**FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University**

MASTER THESIS

Lukáš Račko

Discrete sequential games with random payoffs

Department of probability and mathematical statistics

Supervisor of the master thesis:	doc. RNDr. Ing. Miloš Kopa, PhD.
Study programme:	Probability, mathematical statistics and econometrics
Study branch:	Econometrics

Prague 2023

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date ....

Author's signature

I want to thank doc. Ing. Miloš Kopa, PhD. for his help and guidence trough-out my writting of this thesis. I also want to thank my family and friends for supporting me during my studies at the Charles university.

Title: Discrete sequential games with random payoffs

Author: Lukáš Račko

Department: Department of probability and mathematical statistics

Supervisor: doc. RNDr. Ing. Miloš Kopa, PhD., Department of probability and mathematical statistics

Abstract: In our thesis we consider games with random payoff as a generalizations of the standard concept of games in the game theory. We discuss possible optimality conditions for these types of games. In one of these approaches by the concept of a $\boldsymbol{\alpha}$-Nash equilibria we manage to prove the existence of this generalization of Nash equilibria for the case when the payoff has only finite number of realizations. We then apply those concepts to the case when the game is considered in multiple stages. In the practical part of this thesis we consider an application to a competition of internet providers which we model by a generalized version of the Cornout model of duopoly. We compare results of our optimal strategy with the deterministic approaches to this problem.

Keywords: sequential games, random payoff, non-cooperative games

Contents

1. Introduction

The game theory is a fundamental part of modern-day economics, evolutionary biology and even political science. The main goal of the theory is to mathematically represent decision problems faced by real-world agents and by this to find optimal strategies, that the agents should or rather would take. Game theory in the classical context studies behavior of agents in a deterministic scenarios. This means that the players' payoff is only determined by their actions. In this context, this payoff if observed by an outside observer, may still apear to be somehow random because randomization over several possible actions that may be chosen is most of the times preferential. However, given the players' strategies this payoff is considered deterministic.

This assumption is problematic, because in many situations the actual payoff depends on other factors such as demand on the market or enviroment in which the game is set. In most of the cases authors overcome this by considering these parameters as deterministic and setting them to their mean values. But as we see in the context of stochastic programming this approach may not result in solutions that are optimal and sometimes these solutions are not even feasible for the actual problem.

For this reason another way to model those types of games is by modeling the randomness involved in the game. The most popular model in this context are the stochastic games, which model the randomness by changing states of the game and considering the payoff in a given state as deterministic. This is a generalization of Markov decision process theory.

In our thesis we focus on another approach, that has not yet been considered by many authors, in which we assume that given the players' strategies the payoff is a random variable. We call this type of a game the game with random payoff.

In the second chapter of our thesis we use methods of optimization under uncertainty to develop optimality conditions for a single stage random payoff games. In this chapter we develop a generalization of the concept of a Nash equilibrium, which we call an $\boldsymbol{\alpha}$-Nash equilibrium, that is based on the idea of generalized best responses. Another approach is based on the quantile payoff model, that defines a deterministic equivalent of a game with random payoff, which was first considered in [14]. Lastly generalizations of the minimax problem using stochastic programming, which were first used for the case of matrix games with random payoff in [5], [4],[3] and [6]. We generalize this idea for n-player non-zero sum games.

In the third chapter of our thesis we further explore properties of deterministic sequential games that had been developed by multiple authors and are summarized in [11]. Here we define a sequential game, show how to convert sequential game into a game in the strategic form and consider models with finite and infinite number of stages. We also define the concept of information in the game, behavioral strategies and how they relate to our assumed model with perfect information. Lastly we more deeply consider possible solution models for repeated games and their properties with several approaches that were used to refine the concept of Nash equilibrium for extensive games.

In the fourth chapter of our thesis we discuss the case when randomness is involved in a sequential game. We start with the idea of markov decision process and its properties as were presented in [13] and then discuss the stochastic games as its generalization with their properties presented in [11]. With this we further discuss the case for sequential games with random payoff and their connection to these well known problems. We use criteria developed in the second chapter to present solution models for both the stepwise choise of strategy model and the agregate payoff models which are standardly used for deterministic sequential games. We discuss these approaches more deeply for the cases when the payoff processes have further properties for example if they are homogenous markov chains or are ergodic.

In the fifth chapter we consider the Cornout model of duopoly. We show how to find optimal production for the companies in the standard deterministic version of this model. We further generalize this model into model with dynamic demand and develop its optimal solutions for the case if there is no possibility of risk-free investment for the companies and for the case when the companies have the same risk-free investment oportunity (valuation of time). With this we define a stochastic version of the dynamic model, where we assume that the parameters of the model are random. After that we consider a model scenario of two internet providers who want to determine their optimal allocation of average daily capacity of their network for a 7-day period, where the daily average demand for the network capacity is random. On this model example we compare the optimal solutions from the deterministic approach that uses expected values of parameters of the model with the stochastic approach.

2. Games with random payoff

In this chapter we will take a look at the basic definition and results for mathematical games with random payoff.

2.1 Mathematical games

Definition 1 ([11]). *Let I be a set of players, $\forall i \in I : X_i$ be a set of strategies of the player i and $\forall i \in I : u_i : \bigtimes_{i\in I} X_i \to \mathbb{R}$ be a payoff function of the player i. The triple $(I, \{X_i\}_{i\in I}, \{u_i\}_{i\in I})$ we call a (mathematical) game.*

In contrast to games in classical concept of a game, such as sports, board or computer games, it is notable, that in our definition there is no explicit set of rules of the game itself. Those rules however, are determined implicitly by sets of players, strategies and profiles of the payoff functions. In this context we need to transform explicit rules of a "game" or rather a conflict situation we are studying into implicit rules given by our definition. In some situations those rules may change over time most notably in chess, if we consider one move of a player a single game in the mathematical sense in each round the "rules" change, because the position or number of chess pieces changes by each round and so the sets of strategies change. In other situations the number of players may differ or their respective payoffs. As we will see in the next chapter we can overcome this problem to some extent using methods of sequential games theory.

Let us examine simple example of game of Rock-paper-scissors. This is a game of two players with following meta rules:

- In the same time both players pick one out of three: rock, paper or scissors.
- Rock always beats scissors.
- Paper always beats rock.
- Scissors always beats paper.
- If both players pick the same item it is a tie.

From this list of rules we can now formalize our mathematical representation of this game. It is a game of two players therefore without a loss of generality let $I = \{1, 2\}$. Both players have the same available strategies. For reasons that will be apparent later we will formalize them as $X_1 = X_2 = Conv(\{(1, 0, 0)^T, (0, 1, 0)^T, (0, 0, 1)^T\})$. We will assume that both players have the same insentive to win and we will write their payoffs given each combination of strategies in the following table, where rows denote strategies of player 1 and columns of the player 2 and the first number of each element denotes the payoff of player 1 and the second of the player 2.

	Rock	Paper	Scissors
Rock	0,0	-1,1	1,-1
Paper	1,-1	0,0	-1,1
Scissors	-1,1	1,-1	0,0

We see that in each situation the sum of payoffs of both players is 0, therefore we can replace the table with single matrix

$$A = (a_{ij})_{3\times 3} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix},$$

that denotes the payoffs of the player 1 and the respective payoffs of player 2 are then given by $-A$. With this and our choise of representation of player strategies we can now clearly write

$$\begin{aligned} &\forall \boldsymbol{x} \in X_1, \forall \boldsymbol{y} \in X_2 : u_1(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T A \boldsymbol{y}, \\ &\forall \boldsymbol{x} \in X_1, \forall \boldsymbol{y} \in X_2 : u_2(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T (-A) \boldsymbol{y}. \end{aligned}$$

2.2 Categorization of games

In a previous example we have seen a special case of a game which is called a non-cooperative deterministic zero-sum matrix game. However there are many other different types of games. In this section we will take a closer look at their basic characterization.

We can differ games based on many properties, but the main are:

1. How many players does the game have?
2. Do players have an incentive to cooperate?
3. Is it a game of finite resources?
4. Does the game have multiple stages?
5. Is the outcome of the game deterministic? (It only depends on players' strategies.)

In following paragraphs we will discuss first 4 of the 5 possible characterizations. The distinction between deterministic and stochastic games will be discussed more deeply in the section corresponding to games with random payoff.

