STATIC GAMES
(Simultaneous Decision Games)

LECTURE 1

What is game theory?

• A mathematical formalism for understanding designing and predicting the outcome of games.

• A game is characterized by a number of players (more than 2), assumed to be intelligent and rational that interact with each other by selecting various actions based on their assigned preferences.

• A player (decision maker) is characterized by a set of actions available for him and a set of preference relationships defined for each possible action tuple.
History of game theory

• First studies of games in economic literature deal with models on oligopoly pricing and production: Cournot (1838), Bertrand (1883), Edgeworth (1925)

• The beginning of game theory: John von Neumann and Oskar Morgenstern “Theory of Games and Economic Behaviour”

• Nash (1950) proposed what came to be known as “Nash Equilibrium”


Selten (1965) introduced his concept of a subgame perfection.

Harsanyi (1967-68) introduced his Bayesian Nash Equilibrium to model situations of incomplete information.
Example: **Prisoners’ Dilemma**

- Two crooks are being questioned by the police in connection with a serious crime. They are held in separate cells and cannot communicate. If both of them confess \( <C> \), they will get 4-year sentence, if neither confesses \( <Q> \), they spend 2 years in jail. If only one confesses, then the confessor will be set free and the other spend 5 years in jail.

- Game model:
ANALYSIS LEADS TO A SOLUTION: <C,C>

This is a paradoxical result that shows the difference between non-interactive and interactive decision models.

For example: paying taxes (everyone is better off if he does not pay taxes but then there is no money to provide community services and everyone is worse off).
A solution is said to be **Pareto optimal** (after the Italian economist Vilfredo Pareto) if no players’ payoff can be increased without decreasing the payoff to another player. Such solutions are also called Pareto/socially efficient or just efficient.

Here, **NE** is not Pareto optimal.
1. The set of players indexed by $i \in \{1, \ldots, n\}$.

2. $S_i$ - a pure strategy (or action) set for each player.

3. $r_i : \prod_{i=1}^{n} S_i \to \mathbb{R}$ is a payoff function for Player $i$.

A tabular description of a game using pure strategies is called **normal** form or strategic form of a game.
DEFINITION: A **mixed** strategy $\sigma_i$ for player $i$ gives the probabilities that action/strategy $s \in S_i$ will be played.

Clearly, $\sigma_i$ is an element of $P(S_i)$, the set of probability measures on the set $S_i$.

The payoff for player $i$ when player 1 uses a strategy $\sigma_1$ and player 2 uses a strategy $\sigma_2$ is:

$$
    r_i(\sigma_1, \sigma_2) = \sum_{s_1 \in S_1} \sum_{s_2 \in S_2} p(s_1)q(s_2)r_i(s_1, s_2)
$$

where $p(s_1)$ is the probability of using a pure strategy $s_1$ and $q(s_2)$ is the probability of using a pure strategy $s_2$.  

EXAMPLE:

Let $\sigma_1=(1/3,1/3,1/3)$ and let $\sigma_2=(0,1/2,1/2)$.

Calculate: $r_1(\sigma_1,\sigma_2)$

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>4,3</td>
<td>5,1</td>
<td>6,2</td>
</tr>
<tr>
<td>C</td>
<td>2,1</td>
<td>8,4</td>
<td>3,6</td>
</tr>
<tr>
<td>D</td>
<td>3,0</td>
<td>9,6</td>
<td>2,8</td>
</tr>
</tbody>
</table>
\[
\begin{align*}
    r_1(\sigma_1, \sigma_2) & = \frac{1}{3} \cdot \frac{1}{2} \cdot 5 + \frac{1}{3} \cdot \frac{1}{2} \cdot 8 + \frac{1}{3} \cdot \frac{1}{2} \cdot 8 + \ldots + \frac{1}{3} \cdot \frac{1}{2} \cdot 2 = \\
    & = \frac{1}{6}(5 + 8 + 9 + 6 + 5) = \frac{11}{2}
\end{align*}
\]

**WEAK AND STRICT DOMINANCE**

The strategy \( \sigma_1 \) is weakly dominated by \( \sigma_1' \) if for every \( \sigma_2 \in P(S_2) \), we have \( r_1(\sigma_1', \sigma_2) \geq r_1(\sigma_1, \sigma_2) \).

The strategy \( \sigma_1 \) is strictly dominated by \( \sigma_1' \) if for every \( \sigma_2 \in P(S_2) \), we have \( r_1(\sigma_1', \sigma_2) \geq r_1(\sigma_1, \sigma_2) \) and there exists a strategy \( \hat{\sigma}_2 \in P(S_2) \) such that \( r_1(\sigma_1', \hat{\sigma}_2) > r_1(\sigma_1, \hat{\sigma}_2) \).
Assumption:

**Common Knowledge of Rationality:**

- Players are rational and all know that the others are rational.
- Players all know that the other players know that they are rational, etc.

**ITERATED DOMINANCE ALGORITHM**

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>U</strong></td>
<td>1,0</td>
<td>1,2</td>
<td>0,1</td>
</tr>
<tr>
<td><strong>D</strong></td>
<td>0,3</td>
<td>0,1</td>
<td>2,0</td>
</tr>
</tbody>
</table>
Analysis:

- Player 1 has no dominated strategies.
- R is dominated by M.
- Now D is dominated by U.
- Clearly, Player 2 will play column M.

Remarks:

- There is a problem with the iterated dominance algorithm: the solution may depend on the order in which strategies are eliminated, e.g.,
Order 1: eliminate D, then \(<U,M>\) is the solution.

Order 2: eliminate R, four pairs are left.

- Very few games are solvable by the iterated dominance algorithm.
A Nash Equilibrium for a two-player game is a pair of strategies $(\sigma_1^*, \sigma_2^*) \in P(S_1) \times P(S_2)$ such that

$$r_1(\sigma_1^*, \sigma_2^*) \geq r_1(\sigma_1, \sigma_2^*)$$

for every $\sigma_1 \in P(S_1)$, and

$$r_2(\sigma_1^*, \sigma_2^*) \geq r_2(\sigma_1^*, \sigma_2)$

for every $\sigma_2 \in P(S_2)$. In other words, neither player has a unilateral incentive to change its strategy.

Note: Strictly dominated strategies cannot form a NE.
NASH EQUILIBRIUM PROPERTIES

• **Existence**
  - Pure strategy Nash equilibrium may not exist

• **Uniqueness**
  - Nash equilibrium need not be unique

• **Efficiency**
  - Pareto Optimality? Not always, see Prisoners’ Dilemma.

How to get to the Nash equilibrium?
EXAMPLES:

- A pure strategy Nash Equilibrium can be found by the arrow diagram:

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>1,3</td>
<td>4,2</td>
<td>2,2</td>
</tr>
<tr>
<td>C</td>
<td>4,0</td>
<td>0,3</td>
<td>4,1</td>
</tr>
<tr>
<td>D</td>
<td>3,4</td>
<td>5,6</td>
<td></td>
</tr>
</tbody>
</table>

Is the pair of strategies <D,R> Pareto efficient?

YES!
• Matching Pennies:

- two players simultaneously announce head or tail

- if their choices match, Player 1 wins, otherwise Player 2 wins.

Nash Equilibrium: each player has to randomise his choice with probability 1/2.
• Battle of the Sexes:
  - two players: wife and husband want to spend this evening together;
  - they have their own preferences: the wife (Pl 2) wants to watch a classical ballet and her husband (Pl 1) wants to watch a football match;

Nash Equilibria: <B,B>, <F,F>, a mixed equilibrium: \((\sigma_1^*, \sigma_2^*)\), where \(\sigma_1^* = (1/3, 2/3)\) and \(\sigma_2^* = (2/3, 1/3)\).
• Examples

- Nash equilibrium
- Pareto optimality
STATIC GAMES
(Simultaneous Decision Games)

LECTURE 2

How to obtain Nash Equilibria?

Mathematics

Let us consider a two-player game with the payoff matrix $A$ for player 1 and the payoff matrix $B$ for player 2. Recall that a strategy pair $(x^*, y^*)$ is called a Nash Equilibrium if

$$x^*Ay^* \geq xAy^*$$

for all mixed strategies $x$, and

$$x^*By^* \geq x^*By^*$$

for all mixed strategies $y$. 
THEOREM (NASH, 1950):

Every finite two-player game has at least one Nash equilibrium point.

This result is equivalent the following theorem.

Brouwer’s Fixed Point Theorem:

Suppose that $S$ is non-empty, compact, convex set in $\mathbb{R}^n$. If the function $f : S \rightarrow S$ is continuous, then a fixed point $\tilde{s}$ exists satisfying

$$\tilde{s} = f(\tilde{s}).$$
Proof of the Nash Theorem:

• For a pair of mixed strategies \((x,y)\) define:

\[
\begin{align*}
  c_i &= \max \{A_i \cdot y^t - xAy^t, 0\} \\
  d_j &= \max \{xA_j - xAy^t, 0\}
\end{align*}
\]

and

\[
\begin{align*}
  x_i' &= \frac{x_i + c_i}{1 + \sum_k c_k}, & y_j' &= \frac{y_j + d_j}{1 + \sum_k d_k}.
\end{align*}
\]

• Clearly, \(T(x,y)=(x',y')\) is continuous.

• We prove that \((x,y)=(x',y')\) iff \((x,y)\) is a NE.
• \(\iff\) Indeed, if \((x,y)\) is a NE, then for all \(i\), we have
\[
xA^t y^t \geq A_i y^t.
\]
Hence, \(c_i = 0\). Similarly, \(d_j = 0\) and \(x = x',\ y = y'\).

• \(\implies\) Assume that \((x,y)\) is not a NE. Hence, there exists \(\bar{x}\) such that
\[
\bar{x}A^t y^t > xA^t y^t
\]
(or \(\bar{y}\) such that \(xB\bar{y}^t > xB y^t\)).
We shall only consider the first case, since the second one is analogous.

Note that there is an index \(i\) such that \(A_i y^t > xA^t y^t\).
Consequently, \(c_i > 0\) and \(\sum_k c_k > 0\).

The expression is \(xA^t y^t\) is also a weighted average of \(A_i y^t\) (with \(x_i\) as the weights). Thus, \(A_i y^t \leq xA^t y^t\) for some \(i\) such that \(x_i > 0\).
For this $i$, $c_i=0$ and therefore 

$$x'_i = \frac{x_i}{1 + \sum_k c_k} < x_i.$$ 

Hence, $x \neq x'$.

In the second case, $y \neq y'$.

Hence, $(x,y) = (x',y')$ if and only if a pair $(x,y)$ is a NE.

• The set of all pairs of strategies is bounded, closed and convex, and therefore satisfies the assumptions of Brouwer’s Fixed Point Theorem. Since $T(x,y) = (x',y')$ is continuous, it must possess a fixed point.

This point is a pair of a Nash Equilibrium.
Remark: The proof of the theorem can be easily generalized for $n$-player games.

**BEST RESPONSE STRATEGY:**

A strategy $\hat{\sigma}_1$ is a best response to some fixed strategy $\sigma_2$ for player 2 if

$$\hat{\sigma}_1 \in \arg \max_{\sigma_1 \in P(S_1)} r_1(\sigma_1, \sigma_2)$$

Similarly, $\hat{\sigma}_2$ is a best response to some $\sigma_1$ if

$$\hat{\sigma}_2 \in \arg \max_{\sigma_2 \in P(S_2)} r_2(\sigma_1, \sigma_2)$$
An equivalent definition of NE is the following:

A pair \((\sigma_1^*, \sigma_2^*)\) is a Nash Equilibrium if

\[
\hat{\sigma}_1^* \in \arg \max_{\sigma_1 \in P(S_1)} r_1(\sigma_1, \sigma_2^*)
\]

and

\[
\hat{\sigma}_2^* \in \arg \max_{\sigma_2 \in P(S_2)} r_2(\sigma_1^*, \sigma_2)
\]

We use this definition for drawing the so-called best response diagram.
Example: Matching Pennies

\[
\begin{array}{c|cc}
 & b & 1-b \\
\hline
H & 1,-1 & -1,1 \\
1-a & -1,1 & 1,-1 \\
\end{array}
\]

- If \( b < \frac{1}{2} \), then the best response is to choose \( a = 0 \).
- If \( b > \frac{1}{2} \), then the best response is to choose \( a = 1 \).
- If \( b = \frac{1}{2} \), then the best response is any strategy \( a \).
Similarly,

\[ r_2(\sigma_1, \sigma_2) = -ab + a(1 - b) + b(1 - a) - (1 - a)(1 - b) \]

\[ = (2b - 1)(1 - 2a) \]

- If \( a < \frac{1}{2} \), then the best response is to choose \( b = 1 \).
- If \( a > \frac{1}{2} \), then the best response is to choose \( b = 0 \).
- If \( a = \frac{1}{2} \), then the best response is any strategy \( b \).
THE EQUALITY OF PAYOFFS:

Support of a strategy $\sigma$ is a set $S_{\sigma} \subseteq S$ for which $\sigma$ specifies $p(s)>0$.

**Theorem:** Let the pair $(\sigma_1^*, \sigma_2^*)$ be a Nash Equilibrium and let $S_1^*$ be a support of $\sigma_1^*$. Then

$$r_1(\sigma_1^*, \sigma_2^*) = r_1(s_1, \sigma_2^*)$$
for every $s_1 \in S_1^*$.

**Proof:**

- If $S_1^*$ contains only one strategy, then the theorem is trivially true.
- Suppose that $S_1^*$ contains more than one strategy and suppose
that at least one strategy gives higher payoff for player 1 than 
\( r_1(\sigma_1^*, \sigma_2^*) \): case (i). It is also possible that all \( s \in S_1^* \) gives the 
payoff that is not greater then \( r_1(\sigma_1^*, \sigma_2^*) \): case (ii).

Case (i):
Let \( s' \in S_1^* \) be the strategy that gives the greatest such a payoff. Then

\[
r_1(\sigma_1^*, \sigma_2^*) = \sum_{s \in S_1^*} p(s)r_1(s, \sigma_2^*) = \sum_{s \in S_1^*} p(s)r_1(s, \sigma_2^*) + p(s')r_1(s', \sigma_2^*)
< \sum_{s \in S_1^*} p(s)r_1(s', \sigma_2^*) + p(s')r_1(s', \sigma_2^*) = r_1(s', \sigma_2^*).\]

But this contradicts that \( (\sigma_1^*, \sigma_2^*) \) is a Nash Equilibrium!
Case (ii): Assume that there exists $s' \in S_1^*$ such that

$$r_1(s', \sigma_2^*) < r_1(\sigma_1^*, \sigma_2^*).$$

For the others $s \in S_1^*$ we have

$$r_1(s, \sigma_2^*) \leq r_1(\sigma_1^*, \sigma_2^*).$$

Multiplying these inequalities by $p(s')$ and $p(s)$ and adding them, we have

$$r_1(\sigma_1^*, \sigma_2^*) < r_1(\sigma_1^*, \sigma_2^*).$$

Clearly, the corresponding result also holds for Player 2.
### EXAMPLES:

- **Matching Pennies:**

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1,-1</td>
<td>-1,1</td>
</tr>
<tr>
<td>1-a</td>
<td>-1,1</td>
<td>1,-1</td>
</tr>
</tbody>
</table>

\[
\sigma_2 = (b, 1-b) \quad \text{and} \quad \sigma_1 = (a, 1-a)
\]

\[
r_1(H, \sigma_2^*) = r_1(T, \sigma_2^*) \iff
br_1(H, H) + (1 - b)r_1(H, T) = br_1(T, H) + (1 - b)r_1(T, T)
\]

\[
b + (1 - b)(-1) = -b + (1 - b) \quad \Rightarrow \quad b = 1/2.
\]

The expected payoff equals 0.
- The equality of payoffs:

\[2a + 0(1-a) = 0a + (1-a) \Rightarrow a = \frac{1}{3} \Rightarrow \sigma_1 = (\frac{1}{3}, \frac{2}{3}),\]

\[b + 0(1-b) = 0b + 2(1-b) \Rightarrow b = \frac{2}{3} \Rightarrow \sigma_2 = (\frac{2}{3}, \frac{1}{3}).\]
- The best response diagram:

\[ r_1(\sigma_1, \sigma_2) = ab + 2(1-a)(1-b) = a(3b-2) - 2b + 2 \]

If \( b < \frac{2}{3} \), then the best response is \( a = 0 \).
If \( b > \frac{2}{3} \), then the best response is \( a = 1 \).
If \( b = \frac{2}{3} \), then the best response is any value of \( a \in [0,1] \).

\[ r_2(\sigma_1, \sigma_2) = 2ab + (1-a)(1-b) = b(3a-1) - a + 1 \]

If \( a < \frac{1}{3} \), then the best response is \( b = 0 \).
If \( a > \frac{1}{3} \), then the best response is \( b = 1 \).
If \( a = \frac{1}{3} \), then the best response is any value of \( b \in [0,1] \).
Figure of the best response diagram:

The expected payoffs for a mixed equilibrium equal $\frac{2}{3}$ for both players.
### NASH EQUILIBRIA FOR $2 \times n$ AND $m \times 2$ GAMES

#### PLAYER 1's payoffs

<table>
<thead>
<tr>
<th>q</th>
<th>1-q</th>
<th>(-q + 5(1-q) = 5 - 6q)</th>
<th>(2q + 4(1-q) = 4 - 2q)</th>
<th>(4q - 3(1-q) = -3 + 7q)</th>
<th>(5q - 4(1-q) = -4 + 9q)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1,3</td>
<td>5,-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,1</td>
<td>4,5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4,-2</td>
<td>-3,6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5,10</td>
<td>-4,-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Graph of these four lines:

- the corners, where the payoffs are equalised
Point (I) $\leftrightarrow$ two-player game (row 1 and 2):

<table>
<thead>
<tr>
<th></th>
<th>-1,3</th>
<th>5,-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,1</td>
<td>4,5</td>
<td></td>
</tr>
</tbody>
</table>

$\sigma^*_2 = (1/4, 3/4)$ and $\sigma^*_1 = (4/9, 5/9, 0, 0)$

Point (II) $\leftrightarrow$ two-player game (row 2 and 4):

<table>
<thead>
<tr>
<th></th>
<th>2,1</th>
<th>4,5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,0</td>
<td>-4,-4</td>
<td></td>
</tr>
</tbody>
</table>

$\sigma^*_2 = (8/11, 3/11)$ and $\sigma^*_1 = (0, 7/9, 0, 2/9)$. 
NASH SOLUTION THEOREM

In an $m \times 2$ [$2 \times n$] game at least one of the intersection points of endpoints of the upper of edges of the row [column] player’s expected payoff graph yields a Nash Equilibrium.
EQUILIBRIA FOR n-PLAYER GAMES

Strategy profile: \((\sigma_1, \ldots, \sigma_n)\).

Notation: \((\hat{\sigma}_i, \sigma_{-i}) = (\sigma_1, \ldots, \sigma_{i-1}, \hat{\sigma}_i, \sigma_{i+1}, \ldots, \sigma_n)\).

Definition:

A strategy profile \(\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)\) is a Nash Equilibrium if

\[
 r_i(\sigma_i^*, \sigma_{-i}^*) \geq r_i(\sigma_i, \sigma_{-i}^*)
\]

for all \(\sigma_i \in P(S_i)\) and \(i=1,\ldots,m\).
A **generic game** is a one in which a small change to any one of the payoffs does not introduce new Nash equilibria or remove existing ones.
Most of the games we have considered so far are generic:

- **I** Prisoners’ Dilemma
- **II** Matching Pennies
- **III** Battle of the Sexes

**Oddness Theorem**

A generic game has an *odd* number of Nash Equilibria.

**I** - one pure strategy Nash Equilibrium

**II** - one mixed strategy Nash Equilibrium

**III** - one mixed strategy and two pure strategy Nash Equilibria
Non-generic game:

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>10,0</td>
<td>5,1</td>
<td>4,-2</td>
</tr>
<tr>
<td>1-a</td>
<td>10,1</td>
<td>5,0</td>
<td>1,-1</td>
</tr>
</tbody>
</table>

R is strictly dominated by M. Player 1 would get the same payoff playing U against L. Similarly, playing D against M Player 1 gets the same payoff.

Non-generic games may have infinitely many Nash equilibria. Indeed, let $\sigma_1 = (a, 1-a)$, and $\sigma_2 = (b, 1-b, 0)$.

We shall draw the best response diagram.
\[ r_1(\sigma_1, \sigma_2) = 10ab + 5a(1-b) + 10b(1-a) + 5(1-a)(1-b) = 5 + 5b \]

The best response strategy for Player 1 to a fixed strategy \( \sigma_2 \) of Player 2 is to play \((a,1-a)\) with any \( a \in [0,1] \).

\[ r_2(\sigma_1, \sigma_2) = a(1-b) + b(1-a) = a + b(1-2a) \]

If \( a < 1/2 \), then Player 2 should play \( b=1 \).
If \( a > 1/2 \), then Player 2 should play \( b=0 \).
If \( a = 1/2 \), then Player 2 can play any \( b \in [0,1] \).
The best response diagram for non-generic game:

Nash Equilibria lie on the blue line.
A **zero-sum** game is a game, in which the payoffs add up to zero, for example Matching Pennies game. The interests of the players are exactly opposed: one only wins what the other loses.

