{x}))$ is the matrix of equation (10) of Section II.6. We show that under the sup norm metric, $\|\boldsymbol{x}\| = \max_k\{|x_k| : k = 1, \ldots, N\}$, T is a contraction map with contraction factor $c = 1 - s$, where s is the smallest of the stopping probabilities, assumed to be positive by (6) of Section II.6. Using the Lemma,

$$
\begin{aligned}
\|T\boldsymbol{x} - T\boldsymbol{y}\| &= \max_k |\mathrm{Val}\big(a_{ij}^{(k)} + \sum_{\ell=1}^N P_{ij}^{(k)}(\ell)x_\ell\big) - \mathrm{Val}\big(a_{ij}^{(k)} + \sum_{\ell=1}^N P_{ij}^{(k)}(\ell)y_\ell\big)| \\
&\le \max_{k,i,j} |\sum_{\ell=1}^N P_{ij}^{(k)}(\ell)x_\ell - \sum_{\ell=1}^N P_{ij}^{(k)}(\ell)y_\ell| \\
&\le \max_{k,i,j} \sum_{\ell=1}^N P_{ij}^{(k)}(\ell)|x_\ell - y_\ell| \\
&\le (\max_{k,i,j} \sum_{\ell=1}^N P_{ij}^{(k)}(\ell))\|\boldsymbol{x} - \boldsymbol{y}\| \\
&= (1-s)\|\boldsymbol{x} - \boldsymbol{y}\|
\end{aligned}
$$

Since $c = 1 - s$ is less than 1, the Contraction Mapping Theorem implies there is a unique vector $\boldsymbol{v}$ such that $T\boldsymbol{v} = \boldsymbol{v}$. But this is exactly equation (9) of Theorem 1.

We must now show that the suggested stationary optimal strategies guarantee the value. Let $\boldsymbol{x}^* = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N)$ denote the suggested stationary strategy for Player I where $\boldsymbol{p}_k$ is optimal for him for the matrix $\boldsymbol{A}^{(k)}(\boldsymbol{v})$. We must show that in the stochastic game that starts at state k, this gives an expected return of at least $v(k)$ no matter what Player II does. Let n be arbitrary and consider the game up to stage n. If at stage n, play is forced to stop and the payoff at that stage is $a_{ij}^{(h)} + \sum_{\ell=1}^N P_{ij}^{(h)}(\ell)v(l)$ rather than just $a_{ij}^{(h)}$, assuming the state at that time were h, then $\boldsymbol{x}^*$ would be optimal for this multistage game and the value would be $v(k)$. Hence in the infinite stage game, Player I's expected payoff for the first n stages is at least

$$
v(k) - (1-s)^n \max_{h,i,j} \sum_{\ell=1}^N P_{ij}^{(h)}(\ell)v(\ell) \ge v(k) - (1-s)^n \max_\ell v(\ell).
$$

and for the remaining stages is bounded below by $(1-s)^n M/s$, as in (7). Therefore, Player I's expected payoff is at least

$$
v(k) - (1-s)^n \max_\ell v(\ell) - (1-s)^n M/s.
$$

Since this is true for all n and n is arbitrary, Player I's expected payoff is at least $v(k)$. By symmetry, Player II's expected loss is at most $v(k)$. This is true for all k, proving the theorem.

Appendix 4: Existence of Equilibria in Finite Games

We give a proof of Nash's Theorem based on the celebrated Fixed Point Theorem of L. E. J. Brouwer. Given a set C and a mapping T of C into itself, a point $\boldsymbol{z} \in C$ is said to be a fixed point of T, if $T(\boldsymbol{z}) = \boldsymbol{z}$.

The Brouwer Fixed Point Theorem. *Let C be a nonempty, compact, convex set in a finite dimensional Euclidean space, and let T be a continuous map of C into itself. Then there exists a point $\boldsymbol{z} \in C$ such that $T(\boldsymbol{z}) = \boldsymbol{z}$.*

The proof is not easy. You might look at the paper of Y. Takeuchi and T. Suzuki, "An easily verifiable proof of the Brouwer fixed point theorem", (2011) *arXiv:1109.4604 [math.HO]*.