2.2.1 Types of games by the number of players

Based on the number of players the game has we can differ three cases:

Definition 2 ([11]). *Let $G = (I, \{X_i\}_{i\in I}, \{u_i\}_{i\in I})$ be a game.*

1. *If $|I| = 1$, we call it a single-player game.*
2. *If $|I| = 2$ we call it a two-player game.*
3. *If $|I| > 2$ we call it n-player game or multiple player game, where $n = |I|$ denotes the number of players that the game has.*

An example of a single-player game might be a problem of finding of an optimal portfolio for an investor. In this case the set of players has only one element $I = \{1\}$. Set of strategies is a set of possible portfolios, i.e. if we consider $a \in \mathbb{N}$ assets $X_1 = \{\boldsymbol{x} \in \mathbb{R}^a : \sum_{k=1}^{a} x_k = 1, h_i(\boldsymbol{x}) = 0, \forall i = 1, \ldots, N, g_j(\boldsymbol{x}) \leq 0, \forall j = 1, \ldots, M\}$. Where by functions h_i and g_j we denote any possible requirement the investor may be constraint by. For example if he does not want to borrow money (invest without short sales) he would require only such portfolio weights, that are non-negative. Finally the payoff function might be $u_1(\boldsymbol{x}) = \mathsf{E} \sum_{k=1}^{a} x_k \rho_k = \sum_{k=1}^{a} x_k r_k$, where by ρ_k we denote the return of the k-th asset and r_k the expected return of the k-th asset. As it is well known from the portfolio theory, there are many other payoff functions that we may consider and some of them, for example value at risk, are not linear in $\boldsymbol{x}$ as it was in this case. So in general those are not by any means 'simpler' games. But they are representing in general some non-linear programming problem. This also gives us an intuition that game theory in some sense generalizes standard optimization problems.

We have already seen a two player game in our Rock-paper-scissors example. Let us just note that matrix games are in fact an important class of two-player games as they represent every constant-sum game with finite number of pure strategies. This is a special case of lemma we will prove later.

2.2.2 Types of games by cooperation

Based on the incentive to cooperate we have three cases:

Definition 3 ([11])**.** *Let $G = (I, \{X_i\}_{i \in I}, \{u_i\}_{i \in I})$ be a game.*

1. *If $\forall i \in I : X_i = \{0, 1\}$ and $\forall i \in I : u_i = v : 2^I \to \mathbb{R}$, we call it a cooperative (or a coalition) game.*
2. *If $\forall i \in I : X_i = \{0, 1\}$, we call it a semi-cooperative game.*
3. *If G is not cooperative or semi-cooperative we call it non-cooperative game.*

It is easy to see that every cooperative game is also a semi-cooperative game. Important distinction is in games for which not all payoff functions are the same. That means that different players have a different incentive to cooperate with the rest. It may be that, they even get higher payoff if they do not cooperate with majority. In cooperative games there is single payoff function for all the players. In the theory of cooperative games it is studied how to 'fairly' distribute this payoff among members of the winning coalition. From the general perspective however, players want to maximize this payoff. In both types of games players have only two strategies denoted as 0 and 1, where 0 corresponds to not joining the coalition and 1 corresponds to joining the coalition.

For the rest of our thesis we will only work with non-cooperative and semi-cooperative games. Those games are characteristic in that they study behaviour of individuals rather than coalitions and coalitions occur only 'naturally', that is individuals form a coalition only if it is beneficial to them individually and the coalition is sustainable as long as no one has an incentive to leave it. This kind of behaviour is present in most of economic or political behaviours. For example individuals in market enviroment cooperate only if their personal gain would decrease otherwise or voters support a political party only if its policies are beneficial to them individually and in general do not cooperate as the transactional cost of such cooperation are too high (there is too many players to efficiently coordinate their strategies).

2.2.3 Types of games by available resources

Definition 4 ([11]). *Let $G = (I, \{X_i\}_{i \in I}, \{u_i\}_{i \in I})$ be a game.*

1. *If $\forall \boldsymbol{x} \in \bigtimes_{i \in I} X_i : \sum_{i \in I} u_i(\boldsymbol{x}) = c$ for some constant $c \in \mathbb{R}$ we call G a constant-sum or a c-sum game.*
2. *If G is not a constant-sum game we call it a non-constant-sum game.*

In the case of constant-sum games the overall gain of all players is a given constant. That means that players aim to get majority of given finite amount of resources and so gain of one player is translated to loss of other players. It is easy to see that such games are always non-cooperative. As an example of such game is a competition of companies on a closed market with given number of customers. For example grossery stores, that sell the same bread may increase their profit only by increasing the number of customers that buy bread in their store, which directly translates to decrease in number of customers for their competitors.

In contrast to that in non-constant-sum games players generaly may not lose payoff proportionally to gain of other player. Example of such game may be competition of companies in a market where by specific strategies they may gain new customers, that did not use specific product previously. For example technological companies may gain new customers and so increase their profit if they make their technologies more accessible (cheaper, easier to use or better marketed) for customers. By doing so, profit of their competitors may remain unchanged.

2.2.4 Types of games by the number of stages

Definition 5 ([11]). *Let $G = (I, \{X_i\}_{i \in I}, \{u_i\}_{i \in I})$ be a game.*

1. *If there is a $T \in \mathbb{N} \cup \{\infty\}$ such that $\forall i \in I : X_i = \bigtimes_{t=1}^{T} X_i^t$ and $u_i = \sum_{t=1}^{T} u_i^t$ we call G a sequential, multi-stage or T-stage game.*
2. *If G is a sequential game and $\forall i \in I, \forall s, t \leq T : X_i^s = X_i^t$ and $u_i^s \propto u_i^t$ we call G a repeating game.*
3. *If G is a sequential game and $\forall i \in I, \forall t \leq T, \forall \boldsymbol{x} \in \bigtimes_{i \in I} X_i^t : u_i^t(\boldsymbol{x}) = \alpha_i(t) f_i^t(\boldsymbol{x})$ for some function $f_i^t : \bigtimes_{i \in I} \to \mathbb{R}$ we call G a sequential game with discounting and $\{\alpha_i(t)\}_{t=1}^{T}$ we call a discounting of player i.*
4. *If G is a sequential game with $T = 1$ we call it a single-stage game.*
5. *If G is a T-stage game with $T > 1$ for $t \leq T$ we will call $G_t = (I, \{\bigtimes_{\tau=t}^{T} X_i^\tau\}_{i \in I}, \{\sum_{\tau=t}^{T} u_i^\tau\}_{i \in I})$ a sub-game of G.*

Where $f \propto g$ means that $f = cg$ for some $c \in \mathbb{R}$ constant.

It is apparent that every game is a sequential game as it has at least one stage. However as we will see later there are important distinctions between games with just one stage and games with multiple stages. Especially if there are infinite number of stages. The distinction is in the characterization of payoff and of optimal strategy. Most important feature of games with multiple stages is that we may allow players' strategies to be functions of strategies from the previous stages. This would mean, that players are able to adapt and learn from past experience. They may include their believes about long-term strategies of other players to their decision-making. Another important note is that each stage or subset of stages of multi-stage game may be also thought as a separate game. Later we will ask a question whether optimal solution of a multi-stage game is also optimal for its every sub-game?

We further distinguish between multi-stage games by introducing repeating game and game with discounting. In a repeating game we assume that payoffs of each stage are proportional to each other. This means, that they are same up to a multiplicative constant. This allows us to consider the case of repeating game with discounting as its special case. Repeating games may be thought as single-stage games that repeat T-times. It is however important to note that optimal solution of single-stage variant of the game may very much differ from optimal solution of the multi-stage game. Famously the Prisoner's dilema may be 'overcomed' in the multi-stage variant. For games with discounting the most straightforward motivation comes from games related to finance, where the payoffs are some future yields and discounting rate is used to represent that future yields are less valueable in present time, than current as the future value of money is smaller. But this concept is well aplicable and usefull also in games that are not related to finance, as it quantifies believes of players and how much they value future payoffs in comparison to today's. Therefore a consistent discounting would satisfy $\alpha_i(t) \leq 1$ which means, that every player should value future payoff at most the same way he values current payoff.

We have seen examples of single-stage games in previous discussion. As an example for multi-stage game we may consider a competition of companies on a market, which may occur over a finite or infinite horizont.

For the rest of this chapter we will only consider the case of single-stage games. We will discuss multi-stage games in the next chapter.

2.3 Strategies

Let us now look at different types of strategies, that players may play. In our basic example of a game of 'rock-paper-scissors', both players were able to play either 'rock', 'paper' or 'scissors'. Those are called the pure strategies. They are the single action players may opt to do at the beginning of the game. But as it is well known from the investment problem, sometimes it is preferencial to diversify strategies to minimize the associated risk. Strategies that are composites of multiple pure strategies will be called mixed strategies.

Definition 6 ([11])**.** *Let I be the set of players. $\forall i \in I$ denote P_i the set of distinct actions that player i can opt to play. We call P_i the set of pure strategies.*

Let $\forall i \in I : u_i : \bigtimes_{i\in I} P_i \to \mathbb{R}$ *be a payoff function of player* i*, we call* $G = (I, \{P_i\}_{i\in I}, \{u_i\}_{i\in I})$ *the game in pure strategies.*

Definition 7. *Let* P *be a topological space and* $\mathcal{B}$ *be a Borel* σ*-algebra over* P*. We denote* $\mathcal{D}(P) = \{D : \ D \text{ is a Borel probability measure on } (P, \mathcal{B})\}$*.*

In general we will consider the set of pure strategies to be a subset of $\mathbb{N}$, if there is at most countable different actions that the player may choose or a subset of $\mathbb{R}$, if there is uncountably many of such actions.