Zero-sum games are the first games which were studied formally (the concept of Nash Equilibrium did not exist yet). Such games were solved by **John von Neumann** (The Minimax Theorem).
$v_*$ is the lower value of the game, it is the amount that Player 1 can ensure himself.

$v^* = \sup_{\sigma_1} \inf_{\sigma_2} r(\sigma_1, \sigma_2)$

$v^*$ is the upper value of the game, it is the amount that Player 2 can guarantee himself.

$v^* = \inf_{\sigma_2} \sup_{\sigma_1} r(\sigma_1, \sigma_2)$
Lemma: \( \nu^* \geq \nu_* \).

Proof:

- \( \inf_{\sigma_2} r(\sigma_1, \sigma_2) \leq r(\sigma_1, \sigma_2) \) for all \( \sigma_1 \) and \( \sigma_2 \)
- \( \inf_{\sigma_1} r(\sigma_1, \sigma_2) \leq \sup_{\sigma_1} r(\sigma_1, \sigma_2) \) for all \( \sigma_1 \) and \( \sigma_2 \)
- \( \sup_{\sigma_2} \inf_{\sigma_1} r(\sigma_1, \sigma_2) \leq \sup_{\sigma_1} r(\sigma_1, \sigma_2) \) for all \( \sigma_2 \)
- \( \nu_* = \sup_{\sigma_1} \inf_{\sigma_2} r(\sigma_1, \sigma_2) \leq \inf_{\sigma_2} \sup_{\sigma_1} r(\sigma_1, \sigma_2) = \nu^* \)
If \( v_* = v^* \triangleq v \), then we say that the game has a value \( v \).

The following result says that every finite zero-sum game has a value.

**THE MINIMAX THEOREM**
*(John von Neumann, 1928)*

In every finite zero-sum game there exists a pair of strategies \((\sigma_1^*, \sigma_2^*) \in P(S_1) \times P(S_2)\) such that

\[
r(\sigma_1, \sigma_2^*) \leq r(\sigma_1^*, \sigma_2^*) \leq r(\sigma_1^*, \sigma_2)
\]

for all \( \sigma_1 \) and \( \sigma_2 \).
Note that $v = r(\sigma_1^*, \sigma_2^*)$ is a value of the game. Indeed, the inequality in theorem means that

$$\sup_{\sigma_1} r(\sigma_1, \sigma_2^*) \leq \inf_{\sigma_2} r(\sigma_1^*, \sigma_2),$$

and, therefore

$$v^* = \inf_{\sigma_2} \sup_{\sigma_1} r(\sigma_1, \sigma_2) \leq \sup_{\sigma_1} \inf_{\sigma_2} r(\sigma_1, \sigma_2) = v_*.$$

On the other hand, from the lemma: $v^* \geq v_*$. Hence, it must hold $v^* = v_*$. Clearly,

$$v = r(\sigma_1^*, \sigma_2^*) = v^* = v_*.$$
is an optimal (maxmin) strategy for Player 1, which ensures him an expected gain of at least $v$.

is an optimal (minmax) strategy for Player 2, which ensures him an expected loss of at most $v$.

\[(\sigma^*_1, \sigma^*_2)\] A SADDLE POINT.

**Question:** Let $v$ be a value of the game and let $\hat{\sigma}_1$ and $\hat{\sigma}_2$ be strategies such that

$$v = r(\hat{\sigma}_1, \hat{\sigma}_2).$$

Do the strategies $\hat{\sigma}_1$ and $\hat{\sigma}_2$ have to be optimal?
**Answer:** No! For example, consider the Matching Pennies game:

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1,-1</td>
<td>-1,1</td>
</tr>
<tr>
<td>T</td>
<td>-1,1</td>
<td>1,-1</td>
</tr>
</tbody>
</table>

\[ \hat{\sigma}_1 = (1/2,1/2) \]
\[ \hat{\sigma}_2 = (1,0) \]

The value of the game \( v = 0 \) and \( r(\hat{\sigma}_1, \hat{\sigma}_2) = 0 \).

But \( \hat{\sigma}_2 \) is **not optimal** strategy, because it does not ensure Player 2 the expected loss 0.
Remarks:

• The problem of the existence of a value of the game and a saddle point is a problem of separation of convex sets and its proof is based upon the Separating Hyperplane Theorem. It is not a problem of a fixed point of a certain correspondence as in the case of non-zero sum games.

• In zero-sum games, we call a pair of strategies \((\sigma_1^*, \sigma_2^*)\) optimal, because the interests of the players are exactly opposite and there is only one value. In non-zero-sum games this terminology does not make sense! We may have many Nash Equilibria that lead to different expected payoffs. Which equilibrium is the best?

**Theorem:** A generic zero-sum game has a unique solution, i.e., there exists only one pair of optimal strategies.
A standard form of a Linear Program:

\[
\begin{align*}
\text{max } & \quad b_1y_1 + ... + b_ny_n \\
\text{subject to } & \quad a_{11}y_1 + ... + a_{1n}y_n \leq c_1 \\
& \vdots \\
& \quad a_{m1}y_1 + ... + a_{mn}y_n \leq c_m \\
\text{and } & \quad y_j \geq 0 \text{ for } j=1,...,n.
\end{align*}
\]

The objective function is linear and the constraints are also linear.
Let \( A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \) be the payoff matrix for Player 1. He wants to choose \( \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix} \) that maximize \( \min_{1 \leq j \leq n} \sum_{i=1}^{m} p_i a_{ij} \) subject to the constraints \( p_1 + \ldots + p_m = 1, \ p_i \geq 0 \) for \( i=1,\ldots,m \).

Although the constraints are linear the objective function

\[
\min_{1 \leq j \leq n} \sum_{i=1}^{m} p_i a_{ij}
\]

is not linear!
We add a new variable \( \nu \) to the Player 1’s list of variables, restrict to be less than the objective function \( \nu \leq \min_{1 \leq j \leq n} \sum_{i=1}^{m} p_i a_{ij} \) and make \( \nu \) as large as possible. Hence, we have:

(I)

\[
\begin{align*}
\max & \; \nu \\
\text{subject to} & \; \nu \leq \sum_{i=1}^{m} p_i a_{i1}, \ldots, \nu \leq \sum_{i=1}^{m} p_i a_{in} \\
& \; p_1 + \ldots + p_m = 1, \; p_i \geq 0 \\
& \; \text{for } i=1,\ldots,m.
\end{align*}
\]
In a similar way, one may look at the problem from the minimiser’s point of view and arrive at the following Linear Programm:

choose \( w \) and a vector \( (q_1, \ldots, q_n) \) that satisfy

\[(\text{II})\]

\[
\begin{align*}
\text{min } w \\
\text{subject to } w \geq a_1 q_1 + \ldots + a_n q_n, \\
q_1 + \ldots + q_n = 1, q_j \geq 0
\end{align*}
\]

for \( j = 1, \ldots, n. \)
Problems (I) and (II) can be solved by the SIMPLEX ALGORITHM (see MATLAB, MATHEMATICA). Simplex algorithm was invented by G. Dantzig, although some Linear Programms were solved earlier by the economists, e.g., T. Koopmans.

In Linear Programming there is a theory of duality that says that the programs (I) and (II) are dual. The so-called Duality Theorem implies that dual programs have the same value, i.e., $v = w$. This is exactly the Minimax Theorem.
A GAME WITHOUT VALUE

Let $S_1 = S_2 = \{1, 2, \ldots\}$ be the set of pure strategies for both players and let $A = [a_{ij}]$ be a payoff matrix for Player 1, where

$$a_{ij} = \begin{cases} 
1 & \text{if } i>j \\
0 & \text{if } i=j \\
-1 & \text{if } i<j 
\end{cases}$$

The player who chooses the larger number wins 1.

Let $\sigma_2 = (q_1, q_2, \ldots)$ be a strategy used by Player 2, i.e.,

$$\sum_j q_j = 1 \quad (\bullet)$$

Note that if Player 1 uses a pure strategy, then for every $\sigma_2$,
\[ \sup_{1 \leq i < \infty} \sum_{j} a_{ij} q_j = 1. \]

Indeed, since (0) holds, then for any \( \varepsilon > 0 \) there exists \( N \) such that
\[ \sum_{j=1}^{N} q_j \geq 1 - \varepsilon. \]

Obviously,
\[ v^* = \inf_{\sigma_2} \sup_{\sigma_1} r(\sigma_1, \sigma_2) \geq \sup_{1 \leq i < \infty} \sum_{j} a_{ij} q_j = 1, \]
and since \( \max_{i} a_{ij} = 1 \), then \( v^* = 1 \). Since the situation is ‘symmetric’ for the minimiser, it follows that \( v_* = -1 \).

**Conclusion:** The Minimax Theorem cannot be easily extended to games with infinite pure strategy sets. Additional assumptions are needed (e.g., the compactness of the pure strategy sets).
Graph consists of a finite set $V$ and a set $A$ of unordered pairs.

$\nu \in V$ - vertex or node

$\{\nu_1, \nu_2\} \in A$ - arc, branch or edge

Tree is a graph where any two nodes are connected by exactly one path.
EXAMPLES

- TREE
- Leaves - Terminal Nodes $L(T)$
- Root

This is NOT A TREE
DEFINITION

A game in an extensive form consists of the following objects:

- $N = \{1, \ldots, n\}$ - a set of players
- $T$ - rooted game tree
- $P^0, P^1, \ldots, P^n$ - a partition of the set of non-terminal nodes of $T$ into $(n+1)$ subsets; the members of $P^0$ are called nature/chance nodes, while the members $P^i$ are called nodes of Player $i$
- $P^i$ is divided into $k(i)$ information sets: $U^i_1, \ldots, U^i_{k(i)}$
- $g(t) = (g^1(t), \ldots, g^n(t))$ - n dimensional payoff vector for each terminal node $t \in L(T)$. 
This complete description is common knowledge among the players.

**Definition:**

An *information set* for Player $i$ is a set of decision nodes in game tree such that:

(i) the player concerned is masking a decision (no others)

(ii) the player does not know which node has been reached (only that it is one of the nodes in the set).
EXAMPLES

• (I) Matching Pennies:

\[ P^0 = \emptyset, \ P^1 = U_1^1 = \{a\}, \ P^2 = U_1^2 = \{b, c\}, \ \text{root} = \{a\}, \ \text{payoff vector} \ (g^1(t), g^2(t)) \]
• (II) \( N=\{1,2,3\} \), \( \text{root}=\{a\} \), \( P^0 = \{d\} \), \( P^1 = \{a,e,f\} \), \( U^1_1 = \{a\} \), \( U^1_2 = \{e,f\} \), \( P^2 = U^1_1 = \{b,c\} \), \( P^3 = U^1_3 = \{g\} \)

Note that at his second information set \( U^1_2 \) Player 1 does not recall what his choice was at \( U^1_1 \).
PURE STRATEGIES

- $I^i = \{U^i_1, U^i_2, \ldots, U^i_{k(i)}\}$ - the set of information sets for Player $i$
- $U^i \in I^i$ - a generic element of $I^i$
- $\nu \equiv \nu(U^i)$ - the number of branches going out of each node in $U^i$
- $C(U^i) \triangleq \{1, \ldots, \nu(U^i)\}$ - the set of choices available to Player $i$ at any node in $U^i$
A **pure strategy** $s^i$ is a function $s^i : I^i \rightarrow \{1,2,\ldots\}$ such that $s^i(U^i) \in C(U^i)$ for all $U^i \in I^i$.

In other words, $s^i$ specifies for every information set $U^i \in I^i$ of Player $i$, a choice $s^i(U^i)$ there.

Let $S^i$ be the set of pure strategies of Player $i$ and let $S = S^1 \times \cdots \times S^n$ be the set of pure strategy profiles.

For a strategy profile $s = (s^1,\ldots,s^n) \in S$ the expected payoff for Player $i$ is:

$$h^i(s) = \sum_{t \in L(T)} p_s(t)g^i(t),$$

where $p_s(t)$ is the probability that the game ends at $t$ when the strategy profile $s$ is used.
Example (II) - cont.

• Player 1 has four pure strategies: \(<1,1>, <1,2>, <2,1>, <2,2>\), where \(<i,j>\) means that \(i\) is chosen at \(U_1\) and \(j\) is chosen at \(U_2\).

• Player 2 has two pure strategies: \(<1>, <2>\).

• Player 3 has three pure strategies: \(<1>, <2>, <3>\).

If \(s=(<2,1>,<2>,<3>)\), then \(h^1(s) = 1, h^2(s) = -1, h^3(s) = 1\).

If \(s'=(<1,1>,<1>,<3>)\), then

\[
h(s') = \frac{1}{2}(2,0,0) + \frac{1}{6}(0,2,0) + \frac{1}{3}(0,2,3) = (1,1,1).
\]
EXTENSIVE FORM VS NORMAL FORM

- **Example (III)**

Normal form:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2,0</td>
<td>2,0</td>
</tr>
<tr>
<td>B</td>
<td>0,2</td>
<td>1,1</td>
</tr>
</tbody>
</table>

Then $N = \{1, 2\}$, $S^1 = S^2 = \{A, B\}$ and the extensive form is:

\[
\begin{array}{c|c|c|c|c|}
   & A & B & & \\
\hline
A & & & (2,0) & \\
B & & & (0,2) & \\
\hline
\end{array}
\]

\[
\begin{array}{c|c|c|c|c|}
   & A & B & & \\
\hline
A & & & (2,0) & \\
B & & & (0,2) & \\
\hline
\end{array}
\]

\[
\begin{array}{c|c|c|c|c|}
   & A & B & & \\
\hline
A & & & (2,0) & \\
B & & & (0,2) & \\
\hline
\end{array}
\]

\[
\begin{array}{c|c|c|c|c|}
   & A & B & & \\
\hline
A & & & (2,0) & \\
B & & & (0,2) & \\
\hline
\end{array}
\]

\[
\begin{array}{c|c|c|c|c|}
   & A & B & & \\
\hline
A & & & (2,0) & \\
B & & & (0,2) & \\
\hline
\end{array}
\]

\[
\begin{array}{c|c|c|c|c|}
   & A & B & & \\
\hline
A & & & (2,0) & \\
B & & & (0,2) & \\
\hline
\end{array}
\]

\[
\begin{array}{c|c|c|c|c|}
   & A & B & & \\
\hline
A & & & (2,0) & \\
B & & & (0,2) & \\
\hline
\end{array}
\]

The extensive form contains more information than the normal form!!!
MIXED STRATEGIES

A mixed strategy is

\[ \sigma^i = (\sigma^i(s^i))_{s^i \in S^i} \in P(S^i) \]

It means that Player \( i \) chooses each pure strategy \( s^i \in S^i \) with probability \( \sigma^i(s^i) \).

The expected payoff of Player \( i \) when a strategy profile \( \sigma = (\sigma^1, \ldots, \sigma^n) \) is used equals:

\[ H^i(\sigma) = \sum_{s \in S} \sigma(s)h^i(s), \]

where \( \sigma(s) = \prod_{j=1}^{n} \sigma^j(s^j) \).
THE KUHN THEOREM (1953)

An n-person game in extensive form is of perfect information if all information sets are singletons, i.e., \(|U^i| = 1\).

**THM:** Every finite n-person game of perfect information has an equilibrium point in pure strategies.

Proof by induction.
The resulting equilibrium point is

\[ s = (\langle 2, 2 \rangle, \langle 1, 2 \rangle, \langle 2, 2 \rangle) \]

with payoffs \( h(s) = (4, 1, 2) \).

Remark: Note that equilibrium points constructed in this manner, when restricted to any subgame of the original game, yield equilibria in the subgame as well. Such equilibria are called subgame perfect (see the next lecture).
BEHAVIOUR STRATEGIES

• A pure strategy is a possible plan for the player’s choices at all his information sets.

• A mixed strategy means that the player chooses before the beginning one such a plan at random (according to some probability distribution).

• An alternative approach for the player: to make independent random choice at each one of his information sets. The choices at each information sets are stochastically independent.

These randomisation procedures are called behaviour strategies.
Every **behaviour strategy** is a special kind of **mixed** strategy. When player has “**perfect recall**” then his every **mixed** strategy is equivalent to a **behaviour strategy**.

- $b^i$ - behaviour strategy of Player $i$,
- $B^i$ - the set of all behaviour strategies of Player $i$, i.e.,

$$B^i = \prod_{u^i \in I^i} P(C(U^i)).$$
EXAMPLE (IV)

Two cards marked H and L are dealt at random to Ann and John. The person with a high card obtains 1$ from the person with a low card and decides whether to stop or continue the game. If the games continues, then Paul (not knowing the outcome of the deal) instructs Ann and John either to exchange or keep their cards. Again the holder of the high card receives 1$ from the holder of the low card and the game ends.

Draw a tree for this game and then give its normal form!
THE GAME TREE

H to Ann with pr. 1/2

H to John with pr. 1/2

$U^A_1$

S

(1,-1)

C

$U^J_1$

S

(-1,1)

C

$U^P_1$

K

(2,-2)

E

(0,0)

K

(-2,2)

E

(0,0)
Note that this is a zero-sum game.

- **Ann → Player 1, maximiser**
  

- **John → Player 2, minimiser**
  
  pure strategies: <S>, <C>
<table>
<thead>
<tr>
<th></th>
<th>(&lt;S, K&gt;)</th>
<th>(&lt;C, K&gt;)</th>
<th>(&lt;S, E&gt;)</th>
<th>(&lt;C, E&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 0)</td>
<td>((-\frac{1}{2}, \frac{1}{2}))</td>
<td>((0, 0))</td>
<td>((\frac{1}{2}, -\frac{1}{2}))</td>
<td>((0, 0))</td>
</tr>
<tr>
<td>((\frac{1}{2}, -\frac{1}{2}))</td>
<td>((0, 0))</td>
<td>((\frac{1}{2}, -\frac{1}{2}))</td>
<td>((0, 0))</td>
<td></td>
</tr>
</tbody>
</table>

Strategy \(<S, K>\) is dominated by \(<C, K>\)

Strategy \(<C, E>\) is dominated by \(<S, E>\)

Unique optimal strategies are: \(\left(0, \frac{1}{2}, \frac{1}{2}, 0\right)\) and \(\left(\frac{1}{2}, \frac{1}{2}\right)\);

value of the game: 1/4.
• Consider the behaviour strategy for Player 1: 
  \((\alpha, 1 - \alpha)\) - Ann chooses S with probability \(\alpha\) and C with probability \(1 - \alpha\)

• Paul chooses K with probability \(\beta\) and E with probability \(1 - \beta\)

Player 1’s expected payoff:

• Player 2 plays S:

  \[
  1/2[\alpha + (1 - \alpha)(2\beta + 0(1 - \beta))] + 1/2(-1) = (1 - \alpha)(\beta - 1/2)
  \]

• Player 2 plays C:

  \[
  1/2[\alpha + (1 - \alpha)(2\beta + 0(1 - \beta))] + 1/2[-2\beta + 0] = \alpha(1/2 - \beta)
  \]
The maximum payoff that Player 1 can ensure herself is

$$\max_{0 \leq \alpha, \beta \leq 1} \{ \min \{ (1 - \alpha)(\beta - 1/2), \alpha(1/2 - \beta) \} \} = 0,$$

since either $1/2 - \beta \leq 0$ or $\beta - 1/2 \leq 0$.

**Conclusion:** The behaviour strategies do a poorer job!

The corresponding mixed strategy:

$$\begin{align*}
\langle S, K \rangle, & \quad \langle S, E \rangle, & \quad \langle C, K \rangle, & \quad \langle C, E \rangle \\
\alpha \beta, & \quad \alpha(1 - \beta), & \quad (1 - \alpha)\beta, & \quad (1 - \alpha)(1 - \beta)
\end{align*}$$

**Note:** $\sigma^1 = (0, 1/2, 1/2, 0)$ - the mixed strategy for Player 1
requires randomisation at two information sets that are rather
correlated than independent.
A subgame (subtree) is a part of a game that:

• begins at a decision node (for any player);

• the information set containing the initial decision node contains no other nodes, i.e., players knows all decisions that has been made up until that time

• the subtree contains all the decision nodes that follow the initial node
A **Subgame Perfect Nash Equilibrium** is a Nash Equilibrium in which the behaviour specified in every subgame is a Nash Equilibrium for the subgame.

**Note:** This also applies to subgames that are **NOT** reached during a play of the game.

**Theorem**

Every finite dynamic game has a

**Subgame Perfect Nash Equilibrium.**
EXAMPLE

Two people (husband and wife) are buying items for a dinner party. The husband buys either fish $F$ or meat $M$ for the main course and the wife buys wine: red $R$ or white $W$. Both are conventional and prefer red wine with meat and white wine with fish. However, the husband prefers meat over fish and wife prefers fish over meat.