Now consider a finite n-person game with the notation of Section III.2.1. The pure strategy sets are denoted by $X_1, \ldots, X_n$, with X_k consisting of $m_k \ge 1$ elements, say $X_k = \{1, \ldots, m_k\}$. The space of mixed strategies of Player k is given by X_k^*,

$$X_k^* = \{\boldsymbol{p}_k = (p_{k,1}, \ldots, p_{k,m_k}) : p_{k,i} \ge 0 \text{ for } i = 1, \ldots, m_k, \text{ and } \textstyle\sum_{i=1}^{m_k} p_{k,i} = 1\}. \tag{1}$$

For a given joint pure strategy selection, $\boldsymbol{x} = (i_1, \ldots, i_n)$ with $i_j \in X_j$ for all j, the payoff, or utility, to Player k is denoted by $u_k(i_1, \ldots, i_n)$ for $k = 1, \ldots, n$. For a given joint mixed strategy selection, $(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$ with $\boldsymbol{p}_j \in X_j^*$ for $j = 1, \ldots, n$, the corresponding expected payoff to Player k is given by $g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$,

$$g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n) = \sum_{i_1=1}^{m_1} \cdots \sum_{i_n=1}^{m_n} p_{1,i_1} \cdots p_{n,i_n} u_k(i_1, \ldots, i_n). \tag{2}$$

Let us use the notation $g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n|i)$ to denote the expected payoff to Player k if Player k changes strategy from p_k to the pure strategy $i \in X_k$,

$$g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n|i) = g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_{k-1}, \boldsymbol{\delta}_i, \boldsymbol{p}_{k+1}, \ldots, \boldsymbol{p}_n). \tag{3}$$

where $\boldsymbol{\delta}_i$ represents the probability distribution giving probability 1 to the point i. Note that $g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$ can be reconstructed from the $g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n|i)$ by

$$g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n) = \sum_{i=1}^{m_k} p_{k,i} g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n|i) \tag{4}$$

A vector of mixed strategies, $(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$, is a strategic equilibrium if for all $k = 1, \ldots, n$, and all $i \in X_k$,

$$g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n|i) \le g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n). \tag{5}$$

Theorem. *Every finite n-person game in strategic form has at least one strategic equilibrium.*

Proof. For each k, X_k^* is a compact convex subset of m_k dimensional Euclidean space, and so the product, $C = X_1^* \times \cdots \times X_n^*$, is a compact convex subset of a Euclidean space of dimension $\sum_{i=1}^n m_i$. For $\boldsymbol{z} = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n) \in C$, define the mapping $T(\boldsymbol{z})$ of C into C by

$$T(\boldsymbol{z}) = \boldsymbol{z}' = (\boldsymbol{p}_1', \ldots, \boldsymbol{p}_n') \tag{6}$$

where

$$p_{k,i}' = \frac{p_{k,i} + \max(0, g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n|i) - g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n))}{1 + \sum_{j=1}^{m_k} \max(0, g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n|j) - g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n))}. \tag{7}$$

Note that $p_{k,i} \geq 0$, and the denominator is chosen so that $\sum_{i=1}^{m_k} p_{k,i}' = 1$. Thus $\boldsymbol{z}' \in C$. Moreover the function $f(\boldsymbol{z})$ is continuous since each $g_k(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$ is continuous. Therefore, by the Brouwer Fixed Point Theorem, there is a point, $\boldsymbol{z}' = (\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) \in C$ such that $T(\boldsymbol{z}') = \boldsymbol{z}'$. Thus from (7)

$$q_{k,i} = \frac{q_{k,i} + \max(0, g_k(\boldsymbol{z}'|i) - g_k(\boldsymbol{z}'))}{1 + \sum_{j=1}^{m_k} \max(0, g_k(\boldsymbol{z}'|j) - g_k(\boldsymbol{z}'))}. \tag{8}$$

for all $k = 1, \ldots, n$ and $i = 1, \ldots, m_n$. Since from (4) $g_k(\boldsymbol{z}')$ is an average of the numbers $g_k(\boldsymbol{z}'|i)$, we must have $g_k(\boldsymbol{z}'|i) \leq g_k(\boldsymbol{z}')$ for at least one i for which $q_{k,i} > 0$, so that $\max(0, g_k(\boldsymbol{z}'|i) - g_k(\boldsymbol{z}')) = 0$ for that i. But then (8) implies that $\sum_{j=1}^{m_k} \max(0, g_k(\boldsymbol{z}'|j) - g_k(\boldsymbol{z}')) = 0$, so that $g_k(\boldsymbol{z}'|i) \leq g_k(\boldsymbol{z}')$ for all k and i. From (5) this shows that $\boldsymbol{z}' = (\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n)$ is a strategic equilibrium. ∎

Remark. From the definition of $T(\boldsymbol{z})$, we see that $\boldsymbol{z} = (p_1, \ldots, p_n)$ is a strategic equilibrium if and only if $\boldsymbol{z}$ is a fixed point of T. In other words, the set of strategic equilibria is given by $\{\boldsymbol{z} : T(\boldsymbol{z}) = \boldsymbol{z}\}$. If we could solve the equation $T(\boldsymbol{z}) = \boldsymbol{z}$ we could find the equilibria. Unfortunately, the equation is not easily solved. The method of iteration does not ordinarily work because T is not a contraction map.