If P is a countable set we will consider $\mathcal{T} = 2^P$ to be the topology on P. If P is a uncountable set we will use the standard topology on $\mathbb{R}$. In generality a set of pure strategies may have a higher cardinality than continuum, this would for example happen in a sequential game with continuous time where in each sub-game players have a contiuum of pure strategies. Such a game could be used to represent a price creation on a market. This situation however will not be discussed in this thesis.

Before we discuss the mixed strategies, first lets take a look on a example of a game where players have uncountably many pure strategies. This game is called the 'Keynes beauty contest'. In this game players are asked to choose a number in inteval $[0, 1]$ such that it is closest to the $2/3$ of the mean number selected by all players. That is let $I = \{1, \dots, N\}$ be the finite set of players. Each player i chooses $p_i \in [0, 1]$ and the winner is given as

$$\arg\min_{i\in I} |p_i - \frac{2}{3}\frac{1}{N}\sum_{j\in I} p_j|.$$

Now we will define mixed strategies as follows.

Definition 8 ([11])**.** *Let* I *be the set of players.* $\forall i \in I$ *denote* P_i *the set of pure strategies of* i*. We call* $X_i = \mathcal{D}(P_i)$ *the set of mixed strategies. Let* $\forall i \in I : u_i : \bigtimes_{i\in I} X_i \to \mathbb{R}$ *be a payoff function of player* i*, we call* $G = (I, \{X_i\}_{i\in I}, \{u_i\}_{i\in I})$ *the game in mixed strategies corresponding to game in pure strategies with* $\{P_i\}_{i\in I}$*.*

In our thesis we will use the von Neuman - Morgenstern decision theory axioms. That is, if we consider G as the game in pure strategies and H the corresponding game in mixed strategies and $\hat{u}_i$ be the payoff function of player i in the game G then we will assume that the payoff function of player i in H is given as

$$u_i(\boldsymbol{x}) = \mathsf{E}\ \hat{u}_i(\boldsymbol{x}) = \int \hat{u}_i(\boldsymbol{p}) d\boldsymbol{x}(\boldsymbol{p}),$$

where $\boldsymbol{x} \in \bigtimes_{i\in I} X_i = \{\otimes_{i\in I} x_i : x_i \in X_i, \forall i \in I\}$ here $\mu_1 \otimes \mu_2$ is a product measure of μ_1 and μ_2 and so we assume, that the player strategies are independent. In more generality we could assume that they form a consistent family of probability distributions, that is $\bigtimes_{i\in I} X_i = \{\boldsymbol{x} \in \mathcal{D}(\bigtimes_{i\in I} P_i); \forall j \in I : \boldsymbol{x}_{-j} \in \mathcal{D}(\bigtimes_{i\in I, i\neq j} P_i)\}$ this would be what we call a semi-cooperative game. In a non-cooperative game it satisfies to assume, that player strategies are independent as we consider players to have no incentive to cooperate. This means that in a real-world situation, when agents have some reason to cooperate we consider players of a mathematical non-cooperative game representation of this conflict to be individuals or already formed coallitions of individuals with similar goals. Let us notice that by this we specifically assume that the payoff function of the game in pure strategies is $\mathcal{B}(\bigtimes_{i\in I} P_i)$-measurable.

If $I = \{1, \ldots, N\}$ is a finite set of players and P_i are countable sets of pure strategies this may be simplified to

$$u_i(\boldsymbol{x}) = \sum_{p_1 \in P_1} \cdots \sum_{p_N \in P_N} x_1(p_1) \ldots x_N(p_N) \hat{u}_i(p_1, \ldots, p_N).$$

Futhermore if P_i is a finite set for each player i with $|P_i| = K_i$ we can identify it with $\hat{P}_i = \{\boldsymbol{e}^k, k = 1, \ldots, K_i\}$, where $\boldsymbol{e}^k$ is the k-th standard basis vector of $\mathbb{R}^{K_i}$, that is

$$\boldsymbol{e}^k_m = \begin{cases} 0, & \text{if } m \neq k, \\ 1, & \text{if } m = k. \end{cases}$$

In this case X_i can be identified with $\hat{X}_i = conv(\hat{P}_i)$. In such a case we will for simplicity write P_i and X_i instead of $\hat{P}_i$ and $\hat{X}_i$ as the distinction between those two representations of strategies will come from the context.

Now we will show that sets of mixed strategies are convex.

Lemma 1 ([11])**.** *Let P be a topological space, then $\mathcal{D}(P)$ is a convex set.*

Proof. Let $x, y \in \mathcal{D}(P)$ and $\lambda \in (0, 1)$. $\lambda x + (1 - \lambda)y$ is a measure on $(P, \mathcal{B})$. Consider $B \in \mathcal{B}$ then

$$(\lambda x + (1 - \lambda)y)(B) = \lambda x(B) + (1 - \lambda)y(B) \leq \lambda + (1 - \lambda) = 1$$

and

$$(\lambda x + (1 - \lambda)y)(P) = \lambda + (1 - \lambda) = 1.$$

Lastly $\forall B \in \mathcal{B} : \lambda x(B) + (1 - \lambda)y(B)$ is a convex combination of non-negative numbers and therefore it is non-negative. From this we have that $\lambda x + (1 - \lambda)y$ is a borel probability measure on P and therefore $\lambda x + (1 - \lambda)y \in \mathcal{D}(P)$.

□

We see that each conflict situation may be identified as a game in pure or mixed strategies. Depending on a type of game, the interpretation of mixed strategies may be different. In cases as is the 'rock-paper-scissors' game, player may in one occurence of the game play only a single pure strategy in such a case the interpretation of mixed strategies as probability distributions is better. On the other hand if pure strategies are for example possible marketing strategies for a company, mixed strategies are better interpreted as a convex combination of pure strategies and therefore as chosen split of funds between the possible strategies.

Futhermore, we can also consider the case when players are constrained in possible mixed strategies. For this we will denote the set of possible constrains for player i with following sets of functionals

$$C_i^{\leq} = \{g; g : X_i \to \mathbb{R}, g \text{ is a inequality constraint for player } i\}$$

and

$$C_i^{=} = \{h; h : X_i \to \mathbb{R}, h \text{ is a equality constraint for player } i\}.$$

We then denote $X_i^C = \{x_i \in X_i : g(x_i) \leq 0, \forall g \in C_i^{\leq}, h(x_i) = 0, \forall h \in C_i^{=}\}$. Pension funds investmets are example of a case when it is necessary to consider constrained sets of strategies, as they are required by regulation to have some minimal portion of their investments put in obligations. If some of players in a game G play with a set of constrained strategies we will call G a game with constrained strategies.

Lastly let us note that if I is a countable set of players and $\forall i \in I : P_i$ is a countable set of pure strategies, we will call corresponding games in both mixed and pure strategies as discrete games. Futhermore, if I is a finite set of players and $\forall i \in I : P_i$ is a finite set of pure strategies, we will call corresponding games in mixed and pure strategies as finite. In this thesis we will only consider the case of discrete games.

2.4 Optimal strategies

Before we discuss the case of games with random payoff we will introduce how to classify optimal strategies in single-stage games. Without a loss of generality in this thesis we will assume, that all players want to maximize their payoff, that is, they consider higher payoff preferencial over smaller payoff whenever it is possible to choose. We will also consider that all players are perfectly rational, that means we will suppose that all players are able to consider all possible options before game starts and choose one that benefits them most.

Let us begin by examining our example of Rock-paper-scissors. We have seen that we can represent this game as a matrix game with matrix

$$A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

In this case we considered as our sets of strategies a convex hull of elements of standard basis of $\mathbb{R}^3$. This would be interpreted as that strategy for a player is a probability distribution giving probabilities of whether to play 'rock', 'paper' or 'scissors'. Such strategies are called a mixed strategy in the context of game theory. For start let us observe the payoff of player 1 for an arbitrary combination of such strategies

$$\begin{aligned} u_1(\boldsymbol{x}, \boldsymbol{y}) &= x_2y_1 - x_3y_1 - x_1y_2 + x_3y_2 + x_1y_3 - x_2y_3, \\ &= x_1(y_3 - y_2) + x_2(y_1 - y_3) + x_3(y_2 - y_1), \qquad (2.1) \\ &= y_1(x_2 - x_3) + y_2(x_3 - x_1) + y_3(x_1 - x_2). \qquad (2.2) \end{aligned}$$

From this rewritting we see that if player 1 prefers x_1 that is $x_1 > x_2$ and $x_1 > x_3$ to maximize his payoff, player 2 will have incentive to play $y_3 = 0$, since $u_2 = -u_1$. And so player 1 will have payoff

$$\begin{aligned} u_1(\boldsymbol{x}, \boldsymbol{y}) &= x_1(-y_2) + x_2y_1 + x_3y_2 + x_3(-y_1) \\ &= y_2(x_3 - x_1) + y_1(x_2 - x_3), \end{aligned}$$

where the first term is negative as $x_1 > x_3$ and so player 1 would increase his payoff if he would choose $x_1 = x_3$. Similiar argument may be made for x_2 and x_3 as well and from symetry of the perspective, this can be also argued for the player 2. From this both players would benefit in some cases from chosing a different strategy. We will formalize this in the next definition.