The set of pure strategies for the players:

$$S^h = \{M, F\}, \ S^w = \{RR, RW, WR, WW\}$$
Extensive form of the game

Nash Equilibrium in pure strategies obtained by the backward induction: \((M,RW)\)
**Normal form of the game**

<table>
<thead>
<tr>
<th></th>
<th>&lt;R,R&gt;</th>
<th>&lt;R,W&gt;</th>
<th>&lt;W,R&gt;</th>
<th>&lt;W,W&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>M</strong></td>
<td>(2,1)</td>
<td>(2,1)</td>
<td>(0,0)</td>
<td>(0,0)</td>
</tr>
<tr>
<td><strong>F</strong></td>
<td>(0,0)</td>
<td>(1,2)</td>
<td>(0,0)</td>
<td>(1,2)</td>
</tr>
</tbody>
</table>

We obtain three pure strategy Nash Equilibria: (M,RR), (M,RW), (F,WW).

But only

(M,RW)

is the the Subgame Perfect Nash Equilibrium.
• A **duopoly** is a market in which two firms compete to supply the same set of customers with the same product. The product is infinitely divisible, e.g., petroleum.

• Cournot’s model: the firms choose the production of quantity $q_i$; since the product is infinitely divisible, the action sets are **continuous**.

• The **market price** of product is assumed to be dependent on the total supply $Q = q_1 + q_2$.
Let $P_0$ and $Q_0$ be some positive numbers.
\[ P(Q) = \begin{cases} 
P_0 \left( 1 - \frac{Q}{Q_0} \right), & \text{if } Q < Q_0 \\
0, & \text{if } Q \geq Q_0 
\end{cases} \]

- The production cost for Firm \( i \) is \( C(q_i) = cq_i \), where \( c \) is a cost of production of one unit and \( P_0 > c \).

- The payoff for Firm \( i \):

\[ r_i(q_1, q_2) = q_i P(Q) - C(q_i) \]
SOLUTION:

- Note that it does not payoff to produce more than $Q_0$ for each firm. Hence, we put the action set as $[0, Q_0]$.

- Best response strategy for Firm 1:

$$r_1(q_1, q_2) = q_1P(Q) - cq_1, \quad \text{and} \quad \frac{\partial r_1}{\partial q_1} (\hat{q}_1, q_2) = 0,$$

thus

$$P_0 \left( 1 - \frac{\hat{q}_1 + q_2}{Q_0} \right) + \hat{q}_1 \left( -\frac{1}{Q_0} \right) - c = 0,$$

and, finally
\[ \hat{q}_1 = \frac{Q_0}{2} \left( 1 - \frac{q_2}{Q_0} - \frac{c}{P_0} \right) \]

- Check that it is really the best response:

\( (I) \quad \frac{\partial^2 r_1}{\partial q_1^2}(\hat{q}_1, q_2) = -\left( \frac{P_0}{Q_0} \right) < 0. \)

\( (II) \quad \text{We also need to have } \hat{q}_1 + q_2 \leq Q_0. \) Indeed,

\[ \hat{q}_1 + q_2 = \frac{Q_0 + q_2}{2} - \frac{cQ_0}{2P_0} \leq Q_0 - \frac{cQ_0}{2P_0} < Q_0. \]
• Similarly, we find the best response to a choice of $q_1$ for Firm 2:

\[
\hat{q}_2 = \frac{Q_0}{2} \left( 1 - \frac{q_1}{Q_0} - \frac{c}{P_0} \right)
\]

• A pure strategy Nash Equilibrium is a pair \((q_1^*, q_2^*)\), each of which is the best response to the other. Such a pair can be found by solving simultaneously equations:

\[
q_1^* = \frac{Q_0}{2} \left( 1 - \frac{q_2^*}{Q_0} - \frac{c}{P_0} \right), \quad q_2^* = \frac{Q_0}{2} \left( 1 - \frac{q_1^*}{Q_0} - \frac{c}{P_0} \right).
\]
The solution is:

$$q_C^* \equiv q_1^* = q_2^* = \frac{Q_0}{3} \left( 1 - \frac{c}{P_0} \right)$$

The payoff when the Nash Equilibrium is played is:

$$r_i (q_C^*, q_C^*) = \frac{P_0 Q_0}{9} \left( 1 - \frac{c}{P_0} \right)^2$$
MONOPOLY VS DUOPOLY

A monopolist maximises its profit:

\[ r_m(q) = qP(q) - cq, \]

and therefore his optimal strategy is:

\[ q_m^* = \frac{Q_0}{2} \left( 1 - \frac{c}{P_0} \right). \]

Because \( q_m^* < 2q_c^* \), the price at which goods are sold is higher for the monopoly.

What about his profit?
\[ r_m(q_m^*) = q_m^* \left( P_0 \left( 1 - \frac{q_m^*}{Q_0} \right) - c \right) = \frac{Q_0 P_0}{4} \left( 1 - \frac{c}{P_0} \right)^2. \]

Hence, competition operates to benefit the consumer!

Suppose that two firms agree to form a cartel and use the strategies: \( q_1 = q_2 = \frac{1}{2} q_m^* \).

The profit for each firm is now:

\[ r_1 \left( \frac{1}{2} q_m^*, \frac{1}{2} q_m^* \right) = r_2 \left( \frac{1}{2} q_m^*, \frac{1}{2} q_m^* \right) = \frac{Q_0 P_0}{8} \left( 1 - \frac{c}{P_0} \right)^2. \]
Remarks:

• The **cartel** profit is greater than the **Cournot** payoff.

• The price paid by the consumers in the **cartel** model is the same as they pay under a **monopoly**.

• **Cartel** models give an **unstable** solution, because the best response to a firm producing the **cartel** quantity is to produce:

\[
\hat{q} = \frac{Q_0}{2} \left( 1 - \frac{q_m^*}{2Q_0} - \frac{c}{P_0} \right) = \frac{3}{4} q_m^* > \frac{1}{2} q_m^*.
\]

Check it!
Suppose now that the decisions are made sequentially:

- Firm 1 (**LEADER**) decides on a quantity to produce.
- This decision is observed by Firm 2 (**FOLLOWER**), which then decides on the quantity that it will produce.

We solve this game by **backward induction** to find a

**Subgame Perfect Nash Equilibrium**.

In other words, we have to:
(I) Find the best response of Firm 2: $\hat{q}_2(q_1)$ for every possible choice of quantity $q_1$;

(II) Given that Firm 1 knows Firm 2’s best response to every choice of $q_1$, find a Nash Equilibrium for this game by determining the maximum payoff that Firm 1 can achieve.

(I): Firm 2’s profit is $r_2(q_1,q_2) = q_2[P(Q) - c]$ and the best response to a choice of $q_1$ is found by solving:

$$\frac{\partial r_2}{\partial q_2}(q_1,\hat{q}_2) = 0,$$

which gives that

$$\hat{q}_2(q_1) = \frac{Q_0}{2} \left( 1 - \frac{q_1}{Q_0} - \frac{c}{P_0} \right).$$
(II): If Firm 1 chooses $q_1$ and Firm 2 chooses the best response $\hat{q}_2(q_1)$, then Firm 1’s profit is

$$r_1(q_1, \hat{q}_2(q_1)) = q_1 \frac{P_0}{2} \left( 1 - \frac{q_1}{Q_0} - \frac{c}{P_0} \right).$$

Hence, Firm 1 maximises its profit at

$$\hat{q}_1 = \frac{Q_0}{2} \left( 1 - \frac{c}{P_0} \right).$$

The Nash Equilibrium is:

$$q_1^* = \frac{Q_0}{2} \left( 1 - \frac{c}{P_0} \right) \quad \text{and} \quad q_2^* = \hat{q}_2(q_1^*) = \frac{Q_0}{4} \left( 1 - \frac{c}{P_0} \right).$$
The Subgame Perfect Nash Equilibrium derived above is sometimes called STACKELBERG EQUILIBRIUM.

It is interesting to note that Firm 1 makes the greater profit because \( q_2^* < q_1^* \).

**Remarks:**

This Stackelberg Equilibrium is not the only a Nash Equilibrium in Stackelberg model.

Another Nash Equilibrium is for Firm 1 to produce the Cournot quantity and for Firm 2 to produce the Cournot quantity *regardless* of the production of Firm 1.
Indeed, if \( q_1 = q_C \), then \( \hat{q}_2 = \frac{Q_0}{2} \left( 1 - \frac{q_C}{Q_0} - \frac{c}{P_0} \right) = q_C \).

Hence, Firm 2’s best response to \( q_1 = q_C^* \) is \( \hat{q}_2 = q_C^* \).

If Firm 2 always chooses \( q_2 = q_C^* \), then Firm 1’s profit is

\[
\begin{align*}
    r_1(q_1, q_C^*) &= q_1 \left[ P_0 \left( 1 - \frac{q_1 + q_C^*}{Q_0} \right) - c \right].
\end{align*}
\]

The best response for Firm 1 to \( q_2 = q_C^* \) is found from

\[
\frac{\partial r_1}{\partial q_1}(\hat{q}_1, q_C^*) = 0,
\]

which gives

\[
\hat{q}_1 = \frac{Q_0}{2} \left( 1 - \frac{q_C^*}{Q_0} - \frac{c}{P_0} \right) = q_C^*.
\]
So Firm 1’s best response to $q_2 = q_c^*$ is $\hat{q}_1 = q_c^*$. Because, for both firms the best response to the other firm producing quantity $q_c^*$ is to produce $q_c^*$, the pair $(q_c^*, q_c^*)$ is a Nash Equilibrium.

Although $q_2 = q_c^*$ is a best response to $q_1 = q_c^*$, it is not a best response to $q_1 \neq q_c^*$. Hence, it is not a Nash Equilibrium in each subgame.

Consequently,

the pair $(q_c^*, q_c^*)$ is not a Subgame Perfect Nash Equilibrium.
WAR OF ATTRITION

LECTURE 6

In a War of Attrition two players compete for a resource of value of \( v \), for example, two supermarkets engaged in a price war.

The strategy for each player is a choice of persistence time \( t_i \).

ASSUMPTIONS OF THE MODEL:

• The cost of the contest is related only to its duration, there are no other costs!

• The player that persists the longest gets all of the resource. If both players quit at the same time, then neither gets the resource.
• The cost paid by each player is proportional to the shortest persistence time chosen, i.e., no costs are incurred after one player quits and the contest ends.

• The payoffs for the players are:

\[ r_1(t_1, t_2) = \begin{cases} 
  v - ct_2, & \text{if } t_1 > t_2 \\
  -ct_1, & \text{if } t_1 \leq t_2 
\end{cases} \]

\[ r_2(t_1, t_2) = \begin{cases} 
  v - ct_1, & \text{if } t_2 > t_1 \\
  -ct_2, & \text{if } t_2 \leq t_1 
\end{cases} \]

where \( c \) is some positive constant.
There are two pure strategy Nash Equilibria:

• A pair \( (t_1^*, t_2^*) \) is a NE, where

\[
\begin{align*}
    t_1^* &= \frac{\nu}{c} \quad \text{and} \quad t_2^* = 0
\end{align*}
\]

giving \( r_1 \left( \frac{\nu}{c}, 0 \right) = \nu \) and \( r_2 \left( \frac{\nu}{c}, 0 \right) = 0 \).

Indeed, note that for every \( t_1 > 0 \), we have \( r_1 \left( t_1, 0 \right) = \nu \) and \( r_1 \left( 0, 0 \right) = 0 \). This yields that

\[
    r_1 \left( t_1, t_2^* \right) \leq r_1 \left( t_1^*, t_2^* \right) \quad \text{for every } t_1.
\]
For Player 2, we have $r_2\left(\frac{v}{c}, t_2\right) = -ct_2 < 0$ for every $t_2 \leq \frac{v}{c}$ and $r_2\left(\frac{v}{c}, t_2\right) = 0$ for every $t_2 > \frac{v}{c}$. Hence, 

$$r_2\left(t_1^*, t_2\right) \leq r_2\left(t_1^*, t_2^*\right)$$

for every $t_2$.

• Clearly, a pair $\left(t_1^{**}, t_2^{**}\right)$ is a second pure strategy NE, where $t_1^{**} = 0$ and $t_2^{**} = \frac{v}{c}$ giving $r_1\left(0, \frac{v}{c}\right) = 0$ and $r_2\left(0, \frac{v}{c}\right) = v$. 
**Mixed strategy Nash Equilibrium**

It is convenient to consider strategies based on the costs of the contest, \( x = c t_1 \) and \( y = c t_2 \). The payoffs are

\[
r_1(x, y) = \begin{cases} 
    v - y, & \text{if } x > y \\
    -x, & \text{if } x \leq y 
\end{cases}
\]

and

\[
r_2(x, y) = \begin{cases} 
    v - c t_1, & \text{if } y > x \\
    -c t_2, & \text{if } y \leq x 
\end{cases}
\]

The mixed strategy \( \sigma_1 \) specifies a choice of cost in the range \( x \) to \( x + dx \) with probability \( p(x)dx \); and \( \sigma_2 \) specifies a choice of cost in the range \( y \) to \( y + dy \) with probability \( q(y)dy \).
The expected payoff to Player 1 if he chooses a fixed cost $x$ against a mixed strategy $\sigma_2^*$ is

$$r_1(x, \sigma_2^*) = \int_0^x (v - y)q(y)dy + \int_x^{+\infty} (-x)q(y)dy.$$ 

The first term arises from the probability that Player 2 chooses cost $y \leq x$, and the second term from the probability that Player 2 chooses $y > x$.

By (extension) of the Equality of Payoffs Theorem we must have

$$r_1(x, \sigma_2^*) = const.$$ 

Hence, for fixed $\sigma_2^*$, $r_1(x, \sigma_2^*)$ is independent of $x$, so

$$\frac{\partial r_1(x, \sigma_2^*)}{\partial x} = 0.$$
Now,

\[
\frac{\partial r_1(x, \sigma_2^*)}{\partial x} = \frac{d}{dx} \int_0^x (v - y)q(y)dy - \int_x^{+\infty} q(y)dy - x \frac{d}{dx} \int_x^{+\infty} q(y)dy.
\]

From the fundamental theorem of calculus we obtain

\[
\frac{d}{dx} \int_0^x (v - y)q(y)dy = (v - x)q(x),
\]

and

\[
\frac{d}{dx} \int_x^{+\infty} q(y)dy = \frac{d}{dx} \left[ 1 - \int_0^x q(y)dy \right] = -q(x).
\]
Summing up, we get

\[
\frac{\partial r_1(x, \sigma_2^*)}{\partial x} = (v - x)q(x) - \int x q(y) dy + x q(x)
\]

\[
= v q(x) - \int x q(y) dy \equiv 0.
\]

We identify \( q(y) \) as an exponential probability density, so we put \( q(y) = k e^{-ky} \), where \( k \) is a normalization constant.

Since,

\[
\int x q(y) dy = k \int e^{-ky} dy = e^{-kx},
\]
we have
\[ vke^{-kx} = e^{-kx} \Rightarrow k = \frac{1}{v}. \]

Hence, \( q(y) = \frac{1}{v} e^{-\frac{y}{v}}. \)

The same argument for Player 1 yields the same distribution in terms of costs \( p(x) = \frac{1}{v} e^{-\frac{x}{v}}, \) i.e., the equilibrium is symmetric.

**Nash Equilibrium** in terms of distribution of persistence times \( \tilde{p}(t) \) chosen:

\[
\int_{0}^{\infty} p(x)dx = \int_{0}^{\infty} \frac{1}{v} e^{-\frac{x}{v}}dx = \left\{ \begin{array}{l} x = ct \\ dx = cdt \end{array} \right\} = \int_{0}^{\infty} \frac{c}{v} e^{-\frac{ct}{v}} dt = \int_{0}^{\infty} \tilde{p}(t) dt
\]
Hence, the equilibrium distributions for both players are exponential with mean \( \frac{v}{c} \), i.e.

\[
\tilde{p}(t) = \tilde{q}(t) = c e^{-\frac{ct}{v}}
\]

The expected payoffs for players under this equilibrium can be calculated

• either explicitly from the formula

\[
r_1(\sigma_1^*, \sigma_2^*) = \int_{0}^{+\infty} p(x) \left[ \int_{0}^{x} (v - y)q(y)dy - x \int_{x}^{+\infty} q(y)dy \right] dx,
\]
or by observing that \( r_1(0, \sigma_2^*) = 0 \) and \( \frac{\partial r_1(x, \sigma_2^*)}{\partial x} = 0 \) imply that \( r_1(x, \sigma_2^*) = 0 \) for every \( x \); therefore \( r_1(\sigma_1, \sigma_2^*) = 0 \) for all \( \sigma_1 \), and particularly for \( \sigma_1^* \).

Note: Although \( \tilde{p}(t) \) is the distribution of persistence time, it is not the distribution of contest duration \( D \)

\[
P(D \leq t) = 1 - P(D > t) = 1 - P(\text{neither player has quit before } t).
\]

Now, \( P(\text{Player i does not quit before } t) = \int_{t}^{+\infty} \tilde{p}(\tau)d\tau = e^{-\frac{ct}{v}} \), and

\[
P(D \leq t) = 1 - e^{-\frac{2ct}{v}}.
\]
Kakutani Fixed Point Theorem

Let $K$ be a convex and compact set of an Euclidean space $\mathbb{R}^n$. Let $F$ be a mapping such that

- for every $x \in K$, $F(x)$ is convex and closed in $K$
- the graph of $F$ be closed.

Then there exists a point $x_0$ such that $x_0 \in F(x_0)$.

**NOTE:** Glicksberg (1952) extended this result on infinite dimensional spaces.
**Definition 1:** Consider a non-empty subset $K \subset \mathbb{R}^n$ and a correspondence $F : X \to X$. The graph of $F$ is said to be closed if the following set

$$gph(F) \equiv \{(x, y) \in K \times K : y \in F(X)\}$$

is closed in $K \times K$.

The graph of $F$ is closed if and only if for all sequences $\{x_n\}_{n \in \mathbb{N}}$ and $\{y_n\}_{n \in \mathbb{N}}$ in $K$ satisfying

- $\lim_{n \to \infty} x_n = x$,
- $\lim_{n \to \infty} y_n = y$,
- $y_n \in F(x_n)$ for each $n \in \mathbb{N}$,

we get $y \in F(x)$. 
Definition 2: A subset \( K \subset \mathbb{R}^n \) is **compact** if it is bounded:

\[
\exists M > 0 \quad \forall x \in K \quad \|x\| \leq M
\]

and closed, i.e., for every sequence \( \{x_n\}_{n \in \mathbb{N}} \) in \( K \) converging to some \( x \in \mathbb{R}^n \) the limit vector \( x \) still belongs to \( K \).
Lemma 1: A subset $K \subset \mathbb{R}^n$ is compact if for every sequence $\{x_n\}_{n \in \mathbb{N}}$ in $K$ there exists a subsequence $\{x_{n_k}\}_{k \in \mathbb{N}}$ such that $\lim_{k \to \infty} x_{n_k} = x$.

Nash theorem for compact pure strategy sets

Consider a strategic form of an $n$-person game whose strategy spaces $S_i$, $i = 1, \ldots, n$ are non-empty compact subsets of $\mathbb{R}^n$. If the payoff functions are continuous then there exists a Nash Equilibrium in mixed strategies.
Economists are often interested in a **pure strategy Nash Equilibria**. In order to obtain such an equilibrium, additional assumptions are needed.

**Definition 3:** A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is **quasi-concave** if the sets $\{x : f(x) \geq c\}$ are convex for all $c$.

For every sequence $\{x_n\}_{n \in \mathbb{N}}$ in $K$ converging to some $x \in \mathbb{R}^n$ the limit vector $x$ still belongs to $K$.

**Lemma 2:** Consider a function $u : C \rightarrow \mathbb{R}$, where $C$ is a convex and non-empty subset of $\mathbb{R}^n$. The function $u$ is quasi-concave if and only if

$$u(\alpha x + (1 - \alpha)y) \geq \min \{u(x), u(y)\}$$

for any $x, y \in C$ and $\alpha \in [0, 1]$. 
Lemma 3: Consider a function \( u : C \rightarrow \mathbb{R} \), where \( C \) is a convex and non-empty subset of \( \mathbb{R}^n \). If \( u \) is concave then it is quasi-concave.

Example of a quasi-concave function:
General existence result
(Debreu 1952; Glicksberg 1952; Fan 1952)

Consider a strategic form of an $n$-person game for which

- strategy spaces $S_i, \ i = 1, \ldots, n$ are non-empty compact subsets of $\mathbb{R}^n$,
- the payoff function $r_i$ is continuous on $S = S_1 \times \cdots \times S_n$,
- the payoff function $r_i$ is quasi-concave in $s_i$.