Definition 9 ([11],Definition 4.18)**.** *Let G be a non-cooperative game. We will define the set of best response strategies of player $i \in I$ with respect to other players' strategies $\boldsymbol{x}_{-i} \in X_{-i}$ as*

$$B_i(\boldsymbol{x}_{-i}) = \{x \in X_i : \forall y \in X_i, u_i(x, \boldsymbol{x}_{-i}) \geq u_i(y, \boldsymbol{x}_{-i})\}.$$

The set of best response strategies shows players' insentives to change their strategy and its 'direction'. Interesting case is when for every player it contains their current strategy. In such a case every player is willing to continue to play the same strategy. This means that with perfectly rational players the strategy profile will stop in such a equilibrium point.

Definition 10 ([11])**.** *Let G be a non-cooperative game. We will define the Nash equilibrium of G as such $\boldsymbol{x} \in X$ that $\forall i \in I : x_i \in B_i(\boldsymbol{x}_{-i})$. We will denote the set of Nash equilibria of G as $NE(G)$.*

Nash equlibria of a game were first examined by mathematician John von Neuman, who managed to prove their existence for every finite two-player game in mixed strategies. The famous result about existence of Nash equilibria was done by John Nash Jr., who proved the existence of those equilibria for every finite game in mixed strategies. We formulate this result in the following theorem.

Theorem 2 (The fundamental theorem of game theory, [11])**.** *Let G be a finite non-cooperative game in mixed strategies, then $NE(G) \neq \emptyset$.*

Proof. For the proof we refer to [11].

□

This theorem was further generalized for the case of games with constrained strategies.

Theorem 3 ([11])**.** *Let $G = (I, \{X_i\}_{i \in I}, \{u_i\}_{i \in I})$ be a finite game in constraint strategies, where for each $i \in I$:*

- *The set X_i is a polytope in $\mathbb{R}^{k_i}$.*
- *Function u_i is a multilinear function over the set of pure strategies.*

Then G has a Nash equilibrium.

Proof. For the proof we refer to [11].

□

Lemma 4. *Let $G = (I, X, \{u_i\}_{i \in I})$ be a non-cooperative game. Let $\forall i \in I : K_i \in \mathbb{R}$, $a_i > 0$ then $\boldsymbol{x}^* \in X$ is a Nash equilibrium of G, if and only if $\boldsymbol{x}^*$ is a Nash equilibrium of $H = (I, X, \{a_i u_i + K_i\}_{i \in I})$.*

Proof. To prove this theorem we will show that $\forall i \in I : \forall \boldsymbol{x}_{-i} \in X_{-i}\ B_i^G(\boldsymbol{x}_{-i}) = B_i^H(\boldsymbol{x}_{-i})$. Where $B_i^G(\boldsymbol{x}_{-i})$ and $B_i^H(\boldsymbol{x}_{-i})$ are sets of best response strategies for G and H respectivelly. We have that

$$\forall y \in X_i : u_i(x, \boldsymbol{x}_{-i}) \geq u_i(y, \boldsymbol{x}_{-i}),$$

if and only if

$$\forall y \in X_i : a_i u_i(x, \boldsymbol{x}_{-i}) + K_i \geq a_i u_i(y, \boldsymbol{x}_{-i}) + K_i.$$

This implies that $x \in B_i^G(\boldsymbol{x}_{-i})$, if and only if $x \in B_i^H(\boldsymbol{x}_{-i})$.

$\square$

Special case of this lemma says that every constant-sum game may be represented using zero-sum game.

The main idea of the Nash equilibria is that players will not have a reason to change their current strategy as they would expect their payoff to decrease. Another approach to optimality may be to consider methods of multi-objective programming. That is to find optimal strategies, we want to solve multi-objective problem

$$\max_{\boldsymbol{x} \in X} \boldsymbol{u}(\boldsymbol{x}),$$

where $\boldsymbol{u} = (u_i)_{i \in I}$ is the vector of payoff functions of all players. We will introduce following notation for vector inequalities, if $\boldsymbol{a} \leq \boldsymbol{b}$ and $\boldsymbol{a} \neq \boldsymbol{b}$ we will write $\boldsymbol{a} \lneq \boldsymbol{b}$

Definition 11. *Let $G = (I, X, \boldsymbol{u})$ be a non-cooperative game. Strategy profile $\boldsymbol{x} \in X$ is called efficient if there is no $\hat{\boldsymbol{x}} \in X$ such that $\boldsymbol{u}(\boldsymbol{x}) \lneq \boldsymbol{u}(\hat{\boldsymbol{x}})$. We will denote the set of efficient profiles of game G as $EF(G)$.*

We can also express the set of ideal strategy profiles as

$$ID(G) = \bigcap_{i \in I} \arg\max_{\boldsymbol{x} \in X} u_i(\boldsymbol{x}).$$

For most non-trivial games G however, it holds that $ID(G) = \emptyset$.

We can as well consider players to defend themselves against the worst possible outcome of the game. In such a case player $i \in I$ would want to solve problem

$$v_i^L = \max_{x_i \in X_i} \min_{\boldsymbol{x}_{-i} \in X_{-i}} u_i(x_i, \boldsymbol{x}_{-i}). \tag{2.3}$$

Here v_i^L denotes the lower value of the game for the player i. The lower value of the game is the lowest feasible payoff of the rational player i in the G. This means that the rational player is always able to gain at least v_i^L. Similiarly we will define

$$v_i^U = \min_{\boldsymbol{x}_{-i} \in X_{-i}} \max_{x_i \in X_i} u_i(x_i, \boldsymbol{x}_{-i}) \tag{2.4}$$

the upper value of G for the player i.

Lemma 5 ([11], Theorem 5.40)**.** *Let $G = (I, \{X_i\}_{i \in I}, \{u_i\}_{i \in I})$ be a game, then $\forall i \in I : v_i^U \geq v_i^L$.*

Proof. For the proof of this lemma we refer to [11].

$\square$

Definition 12. *Let G be a non-cooperative game in mixed strategies. Strategy $x_i \in X_i$ that solves (2.3) is called the minimax strategy of player $i \in I$. We denote $MM_i(G)$ the set of minimax strategies for the player i and $MM(G) = \{(x_i)_{i\in I} \in X : x_i \in MM_i(G)\}$ the set of minimax strategy profiles.*

We will say that two mathematical programs M_1 and M_2 are congruent, if $x_1^* = x_2^*$, where x_1^* is the optimal solution of M_1 and x_2^* is the optimal solution of M_2 and $f_1(x_1^*) = f_2(x_2^*)$, where f_1 and f_2 are objective functions of M_1 and M_2 respectively. We write $M_1 \equiv M_2$.

We propose a following claim.

Theorem 6. *Let G be a non-cooperative game in mixed strategies such that $NE(G) \neq \emptyset$, then*

$$ID(G) \subseteq NE(G), \tag{2.5}$$
$$ID(G) \subseteq EF(G). \tag{2.6}$$

If

$$\forall i \in I : M_i = \max_{x_i \in X_i} \min_{\boldsymbol{x}_{-i} \in X_{-i}} u_i(x_i, \boldsymbol{x}_{-i}) \equiv N_i = \min_{\boldsymbol{x}_{-i} \in X_{-i}} \max_{x_i \in X_i} u_i(x_i, \boldsymbol{x}_{-i})$$

then also

$$MM(G) \subseteq NE(G). \tag{2.7}$$

Proof. If $\boldsymbol{x} \in ID(G)$ then it must be a Nash equilibrium because $\forall i \in I$: $\boldsymbol{x} \in \arg\max_{\boldsymbol{x}\in X} u_i(\boldsymbol{x})$. This means that $\forall i \in I, \forall \boldsymbol{y} \in X : u_i(\boldsymbol{y}) \leq u_i(\boldsymbol{x})$ for which a special case is $\forall i \in I, \forall y_i \in X_i : u_i(y_i, \boldsymbol{x}_{-i}) \leq u_i(x_i, \boldsymbol{x}_{-i}) = u_i(\boldsymbol{x})$ and this implies that $\forall i \in I : x_i \in BR_i(\boldsymbol{x}_{-i})$. Therefore $\boldsymbol{x} \in NE(G)$ and also $\boldsymbol{x} \in EF(G)$. Now let $\boldsymbol{x} \in MM(G)$ and take $i \in I$. Denote $(x_i, \boldsymbol{y}_{-i})$ the optimal solution of (2.3). From our assumption it is also the optimal solution of the program N_i. This implies that

$$\forall \boldsymbol{z}_{-i} \in X_{-i}, \forall z_i \in X_i : u_i(x_i, \boldsymbol{z}_{-i}) \geq u_i(x_i, \boldsymbol{y}_{-i}) \geq \max_{z_i \in X_i} u_i(z_i, \boldsymbol{z}_{-i}) \geq u_i(z_i, \boldsymbol{z}_{-i}).$$

Which specially holds true for $\boldsymbol{z}_i = \boldsymbol{x}_{-i}$ from which we get $\forall z_i \in X_i$: $u(x_i, \boldsymbol{x}_{-i}) \geq u_i(z_i, \boldsymbol{x}_{-i})$ or in other words $x_i \in BR_i(\boldsymbol{x}_{-i})$. Because we chose an arbitrary $i \in I$ it must be that $\boldsymbol{x} \in NE(G)$. But there may exist $i \in I$ and $\hat{\boldsymbol{x}}_{-i} \in X_{-i}$ such that $u_i(x_i, \hat{\boldsymbol{x}}_{-i}) > u_i(x_i, \boldsymbol{x}_{-i})$, so in general $MM(G) \neq EF(G)$ which also implies that $MM(G) \neq NE(G)$ and $EF(G) \neq NE(G)$.

$\square$

2.5 Dominance

Now we will propose two notions of dominance. First is based on dominance in multi-objective programming and the second one is based on the minimax idea, that player wants to defend against the worst possible situation.

Definition 13. *Let $\boldsymbol{x}, \boldsymbol{y} \in X$. If $\boldsymbol{u}(\boldsymbol{x}) \gneq \boldsymbol{u}(\boldsymbol{y})$ we say that strategy profile $\boldsymbol{x}$ dominates strategy profile $\boldsymbol{y}$ (or that $\boldsymbol{y}$ is dominated by $\boldsymbol{x}$) and write $\boldsymbol{x} \gg \boldsymbol{y}$.*

Definition 14 ([11])**.** *Let $x_i, y_i \in X_i$, $S \subseteq X_{-i}$. If $\forall s \in S : u_i(x_i, s) \geq u_i(y_i, s)$ we say that strategy x_i (very) weakly dominates strategy y_i with respect to S (or that y_i is weakly dominated by x_i with respect to S) and write that $x_i \succeq_S y_i$. If $\forall s \in S : u_i(x_i, s) > u_i(y_i, s)$ we say that strategy x_i strictly dominates strategy y_i with respect to S(or that y_i is strongly dominated by x_i with respect to S) and write that $x_i \succ_S y_i$.*

By using the dominance based on strategy profiles we always reduce the set of strategy profiles to a unique set of undominated profiles, which are called efficient. Properties of a so-called 'maximal reduction' under iterated removal of weakly or strictly dominated strategies had been studied by multiple authors. Existence and uniqness of such a reduction was discussed for example in [7] or [8] and more general results about iterative elimination procedures in general choice problems were shown in [9].

2.6 Random payoff

So far we discussed only the case of deterministic games, that is games for which only player actions may have resulted in the change of payoff. In this section we will take a look on the case when randomness is involved. This means that now given players' strategy profile the payoff function will be a random variable.

Definition 15. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let I be the set of players, $\forall i \in I$ let X_i be the set of strategies of player i and $u_i : \bigtimes_{i \in I} X_i \times \Omega \to \mathbb{R}$ such that $u_i(x)$ is $\mathcal{A}$-measurable for every $x \in \bigtimes_{i \in I} X_i$. We say that the triple $G = (I, \{X_i\}_{i \in I}, \{u_i\}_{i \in I})$ is a game with random payoff. For a given $\omega \in \Omega$ we will write $G(\omega)$ and mean $(I, \{X_i\}_{i \in I}, \{u_i(\omega)\}_{i \in I})$ the realization of G for scenario ω. We say that payoff function $\{u_i(\boldsymbol{x})\}_{i \in I}$ has a distribution $D_{\boldsymbol{x}}$, if $D_{\boldsymbol{x}}(B) = P_{\boldsymbol{u}(\boldsymbol{x})}(B) = \mathbb{P}(\boldsymbol{u}(\boldsymbol{x}) \in B), \forall B \in \mathcal{A}$.*

First thing for us to understand is that randomness in the Definition 15 is different from the randomness that comes from players playing mixed strategies. Under the Von Neuman - Morgenstern axioms the payoff function of a game in mixed strategies is the expected payoff from playing each pure strategy with a given probability, however in reality players observe the outcome for a certain combination of pure strategies, which is given by a realization of a combination of mixed strategies played by each player. In a game with random payoff we consider situation when, even if all players knew strategies of their oponents before the occurance of game, they would still observe randomness in their payoff. This randomness may be naturally occuring for example in a competition of companies on certain market, where the demand for their products is random or it may be as a result of their incomplete knowlege of the game, i.e. when there are some unknown players in the game. One interesting statistical question to ask in a game with random payoff with multiple stages is how to determine whether the change in payoff distribution observed by players is the result of different distribution of the payoff function or whether it is the result of other players choosing different strategy?

First let us discuss a case of unknown players in the game.

Let $G = (I, \{X_i\}_{i\in I}, \{u_i\}_{i\in I})$ be a deterministic game in mixed strategies and suppose that we are only able to observe game dynamics for $J \subsetneq I$ set of players then $\hat{G} = (J, \{X_j\}_{j\in J}, \{u_j\}_{j\in J})$ is a game with random payoff. Let P_i be the set of pure strategies for the player $i \in I$, then we can set $\Omega = \bigtimes_{i\in I\setminus J} P_i$, $\mathcal{A} = \mathcal{B}(\Omega)$ and $\mathbb{P} = \bigotimes_{i\in I\setminus J} x_i$. We need to show that $u_i(\boldsymbol{x}) = \int_{\bigtimes_{j\in J} P_j} \hat{u}_i(\boldsymbol{p}) d\boldsymbol{x}(\boldsymbol{p})$ is $\mathcal{A}$-measurable for all $\boldsymbol{x} \in \bigtimes_{j\in J} X_j$, where $\hat{u}_i$ is a payoff function from the corresponding game in pure strategies. We have that $\hat{u}_i$ is $\mathcal{B}(\bigtimes_{i\in I} P_i)$-measurable function and so it is also measurable with respect to $\mathcal{M} = \sigma(\{B \times \bigtimes_{j\in J} P_j; B \in \mathcal{B}(\Omega)\}) \subset \mathcal{B}(\bigtimes_{i\in I} P_i)$ from this we get that $u_i(\boldsymbol{x})$ is $\mathcal{B}(\Omega)$-measurable for all $\boldsymbol{x} \in \bigtimes_{j\in J} X_j$. This motivates us to think about the random elements in the game with random payoff as "nature's" strategy. In this model nature chooses a strategy $\omega \in \Omega$, that is unknown to players and then they want to find their optimal strategies.

In games with random payoff we need to find a new way to clasify optimal strategies. We can consider all of the definitions of optimal strategies to hold almost surely. That is we could use similiar criteria on optimality with addition that there is $N \subset \Omega$ such that those definitions hold for $G(\omega), \forall \omega \in \Omega \setminus N$ and $\mathbb{P}(N) = 0$. This however in general may be too restricting, especially in the case of the Nash equilibrium where there is so far no known result about its existence in a general game with random payoff. Therefore we will require those definitions to hold on some prescribed level of confidence.

2.6.1 Generalizing Nash equilibria

First approach to optimality we will be discussing is generalizing the idea of Nash equilibrium as a best response to itself. We will start by defining the set of α-best response strategies, where $\alpha \in [0, 1]$ is a prescribed level of confidence. Using this we may define an $\boldsymbol{\alpha}$-Nash equilibrium as a strategy profile, such that it is a α_i-best response to itself in every player's strategy. The sets of α-best responses may be defined using individual or joint probabilistic constraints. We will start by considering the case with individual constraints.

Definition 16. *Let G be a game with random payoff and $\alpha_i \in [0, 1]$. We define the set of best response strategies on a confidence level of α for the player i, given strategy profile of other players $\boldsymbol{x}_{-i} \in X_{-i}$ as*

$$B_i^{\alpha_i}(\boldsymbol{x}_{-i}) = \{x_i \in X_i; \forall y \in X_i, \mathbb{P}(u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(y, \boldsymbol{x}_{-i})) \geq \alpha_i\}.$$

Similiarly we will define a Nash equilibrium.

Definition 17. *Let G be a game with random payoff and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots) \in [0, 1]^I$. If $\forall i \in I : x_i \in B_i^{\alpha_i}(\boldsymbol{x}_{-i})$, we say that $\boldsymbol{x} \in X$ is a Nash equilibrium of G on confidence levels $\boldsymbol{\alpha}$ or that it is a $\boldsymbol{\alpha}$-Nash equilibrium of G. We denote the set of $\boldsymbol{\alpha}$-Nash equilibria of G as $\boldsymbol{\alpha}$-$NE(G)$. Vector $\boldsymbol{\alpha}$ is called the confidence level vector.*

The question we may ask is whether those sets are also convex? This holds true for a special case of payoff function, which is a standard result of stochastic programming.