Then there exists at least one pure strategy Nash Equilibrium.
The proof is based on verifying that the continuous payoffs imply non-empty, closed graph of best response correspondence, and that quasi-concavity in players’ own action implies convex-valued best response correspondence.

Remark: If the payoff functions are not continuous, then the best response correspondence can fail to have a closed graph and/or fail to be non-empty. The latter problem arises, because discontinuous function need not attain a maximum, for example,

\[
f(x) = \begin{cases} 
-|x|, & x \neq 0 \\
-1, & x = 0
\end{cases}
\]
Example: Consider the following two-player game with strategy sets $S_1 = S_2 = [0, 1]$ and payoff functions:

$$r_1(s_1, s_2) = -(s_1 - s_2)^2$$

$$r_2(s_1, s_2) = \begin{cases} 
- \left( s_1 - s_2 - \frac{1}{3} \right)^2, & s_1 \geq \frac{1}{3} \\
- \left( s_1 - s_2 + \frac{1}{3} \right)^2, & s_1 < \frac{1}{3}.
\end{cases}$$

Here each player’s payoff is strictly concave in his own strategy, and a best response exists (and is unique) for each strategy of the opponent.
However, the game does not have a pure strategy Nash Equilibrium:

- Player 1’s best response is always to choose $s_1 := s_2$;
- Player 2’s best response is to choose $s_2 := s_1 - \frac{1}{3}$ for $s_1 \geq \frac{1}{3}$ and $s_1 + \frac{1}{3}$ for $s_1 < \frac{1}{3}$.

Best response functions do not intersect.
Robert Aumann in the article [J. Math. Econ. 1 (1974)]:

“Subjectivity and correlation in randomized strategies”,

proposed a mechanism that permits the players to enter into preplay arrangements so that their strategy choices could be correlated instead of independently chosen.
Example proposed by Aumann:

The game has two pure strategy and one mixed strategy Nash Equilibrium:

$$(U,L), \ (D,R) \text{ and } \sigma_1^* = \sigma_2^* = \left( \frac{1}{2}, \frac{1}{2} \right).$$

The corresponding payoffs are:

- (5,1),
- (1,5),
- (2.5,2.5).
Now, if the players agree to play by jointly observing a coin flip and playing (U, L) if the result is head and (D,R) if the result is tail, then the expected outcome is (3,3), which is a convex combination of two pure strategy Nash Equilibrium outcomes.

Observe that this way of playing the game defines an equilibrium for the following extensive game:
We now push the example one step further by assuming that the players agree to play according to the following mechanism:

* *a random device selects one cell in the game matrix with the probabilities*

\[
\begin{array}{cc}
  & L & R \\
 U & 1/3 & 0 \\
 D & 1/3 & 1/3 \\
\end{array}
\]

When a cell is selected, then each player is told to play the corresponding pure strategy. The trick is that a player is told what to play but is not told what the recommendation to the other player is. The information is *not public* any more!
When Player 1 receives the signal “play D” he only knows that with probability 1/2 the other player has been told to play R. When Player 1 receives the signal “play U” he knows that the other will play L with probability 1.

Consider now what Player 1 can do if he assumes that Player 2 plays according to the recommendation.

- If Player 1 has been told “play D” and if he plays so, he expects: 
  \[ \frac{1}{2} \times 4 + \frac{1}{2} \times \frac{5}{2} = \frac{5}{2} \], but if he plays U instead, he expects: 
  \[ \frac{1}{2} \times 5 + \frac{1}{2} \times 0 = \frac{5}{2} \].

  Hence, he **cannot improve** his expected payoff.

- If Player 1 has been told “play U” and if he plays so, he gets 5, whereas he plays D he expects 4.
**Conclusion:** For Player 1 obeying to the recommendation is the best reply to Player 2’s behaviour when he himself plays according to the suggestion of the signalling scheme.

Now we can repeat the verification for Player 2.

- If he has been told “play L” he expects $\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 4 = \frac{5}{2}$, but if he plays R instead, he expects $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 5 = \frac{5}{2}$. Hence, cannot improve his expected payoff.

- If he has been told “play R” he expects 5, otherwise 4.

Thus, we have checked that an equilibrium property holds for this scheme of playing the game.

**Each player expects:** $\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 5 + \frac{1}{3} \cdot 4 = 3.3333...$
Extensive form of game leading to the correlated equilibrium:
Remarks: By expanding the game via the adjunct of a first stage where the Nature plays and gives information to the players, a new class of equilibria can be reached, which dominate some of the original Nash Equilibria.

(1) If the random device gives information which is common to all players, then it permits a mixing of different pure strategy Nash Equilibria and the outcome is in the convex hull of the Nash Equilibrium outcomes. Such an equilibrium is sometimes called SUNSPOT EQUILIBRIUM.
(2) If the random device gives information which can be different from one player to the other, then the correlated equilibrium can have an outcome which lies outside of the convex hull. Such an equilibrium is called CORRELATED EQUILIBRIUM.
**Definition:** A correlated equilibrium in $n$-player game is any probability distribution $p(\cdot)$ over the set of pure strategies of the players, i.e., $S = S_1 \times \cdots \times S_n$ such that for every player $j$ and any mapping $d_j : S_j \to S_j$, it holds

$$
\sum_{s \in S} p(s) r_j(s_j, s_{-j}) \geq \sum_{s \in S} p(s) r_j(d(s_j), s_{-j})
$$

where $(\tilde{s}_j, s_{-j}) = (s_1, \ldots, s_{j-1}, \tilde{s}_j, s_{j+1}, \ldots, s_n)$.
In a game of incomplete information some players do not know exactly what the characteristics of other players are.

Example: Player 1 can be of two types, called $\theta_1$ and $\theta_2$. The payoffs matrices are the following:
Extensive form of this game:
J. Harsanyi proposed a transformation of a game with incomplete information into a game with *imperfect information*. This transformation introduces a preliminary “chance move” played by the Nature which decides randomly the type $\theta_i$ of Player 1. The probability of each type is $p_1$ and $p_2 = 1 - p_1$, and represents the beliefs of Player 2. One assumes that Player 1 also knows the beliefs, thus the prior probabilities are common knowledge.

- Let $x_i$ be a probability of choosing U and $1 - x_i$ be a probability of choosing D when the Player 1 is of type $\theta_i$.

- Let $y$ be a probability of choosing L and $1 - y$ be a probability of choosing R. In other words, $(y, 1 - y)$ is a mixed strategy for Player 2.
We can define the best response of Player 1 to a mixed strategy \((y, 1-y)\) of Player 2.

- If Player 1 is of type \(\theta_1\), then \(\max_{i=1,2} a_{i1}^{\theta_1} y + a_{i2}^{\theta_1} (1-y)\) is a best response.

- If Player 1 is of type \(\theta_2\), then \(\max_{i=1,2} a_{i1}^{\theta_2} y + a_{i2}^{\theta_2} (1-y)\) is a best response.

Here, \(a_{ij}^{\theta_l}\), \(b_{ij}^{\theta_l}\) are the payoffs for Player 1 and Player 2, when Player 1 is of type \(\theta_l\). From above points we get:

- \(\theta_1 \implies \max \{0y + 2(1-y), 2y + 3(1-y)\}\) and
- \(\theta_2 \implies \max \{3/2 y + 7/2 (1-y), 2y + 3(1-y)\}\)
We can now define the optimal response of Player 2 to the pair of mixed strategy \((x_i, 1 - x_i)\) of Player 1 by solving

\[
\max_{j=1,2} \ p_1 (x_1 b_{1j}^{\theta_1} + (1 - x_1) b_{2j}^{\theta_1}) + p_2 (x_2 b_{1j}^{\theta_2} + (1 - x_2) b_{2j}^{\theta_2})
\]

Note that if Player 1 is of type \(\theta_1\), he always chooses a strategy D, i.e., \(x_1 = 0\) (see Figure).
Therefore, the best reply conditions for Player 2 can be written as follows:

$$\max \left\{ 1p_1 + p_2(x_2(-1) + 1(1-x_2), p_2(0x_2 + 0(1-x_2)) + 0p_1 \right\}$$

which is $$\max \left\{ 1 - 2(1-p_1)x_2, 0 \right\}$$.

Hence, Player 2 chooses

- strategy L, if $$\frac{1}{2(1-p_1)} > x_2$$,
- strategy R, if $$\frac{1}{2(1-p_1)} < x_2$$,
- any mixed strategy $$y \in [0,1]$$, if $$\frac{1}{2(1-p_1)} = x_2$$. 

Conclusion: The equilibria of this game are characterized as follows

- $x_1 \equiv 0$,

- if $p_1 \leq \frac{1}{2}$, then $x_2 = 0$, $y = 1$ or $x_2 = 1$, $y = 0$ or $x_2 = \frac{1}{2(1 - p_1)}$, $y = \frac{1}{2}$,

- if $p_1 > \frac{1}{2}$, then $x_2 = 0$, $y = 1$. 
**Definition:** Each player \( j \) (in \( n \)-player game) may be of different possible types. Let \( \Theta_j \) be the finite set of types for Player \( j \). Whatever his type is, Player \( j \) has the same set of pure strategies \( S_j \). If \( \theta = (\theta_1, \ldots, \theta_n) \in \Theta = \Theta_1 \times \cdots \times \Theta_n \) is a type specification for every player, then the normal form of the game is specified by the payoff functions:

\[
r_j(\theta; \bullet): S_1 \times \cdots \times S_n \to \mathbb{R}, \quad j = 1, \ldots, n. \quad (*)
\]

A prior probability distribution \( p(\theta_1, \ldots, \theta_n) \) on \( \Theta \) is given as common knowledge. Furthermore, we shall assume that all the marginal probabilities are nonzero, i.e., \( p_j(\theta_j) > 0 \) for all \( j = 1, \ldots, n \).
When Player $j$ observes that he is of type $\theta_j \in \Theta_j$, he can construct his revised conditional probability distribution on the types $\theta_{-j}$ of the other players through the Bayes formula:

$$p(\theta_{-j} | \theta_j) = \frac{p([\theta_j, \theta_{-j}])}{p_j(\theta_j)}.$$

We can now introduce an expanded game where the Nature draws randomly, according to the prior probability distribution $p(\bullet)$, a type vector $\theta \in \Theta$ for all players. Player $j$ observes only his type and picks a strategy from his strategy set $S_j$. The outcome is then defined by (*).
A **BAYESIAN EQUILIBRIUM** is a Nash Equilibrium in the expanded game, in which each player’s pure strategy $g_j$ is a map from $\Theta_j$ to $S_j$.

The payoff to Player $j$ associated with a strategy profile $g = (g_1, ..., g_n)$ is given by

$$R_j (g) = \sum_{\theta_j \in \Theta_j} p_j (\theta_j) \sum_{\theta_{-j} \in \Theta_{-j}} p(\theta_{-j} | \theta_j) r_j ([\theta_j, \theta_{-j}]; g_1(\theta_1), ..., g_n(\theta_n)).$$

As usual a **Nash Equilibrium** is a strategy profile $g^*$ such that for every $g_j$ and $j = 1, ..., n$ it holds

$$R_j (g^*) \geq R_j (g_j, g^*_j).$$
BAYESIAN EQUILIBRIUM

LECTURE 8

BARGAINING UNDER INCOMPLETE INFORMATION


THE MODEL

Single seller of an indivisible good faces a single potential buyer. A bargain is concluded if and only if the good is transferred at a mutually acceptable price.
Let
• \( v_s \in [0,1] \) denote the seller’s reservation price (the smallest monetary sum he will accept in exchange for the good),
• \( v_b \in [0,1] \) denote the buyer’s reservation price (the greatest sum he is willing to pay for the good).

The bargain is struck when the sale price \( P \) satisfies
\[
    v_s \leq P \leq v_b.
\]

**Incomplete information** is modeled by the following situation: Each party knows his reservation price, but is uncertain about his adversary’s, assessing a subjective
probability distribution over the range of possible values that his opponent may hold. Specifically, the buyer regards \( v_s \) as a random variable possessing a c.d.f. \( F_b(v_s) \) satisfying \( F_b(0) = 0 \) and \( F_b(1) = 1 \). Moreover, \( F_b(\bullet) \) is strictly increasing on \([0,1]\) and has a density function. Similarly, the seller’s knowledge of \( v_b \) is summarized by \( F_s(v_b) \) which satisfies \( F_s(0) = 0 \) and \( F_s(1) = 1 \). \( F_s(\bullet) \) is also strictly increasing on \([0,1]\) and has a density function. These distribution functions are common knowledge, i.e., each side knows these distributions, knows that the latter knowledge is known, etc.
**Bargaining Rule**

**Seller** and **buyer** submit sealed offers, $s$ and $b$, respectively. If

- $s \leq b$, a bargain is enacted and the good is sold at price, $P = kb + (1 - b)s$, where $0 \leq k \leq 1$, and if
- $s > b$, there is no sale and no money transfer.

If $k = \frac{1}{2}$, then the rule determines a final sale price by splitting the difference between the player offers. Both offers carry equal weight in determining the sale price.

We shall consider this case!
**Objective:** Characterize Bayesian equilibrium solutions in this one-stage bargain game.

- In the event of agreement the seller has a profit $P - v_s$ and the buyer $v_b - P$.
- In the event of no agreement each side earns 0.

Each bargainer makes the offer to maximize his expected profit.

Notation:

$b = B(v_b)$ and $s = S(v_s)$ indicate that the players’ offers depend on their reservation prices. The functions $B(\bullet)$ and $S(\bullet)$ are referred to as the player offer strategies.
• The expected profit for the buyer is given by

\[ r_b(b, v_b) = \begin{cases} 
\int_0^b \left( v_b - \frac{b + s}{2} \right) g_b(s) ds, & b > 0 \\
0, & b \leq 0.
\end{cases} \]

Here, \( g_b(\cdot) \) is a density function of the seller induced by \( S(\cdot) \) and \( F_b(\cdot) \), i.e.,

\[ G_b(\zeta) = \Pr(S(v_s) \leq \zeta) = \Pr(v_s \leq S^{-1}(\zeta)) = F_b(S^{-1}(\zeta)). \]
The expected profit for the seller is given by

\[
    r_s(s, v_s) = \begin{cases} 
        \int_s^1 \left( \frac{b + s}{2} - v_s \right) g_s(b) db, & s < 1 \\
        0, & s \geq 1.
    \end{cases}
\]

Here, \( g_s(\cdot) \) is a density function of the seller induced by \( B(\cdot) \) and \( F_s(\cdot) \), i.e.,

\[(2)\]

\[
    G_s(\beta) = \Pr(B(v_b) \leq \beta) = \Pr(v_b \leq B^{-1}(\beta)) = F_s(B^{-1}(\beta)).
\]
We look for a pure strategy equilibrium.
Conditional on $v_s$ the seller’s offer $s^*$ is a best response against $B(\bullet)$ if $r_s(s^*, v_s) \geq r_s(s, v_s)$ for all $s \in [0,1]$. Similarly, the buyer of type $v_b$ makes offer $b^*$ that is a best response against $S(\bullet)$ if $r_b(b^*, v_b) \geq r_b(b, v_b)$ for all $b \in [0,1]$. A pair of best response offer strategies constitutes a Bayesian (Nash) equilibrium.

**Proposition**

Under the sealed offer bargaining rule, the equilibrium bargaining strategies of the buyer and seller are increasing in the respective reservation prices.
Prove this proposition.

**Hint:** Note that, for example, for the buyer we have the following inequalities: \( r_b(b, v_b) \geq r_b(b', v_b) \) and \( r_b(b', v_b') \geq r_b(b, v_b') \).

Let us now consider the maximization problem for the seller of type \( v_s \), i.e.,

\[
\max_{s \in [0,1]} r_s(s, v_s) = \max_{s \in [0,1]} \int_s^1 \left( \frac{s + b}{2} - v_s \right) G_s(db).
\]

Note that

\[
\int_s^1 \left( \frac{s + b}{2} - v_s \right) G_s(db) = \left[ \left( \frac{s + b}{2} - v_s \right) G_s(b) \right]_{b=s}^{b=1} - \frac{1}{2} \int_s^1 G_s(b)db.
\]
Hence,

\[
\int_{s}^{1} \left( \frac{s + b}{2} - v_s \right) G_s(db) = \left( \frac{s + 1}{2} - v_s \right) - (s - v_s) G_s(s) - \frac{1}{2} \int_{s}^{1} G_s(b) db.
\]

First order condition \( \frac{\partial r_s(s, v_s)}{\partial s} = 0 \) implies that

\[
\frac{1}{2} - G_s(s) - (s - v_s) g_s(s) + \frac{1}{2} G_s(s) = 0.
\]

Finally, we obtain that
\[
\frac{1}{2} (1 - G_s(s)) = (s - v_s) g_s(s). \quad (*)
\]

Let us now consider the maximization problem for the buyer of type \( v_b \), i.e.,

\[
\max r_b(b, v_b) = \max_{b \in [0,1]} \int_0^b \left( v_b - \frac{s + b}{2} \right) G_b(ds).
\]

Note that

\[
\int_0^b \left( v_b - \frac{s + b}{2} \right) G_b(ds) = \left[ \left( v_b - \frac{s + b}{2} \right) G_b(s) \right]_{s=0}^{s=b} + \frac{1}{2} \int_0^b G_b(s)ds.
\]
Hence,
\[
\int_{0}^{b} \left( v_b - \frac{s + b}{2} \right) G_b(ds) = (v_b - b)G_b(b) + \frac{1}{2} \int_{0}^{b} G_b(s)ds.
\]

First order condition \( \frac{\partial r_b(b, v_b)}{\partial b} = 0 \) implies that

\[
(v_b - b)g_b(b) - G_b(b) + \frac{1}{2} G_b(b) = 0.
\]

Finally, we obtain the following condition

\[
(v_b - b)g_b(b) = \frac{1}{2} G_b(b). \quad (**)
\]
Let us further assume that $F_b(\cdot)$ and $F_s(\cdot)$ are uniform distributions on $[0,1]$. We look for a linear strategies:

$$S(v_s) = \alpha_1 + \beta_1 v_s = s,$$

$$B(v_b) = \alpha_2 + \beta_2 v_b = b,$$

where $\beta_i \neq 0$ for $i=1,2$.

Then, by (1) we get that

$$G_b(s) = F_b(S^{-1}(s)) = F_b\left(\frac{s - \alpha_1}{\beta_1}\right) = \frac{s - \alpha_1}{\beta_1},$$

which implies that $g_b(s) = \frac{1}{\beta_1}$ for $s \in [\alpha_1, \alpha_1 + \beta_1]$. 
Analogously, by (2) we have that

\[ G_s(b) = F_s(B^{-1}(b)) = F_s\left(\frac{b - \alpha_2}{\beta_2}\right) = \frac{b - \alpha_2}{\beta_2}, \]

which implies that \( g_s(b) = \frac{1}{\beta_2} \) for \( b \in [\alpha_2, \alpha_2 + \beta_2] \).

We plug \( g_s(\bullet) \) into (*) and \( g_b(\bullet) \) into (**) and have:

\[ \frac{1}{2}\left(1 - \frac{s - \alpha_2}{\beta_2}\right) = \frac{1}{\beta_2}(s - v_s) \]

\[ \frac{1}{\beta_1}\left(v_b - (\alpha_2 + \beta_2 v_b)\right) = \frac{1}{2}\frac{b - \alpha_1}{\beta_1}. \]
After some rearrangements we obtain

\[
\begin{cases}
  v_b (2 - 3\beta_2) = 3\alpha_2 - \alpha_1 \\
  v_s (-2 + 3\beta_1) = -3\alpha_1 + \alpha_2 + \beta_2 .
\end{cases}
\]

Since \( v_b \) and \( v_s \) vary from 0 to 1, then it must hold that

- \( 2 - 3\beta_2 = 0 \implies \beta_2 = \frac{2}{3}, \ 3\beta_1 - 2 = 0 \implies \beta_1 = \frac{2}{3} , \)

- \( 3\alpha_2 = \alpha_1, \ 3\alpha_1 - \alpha_2 = \frac{2}{3} , \) which yields that \( \alpha_1 = \frac{1}{4} \) and \( \alpha_2 = \frac{1}{12} . \)
Hence, we have obtained that

\[ S(v_s) = \frac{1}{4} + \frac{2}{3} v_s \]

\[ B(v_b) = \frac{1}{12} + \frac{2}{3} v_b \]

In equilibrium the parties trades if and only if it holds

\[ S(v_s) \leq B(v_b) \Rightarrow v_s + \frac{1}{4} \leq v_b \]

.
Moreover, due to rational assumptions made at the beginning: $S(v_s) \geq v_s$ and $B(v_b) \leq v_b$. These inequalities imply that $v_s \leq 3/4$ and $v_b \geq 1/4$. Trade is enacted, if $v_b \geq v_s$ - but this inequality holds (see above).