Lemma 7 (Convexity of sets of best response strategies.)**.** *If* $\forall \omega \in \Omega$, $\forall y \in X_i$: $u_i(x, \boldsymbol{x}_{-i}, \omega) - u_i(y, \boldsymbol{x}_{-i}, \omega) = f_i(x, y) - g(\omega)$, *where* $f_i(x, y)$ *is a quasiconcave function of* $x \in X_i$ *then* $B_i^\alpha(\boldsymbol{x}_{-i})$ *is a convex set for all* $\alpha \in [0, 1]$.

Proof. For the proof of convexity of sets with probabilistic constrains we refer to [2].

□

Now lets take a look on how both sets of best response strategies and of Nash equilibria behave when we change the confidence levels on which we consider the game.

Lemma 8 (Monotonicity in confidence levels)**.** *Let* G *be a game with random payoff. Then:*

1. $\forall i \in I : \forall \boldsymbol{x}_{-i} \in X_{-i} : B_i^\alpha(\boldsymbol{x}_{-i})$ *is non-increasing in* α. *That is* $\forall \alpha \geq \beta$: $B_i^\alpha(\boldsymbol{x}_{-i}) \subseteq B_i^\beta(\boldsymbol{x}_{-i})$.

2. $\forall i \in I$: $\boldsymbol{\alpha}$-$NE(G)$ *is non-increasing in* $\boldsymbol{\alpha}$. *That is* $\forall \boldsymbol{\alpha} \geq \boldsymbol{\beta}$ *we have that* $\boldsymbol{\alpha}$-$NE(G) \subseteq \boldsymbol{\beta}$-$NE(G)$.

Proof. 1. Let $\alpha \geq \beta$, $\boldsymbol{x}_{-i} \in X_{-i}$ and $x_i \in B_i^\alpha(\boldsymbol{x}_{-i})$ then we have that $\forall y \in X_i$: $\mathbb{P}(u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(y, \boldsymbol{x}_{-i})) \geq \alpha \geq \beta$ and so $x \in B_i^\beta(\boldsymbol{x}_{-i})$.

2. Let $\boldsymbol{\alpha} \geq \boldsymbol{\beta}$ and $\boldsymbol{x} \in \boldsymbol{\alpha}$-$NE(G)$ then $\forall i \in I : x_i \in B_i^{\alpha_i}(\boldsymbol{x}_{-i})$ and from the first part of this proof we have that $B_i^{\alpha_i}(\boldsymbol{x}_{-i}) \subseteq B_i^{\beta_i}(\boldsymbol{x}_{-i})$ therefore $\forall i \in I : x_i \in B_i^{\beta_i}(\boldsymbol{x}_{-i})$ and so $\boldsymbol{x} \in \boldsymbol{\beta}$-$NE(G)$.

□

Clearly we have that $B_i^0(\boldsymbol{x}_{-i}) = X_i$ and $\mathbf{0}$-$NE(G) = X$. By this lemma, if $B_i^1(\boldsymbol{x}_{-i}) \neq \emptyset$ then $B_i^\alpha(\boldsymbol{x}_{-i}) \neq \emptyset, \forall \alpha \in (0, 1)$. Similiarly, if $\mathbf{1}$-$NE(G) \neq \emptyset$ then $\boldsymbol{\alpha}$-$NE(G) \neq \emptyset, \forall \boldsymbol{\alpha} \in (0, 1)^I$. This implies that our definition of best response strategies and set of Nash equilibria is consistent with the almost surely definition in a sense that if a strategy profile is a Nash equilibrium almost surely, then it is also $\boldsymbol{\alpha}$-Nash equilibrium. So far there has been to the best of our knowledge no results on when there $\exists \boldsymbol{\alpha}$ such that $\boldsymbol{\alpha}$-$NE(G) \neq \emptyset$.

We propose a following existence theorem which is based on the idea of unknown players.

Theorem 9. *Let* G *be a game with random payoff in mixed strategies on a probability space* $(\Omega, \mathcal{A}, \mathbb{P})$. *If there is a deterministic game* H *such that following holds:*

1. *the set of pure strategy profiles of* H *is* $P_H = P_G \times \Omega \times P_O$, *where* P_G *is the set of pure strategy profiles of* G *and* P_O *is the set of pure strategy profiles of the uknown players which does not influence the payoffs of players in* G *(*G *is a game with unknown players.),*

2. $\forall i \in I_G : X_i$ *is a compact metric space,*

3. $\forall i \in I_G : 0 \leq u_i^G(\omega) \leq K_i < \infty$ $, \forall \omega \in \Omega$, *where* u_i^G *denotes the payoff function of player* i *in* G *(payoff functions are surely bounded.),*

4. $\forall i \in I_G : u_i^G$ *is a constant function of* P_O *(payoff functions are independent of the orthogonal players),*

5. $g_i(y, \omega) = u_i^G(y, \boldsymbol{x}_{-i}, \omega)$ *is a continuous function of* y *for every fixed* $\boldsymbol{x}_{-i} \in X_{-i}$ *and* $\forall \omega \in \Omega$ *(*g_i *is a surely continuous function),*

6. there exists a Nash equilibrium of H *such that* $\boldsymbol{x}_H^* = \boldsymbol{x}_G^* \otimes \mathbb{P} \otimes \boldsymbol{x}_O^*$ *(*$\mathbb{P}$ *is an optimal strategy profile of uknown players).*

Then there exists $\boldsymbol{\alpha} \in (0,1]^{I_G}$ *confidence levels such that* $\boldsymbol{x}_G^*$ *is a* $\boldsymbol{\alpha}$*-Nash equilibrium of* G*.*

Proof. Let I_G and I_H denote the sets of players in G and H respectivelly. Because $\boldsymbol{x}_H^*$ is a Nash equilibrium and $I_G \subset I_H$ we have that

$$\forall i \in I_G, \forall y \in X_i : u_i^H(\boldsymbol{x}_G, \mathbb{P}) \geq u_i^H(y, \boldsymbol{x}_{G,-i}^*, \mathbb{P}).$$

From Von Neuman - Morgenstern axioms and Fubini theorem we have following inequalities

$$\begin{gathered}
\forall i \in I_G, \forall y \in X_i : u_i^H(\boldsymbol{x}_G, \mathbb{P}) \geq u_i^H(y, \boldsymbol{x}_{G,-i}^*, \mathbb{P}), \\
\forall i \in I_G, \forall y \in X_i : \int u_i^H(\boldsymbol{p}_G, \omega) d(\boldsymbol{x}_G^* \otimes \mathbb{P}) \geq \int u_i^H(p_i, \boldsymbol{p}_{-i}, \omega) d(y \otimes \boldsymbol{x}_{G,-i}^* \otimes \mathbb{P}), \\
\forall i \in I_G, \forall y \in X_i : \int \int u_i^H(\boldsymbol{p}_G, \omega) d(\boldsymbol{x}_G^*) d\mathbb{P} \geq \int \int u_i^H(p_i, \boldsymbol{p}_{-i}, \omega) d(y \otimes \boldsymbol{x}_{G,-i}^*) d\mathbb{P}, \\
\forall i \in I_G, \forall y \in X_i : \int u_i^G(\boldsymbol{x}_G^*, \omega) d\mathbb{P} \geq \int u_i^G(y, \boldsymbol{x}_{G,-i}^*, \omega) d\mathbb{P}.
\end{gathered}$$

Now denote $f(\omega) = u_i^G(\boldsymbol{x}_G^*, \omega)$ and $g_i(y, \omega) = u_i^G(y, \boldsymbol{x}_{G,-i}^*, \omega)$. From above and linearity of the Lebesgue integral we have that

$$\forall i \in I_G, \forall y \in X_i : \int f - g_i(y) d\mathbb{P} \geq 0,$$

which means that

$$\forall i \in I_G, \forall y \in X_i : \int_{\{f > g_i(y)\}} f - g_i(y) d\mathbb{P} \geq \int_{\{f < g_i(y)\}} g_i(y) - f d\mathbb{P}. \tag{2.8}$$

For the right hand side of ineaquality we get

$$\forall i \in I_G, \forall y \in X_i : \int_{\{f < g_i(y)\}} g_i(y) - f d\mathbb{P} \geq 0.$$

Take $i \in I_G$, $y \in X_i$ and first suppose

$$\int_{\{f < g_i(y)\}} g_i(y) - f d\mathbb{P} = 0,$$

this means that $\mathbb{P}(\{f < g_i(y)\}) = 0$ or in other words that $\mathbb{P}(u_i^G(\boldsymbol{x}_G^*) \geq u_i^G(y, \boldsymbol{x}_{G,-i}^*)) = 1$.