The above conditions mean that the seller never sells his good below his own reservation price, and the buyer never buys the product for the higher price than his own reservation price. Clearly, $v_b \in [1/4,1]$ and $v_s \in [0,3/4]$. 
Figure: Equilibrium strategies for the seller and the buyer

\[ S(v_s) \quad \text{and} \quad B(v_b) \]
What is the probability that the bargain will be reached?

This probability is: \( \Pr \left( v_b \geq \frac{1}{4} + v_s, \, v_s \leq \frac{3}{4}, \, v_b \geq \frac{1}{4} \right) = (*) \).

Clearly, \( v_b, v_s \) are independent random variables. Hence,

\[
\Pr \left( v_b \geq \frac{1}{4} + v_s, \, v_s \leq \frac{3}{4}, \, v_b \geq \frac{1}{4} \right) = \frac{9}{32}
\]
The expected profit for the buyer:
\[ \int_0^{b^*} \left( v_b - \frac{s + b^*}{2} \right) G_b(ds), \]
where \( G_b(s) = \frac{s - 1/4}{2/3} \) for \( s \in [1/4, 11/12] \). Hence,
\[ \frac{2}{3} v_b + \frac{1}{12} = \int_{1/4}^{s + 2/3 v_b + 1/12} \left( v_b - \frac{s + 2/3 v_b + 1/12}{2} \right) \frac{3}{2} ds + \int_0^0 ds = \frac{1}{2} \left( v_b - \frac{1}{4} \right)^2 \]
and \[ \int_{1/4}^{1} \frac{1}{2} \left( v_b - \frac{1}{4} \right)^2 dv_b = \frac{9}{128}. \]
Calculate the profit for the seller.
• Recall the payoff matrix for one-stage Prisoners’ Dilemma:

\[
\begin{array}{c|cc}
& Q & C \\
\hline
Q & (3,3) & (0,5) \\
C & (5,0) & (1,1) \\
\end{array}
\]

Q-Keep Quiet
C-Confess

• In two-stage game the pure strategy sets
are: \{QQ,CQ,QC,CC\}.

We solve this game by backward induction: in the final stage there is no further interaction so the only consequence is the payoff to be gained: (1,1). Now consider the first stage; the payoffs in this stage can be calculated by adding the payoffs for the Nash Equilibrium in the second stage (we are only interested in a Subgame Perfect Nash Equilibrium):

\[
\begin{array}{ccc}
\text{QC} & \text{CC} \\
\text{QC} & (4,4) & (1,6) \\
\text{CC} & (6,1) & (2,2)
\end{array}
\]

Q-Keep Quiet  C-Confess

\[NE \Rightarrow \{CC, CC\}\]
• Consider now an infinite number of repeats, indexed by \( t=0,1,\ldots \).

If there is no end of the game, then there is no last stage to work backwards from. Since all subgames look the same, this suggests that we should look for stationary strategies.

**Definition:** A *stationary strategy* is a strategy in which the rule for choosing an action is the same in every stage.

Note that this does not necessarily mean that the action chosen in each stage must be the same!

Examples: (a) Play Q in every stage or (b) Play Q if the other player has never played C and play C otherwise.
Let $s$ be a profile of strategies used by the players and $r_i(t)$ is the payoff to Player $i$ received at the stage $t$ of the game.

Then the total payoff is: $\sum_{t=0}^{\infty} r_i^s(t)$. But observe that

$$\sum_{t=0}^{\infty} r_i^{s_QsQ}(t) = \sum_{t=0}^{\infty} 3 = +\infty,$$
where $s_Q$ - play always $Q$, and

$$\sum_{t=0}^{\infty} r_i^{sCsC}(t) = \sum_{t=0}^{\infty} 1 = +\infty,$$
where $s_C$ - play always $C$.

**IDEA:** TO DISCOUNT FUTURE PAYOFFS BY A DISCOUNT FACTOR $\delta \in (0,1)$. 

Then the discounted payoff to Player $i$, when the profile $s = (s_1, s_2)$ is used, equals: $R_i(s_1, s_2) = \sum_{t=0}^{\infty} \delta^t r^s_i(t)$.

Depending on the situation the discount factor represents, for example, the uncertainty about whether the game will continue (with the probability $\delta$) or $\delta = \frac{1}{1+r}$, where $r$ is an interest rate. In the latter case, when $V$ is an initial amount, then in consecutive periods we get $V(1+r)$, $V(1+r)^2 = (V(1+r))(1+r)$ and so on. Hence, $PV = \frac{FV_T}{(1+r)^T}$. 
Examples:  
\[ R_i(s_Q, s_Q) = \sum_{t=0}^{\infty} 3\delta^t = \frac{3}{1-\delta}, \]
\[ R_i(s_C, s_C) = \sum_{t=0}^{\infty} \delta^t = \frac{1}{1-\delta}. \]

Let us now consider the *trigger* strategy (sometimes it is called **GRIM**): \( s_G \) - start by playing \( Q \) and continue until the other player confesses, then use \( C \) forever after. If both players adopt this strategy then permanent cooperation gives that \( R_i(s_G, s_G) = \frac{3}{1-\delta} \).

Is the strategy pair \((s_G, s_G)\) a **Nash Equilibrium**?
Let us restrict the pure strategy sets for both players to the set:

\[ S = \{ s_G, s_Q, s_C \}. \]

- Suppose that Player 1 uses strategy \( s_Q \) instead. Then,

\[ R_1(s_Q, s_G) = \frac{3}{1 - \delta}. \]

Hence, Player 1 cannot do better by switching to \( s_Q \).

- Suppose that Player 1 uses strategy \( s_C \) instead. Then,

\[ R_1(s_C, s_G) = 5 + \delta + \delta^2 + ... = 5 + \frac{\delta}{1 - \delta}, \]

because the sequence of actions is as follows:
Hence, Player 1 cannot do better by switching to $s_C$, if

$$5 + \frac{\delta}{1-\delta} \leq \frac{3}{1-\delta} \quad \Rightarrow \quad \delta \geq \frac{1}{2}. $$

Thus, the pair of strategies $(s_G, s_G)$ is a Nash Equilibrium for sufficiently large discount factor. However, $(s_G, s_G)$ is not a Subgame Perfect Nash Equilibrium.
To show this claim consider the possible outcomes:

(i) neither player has played C,
(ii) both players have played C,
(iii) Player 1 used C in the last stage but Player 2 did not,
(iv) Player 2 used C in the last stage but Player 1 did not.

Note that in cases (i) and (ii) there is no conflict with the concept of **SPNE**. Indeed, for example, in (ii) both players have played C so the NE pair \((s_G, s_G)\) specifies that each player should play C forever. Hence, the strategy pair \((s_C, s_C)\) adopted in this case is a NE of the subgame, because it is a NE of the entire game.
In case (iii) $s_Q$ specifies that Player 2 should switch to using $C$ forever. However, Player 2 has not yet played $C$, so Player 1 should continue to use $s_G$. Thus, a NE for the whole game specifies that the strategy pair $(s_G, s_C)$ should be adopted in the subgame. However, this pair is not a NE for the subgame, because Player 1 could obtain a greater payoff by using $s_C$. Indeed, the corresponding payoffs in this subgame are $R_1(s_G, s_C) = 0 + \delta + \delta^2 + ... = \frac{\delta}{1 - \delta}$ and $R_1(s_C, s_C) = 1 + \delta + \delta^2 + ... = \frac{1}{1 - \delta}$. 
Although \((s_G, s_G)\) is not a SPNE, a very similar strategy lead to a SPNE, when it is adopted by both players:

\[
\begin{align*}
\text{\textcolor{red}{s^*}} & - \text{ start by being quiet and continue to play } Q \text{ until either player confesses, then play } C \text{ forever after.}
\end{align*}
\]

The pair \((s^*, s^*)\) is a Subgame Perfect Nash Equilibrium, because it specifies that the players should play \((s_C, s_C)\) in the subgames (iii) and (iv).
So far, we have allowed repeated games to have only limited set of strategies. Will the Subgame Perfect Nash Equilibrium change, if we allow more general strategies?

**Definition:** A pair of strategies \((\sigma_1, \sigma_2)\) satisfies the **ONE-STAGE DEVIATION CONDITION** if neither player can increase his payoff by deviating (unilaterally) from his strategy in any single stage and returning to the specified strategy thereafter.

Example (cont.) Does a strategy pair \((s^*, s^*)\) satisfies one-stage deviation principle?
The game can be in one of two classes of subgame:

(I) both players have always used $Q$, or

(II) at least one player has used $C$ in the previous round.

\[
\begin{array}{c|ccccc}
& t-1 & t & t+1 & t+2 & t+3 \\
\hline
\text{Player 1} & Q & C & \varnothing & C & Q & C & Q & \ldots \\
\text{Player 2} & Q & Q & C & Q & C & Q & C & Q & \ldots \\
\end{array}
\]

\[
5 + \frac{\delta}{1 - \delta} \leq R_1(s_Q, s_Q) = \frac{3}{1 - \delta}
\]

for \( \delta \geq \frac{1}{2} \)
Hence, \((s^*, s^*)\) satisfies one-stage deviation condition provided that \(\frac{1}{2} \leq \delta < 1\).
THEOREM

A pair of strategies is a **Subgame Perfect Nash Equilibrium** for a discounted repeated game if and only if it satisfies **one-stage deviation condition**.

This theorem is called **ONE-STAGE DEVIATION PRINCIPLE**.
**Proof:** For finitely repeated games, the equivalence of **DP** and **DC** is obtained via backward induction (it shows that a **SPNE** payoff cannot be improved by deviating in any finite number of stages).

Because the definition of subgame perfection implies the one-stage deviation condition for finitely and infinitely repeated games, it only remains to prove that the one-stage deviation condition implies **SPNE** in an infinitely repeated games.

The proof of this fact is by a contradiction.
On the contrary, suppose that a strategy pair \((\sigma_1, \sigma_2)\) satisfies \textbf{DC} but it is not a \textbf{SPNE}. It follows that there is some stage \(t\) at which it would be better for one of the players, say Player 1, to adopt a different strategy \(\hat{\sigma}_1\).

This means that there exists \(\varepsilon > 0\) such that
\[
R_1(\hat{\sigma}_1, \sigma_2) - R_1(\sigma_1, \sigma_2) > 2\varepsilon.
\]

Now consider a strategy \(\sigma'_1\) defined as follows:

\begin{itemize}
  \item \(\sigma'_1 = \hat{\sigma}_1\) from stage \(t\) to stage \(T\), and
  \item \(\sigma'_1 = \sigma_1\) from stage \(T\) onwards.
\end{itemize}
Because the future payoffs are discounted, then

$$|R_1(\hat{\sigma}_1, \sigma_2) - R_1(\sigma'_1, \sigma_2)| \propto \delta^{T-t},$$

so we can choose $T$ sufficiently large such that

$$R_1(\hat{\sigma}_1, \sigma_2) - R_1(\sigma'_1, \sigma_2) < \varepsilon$$

and

$$R_1(\sigma'_1, \sigma_2) - R_1(\hat{\sigma}_1, \sigma_2) > -\varepsilon.$$
Combining these inequalities we get that

\[ R_1(\sigma'_1, \sigma_2) - R_1(\sigma_1, \sigma_2) > \varepsilon. \]

But \( \sigma_1 \) and \( \sigma'_1 \) differ at only finite number of stages. Therefore, this inequality contradicts the one-stage deviation principle for finitely repeated games.

Hence, \((\sigma_1, \sigma_2)\) cannot satisfy one-stage deviation principle without being a Subgame Perfect Nash Equilibrium.
STOCHASTIC GAMES

LECTURE 10

PRELIMINARIES

• $X$ - a Borel space, i.e., a Borel subset of complete separable metric space;

• $\{X_n\}$ - a sequence of Borel spaces $X_n \neq \emptyset$ for every $n \in \mathbb{N}$; put $X^\infty = X_1 \times X_2 \times \cdots = \prod_{n=1}^{\infty} X_n$;
obviously, \( x = \{ x_n \} \in X^\infty, \ x_n \in X_n \) for every \( n \in \mathbb{N} \);

- \( \mathcal{B}(X_n) \) - Borel sets in \( X_n \);

- **Cylinder sets** in \( X^\infty \):

\[
C_k = D_1 \times \cdots \times D_k \times X_{k+1} \times X_{k+2} \times \cdots \quad (*)
\]

\( D_i \in \mathcal{B}(X_i) \) for \( i = 1, \ldots, k \);

- \( \mathcal{F}(X^\infty) \) - the smallest algebra in \( X^\infty \) containing the cylinder sets.
DATA:

\[ \{ p_n \} - \text{a sequence of conditional probabilities,} \]

where \( p_1 \) is a probability measure on \( X_1 \), and

\( p_n (\bullet | x_1, \ldots, x_{n-1}) \) is a probability measure on \( X_n \) given

\( (x_1, \ldots, x_{n-1}) \) for \( n \geq 2 \).

For every \( D \in \mathcal{B}(X_n) \), \( p_n \left( D \mid x_1, \ldots, x_{n-1} \right) \) is a measurable

function on \( X_1 \times \cdots \times X_{n-1} \).
THE IONESCU-TULCEA THEOREM

For a sequence of conditional probability measures

\( p = \{ p_n \} \) there exists a unique probability measure \( P^p \) on \( \mathcal{F}(X^\infty) \) such that for any cylinder set \( C_k \subset X^\infty \) defined in (*) it holds

\[
P^p(C_k) = \int \cdots \int_{D_{k-1}} p_k \left( D_k | x_1, \ldots, x_{k-1} \right) p_{k-1} \left( dx_{k-1} | x_1, \ldots, x_{k-2} \right) \cdots p_1 \left( dx_1 \right).
\]

For example, if \( k = 2 \), then

\[
P^p(C_2) = \int \int_{x_1 x_2} 1_{D_2} (x_2) 1_{D_1} (x_1) p_2 \left( dx_2 | x_1 \right) p_1 \left( dx_1 \right).
\]
THE MODEL OF A STOCHASTIC GAME

- \( S \) - a countable state space;
- \( A, B \) - finite sets of actions for Player 1 and 2, respectively;
- \( q(z|s,a,b) \) - a transition probability from state \( s \) to \( z \), when Player 1 picks an action \( a \), and Player 2 chooses \( b \);
- \( r_i : S \times A \times B \rightarrow \mathbb{R} \) - one-step bounded payoff function for Player \( i \).

Evolution of the game: The players observe an initial state \( s_1 \in S \) and selects independently actions \( a_1 \in A \) and \( b_1 \in B \). The actions
can be chosen by a random device with assigned probabilities. Then Player $i$ receives the payoff $r_i(s_1,a_1,b_1)$. If $r_i(s_1,a_1,b_1) < 0$, it means that it is a cost he has to pay. The game moves to a next state $s_2$ according to $q(s_2|s_1,a_1,b_1)$. Remembering $s_1,a_1,b_1$ and observing $s_2$, the players pick $a_2 \in A$ and $b_2 \in B$, Player $i$ receives $r_i(s_2,a_2,b_2)$, and the game moves to a next state $s_3$ according to $q(s_3|s_2,a_2,b_2)$. In this way we obtain a trajectory of a stochastic process $h = (s_1,a_1,b_1,s_2,a_2,b_2,...)$. Let $h_n = (s_1,a_1,b_1,...,s_{n-1},a_{n-1},b_{n-1},s_n)$ be a history up to the $n$th state for $n \geq 2$ and put $h_1 = (s_1)$. 
STRATEGY FOR PLAYER 1 is a sequence $\pi = \{\pi_n\}$, where $\pi_n (\cdot | h_n) \in \text{Pr}(A)$ for every $n \in \mathbb{N}$. 
STRATEGY FOR PLAYER 2 is a sequence \( \gamma = \{ \gamma_n \} \), where \( \gamma_n (\cdot | h_n) \in \Pr(B) \) for every \( n \in \mathbb{N} \).

Denote by \( \Pi \) the set of all strategies for Player 1 and \( \Gamma \) the set of all strategies for Player 2.

The existence of a process is guaranteed by the Ionescu-Tulcea Theorem. For any initial state \( s := s_1 \in S \) and strategies \( \pi = \{ \pi_n \} \) and \( \gamma = \{ \gamma_n \} \), there exists a unique probability measure \( P_{s}^{\pi\gamma} \) on \( (\Omega, \mathcal{F}(\Omega)) \), where \( \Omega = (S \times A \times B)^{\infty} \) and \( \mathcal{F}(\Omega) \) is a \( \sigma \)-algebra generated by the cylinder ("history") sets in \( \Omega \).
$C_k$ - cylinder set in $\Omega$, the event that the game has a
history $h_k$, i.e.,

$$C_k = \left\{ s_1, a_1, b_1, \ldots, s_{k-1}, a_{k-1}, b_{k-1}, s_k, A, B, S, A, B, S \ldots \right\}.$$

Then,

$$P_s^{\pi \gamma} (C_k) = q(s_k | s_{k-1}, a_{k-1}, b_{k-1}) \pi_{k-1}(a_{k-1} | h_{k-1}) \gamma_{k-1}(b_{k-1} | h_{k-1})$$

$$\ldots q(s_2 | s_1, a_1, b_1) \pi_1(a_1 | s_1) \gamma_1(b_1 | s_1).$$
$r_i^n(s, \pi, \gamma)$ - the expected payoff for Player $i$ at the $n$th stage of the game, when players use strategies $\pi$ and $\gamma$.

For example,

$$r_i^3(s, \pi, \gamma) =$$

$$\sum_{a_1, b_1} \sum_{s_2} \sum_{a_2, b_2} \sum_{s_3} \sum_{a_3, b_3} \sum_{s_1} \sum_{a_1, b_1} \sum_{s_1} r_i(s_3, a_3, b_3) \pi_3(a_3| h_3) \gamma_3(b_3| h_3)$$

$$q(s_3| s_2, a_2, b_2) \pi_2(a_2| h_2) \gamma_2(b_2| h_2)$$

$$q(s_2| s_1, a_1, b_1) \pi_1(a_1| s_1) \gamma_1(b_1| s_1).$$

Here, $s = s_1$. 
(I) **FINITE HORIZON MODEL**

\[ J_i^N (s, \pi, \gamma) = \sum_{n=1}^{N} \delta^{n-1} r_i^n (s, \pi, \gamma), \]

where \( \delta \in (0, 1] \) is a discount coefficient.

(II) **INFINITE HORIZON MODEL**

\[ J_i (s, \pi, \gamma) = \sum_{n=1}^{\infty} \delta^{n-1} r_i^n (s, \pi, \gamma). \]

Let us now get rid of one player, say Player 2, and examine the model with a one decision maker. It is known as **MARKOV DECISION PROCESS**, abbr. MDP.
**MDP**: $(S, A, q, r)$

- $S$ - a denumerable state space;
- $A$ - finite set of action for the decision maker;
- $q(z|s,a)$ - a transition probability from state $s$ to $z$;
- $r : S \times A \rightarrow \mathbb{R}$ - one-step bounded payoff function.

**Evolution of the process** develops analogously as in the game.

Let $f : S \rightarrow A$ be a decision function. A strategy $\pi = \{f, f, \ldots\}$ is called **stationary**. We shall further identify $\pi = \{f, f, \ldots\}$ with its member $f$. 
In other words, a decision maker, when uses a stationary strategy, looks only at the current state of the system and based on this knowledge he chooses an action. We shall consider non-randomised strategies, which will turn out to be sufficient for studying MDP models.

**Definition:** Let $u$ be a bounded function on $S$. For any stationary strategy $f$ define the operator $L_f$ as follows

$$
(L_f u)(s) = r(s, f(s)) + \delta \sum_{z \in S} u(z)q(z | s, f(s)).
$$

Let us now examine consecutive composition of $L_f$ with itself.
\[(L_f u)(s) = r(s, f(s)) + \delta \sum_{z \in S} (L_f u)(z)q(z|s, f(s)) = r(s, f(s)) + \delta \sum_{z \in S} \left( r(z, f(z)) + \delta \sum_{z' \in S} u(z')q(z'|z, f(z)) \right) q(z|s, f(s)) \]

\[= J^2(s, f) + \delta^2 \sum_{z \in S} \sum_{z' \in S} u(z')q(z'|z, f(z))q(z|s, f(s)) \]

\[= J^2(s, f) + \delta^2 \sum_{z \in S} u(z)q(z|s, f(s))^2. \]

Hence, for an arbitrary \( n \in \mathbb{N} \) we have

\[\left( L^n_f u \right)(s) = J^n(s, f) + \delta^n \sum_{z \in S} u(z)q^n(z|s, f(s)).\]
Here, $q^n$ denotes the transition after $n$ periods when a stationary strategy $\pi = \{f, f, \ldots\}$ is used and $J^n(s, f)$ represents the corresponding expected discounted cost in $n$-step model.

Since, $|u(z)| < K$ for all $z \in S$ and some positive constant $K$, then

$$\delta^n \sum_{z \in S} u(z) q^n(z \mid s, f(s)) < \delta^n K \to 0 \text{ as } n \to \infty.$$ 

Hence,

$$\lim_{n \to \infty} (L^n_f u)(s) = \lim_{n \to \infty} J^n(s, f) = J(s, f) = \sum_{n=1}^{\infty} \delta^n r^n(s, f).$$

$J(s, f)$ is the expected discounted cost in infinite horizon model.
CLAIM 1: Let \( w_i : S \times A \rightarrow \mathbb{R}, i=1,2, \) be bounded functions. Then,
\[
\max_{a \in A} w_1(s, a) - \max_{a \in A} w_2(s, a) \leq \max_{a \in A} |w_1(s, a) - w_2(s, a)|.
\]

Proof: Note that
\[
\max_{a \in A} w_1(s, a) = \max_{a \in A} \left( w_1(s, a) - w_2(s, a) + w_2(s, a) \right) \\
\leq \max_{a \in A} \left( w_1(s, a) - w_2(s, a) \right) + \max_{a \in A} w_2(s, a).
\]

Hence, \( \max_{a \in A} w_1(s, a) - \max_{a \in A} w_2(s, a) \leq \max_{a \in A} |w_1(s, a) - w_2(s, a)|. \)

The rest follows from symmetry. \( \square \)
THE BANACH FIXED POINT THEOREM

Let $X$ be a complete metric space and $T : X \rightarrow X$ be a contractive operator, i.e., there exists a constant $\delta \in (0,1)$ such that

$$d(T(x), T(y)) \leq \delta d(x, y), \text{ for all } x, y \in X.$$  

Clearly, $d$ is a metric in $X$. Then,

(I) there exists a unique point $x^* \in X$ such that $x^* = T(x^*)$, and

(II) $x^* = \lim_{n \to \infty} T^n x_0$, for any $x_0 \in X$.  

Let $B(S)$ be the set of all bounded functions on $S$ (bounded sequences) equipped with the supremum metric:

$$d(u_1, u_2) = \sup_{s \in S} |u_1(s) - u_2(s)|.$$ 

Then $B(S)$ is a Banach space. Let $T : B(S) \rightarrow B(S)$ be an operator defined as follows

$$(Tu)(s) = \max_{a \in A} \left[ r(s, a) + \delta \sum_{z \in S} u(z)q(z | s, a) \right].$$
CLAIM 2: $T$ is a contraction operator.

Proof: Note that by virtue of Claim 1 we have that

$$
\left| (Tu_1)(s) - (Tu_2)(s) \right| \leq \max_{a \in A} \left| \delta \sum_{z \in S} [u_1(z) - u_2(z)] q(z|s,a) \right|.
$$

Further, we get that

$$
\left| (Tu_1)(s) - (Tu_2)(s) \right| \leq \max_{a \in A} \delta \sum_{z \in S} \sup_{z \in S} \left| u_1(z) - u_2(z) \right| q(z|s,a)
\leq \delta d(u_1,u_2).
$$

Hence, $d(Tu_1,Tu_2) \leq \delta d(u_1,u_2)$. \qed
From the Banach Fixed Point Theorem there exists a unique function $u^* \in B(S)$ such that it holds

$$u^*(s) = \max_{a \in A} \left[ r(s,a) + \delta \sum_{z \in S} u^*(z) q(z|s,a) \right]$$

for all $s \in S$. This equation is called

**THE BELLMAN EQUATION**, and the operator $T$ **THE DYNAMIC PROGRAMMING OPERATOR.**
STOCHASTIC GAMES

LECTURE 11

Recall the Bellman equation:

\[ u^*(s) = \max_{a \in A} \left[ r(s, a) + \delta \sum_{z \in S} u^*(z) q(z \mid s, a) \right]. \]

A is finite \( \Rightarrow \exists \) a decision rule \( f : S \to A \) such that

\[ u^*(s) = \left[ r(s, f(s)) + \delta \sum_{z \in S} u^*(z) q(z \mid s, f(s)) \right] = (L_f u^*) (s). \]
Thus, $u^*(s) = \left( L^n_f u^* \right)(s)$ for every $n \geq 1$ and all $s \in S$.

From the previous lecture we know that

$$u^*(s) = \lim_{n \to \infty} \left( L^n_f u^* \right)(s) = J(s, f),$$

where $J(s, f)$ is a discounted payoff in infinite horizon time, where a stationary strategy $\{f, f, \ldots\}$ is used.

On the other hand, for any $a \in A$, we may write

$$u^*(s) \geq r(s, a) + \delta \sum_{z \in S} u^*(z)q(z|s,a).$$

Iterating this inequality we get that
$u^*(s) \geq r(s, a_1) + \delta \sum_{s_2 \in S} \left[ r(s_2, a_2) + \delta \sum_{s_3 \in S} u^*(s_3)q(s_3 | s_2, a_2) \right] q(s_2 | s, a_1).$

Hence,

$u^*(s) \geq J^2(s, \pi) + \delta^2 \sum_{s_2, s_3} u^*(s_3)q(s_3 | s_2, a_2)q(s_2 | s, a_1),$

and for any $n \geq 1,$

$u^*(s) \geq J^n(s, \pi) + \delta^n \sum_{s_2, \ldots, s_{n+1}} u^*(s_{n+1})q(s_{n+1} | s_n, a_n) \cdots q(s_2 | s, a_1).$
Clearly, the second term tends to 0 ($u^* \in B(S)$).
Letting $n \to \infty$, we conclude that $u^*(s) \geq J(s, \pi)$ for each $s \in S$ and any strategy $\pi \in \Pi$.

Summing up,

$$J(s, f) = u^*(s) \geq J(s, \pi) \quad \text{for any } \pi \in \Pi \text{ and all } s \in S;$$

$$u^*(s) = \max_{\pi \in \Pi} J(s, \pi) = J(s, f)$$ -optimal payoff function

$$\{f, f, f, f, f, \ldots\}$$ -optimal stationary strategy
From the Banach Fixed Point Theorem (see point (II)), we know that \( u^*(s) = \lim_{n \to \infty} \left( T^n u_0 \right)(s) \), \( s \in S \), for any \( u_0 \in B(S) \).

Let us put \( u_0 \equiv 0 \) and examine \( T^n u_0 \).

- \( (Tu_0)(s) = \max_{a \in A} r(s, a) \) - the optimal payoff in 1-step model;

  \[ \exists f_1 : S \to A \quad (Tu_0)(s) = \left( L_{f_1} u_0 \right)(s) = r(s, f_1(s)). \]

- \( (T^2u_0)(s) = \max_{a \in A} \left[ r(s, a) + \delta \sum_{z \in S} (Tu_0)(z)q(z|s, a) \right] \) - the optimal payoff in 2-step model;
∃ f_2 : S → A \quad (T^2 u_0)(s) = \\
\max_{a \in A} \left[ r(s, f_2(s)) + \delta \sum_{z \in S} (T u_0)(z) q(z | s, f_2(s)) \right],

(T^2 u_0)(s) = \left( L_{f_2} \left( L_{f_1} u_0 \right) \right)(s) = J^2 \left( s, \pi^* \right), \text{ where } \pi^* = \{ f_2, f_1 \}

is an optimal strategy for two-step game.

\vdots

\cdot (T^n u_0)(s) = \left( L_{f_n} \ldots \left( L_{f_2} \left( L_{f_1} u_0 \right) \right) \right)(s) = J^n \left( s, \pi^* \right) - the optimal payoff in n-step model, where \pi^* = \{ f_n, \ldots, f_1 \} is an optimal strategy in this model.
CLAIM 3: $\sup_{\pi \in \Pi} J^n(s, \pi) \rightarrow \sup_{\pi \in \Pi} J(s, \pi)$ as $n \rightarrow \infty$.

Proof: By Claim 1, it follows that

$|\sup_{\pi \in \Pi} J^n(s, \pi) - \sup_{\pi \in \Pi} J(s, \pi)| \leq \sup_{\pi \in \Pi} |J^n(s, \pi) - J(s, \pi)|$.

The right-hand side equals:

$\sup_{\pi \in \Pi} \left| \sum_{k=n+1}^{\infty} \delta^{k-1} r^k(s, \pi) \right| \leq \sup_{\pi \in \Pi} \left| \sum_{k=n+1}^{\infty} \delta^{k-1} C \right| = \frac{C \delta^n}{1-\delta} \rightarrow 0$.

Here, $C$ is a constant such that $|r(s, a)| \leq C$ for any $s \in S$, $a \in A$ and consequently $|r^k(s, \pi)| \leq C$ for any $k \in \mathbb{N}$. ☐
CONCLUDING REMARKS FOR MDP MODELS

For a discounted MDP with infinite horizon time the following facts hold:

(I) There exists a solution $u^*$ to the Bellman Equation:

$$u^*(s) = \max_{a \in A} \left[ r(s,a) + \delta \sum_{z \in S} u^*(z) q(z|s,a) \right],$$

where $u^*(s) = (Tu^*)(s)$; moreover, $u^*$ is an optimal discounted payoff, i.e., $u^*(s) = \sup_{\pi \in \Pi} J(s,\pi)$;
There exists an optimal stationary strategy $\tilde{\pi}$, where $\tilde{\pi} = \{f, f, \ldots\}$, and a decision function $f : S \rightarrow A$ attains maximum on the right-hand side in the Bellman equation;

The infinite horizon model can be approximated by the finite models:

$$u^*(s) = \lim_{n \rightarrow \infty} \left( T^n u_0 \right)(s), \quad u_0 \equiv 0,$$

optimal discounted payoff in $n$-step model

This algorithm is known as value iteration, since the function $u^*$ is also named value function;
For every $n$-step horizon model there exists an optimal strategy $\pi^*$ obtained by the backward induction. Each decision rule in $\pi^*$ depends on the current state and the moment/epoch at which the action is selected. Such a strategy is called Markov strategy.

The beginning of the theory of stochastic games as well as Markov decision processes is dated back the L.S. Shapley’ paper: “Stochastic games”, Proc. Nat. Acad. Sci. USA, vol. 39, pp. 1095-1100. Shapley defined an infinite game, for which the probability of its continuation is $\delta$ (our discount coefficient).
ZERO-SUM DISCOUNTED STOCHASTIC GAMES

\((S, A, B, q, r)\)

- state space
- action spaces
- transition probability

\(r : S \times A \times B \to \mathbb{R}\) - one-step bounded reward function for Player 1 and cost function for Player 2;

Since we deal with a game, we allow the players to randomize their choices! Recall that

\[h_n = (s_1, a_1, b_1, \ldots, s_{n-1}, a_{n-1}, b_{n-1}, s_n).\]
**Player 1 → Maximizer:**
strategy $\pi = (\pi_1, \pi_2, \ldots) \in \Pi$, where $\pi_n (\cdot | h_n) \in \Pr(A)$,
stationary strategy $\pi = (f, f, \ldots) \in F$, where $f (\cdot | s) \in \Pr(A)$

**Player 2 → Minimizer:**
strategy $\gamma = (\gamma_1, \gamma_2, \ldots) \in \Gamma$, where $\gamma_n (\cdot | h_n) \in \Pr(B)$,
stationary strategy $\gamma = (g, g, \ldots) \in G$, where $g (\cdot | s) \in \Pr(B)$

The discounted payoff function for Player 1 (cost function for Player 2): $J (s, \pi, \gamma) = \sum_{n=1}^{\infty} \delta^{n-1} r^n (s, \pi, \gamma)$, where $r^n (s, \pi, \gamma)$ is the expected reward (for Player 1) in the $n$th step.
THEOREM

(I) A discounted zero-sum stochastic game has a value $V$, i.e.,

$$V(s) := \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J(s, \pi, \gamma) = \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J(s, \pi, \gamma).$$

(II) Moreover, each player possesses an optimal stationary strategy, that is,

$\exists f^* \in F$ such that $V(s) = \inf_{\gamma \in \Gamma} J(s, f^*, \gamma)$, and

$\exists g^* \in G$ such that $V(s) = \sup_{\pi \in \Pi} J(s, \pi, g^*)$ for all $s \in S$. 
**Proof:**

- Define the operator $T : B(S) \rightarrow B(S)$ as follows:

$$(Tu)(s) := \max_{\mu \in \Pr(A)} \min_{\rho \in \Pr(B)} \left[ r(s, \mu, \rho) + \delta \sum_{z \in S} u(z) q(z \mid s, \mu, \rho) \right].$$

By the von Neumann Theorem we may replace the order of min and max and write this definition as

$$(Tu)(s) = \text{val} \left[ r(s, \bullet, \bullet) + \delta \sum_{z \in S} u(z) q(z \mid s, \bullet, \bullet) \right].$$

Here, $r(s, \mu, \rho) = \sum_{a \in A} \sum_{b \in B} r(s, a, b) \rho(b) \mu(a)$, where $\mu(a)$ is a
probability of choosing an action $a$, and $\rho(b)$ is a probability of selecting $b$. Similar formula applies to $q(\cdot|s, \mu, \rho)$.

- Next we have to show that $T$ is contractive, that is, for any $u_1, u_2 \in B(S)$ it holds

$$\sup_{s \in S} \left| (Tu_1)(s) - (Tu_2)(s) \right| \leq \delta \sup_{s \in S} \left| u_1(s) - u_2(s) \right|.$$ 

Proof of this fact is similar to the case of MDP, and follows from the fact that $\min \varphi = -\max(-\varphi)$.

The Banach Fixed Point Theorem implies that there exists $v \in B(S)$ such that $v(s) = (Tv)(s)$ for all $s \in S$. 
• We choose $f(\cdot|s) \in \Pr(A)$ such that

$$v(s) = \min_{\rho \in \Pr(B)} \left[ r(s, f(\cdot|s), \rho) + \delta \sum_{z \in S} v(z)q(z|s, f(\cdot|s), \rho) \right].$$ \quad (*)$$

Hence, $f(\cdot|s)$ realizes the maximum of the minimum.

Let us now consider MDP, for Player 2, defined by the objects: $(S, \Pr(B), \tilde{q}, \tilde{r})$, where

$$\tilde{q}(\cdot|s, \rho) = q(\cdot|s, f(\cdot|s), \rho) = \sum_{a,b} q(\cdot|s,a,b) f(a|s) \rho(b),$$

and
\[ \tilde{r}(s, \rho) = r\left(s, f(\cdot|s), \rho\right) = \sum_{a, b} r(s, a, b) f(a|s) \rho(b). \]

Hence, (*) takes the form

\[
\begin{align*}
  v(s) &= \min_{\rho \in \Pr(B)} \left[ \tilde{r}(s, \rho) + \delta \sum_{z \in S} v(z) \tilde{q}(z|s, \rho) \right] \\
  &\leq \left[ \tilde{r}(s, \rho) + \delta \sum_{z \in S} v(z) \tilde{q}(z|s, \rho) \right] \quad \forall \rho \in \Pr(B).
\end{align*}
\]

Iterating the last inequality, we infer that

\[ v(s) \leq \tilde{J}(s, \gamma) = J(s, f, \gamma) \text{ for every } \gamma \in \Gamma \text{ and } s \in S. \]

This means that \[ v(s) \leq \inf_{\gamma \in \Gamma} J(s, f, \gamma) \leq \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J(s, \pi, \gamma). \]
Now we choose \( g(\cdot|s) \in \Pr(B) \) such that
\[
\nu(s) = \max_{\mu \in \Pr(A)} \left[ r(s, \mu, g(\cdot|s)) + \delta \sum_{z \in S} v(z) q(z|s, \mu, g(\cdot|s)) \right]. \tag{\Delta}
\]
Hence, \( g(\cdot|s) \) realizes the minimum of the maximum. Let us now consider \textbf{MDP}, for Player 1, defined by the objects: \((S, \Pr(A), \hat{q}, \hat{r})\), where
\[
\hat{q}(\cdot|s, \mu) = q(\cdot|s, \mu, g(\cdot|s)) = \sum_{a,b} q(\cdot|s,a,b) \mu(a) g(b|s),
\]
and
\[ \hat{r}(s, \mu) = r(s, \mu, g(\cdot|s)) = \sum_{a,b} r(s,a,b)\mu(a)g(b|s). \]

Hence, \((\Delta)\) takes the form

\[
\nu(s) = \max_{\mu \in \Pr(A)} \left[ \hat{r}(s, \mu) + \delta \sum_{z \in S} v(z)\hat{q}(z|s, \mu) \right]
\geq \left[ \hat{r}(s, \mu) + \delta \sum_{z \in S} v(z)\hat{q}(z|s, \mu) \right] \quad \forall \mu \in \Pr(A).
\]

Iterating the last inequality, we deduce that

\[ \nu(s) \geq \hat{J}(s, \pi) = J(s, \pi, g) \text{ for every } \pi \in \Pi \text{ and } s \in S. \]

This means that

\[ \nu(s) \geq \sup_{\pi \in \Pi} J(s, \pi, g) \geq \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J(s, \pi, \gamma). \]
Summing up, we obtain that
\[
\inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J(s, \pi, \gamma) \leq \sup_{\pi \in \Pi} J(s, \pi, g) \leq \nu(s)
\]
\[
\leq \inf_{\gamma \in \Gamma} J(s, f, \gamma) \leq \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J(s, \pi, \gamma) \quad (\bigcirc).
\]

But, it is always true that
\[
\inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J(s, \pi, \gamma) \geq \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J(s, \pi, \gamma),
\]
thus we must have equalities in all places in \((\bigcirc)\). Hence, \(\nu\) is a \textbf{value} of the game, and optimal stationary strategies for Player 1 and 2 are: \(f^* := f\), \(g^* := g\), which are obtained from the corresponding Bellman equations.
Further extensions

Instead of zero-sum discounted stochastic games one may consider games with ergodic payoff:

\[ w(s, \pi, \gamma) := \limsup_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} r^n(s, \pi, \gamma) \]

with a finite state space \( S \). Neyman and Mertens in 1981 proved that \( \lim_{\delta \to 1^-} (1 - \delta)v_\delta(s) \) is a value of the game with ergodic payoff, where \( v_\delta \) is a value of a discounted game. Earlier Bewley and Kohlberg (1976) showed that this limit always exists.
We consider the following non-zero sum stochastic game:

- There are two-players, who extract renewable resource, e.g., fish;
- The resource stock develops over time according to the
growth rule:

\[ x_{t+1} = x_t^\alpha, \quad 0 < \alpha < 1, \quad t = 1, 2, \ldots \]

\[ x_t \in X := [0,1] \text{-state space;} \]

• The utility (payoff) function for Player \( i \) (\( i=1,2 \)) is:

\[ u^i(c) = \ln c, \quad c \in X \]

• Let \( c_t^i \) be a resource consumption of Player \( i \) in period \( t \).

The aim of Player \( i \) is:

\[ \max U^i(c_1^i, c_2^i, \ldots) \text{ subject to } x_{t+1} = \left( x_t - c_1^i - c_2^i \right)^\alpha. \]

The strategies for the players are defined in usual manner.
Let $\pi = (\pi^1, \pi^2)$ be any strategy profile.

We shall consider discounted utilities:

(a) in finite horizon game

$$U^i_N(x, \pi) = \sum_{t=1}^{N} \delta^{t-1} \ln c^i_t,$$

(b) in infinite horizon game

$$U^i(x, \pi) = \sum_{t=1}^{\infty} \delta^{t-1} \ln c^i_t.$$

**AIM:** Find a Nash Equilibrium in a game (a) and (b).
Theorem 1:

The finite $N$-step game has a symmetric Nash Equilibrium $\pi^* = \{c_N, c_{N-1}, \ldots, c_1\}$:

- $c_1(x) = \frac{x}{2}$, $c_t(x) = \frac{x}{2 + \alpha \delta + \ldots + (\alpha \delta)^{t-1}}$, $t = 2, \ldots, N$.

The equilibrium function in the $k$-period $k \in \{1, \ldots, N\}$ game is: $V_k(x) = U_k^i(x, \pi^*)$, where

- $V_1(x) = \ln \frac{x}{2}$,
\[ V_{k+1}(x) = \left(1 + \cdots + (\alpha \delta)^k\right) \ln x + \ln B_{k+1} \quad (\Delta) \]

and \( B_1 = \frac{1}{2}, \)

\[
B_{k+1} = \frac{B_{k}^\delta (\alpha \delta + \cdots + (\alpha \delta)^k)^{\alpha \delta + \cdots + (\alpha \delta)^k}}{(2 + \alpha \delta + \cdots + (\alpha \delta)^k)^{1+\alpha \delta + \cdots + (\alpha \delta)^k}}. 
\]

Moreover, for \( n \in \{1, \ldots, N - 1\} \)

\[
V_{n+1}(x) = \ln c_{n+1}(x) + \delta V_n \left((x - 2c_{n+1}(x))^\alpha\right) \\
= \max_{0 \leq c \leq x} \left[ \ln c + \delta V_n \left((x - c_{n+1}(x) - c)^\alpha\right) \right].
\]
Proof: Proof is by backward induction:

- \( V_1(x) = \max_{0 \leq c \leq x/2} \left[ \ln c \right] = \ln \frac{x}{2} \) - there is nothing afterwards and the players divide equally the amount \( x \) between them. Thus, \( c_1(x) = \frac{x}{2} \).