Now suppose that

$$\int_{\{f < g_i(y)\}} g_i(y) - f d\mathbb{P} \geq c_i(y) > 0,$$

because $\mathbb{P}$ is a probability measure there exists $a_i(y) > c_i(y)$ such that

$$\int_{\{f<g_i(y)\}} g_i(y) - f d\mathbb{P} = a_i(y)\mathbb{P}(\{f < g_i(y)\})$$

and because X_i is a compact set there exists $K_i \geq a_i = \min_{y\in X_i} a_i(y) > 0$, where we set $a_i(y) = K_i$, if $\mathbb{P}(\{f < g_i(y)\}) = 0$. This means that

$$\forall i \in I_G : \forall y \in X_i : a_i(y)\mathbb{P}(\{f < g_i(y)\}) \geq a_i\mathbb{P}(\{f < g_i(y)\}).$$

We set $Y_i = \{y \in X_i; \mathbb{P}(f < g_i(y)) > 0\}$, $Y_i \subset X_i$ compact, therefore Y_i is relatively compact set, furthermore we show that Y_i is closed and therefore it is compact subset of X_i. Take Cauchy sequence $\{y_n\}_{n\in\mathbb{N}} \subset Y_i \subset X_i$, so there exists $y \in X_i$ such that $y_n \to y, n \to \infty$. Therefore $\forall \epsilon > 0, \exists n_0 \in \mathbb{N} : \forall n \geq n_0 : y_n \in \mathcal{U}_\epsilon(y)$. Because $g_i(y)$ is surely continuous in y, $\forall y_n : \exists \delta_n > 0 : \forall z \in \mathcal{U}_{\delta_n}(y_n) : \mathbb{P}(f < g_i(z)) > 0$ and consequently for $n \geq n_0 : \mathcal{U}_{\delta_n}(y_n) \cap \mathcal{U}_\epsilon(y) \neq \emptyset$ and so $\bigcup_{n\geq n_0} \mathcal{U}_{\delta_n}(y_n) \cap \mathcal{U}_\epsilon(y) \neq \emptyset$ holds for every $\epsilon > 0$, therefore $y \in \bigcup_{n\in\mathbb{N}} \mathcal{U}_{\delta_n}(y_n)$ from which $\mathbb{P}(f < g_i(y)) > 0$ and so $y \in Y_i$. This means Y_i is closed relativelly compact set in a compact metric space so it is compact.

Now define $g_i(y^*) = \min_{y\in Y_i} g_i(y)$ it holds that $\mathbb{P}(f < g_i(y^*)) > 0$ and we get that

$$\forall y \in Y_i : \mathbb{P}(f < g_i(y)) \geq \mathbb{P}(f < g_i(y^*)) > 0$$

Now lets take a look on the left side of (2.8), it holds

$$\begin{aligned}\forall i \in I_G, \forall y \in X_i : \int_{\{f>g_i(y)\}} f - g_i(y) d\mathbb{P} &\leq \int_{\{f>g_i(y)\}} f d\mathbb{P} \leq \\ &\leq \int_{\{f>g_i(y)\}} \sup_{\omega\in\Omega} f d\mathbb{P} = \sup_{\omega\in\Omega} f\mathbb{P}(\{f > g_i(y)\}) \leq K_i\mathbb{P}(\{f \geq g_i(y)\})\end{aligned}$$

By comparing the two sides of inequality we get that

$$\forall i \in I_G, \forall y \in Y_i : \mathbb{P}(\{f \geq g_i(y)\}) \geq \frac{a_i}{K_i}\mathbb{P}(\{f < g_i(y^*)\}).$$

Which gives us that $\forall i \in I_G$, $\forall y \in X_i$ either $\mathbb{P}(\{f < g_i(y)\}) = 0$ and then $\mathbb{P}(u_i^G(\boldsymbol{x}_G^*) \geq u_i^G(y, \boldsymbol{x}_{G,-i}^*)) = 1 \geq \alpha_i$, $\forall \alpha_i \in [0, 1]$ or $\mathbb{P}(u_i^G(\boldsymbol{x}_G^*) \geq u_i^G(y, \boldsymbol{x}_{G,-i}^*)) \geq \frac{a_i}{K_i}\mathbb{P}(\{f < g_i(y^*)\})$, where $0 < \frac{a_i}{K_i}\mathbb{P}(\{f < g_i(y^*)\}) \leq 1$ and so for $\alpha_i = \frac{a_i}{K_i}\mathbb{P}(\{f < g_i(y^*)\})$, $x_i^* \in B_i^{\alpha_i}(\boldsymbol{x}_{G,-i}^*)$, therefore $\boldsymbol{x}_G^* \in \boldsymbol{\alpha}\text{-}NE(G)$.

□

With Lemma 4 we can weaken the condition of surely non-negative and bounded in Theorem 9 to surely bounded.

Corollary. Let G be a game with random payoff in mixed strategies on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. If there is a deterministic game H such that following holds:

1. the set of pure strategy profiles of H is $P_H = P_G \times \Omega \times P_O$, where P_G is the set of pure strategy profiles of G and P_O are the pure strategies of orthogonal players (G is a game with unknown players.),
2. $\forall i \in I : X_i$ is a compact metric space,

3. $\forall i \in I : |u_i^G(\omega)| \leq K_i < \infty$, $\forall \omega \in \Omega$, where u_i^G denotes the payoff function of player i in G (payoff functions are surely bounded.),

4. $\forall i \in I_G : u_i^G$ is a constant function of P_O (payoff functions are independent of the orthogonal players),

5. $g_i(y, \omega) = u_i^G(y, \boldsymbol{x}_{-i}, \omega)$ is a continuous function of y for every fixed $\boldsymbol{x}_{-i} \in X_{-i}$ and $\forall \omega \in \Omega$ (g_i is a surely continuous function),

6. there exists a Nash equilibrium of H such that $\boldsymbol{x}_H^* = \boldsymbol{x}_G^* \otimes \mathbb{P}$ ($\mathbb{P}$ is an optimal strategy profile of uknown players).

Then there exists $\boldsymbol{\alpha} \in (0,1]^{I_G}$ confidence levels such that $\boldsymbol{x}_G^*$ is a $\boldsymbol{\alpha}$-Nash equilibrium of G.

Proof. By variant of lemma 4 for games with random payoff, game G has a Nash equilibrium, if and only if $\hat{G}$ has a Nash equilibrium, where $u_i^{\hat{G}} := u_i^G + K_i$. Such $\hat{G}$ satisfies the assumptions of theorem 9 and therefore there exists $\boldsymbol{\alpha} \in (0,1]^I$ such that $\boldsymbol{x}_G^*$ is a $\boldsymbol{\alpha}$-Nash equilibrium of $\hat{G}$.

□

With this we will now show that all of the assumptions of Theorem 9 are satisfied in the case when we consider a finite game with finite number of realizations. To do this we have to show that for every such game with random payoff there exists a game with unknown players with $\mathbb{P}$ being the optimal strategy of the unknown players. To show this we first need to prove following lemma.

Lemma 10. *Let P_i be a finite set of pure strategies, then $\forall x \in X_i$ mixed strategies, there exists a game G, such that x is a unique Nash equilibrium strategy of G for the player i.*

Proof. We will show that there exists a matrix game such that x is a unique Nash equilibrium strategy of such a game for the first player. Let $|P_i| = n \in \mathbb{N}$ this means that $x \in X_i$ may be represented as a vector $\boldsymbol{x} = (x_1, \dots, x_n)^T \geq \mathbf{0}^T$ such that $\sum_{k=1}^n x_k = 1$. Denote $J = \{k; x_k = 0\}$ and define

$$B = (b_{kl})_{k=1,l=1}^{n,n},$$

where

$$b_{kl} = \begin{cases} 1, \text{ if } k \neq l \text{ and } k \notin J, \\ \frac{-\sum_{m \neq k} x_m}{x_k}, \text{ if } k = l \text{ and } k \notin J, \\ 0, \text{ if } k \neq l \text{ and } k \in J, \\ -1, \text{ if } k = l \text{ and } k \in J. \end{cases}$$

Now let us consider a matrix game with payoff matrix $A = B^T$, so that both players have n pure strategies. It is a common result that in such a game Nash equilibrium exists and it is equal to solving the minimax problem (see for example [10])

$$\max_{\hat{\boldsymbol{x}} \in X} \min_{\boldsymbol{y} \in Y} \hat{\boldsymbol{x}}^T A \boldsymbol{y},$$

which is equivalent for the first player to solve

$$\max_{\hat{\boldsymbol{x}} \in X, v \in \mathbb{R}} v, \qquad (2.9)$$
$$\text{s.t. } A^T \hat{\boldsymbol{x}} \geq v\mathbf{1}.$$