- \( V_2(x) = \)

\[
= \max_{0 \leq c \leq x} \left[ \ln c + \delta V_1(y) \right] = \max_{0 \leq c \leq x} \left[ \ln c + \delta \ln \left( \frac{(x - c - c^*)^\alpha}{2} \right) \right].
\]
First order condition gives: \[
\frac{1}{c} + \frac{-\alpha \delta}{x - c - c^*} = 0.
\]
Similarly, by symmetry for Player 2 we have that \[
\frac{1}{c^*} + \frac{-\alpha \delta}{x - c - c^*} = 0.
\]
Since, we look for a symmetric equilibrium, we get that \[
c_2(x) := c = c^* = \frac{x}{2 + \alpha \delta} < \frac{x}{2}.
\]
Thus,
\[
V_2(x) = \ln \frac{x}{2 + \alpha \delta} + \alpha \delta \ln \left( \frac{\alpha \delta x}{2 + \alpha \delta} \right) + \ln \left( \frac{1}{2} \right)^{\delta} = \]

\[(1 + \alpha \delta) \ln x + \ln \left(\frac{(\alpha \delta)^{\alpha \delta} (1/2)^{\delta}}{B_2 (2 + \alpha \delta)^{1+\alpha \delta}}\right).\]

• \(V_3(x) = \max_{0 \leq c \leq x} \left[ \ln c + \delta V_1(y) \right] = \)

Applying again FOC and the fact that \(c = c^*\), we get that
\[
\frac{1}{c} = \frac{\alpha \delta(1 + \alpha \delta)}{x - 2c}
\]
and, consequently,
\[ c_3(x) := c = \frac{x}{2 + \alpha \delta (1 + \alpha \delta)}. \]

Putting this formula into the above equation we obtain

\[ V_3(x) = \ln \frac{x}{2 + \alpha \delta + (\alpha \delta)^2} + \alpha \delta (1 + \alpha \delta) \ln \left( \frac{(\alpha \delta + (\alpha \delta)^2) x}{2 + \alpha \delta + (\alpha \delta)^2} \right) + \delta \ln B_2 = \]
\[
= \left(1 + \alpha \delta + (\alpha \delta)^2 \right) \ln x + \ln \frac{B_2^\delta}{2 + \alpha \delta + (\alpha \delta)^2}
\]

\[
+ \ln \left(\frac{\alpha \delta + (\alpha \delta)^2}{2 + \alpha \delta + (\alpha \delta)^2}\right)^{\alpha \delta + (\alpha \delta)^2} = \left(1 + \alpha \delta + (\alpha \delta)^2 \right) \ln x + \ln B_3.
\]

Continuing this procedure, we obtain the expression for each \( V_k \), \( 1 \leq k \leq N \).

\[\Box\]

**Remark:** In order to obtain the result we have used backward induction, which means that the equilibrium
function for $N$ period has the number 1, for $N-1$ has the number 2, etc.

\[
\begin{align*}
&V_N, c_N \\
&V_{N-1}, \quad c_{N-1} \\
&V_{N-2}, \quad c_{N-2}
\end{align*}
\]
CLAIM 1: \( \{ B_k \} \) is decreasing.

**Proof:** Recall that \( B_1 = \frac{1}{2} \) and for \( k \geq 1 \)

\[
B_{k+1} = \frac{B_k^\delta (\alpha \delta + ... + (\alpha \delta)^k)^{\alpha \delta + ... + (\alpha \delta)^k}}{(2 + \alpha \delta + ... + (\alpha \delta)^k)^{1 + \alpha \delta + ... + (\alpha \delta)^k}}.
\]

Define \( \phi : (0,1) \rightarrow (0,\infty) \) as follows

\[
\phi(y) = \left( \frac{y}{2 + y} \right)^y.
\]

Then,
\[
\frac{\phi'(y)}{\phi(y)} = (\ln \phi(y))' = \ln \frac{y}{2+y} - \frac{y}{2+y} + 1 \leq 0,
\]
because \( \ln s - s + 1 \leq 0 \), for \( s \in (0,1) \). Hence, \( \phi \) is decreasing and, consequently, \( \psi(y) = \frac{\phi(y)}{2+y} \) is decreasing. Put

\[
y_k := \frac{(\alpha \delta + \ldots + (\alpha \delta)^k)^{\alpha \delta + \ldots + (\alpha \delta)^k}}{(2 + \alpha \delta + \ldots + (\alpha \delta)^k)^{1 + \alpha \delta + \ldots + (\alpha \delta)^k}}
\]
and observe that \( y_n > y_{n+1} \) for every \( n \in \mathbb{N} \).
Now by induction we show that \( \{B_n\} \) is decreasing.

1° \[ B_2 = \frac{1}{2^\delta (2 + \alpha \delta)} \frac{(\alpha \delta)^{\alpha \delta}}{(2 + \alpha \delta)^{\alpha \delta}} < \frac{1}{2^\delta (2 + \alpha \delta)} \]

\[ < \frac{1}{2^\delta 2} < \frac{1}{2} = B_1. \]

2° Suppose that \( B_{n+1} < B_n \) for some \( n \in \mathbb{N} \). Since \( y_{n+1} < y_n \), we obtain \( B_{n+2} = B_{n+1} y_{n+1} < B_n^\delta y_n = B_{n+1}, \) which completes the induction step. \( \square \)
CLAIM 2: \( \{B_k\} \) is bounded from below by some positive constant.

**Proof:** Observe that

\[
B_{n+1}^{1-\delta} = \frac{B_{n+1}}{B_n^\delta} > \frac{B_{n+1}}{B_n^\delta} = y_n > \lim_{k \to \infty} y_k > 0.
\]

Therefore, \( l := \lim_{n \to \infty} B_{n+1} \) exists and \( l = \inf_{n \in \mathbb{N}} B_{n+1} > 0. \)

Now, since \( B_{n+1} = B_n^\delta y_n \), it follows that \( l = \frac{l_1}{l_2} \), where
\[ l_1 = \left( \frac{\alpha \delta}{1 - \alpha \delta} \right)^{1/(1 - \alpha \delta)(1 - \delta)}, \quad l_2 = \left( \frac{2 - \alpha \delta}{1 - \alpha \delta} \right)^{1/(1 - \alpha \delta)(1 - \delta)}. \]

Next, we observe that

\[ \lim_{N \to \infty} c_N(x) = \frac{x}{2 + \alpha \delta + \ldots + (\alpha \delta)^{N-1}} = \frac{x(1 - \alpha \delta)}{2 - \alpha \delta} =: f_*(x). \]

Set \( v_\delta := \lim_{k \to \infty} V_{k+1} \). By above discussion \( v_\delta \) is well-defined, i.e., \( v_\delta(x) < \infty \) for every \( x \in X \). Letting \( k \to \infty \) (and at
the same time $N \to \infty$) in $(\Delta)$,

$$v_\delta(x) = \lim_{k \to \infty} \left( \left(1 + \cdots + (\alpha \delta)^k\right) \ln x + \ln B_{k+1} \right) =$$

$$\frac{\ln x}{1 - \alpha \delta} + \frac{\alpha \delta \ln(\alpha \delta) + (1 - \alpha \delta) \ln(1 - \alpha \delta) - \ln(2 - \alpha \delta)}{(1 - \alpha \delta)(1 - \delta)}$$

for every $x \in X$. 
Theorem 2:

The infinite game has a symmetric \textit{stationary} Nash Equilibrium

\[
f^*_*(x) = \frac{x(1 - \alpha \delta)}{2 - \alpha \delta}.
\]

The equilibrium function is \(v_\delta\) defined in (*).
This function for every \(x \in X\) satisfies the following equations:
\[ v_\delta(x) = \ln f_*(x) + \delta v_\delta \left( (x - 2f_*(x))^\alpha \right) \quad (**) \]

\[ = \max_{0 \leq c \leq x} \left[ \ln c + \delta v_\delta \left( (x - f_*(x) - c)^\alpha \right) \right]. \]

**Remark:** The above equation is called a Bellman equation. Since we deal with a non-zero sum game, we get the Bellman equation for each player. However, we do not write the upper index to indicate the player, because we look for symmetric equilibrium and, consequently, the equilibrium function is the same for both players.
**Proof:** Recall that from Theorem 1, we have that

\[ V_{k+1}(x) = \ln c_{k+1}(x) + \delta V_k \left( \left( x - 2c_{k+1}(x) \right)^\alpha \right) \]

for \( 1 \leq k \leq N - 1 \) and \( x \in X \). Hence, letting \( k, N \to \infty \), we obtain the first inequality in (**).

Now from Theorem 1 we also obtain that

\[ V_{k+1}(x) = \max_{0 \leq c \leq x} \left[ \ln c + \delta V_k \left( \left( x - c_{k+1}(x) - c \right)^\alpha \right) \right]. \]

Clearly, \( V_{k+1} \to v_\delta \) as \( k \to \infty \).

Next we show that the second equality in (**) holds. Indeed, it is easy to see that when we insert formulae for
\( \nu_\delta \) and \( f_* \) in (**) and making use of the first order condition, we get that 
\[
\frac{1}{c} - \frac{\alpha \delta}{1 - \alpha \delta} \frac{1}{x - f_*(x) - c} = 0,
\]
which implies that 
\[
\frac{x}{2 - \alpha \delta} - c = \frac{\alpha \delta}{1 - \alpha \delta} c \quad \text{and},
\]
consequently, 
\[
c = \frac{(1 - \alpha \delta) x}{2 - \alpha \delta}.
\]
Putting this expression in the second equality of (**), we get exactly the right-hand side of the first equality of (**). This completes the proof.
Supermodular games are games in which each player’s marginal utility of increasing his strategy rises with increases in his rivals’ strategies. In such games the best response correspondences are increasing.

Supermodular games are well behaved and possess pure strategy Nash equilibria. The simplicity of supermodular games makes convexity and differentiability assumptions unnecessary, although they are satisfied in most applications.
• $S_i$ Player $i$’s strategy set; it is a subset (not necessarily compact and convex) of a finite dimensional Euclidean space $\mathbb{R}^{m_i}$;

• $S = S_1 \times \cdots \times S_n$ is a subset of $\mathbb{R}^m$, where $m = m_1 + \cdots + m_n$.

Let us introduce the \textit{partial order} ‘$\geq$’ in $\mathbb{R}^K$:

• $x \geq y$, if $x_k \geq y_k$ for all $k = 1, \ldots, K$;

• $x > y$, if $x \geq y$ and there exists $k \in \{1, \ldots, K\}$ such that $x_k > y_k$;
Remark: If a vector dominates another in one component but is dominated in another component, the vectors cannot be compared.

- **meet** $\iff x \land y \triangleq (\min\{x_1, y_1\}, \ldots, \min\{x_K, y_K\})$
- **join** $\iff x \lor y \triangleq (\max\{x_1, y_1\}, \ldots, \max\{x_K, y_K\})$

$S$ is **sublattice** of $\mathbb{R}^m$ if $s \in S$ and $\tilde{s} \in S$ imply that $\tilde{s} \land s \in S$ and $\tilde{s} \lor s \in S$.

A set $S$ has a **greatest** element $\bar{s}$ (a **least** element $\underline{s}$) if $s \leq \bar{s}$ ($s \geq \underline{s}$) for all $s \in S$. 
A topological result (see Birkhoff 1967) guarantees that $S$ has a **greatest** element and a **least** element, if $S$ is non-empty and compact sublattice of $\mathbb{R}^m$.

**Definition:** $r_i(s_i, s_{-i})$ has **increasing differences** in $(s_i, s_{-i})$, if for all $(s_i, \tilde{s}_i) \in S_i^2$ and $(s_{-i}, \tilde{s}_{-i}) \in S_{-i}^2$ such that $s_i \geq \tilde{s}_i$ and $s_{-i} \geq \tilde{s}_{-i}$,

$$r_i(s_i, s_{-i}) - r_i(\tilde{s}_i, s_{-i}) \geq r_i(s_i, \tilde{s}_{-i}) - r_i(\tilde{s}_i, \tilde{s}_{-i});$$

$r_i(s_i, s_{-i})$ has strictly increasing differences in $(s_i, s_{-i})$, if for all $(s_i, \tilde{s}_i) \in S_i^2$ and $(s_{-i}, \tilde{s}_{-i}) \in S_{-i}^2$ such that $s_i > \tilde{s}_i$.
and $s_{-i} > \tilde{s}_{-i}$,

$$r_i(s_i, s_{-i}) - r_i(\tilde{s}_i, s_{-i}) > r_i(s_i, \tilde{s}_{-i}) - r_i(\tilde{s}_i, \tilde{s}_{-i}).$$

**Remark:** Increasing differences says that an increase in strategies of Player $i$’s rivals raises the desirability of playing a high strategy for Player $i$.

**Definition:** $r_i(s_i, s_{-i})$ is *supermodular* in $s_i$, if for each $s_{-i}$

$$r_i(s_i, s_{-i}) + r_i(\tilde{s}_i, s_{-i}) \leq r_i(s_i \land \tilde{s}_i, s_{-i}) + r_i(s_i \lor \tilde{s}_i, s_{-i})$$
for all \((s_i, \tilde{s}_i) \in S_i^2\).

\(r_i(s_i, s_{-i})\) is **strictly supermodular** in \(s_i\), if this inequality is strict whenever \(s_i\) and \(\tilde{s}_i\) cannot be compared with respect to ‘\(\geq\)’.

Note that supermodularity is automatically satisfied if \(S_i\) is single-dimensional. The supermodularity will be needed in the case of multi-dimensional strategy spaces to prove that each player’s best responses are increasing with his rivals’ strategies.
Topkis (1979) showed that if $S_i = \mathbb{R}^{m_i}$ and if $r_i$ is twice continuously differentiable in $s_i$, then $r_i$ is supermodular in $s_i$, iff for any two components $s_{i\kappa}$ and $s_{i\ell}$ (with $\kappa \neq \ell$)

$$\frac{\partial^2 r_i}{\partial s_{i\kappa} \partial s_{i\ell}} \geq 0.$$ 

Supermodular Game is such that for each $i$, $S_i$ is a sublattice of $\mathbb{R}^{m_i}$, $r_i$ has increasing differences in $(s_i, s_{-i})$ and $r_i$ is supermodular in $s_i$.

Example (Cournot game) Consider a duopoly. Firm $i \in \{1, 2\}$ chooses a quantity $q_i \in [0, \bar{q}_i]$. Suppose that inverse demand functions $P_i(q_i, q_j)$ are twice continuously differentiable, and that $P_i$ and firm $i$’s marginal revenue, i.e., $P_i + q_i \frac{\partial P_i}{\partial q_i}$ are decreasing in $q_j$. 
Firm $i$’s cost $C_i(q_i)$ is assumed differentiable. The payoffs are

$$r_i(q_i,q_j) = q_i P_i(q_i,q_j) - C_i(q_i).$$

If $s_1 \equiv q_1$ and $s_2 \equiv -q_2$, the transformed payoffs satisfy

$$\frac{\partial^2 r_i}{\partial s_i \partial s_j} \geq 0$$

for all $i \neq j$. Thus the game is supermodular.

Note that the above-mentioned transformation works only for a two-player game.
THEOREM (Tarski 1955)

If $S$ is non-empty compact sublattice of $\mathbb{R}^m$ and $f : S \rightarrow S$ is increasing (i.e., $f(x) \leq f(y)$ for $x \leq y$), then $f$ has a fixed point in $S$.

**Intuition:** Consider a single-dimensional case with $S = [0,1]$. In order *not* to have a fixed point, the function $f$ (see the figure below) would need to “escape” the area above the diagonal and “jump” into the area below the diagonal; but increasing functions do not jump down.
In the multi-dimensional case the intuition is the same, as no component of $f(s)$ jumps down when an arbitrary component of $s$ increases.
**Definition:** A function $r(\bullet)$ on $S$ is *upper semicontinuous* at $s$, if for any sequence $s^k \to s$

$$\limsup_{k \to \infty} r(s^k) \leq r(s).$$

**Upper semicontinuous functions**
THEOREM 1 (Topkis 1979)

Let $S_i$ be compact, $r_i$ be upper semicontinuous in $s_i$ for each $s_{-i}$ and all $i \in \{1, \ldots, n\}$. Suppose that the game is supermodular. Then, the set of pure strategy Nash equilibria is a non-empty and possesses greatest and least equilibrium points $\bar{s}$ and $\underline{s}$.

_Sketch of the proof:_ Let $b_i(s_{-i})$ be the set of best responses of Player $i$ to a strategy profile $s_{-i}$, i.e., the set
of $s_i \in S_i$ that maximize $r_i(\bullet, s_{-i})$.

**Claim:** If $S_i$ is compact and $r_i$ is upper semicontinuous in $s_i$, then $b_i$ is non-empty, since $r_i(s_i, s_{-i})$ attains the maximum in $s_i$ on $S_i$.

Indeed, consider a sequence $s_{i}^{k}$ such that

$$\sup_{s_i \in S_i} r_i(s'_i, s_{-i}) = \lim_{k \to \infty} r_i(s_i^k, s_{-i}) .$$

Compactness implies that there exists a converging subsequence of $s_i^k \to s_i$ (w.l.g. we may assume that $s_{i}^{k}$ is this subsequence). Upper
semicontinuity implies that \( r_i(s_i', s_{-i}) \geq \limsup_{k \to \infty} r_i(s_i^k, s_{-i}) \), so that \( s_i \) is indeed a best response to \( s_{-i} \). \( \square \)

To show that \( b_i(s_{-i}) \) is a sublattice for each \( s_{-i} \), suppose that \( s_i \) and \( \tilde{s}_i \) belong to \( b_i(s_{-i}) \) and

\[
r_i(s_i \land \tilde{s}_i, s_{-i}) < r_i(s_i, s_{-i}) = r_i(\tilde{s}_i, s_{-i}).
\]

Supermodularity of \( r_i(\bullet, s_{-i}) \) implies that

\[
\begin{align*}
r_i(s_i \lor \tilde{s}_i, s_{-i}) &\geq r_i(s_i, s_{-i}) + r_i(\tilde{s}_i, s_{-i}) - r_i(s_i \land \tilde{s}_i, s_{-i}) \\
&> r_i(s_i, s_{-i}) + r_i(\tilde{s}_i, s_{-i}) - r_i(\tilde{s}_i, s_{-i}) = r_i(s_i, s_{-i}),
\end{align*}
\]
which contradicts the assumption that $s_i$ and $\tilde{s}_i$ are best responses to $s_{-i}$. The same reasoning applies to the join.

Since $b_i(s_{-i})$ is non-empty compact sublattice of $\mathbb{R}^{m_i}$, it has a greatest element $\bar{s}_i(s_{-i})$.

**Claim:** $\bar{s}_i(\bullet)$ is non-decreasing, i.e.,

$$s_{-i} \geq \tilde{s}_{-i} \Rightarrow \bar{s}_i(s_{-i}) \geq \bar{s}_i(\tilde{s}_{-i}).$$

Indeed, let $s_i \triangleq \bar{s}_i(s_{-i})$ and $\tilde{s}_i \triangleq \bar{s}_i(\tilde{s}_{-i})$. Then,
0 \leq r_i(s_i, s_{-i}) - r_i(s_i \lor \tilde{s}_i, s_{-i}) \leq r_i(s_i, \tilde{s}_{-i}) - r_i(s_i \lor \tilde{s}_i, \tilde{s}_{-i}) \\
\leq r_i(s_i \land \tilde{s}_i, \tilde{s}_{-i}) - r_i(\tilde{s}_i, \tilde{s}_{-i}) \leq 0.

The first and fourth inequalities result from optimality of $s_i$ against $s_{-i}$ and $\tilde{s}_i$ against $\tilde{s}_{-i}$, the second from the increasing differences and $s_i \leq s_i \lor \tilde{s}_i$, and the third from supermodularity in Player $i$’s strategy. Since $s_i$ is the greatest element of $b_i(s_{-i})$, then $s_i \geq s_i \lor \tilde{s}_i$. Therefore, $s_i = s_i \lor \tilde{s}_i$, and consequently $s_i \geq \tilde{s}_i$. \hfill \Box

We can now apply Tarski’s theorem to $f(\tilde{s}) =$
\( (\bar{s}_1(\bar{s}),...,\bar{s}_n(\bar{s})) \). By construction a fixed point \( \bar{s} \) of \( f(\cdot) \) (which exists) is a pure strategy \textbf{Nash Equilibrium}. It can be shown that \( \bar{s} \) is the greatest element in the set of Nash equilibria. The intuition is again that a higher strategy triggers a higher best response. Finally, by symmetry, the analysis applies to lower bounds as well.

\( \square \)

For supermodular games with \textit{strictly} increasing differences, one can prove monotonicity of the entire reaction correspondence.
THEOREM 2 (Topkis 1979)

Consider a supermodular game with strictly increasing differences. If \( s_i \in b_i(s_{-i}) \), \( \tilde{s}_i \in b_i(\tilde{s}_{-i}) \) and \( s_{-i} \geq \tilde{s}_{-i} \), then \( s_i \geq \tilde{s}_i \).

**Proof:** The assertion follows from the following chain of inequalities:

\[
0 \leq r_i(s_i, s_{-i}) - r_i(s_i \lor \tilde{s}_i, s_{-i}) \leq r_i(s_i, \tilde{s}_{-i}) - r_i(s_i \lor \tilde{s}_i, \tilde{s}_{-i}) \\
\leq r_i(s_i \land \tilde{s}_i, \tilde{s}_{-i}) - r_i(\tilde{s}_i, \tilde{s}_{-i}) \leq 0.
\]

The first and fourth inequalities result from optimality of
\( s_i \) against \( s_{-i} \) and \( \tilde{s}_i \) against \( \tilde{s}_{-i} \), the second from the increasing differences and \( s_i \leq s_i \lor \tilde{s}_i \), and the third from supermodularity in Player \( i' \)’s strategy. Finally, note that if \( s_i \) is not greater or equal \( \tilde{s}_i \), then \( s_i < s_i \lor \tilde{s}_i \), and strictly increasing differences implies that the second inequality is strict, which is contradiction. \( \square \)
A **COOPERATIVE GAME** is a game in which the players have complete freedom of preplay communication to make joint *binding* agreements. Such a game is described by a triplet

\((N, P(N), v)\)

where
\( N = \{1, \ldots, n\} \) - set of \( n \) players.

\( P(N) = 2^N \) - set of all coalitions;
Any subset \( S \subseteq N \) is called a coalition.

\( v : P(N) \rightarrow \mathbb{R} \) - the characteristic function; in other words, it is a function such that \( v(S) \) represents the largest joint payoff which the coalition \( S \subseteq N \) is guaranteed.

We define \( v(\emptyset) = 0 \).
AIM OF THE GAME

To find a “just” distribution of $v(N)$ among players, who can form coalitions $S \subseteq N$.
A value of a coalition $S$ is a number $v(S)$.

The following property, possessed by any characteristic function, is called *supperadditivity*.

**Proposition 1**

For any finite cooperative game

$$v(S \cup T) \geq v(S) + v(T), \quad S, T \subseteq N, \quad S \cap T = \emptyset.$$
The strongest of supperadditivity is \textit{additivity}:
\[ v(S \cup T) = v(S) + v(T), \quad S, T \subseteq N, \quad S \cap T = \emptyset. \]
In such a game there is plainly no positive incentive for coalitions involving one player to form.

\textbf{Definition:} A cooperative game with additive characteristic function is called \textit{inessential}. Other cooperative games are called \textit{essential}. 
**Imputation** is a possible distribution of available payoff:

\[ x = (x_1, \ldots, x_n) \]

that satisfies the following conditions:

- **individual rationality:** \( x_i \geq v(\{i\}) \) for all \( i \in N \) (no member of coalition will consent to receive less than he or she can obtain individually),

- **collective rationality:** \( \sum_{i \in N} x_i = v(N) \).
Proposition 2

A finite cooperative game is essential if and only if

\[ \sum_{i \in N} x_i = v(N). \]

The set of imputations is never empty, since from the supperadditivity of \( v \), we have \( \sum_{i \in N} v(\{i\}) \leq v(N) \). For example, one imputation is given by \( x = (x_1, \ldots, x_n) \), where \( x_i = v(\{i\}) \) for \( i = 1, \ldots, n - 1 \) and \( x_n = v(N) - \sum_{i=1}^{n-1} v(\{i\}) \).
This is an imputation most preferred by Player $n$. In fact the set of imputations is exactly the simplex consisting of the convex hull of the $n$ points obtained by letting $x_i = v(\{i\})$ for all $x_i$ except one, which is then chosen to satisfy $\sum_{i \in N} x_i = v(N)$.

For example, if $v(\{1\}) = \frac{1}{2}$, $v(\{2\}) = 0$, $v(\{3\}) = 1$, and $v(\{N\}) = 9$. The set of imputations is
\left\{ (x_1, x_2, x_3) : x_1 + x_2 + x_3 = 9, x_1 \geq \frac{1}{2}, x_2 \geq 0, x_3 \geq 1 \right\}.

This is a triangle each of whose vertices satisfy two of the three inequalities with equality, namely

\( (8, 0, 1) \) - imputation most preferred by Player 1

\( \left( \frac{1}{2}, \frac{15}{2}, 1 \right) \) - imputation most preferred by Player 2

\( \left( \frac{1}{2}, 0, \frac{17}{2} \right) \) - imputation most preferred by Player 3.
THE CORE

Suppose some imputation $x$ is being proposed as a division of $v(N)$ among the players. If there exists a coalition $S$, whose total return from $x$ is less than what the coalition can achieve acting by itself, that is, $\sum_{i \in S} x_i < v(S)$, then it will be a tendency for a coalition $S$ to form and upset the proposed $x$ because such a coalition could guarantee each of its members more than they would receive from $x$. Such an imputation has an inherent instability.
Definition: An imputation $x$ is said to be **unstable** through a coalition $S$ if $\sum_{i \in S} x_i < v(S)$. We say $x$ is **unstable** if there is a coalition $S$ such that $x$ is unstable through $S$, and we say $x$ is **stable** otherwise.

The set $C$ of stable imputations is called the **core**, 

$$C = \{ x = (x_1, \ldots, x_n) : \sum_{i \in N} x_i = v(N), \sum_{i \in S} x_i \geq v(S) \text{ for all } S \subset N \}.$$
The core can consist of many points, but the core can also be empty. One may take the size of the core as a measure of stability, or of how likely it is that a negotiated agreement is prone to be upset.

**Definition:** A game in coalitional form is said to be **constant-sum** if $v(S) + v(S^c) = v(N)$ for all coalitions $S \in 2^N$.

**Proposition 3**

The core of an essential constant-sum game is empty.
Proof: Let $x$ be an imputation. Since the game is essential, we have $\sum_{i \in N} v({i}) = v(N)$. Then there must be a player $k$ such that $x_k > v({k})$, for otherwise

$$v(N) = \sum_{i \in N} x_i \leq \sum_{i \in N} v({i}) < v(N).$$

Moreover, since the game is constant-sum, we have that
\( v(N \setminus \{k\}) + v(\{k\}) = v(N) \). But then \( x \) must be unstable through the coalition \( N \setminus \{k\} \), because

\[
\sum_{i \neq k} x_i = \sum_{i \in N} x_i - x_k < v(N) - v(\{k\}) = v(N \setminus \{k\}). \]

**Example 1.** Consider the game with characteristic function \( v \) given by

\[
\begin{align*}
v(\{1\}) &= 1 & v(\{1, 2\}) &= 4 \\
v(\{2\}) &= 0 & v(\{1, 3\}) &= 3 & v(\{1, 2, 3\}) &= 8 \\
v(\{3\}) &= 1 & v(\{3, 2\}) &= 5 & \text{grand coalition}
\end{align*}
\]
The imputations are the points \((x_1, x_2, x_3)\) such that 
\[x_1 + x_2 + x_3 = 8\] and \(x_1 \geq 1, x_2 \geq 0, x_3 \geq 1\). This set is the triangle with vertices \((7, 0, 1), (1, 6, 1), (1, 0, 7)\).

We plot this triangle in \textit{barycentric coordinates}. This is done by pretending that the plane of the plot is 
\[x_1 + x_2 + x_3 = 8\] and giving each point on the plane three coordinates which add to 8. Then, it is easy to draw the line \(x_1 = 1\) or the line \(x_1 + x_3 = 3\). Note that the latter line is the same as the line \(x_2 = 5\). It is apparent that the set of imputations is an equilateral triangle.
$x_1 + x_3 = 3$  \( (3,4,1) \)

$y_1 + x_3 = 3$  \( (3,4,1) \)

$z_1 + x_3 = 3$  \( (3,4,1) \)

Unstable through \( \{2,3\} \)
Let us explain which imputations are unstable. The coalition can guarantee itself $\nu(\{2,3\}) = 5$, so all points $(x_1,x_2,x_3)$ with $x_2 + x_3 < 5$ are unstable through $\{2,3\}$. These are the points below the line $x_2 + x_3 = 5$ in the diagram. Since $\{1,2\}$ can ensure itself $\nu(\{1,2\}) = 4$, all points below and to the right of the line $x_2 + x_1 = 4$ are unstable. Finally since $\{1,3\}$ can guarantee itself $\nu(\{1,3\}) = 3$, all points below the line $x_3 + x_1 = 3$ are unstable. The core is the remaining set of points in the set of imputations given by the 5-sided figure, including the boundary.
Example 2. A certain object is worth $w_i$ to Player $i$ for $i=1,2,3$. We assume that $w_1 < w_2 < w_3$, which means that Player 3 values the object most. But Player 1 owns the object, so $v(\{1\}) = w_1$. Player 2 and 3 themselves can do nothing, so $v(\{2\}) = v(\{3\}) = 0$ and $v(\{2,3\}) = 0$. If Players 1 and 2 come together, the joint worth is $w_2$, and therefore $v(\{2,1\}) = w_2$. Similarly, $v(\{3,1\}) = w_3$. If all three get together, the object is still only worth $w_3$, so $v(\{3,2,1\}) = w_3$.

Find the core of this game.
The core consists of all vectors \((x_1, x_2, x_3)\) satisfying

\[
\begin{align*}
x_1 & \geq w_1 & \quad x_1 + x_2 & \geq w_2 \\
x_2 & \geq 0 & \quad x_1 + x_3 & \geq w_3 & \quad x_1 + x_2 + x_3 & = w_3 \\
x_3 & \geq 0 & \quad x_2 + x_2 & \geq 0
\end{align*}
\]

It follows from \(x_2 = w_3 - x_1 - x_3 \leq 0\) and \(x_2 \geq 0\) that \(x_2 = 0\) for all points of the core. Then we obtain that \(x_1 \geq w_2\) and \(x_3 = w_3 - x_1\). Hence, the core is
This fact indicates that the object will be purchased by Player 3 at some purchase price $x$ between $w_2$ and $w_3$. Hence, Player 1 ends up with the amount $x$ and Player 3 ends up with the object and the amount minus $x$. Player 2 plays no active role in this bargain, but without Player 2 around, Player 3 might hope to get the object for less than $w_2$ (but not less than $w_1$).
Lloyd Shapley in „A value for n-person games”
proposed the following axioms of distribution of $v(N)$:

(S1) Axiom of effectiveness: $\sum_{i \in N} \varphi_i[v] = v(N)$

(S2) Axiom of “dummy player”:
$\forall S \subseteq N \quad v(S) = v(S \setminus \{i\}) \quad \Rightarrow \quad \varphi_i[v] = 0$. 
(S3) **Axiom of symmetry:** for any $S \subseteq N \setminus \{i, j\}$, if
$$v(S \cup \{i\}) = v(S \cup \{j\}) \implies \phi_i[v] = \phi_j[v]$$

(S4) **Axiom of aggregation:** if $v$ and $w$ are characteristic functions of two games with the same set of players, then
$$\forall i \in N \quad \phi_i[v + w] = \phi_i[v] + \phi_i[w],$$
where $(v + w)(S) = v(S) + w(s)$.

Axiom (S4) was criticized, since it refers to the sum of characteristic functions of two games.
Theorem (Shapley (1953))

The unique vector $\varphi[v]$, which satisfies the axioms (S1)- (S4) is given by:

$$
\varphi_i[v] = \sum_{i \in S} \frac{(n-|S|)!(|S|-1)!}{n!} (v(S) - v(S \setminus \{i\}))
$$

$$
= \sum_{T \subseteq N \setminus \{i\}} \frac{(n-1-|T|)!|T|!}{n!} (v(T \cup \{i\}) - v(T))
$$

$$
\underbrace{P_T}
$$
**Interpretation:** Since $p_T$ is the probability of forming coalition $T \left( \sum_{T \subseteq N \setminus \{i\}} p_T = 1 \right)$, then $\varphi_i [v]$ is the expected value of marginal distributions of Player $i$ to each coalition.

Thus the theorem demonstrates that Shapley’s axioms are equivalent to the proposal that Player $i$ should receive the average of all his or her contributions to coalitions $S$, which contain $i$. 
Remark: H.P. Young (1985) replaced the controversial aggregation axiom with the following one:

(S4') if 
\[ \forall S \subseteq N \setminus \{i\} \quad v(S \cup \{i\}) - v(s) = w(S \cup \{i\}) - w(s), \]
then \( \varphi_i[v] = \varphi_i[w] \).
If Player $i$ has the same marginal distribution in two games, then his/her payoffs in these games are equal.

**Theorem (Young (1985))**

If a game satisfies the axioms (S1)-(S3)+(S4'), then there exists a unique imputation $\phi[v]$ which is a Shapley value.
VOTING GAMES

Let us consider a game, in which characteristic function equals $0$ for losing coalitions and $1$ for winning coalitions.

**Definition:** Let $W$ be a set of winning coalition. A game is called *simple*, if:

- $\emptyset \not\in W$
- $N \in W$
- if $S \in W$ and $S \subseteq T$, then $T \in W$ (each coalition including winning coalition is also winning).
Weighted voting games:

\[ [q; w_1, ..., w_n] \]

\( q \) - number of votes to pass a bill,
\( w_i \) - number of votes of Player \( i \).

\[ S \in \mathcal{W} \iff \sum_{\{i\} \in S} w_i \geq q \]
Example: United Nations Security Council. In 1978 U.N. Security Council had 15 members, 5 of whom had vetoes. For a resolution to pass it is necessary to have 9 affirmative votes and no vetoes.
To treat this as a 15-player simple game take $\mathcal{N} = \{1, \ldots, 15\}$ and suppose that the first five have the veto. We say that coalition $S$ is winning, i.e., $v(S) = 1$, if it can defeat a resolution, otherwise $v(S) = 0$. Thus the characteristic function is:
- if $i \in S$ for some $1 \leq i \leq 5$, then $v(S) = 1$
- otherwise,
\[ v(S) = \begin{cases} \ 1, & \text{if } |S| \geq 7 \\ \ 0, & \text{if } |S| < 7 \end{cases} \]

Fortunately, it is only necessary to calculate \( \varphi_1 \), since the symmetries of the game will then permit easy computation of \( \varphi_i \) for \( i > 5 \). Then,

\[
\varphi_1 = \sum_{1 \in S} \frac{(n-|S|)!(|S|-1)!}{n!} (v(S) - v(S \setminus \{i\})) =
\]

\[
\frac{14!0!}{15!} (v(\{1\}) - v(\emptyset)) + \frac{13!1!}{15!} (v(\{16\}) - v(6)) \left(10\right) +
\]

\[
\frac{12!2!}{15!} (v(\{167\}) - v(67)) \binom{10}{2} + \frac{11!3!}{15!} (v(\{1678\}) - v(678)) \binom{10}{3}
\]

\[
+ \frac{10!4!}{15!} (v(\{16789\}) - v(6789)) \binom{10}{4} + \frac{9!5!}{15!}
\]

\[
(v(\{1678910\}) - v(678910)) \binom{10}{5} + \frac{8!6!}{15!} (v(\{167891011\}) - v(67891011)) \binom{10}{6} = 0.19627...
\]
Hence,

\[
\varphi_i = \begin{cases} 
0.19627... & \text{if } 1 \leq i \leq 5 \\
0.00186... & \text{if } 6 \leq i \leq 15.
\end{cases}
\]

From this it is apparent that 98.1% of the power is in the hands of five permanent members and individually a permanent member is more than 105 times as powerful as an ordinary member.
Assumption: Player $i$ possesses a **Banzhaf power index**, which is proportional to the number of coalitions, in which he/she is critical (i.e., if he decides to leave the winning coalition, then this coalition becomes losing):

\[
\beta_i := \frac{\text{a number of winning coalition with Player } i}{\text{a number of all coalitions with Player } i}
\]

\[
= \sum_{S \subseteq N \setminus \{i\}} \frac{1}{2^{n-1}} [v(S \cup \{i\}) - v(S)]
\]

**Shapley Value ↔ Banzhaf Index**

Let us consider the following weighted voting game:

\[ [6; 4, 3, 2, 1] \]

Then, the **Shapley value** is: \( \varphi_i = \)

\[
\frac{1}{n!} \sum_{i \text{ is critical in } S} (|S| - 1)! (n - |S|)! \rightarrow \left( \frac{5}{12}, \frac{3}{12}, \frac{3}{12}, \frac{1}{12} \right)
\]

\( \)
Write all permutations of players and denote the critical players:

Player B is critical.

Player A as the first enters the grand coalition

Player B as the second enters the grand coalition

A\(\text{B}\) D C, A\(\text{C}\) B D, A\(\text{C}\) D B, A D\(\text{B}\) C, A D\(\text{C}\) B,...etc.
Sometimes the Shapley value in voting games is called a *Shapley-Shubik power index*;


**The Banzhaf power index** is:

\[
\beta = \left( \frac{5}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8} \right)
\]

Write all winning coalitions and denote the critical players (the order does not play a role):
Remark: Generally, $\sum_{i \in N} \beta_i \neq \nu(N) = 1$. Hence, the Banzhaf index cannot be used to the division of $\nu(N)$ among players, but it can be used to a study of importance of each player.
AXIOMATIZATIONS OF THE BANZHAF POWER INDEX

- P. Dubey, L.S. Shapley (1979): the axioms that give some semi-values;
- G. Owen (1982): the axioms, that do not give the Banzhaf index uniquely;
THE AXIOMS OF EHUD LEHRER:

(L1) \[ \forall S \subseteq N \setminus \{i\} \quad v(S \cup \{i\}) = v(S) + v(\{i\}) \Rightarrow \beta_i = v(\{i\}); \]

(L2) \[ \forall S \subseteq N \setminus \{i, j\} \quad v(S \cup \{i\}) = v(S \cup \{j\}) \Rightarrow \beta_i = \beta_j; \]

(L3) Let \( (v \land w)(S) = \min\{v(S), w(S)\} \) and \( (v \lor w)(S) = \max\{v(S), w(S)\} \); Then,
\[ \beta[v \land w] + \beta[v \lor w] = \beta[v] + \beta[w]; \]
The last axiom (L4) follows from this observation: Let us consider the following weighted voting game

\[ v = [1; \frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \]

Then,

\[ \beta[v] = \left( \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right) \quad \text{and} \quad \varphi[v] = \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right). \]

Now let us consider a new game

\[ w = [1; \frac{2}{3}, \frac{1}{3}] \], in which Player 1 and Player 2 are one player.
Then,

\[ \beta[v] = \left( \frac{1}{2}, \frac{1}{2} \right) = \varphi[v]. \]

Note that from the Shapley value’s point of view the unification does not pay off, because \( \frac{1}{3} + \frac{1}{3} > \frac{1}{2} \), but from the Banzhaf power index’ point of view the unification is never harmful, because

\[ \frac{1}{4} + \frac{1}{4} \leq \frac{1}{2}. \]
(L4) Let $v_p$ be $(n-1)$-player game, in which the players $\{i, j\}$ are one player $\{p\}$:

$$\beta_p[v_p] \geq \beta_i[v] + \beta_j[v].$$

**Theorem (Lehrer (1988))**

If a game satisfies the axioms (L1) - (L4), then the vector $\beta$ is a **Banzhaf power index**.
**Remark:** A.S. Nowak (1997) replaced the axiom \((L3)\) with an axiom similar to the one given by Young in the axiomatization of the Shapley value:

\[(L3')\] if for every \(S \subseteq N \setminus \{i\}\) \(v(S \cup \{i\}) - v(S) = w(S \cup \{i\}) - w(S)\), then \(\beta_i[v] = \beta_i[w]\).

**Theorem (Nowak (1997))**

If a game satisfies the axioms \((L1, L2, L4)\) and \((L3')\), then the vector \(\beta\) is a *Banzhaf power index*. 
BIBLIOGRAPHY TO THE COURSE ON “GAME THEORY”

• LECTURE 1

• LECTURE 2
• LECTURE 3
- T.S. Ferguson, *Game Theory*, 2005,

• LECTURE 4
• LECTURE 5

• LECTURE 6

• LECTURE 7
- A. Haurie, J.B. Krawczyk, *An Introduction to Dynamic Games*, 2000,
  

- **LECTURE 8**


- **LECTURE 9**

• LECTURE 10

• LECTURE 11
- A. Haurie, J.B. Krawczyk, *An Introduction to Dynamic Games*, 2000,
  
• LECTURE 12

• LECTURE 13

• LECTURE 14
- T.S. Ferguson, Game Theory, 2005,
  http://www.math.ucla.edu/~tom/Game_Theory/Contents.html
• **LECTURE 15**

- T.S. Ferguson, *Game Theory*, 2005,  