Now we will show that our $\boldsymbol{x}$ is a unique optimal solution of (2.9). First notice that condition $A^T\hat{\boldsymbol{x}} \geq v\mathbf{1}$ is equivalent with $\min(A^T\hat{\boldsymbol{x}}) \geq v$. Further note that $A^T\boldsymbol{x} = B\boldsymbol{x} = \mathbf{0}$ because

$$(B\boldsymbol{x})_k = \begin{cases} -x_k = 0, \text{ if } k \in J, \\ -\sum_{m \neq k} x_m + \sum_{m \neq k} x_m = 0, \text{ if } k \notin J. \end{cases}$$

and so $(\boldsymbol{x}^T, 0)^T$ is feasible for (2.9). For the optimality we will show that, $\forall \hat{\boldsymbol{x}} \in X, \hat{\boldsymbol{x}} \neq \boldsymbol{x} : \min(B\hat{\boldsymbol{x}}) < 0$. Let $\hat{\boldsymbol{x}} \neq \boldsymbol{x}$ then there exists $k \in \{1, \ldots, n\}$ such that $\hat{x}_k > x_k$ and so

$$(B\hat{\boldsymbol{x}})_k = \begin{cases} -\frac{\hat{x}_k}{x_k} \sum_{m \neq k} x_m + \sum_{m \neq k} \hat{x}_m < 0, \text{ if } k \notin J, \\ -\hat{x}_k < 0, \text{ if } k \in J. \end{cases}$$

where the first case is given by the fact that $\frac{\hat{x}_k}{x_k} > 1$ and $\sum_{m \neq k} x_m > \sum_{m \neq k} \hat{x}_m$ and the second case is given by $\hat{x}_k > 0$. Therefore $\min(B\hat{\boldsymbol{x}}) < 0$ and so $\boldsymbol{x}$ is optimal for (2.9) and so it is a unique Nash equilibrium strategy for the first player.

□

Using the previous lemma we can now weaken the existence criteria to th following form of a Corollary of the Theorem 9.

Corollary. Let G be a game with random payoff on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where Ω is a finite set. If $\forall i \in I_G, \forall \omega \in \Omega : u_i^{G(\omega)}$ is continuous and bounded and the set of pure strategy profiles of G is finite, then there exists non-trivial confidence levels $\boldsymbol{\alpha} \in (0,1]^{I_G}$ such that $\boldsymbol{\alpha}$-$NE(G) \neq \emptyset$.

Proof. For the proof of this claim we will show that there exists a game H in which $\mathbb{P} = \sum_{\omega \in \Omega} a_\omega \delta_\omega$, where $\sum_{\omega \in \Omega} a_\omega = 1$, $a_\omega \geq 0$ is an optimal strategy profile of unknown players. Without a loss of generality assume that $\Omega = \{1, \ldots, n\}$ for some $n \in \mathbb{N}$. Define the set of pure strategies of H as $P_H = P_G \times \Omega$ and let $I_H = I_G \cup \{0, -1\}$, where 0 and -1 are the unknown players of G.

Denote $\boldsymbol{a} = (a_1, \ldots, a_n)^T \in \mathbb{R}^n$ the vector of probabilities of each realization of ω and take $u_0^H(\boldsymbol{x}_G, \hat{\boldsymbol{a}}, \boldsymbol{b}) = \hat{\boldsymbol{a}}^T A\boldsymbol{b}$, $u_{-1}^H(\boldsymbol{x}_G, \hat{\boldsymbol{a}}, \boldsymbol{b}) = -\hat{\boldsymbol{a}}^T A\boldsymbol{b}$, where A defines a matrix game from the proof of the Lemma 10 in which $\boldsymbol{a}$ is a Nash equilibrium strategy. H is a finite deterministic game in mixed strategies, therefore there exists a Nash equilibrium $\boldsymbol{x}_G^* \otimes P^* \otimes Q^* \in X_G \times \mathcal{D}(\Omega) \times \mathcal{D}(\Omega)$. Now suppose that $P^* = \sum_{\omega \in \Omega} p_\omega \delta_\omega \neq \mathbb{P}$, then $\boldsymbol{p} = (p_1, \ldots, p_n)^T \neq (a_1, \ldots, a_n)^T = \boldsymbol{a}$ and so $u_0^H(\boldsymbol{x}_G^*, P^*, Q^*) = \boldsymbol{p}^T A\boldsymbol{q} < \boldsymbol{a}^T A\boldsymbol{q} = u_0^H(\boldsymbol{x}_G^*, \mathbb{P}, Q^*)$, where $Q^* = \sum_{\omega \in \Omega} q_\omega \delta_\omega$. This is a contradiction with the Nash equilibrium property. Therefore $P^* = \mathbb{P}$. We have constructed a game H which satisfies the assumptions of the Theorem 9, therefore $\exists \boldsymbol{\alpha} \in (0,1]^{I_G}$ such that $\boldsymbol{x}_G^* \in \boldsymbol{\alpha}$-$NE(G)$.

□

In this proof we have also shown a way how to approach finding of such a equilibrium point in the case of games with random payoff with finite scenarios, where the probability of all scenarios is known. In this case we want to find the Nash equilibrium of H as defined above.

Based on the previous results and theorems, we are able to prove following theorem, which is stochastic equivalent of John Forbes Nash, Jr.'s theorem on deterministic games.

Theorem 11 (Fundamental theorem of stochastic game theory)**.** *Let G be a finite non-cooperative game with random payoff on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where Ω is a finite set. Then there exists a non-trivial confidence levels $\boldsymbol{\alpha} \in (0,1]^{I_G}$ such that $\boldsymbol{\alpha}$-$NE(G) \neq \emptyset$.*

Proof. To prove this theorem we will show that every payoff function in such a game is continuous. We have seen in the Section 3 of this chapter, that in a case of a finite game, mixed strategies, may be represented as a closed convex hull of standard basis vectors of a finite vector space over the real numbers. Therefore, they form a compact metric space. From which we get that continuous functions are bounded on the set of mixed strategies. Now let P be the set of pure strategy profiles of G and $\hat{u}_i^G(\boldsymbol{p}, \omega)$ denote the payoff of player $i \in I_G$ given the pure strategy profile $\boldsymbol{p} \in P$ in the scenario $\omega \in \Omega$, then for a $\boldsymbol{x} \in X$ mixed strategy we have that

$$u_i^G(\boldsymbol{x}, \omega) = \sum_{\boldsymbol{p} \in P} \boldsymbol{x}(\boldsymbol{p}) \hat{u}_i^G(\boldsymbol{p}, \omega)$$

and so let $\{\boldsymbol{x}_n\}_{n \in \mathbb{N}} \subset X$ be an arbitrary sequence convergent to $\boldsymbol{x} \in X$ then for $\omega \in \Omega$ we have

$$\lim_{n\to\infty} u_i^G(\boldsymbol{x}_n, \omega) = \lim_{n\to\infty} \sum_{\boldsymbol{p} \in P} \boldsymbol{x}_n(\boldsymbol{p}) \hat{u}_i^G(\boldsymbol{p}, \omega) = \sum_{\boldsymbol{p} \in P} \lim_{n\to\infty} \boldsymbol{x}_n(\boldsymbol{p}) \hat{u}_i^G(\boldsymbol{p}, \omega) =$$
$$= \sum_{\boldsymbol{p} \in P} \boldsymbol{x}(\boldsymbol{p}) \hat{u}_i^G(\boldsymbol{p}, \omega) = u_i^G(\boldsymbol{x}, \omega).$$

Therefore, an arbitrary payoff function is surely continuous in mixed strategies, which implies that it is surely bounded on the set of mixed strategies and so the assumptions of Theorem 9 are satisfied and therefore there exists $\boldsymbol{\alpha} \in (0,1]^{I_G}$ such that $\boldsymbol{\alpha}$-$NE(G) \neq \emptyset$.

□

Definition 18. *We will call a game G with random payoff on a finite probability space with finite number of pure strategy profiles a finite game with random payoff. We will denote $\mathcal{C}(G) = \{\boldsymbol{\alpha} \in [0,1]^{I_G} : \boldsymbol{\alpha}\text{-}NE(G) \neq \emptyset\}$ the set of confidence levels for which G is a Nash-solvable game.*

We have shown that in a finite game with random payoff $\mathcal{C}(G) = [\boldsymbol{0}, \boldsymbol{\lambda}] := [0, \lambda_1] \times \cdots \times [0, \lambda_{|I_G|}]$ for some $\boldsymbol{1} \geq \boldsymbol{\lambda} > \boldsymbol{0}$. This motivates us to define the highest possible confidence level for G and the most probable Nash equilibrium of the game.

Definition 19. *Let G be a game with random payoff and let $\boldsymbol{\lambda} = \sup \mathcal{C}(G)$ be the highest confidence levels for which G is a Nash-solvable game, we will shortly write $NE(G) := \boldsymbol{\lambda}$-$NE(G)$ and call its elements the most probable Nash-equilibria of G.*

Now the question is how to find the highest confidence levels on which the finite game with random payoff is Nash-solvable? In our proof of the Fundamental theorem of stochastic game theory we have shown how to find a game with unknown players for which a Nash equilibrium exists. From this we may construct a Nash equilibrium on some $\boldsymbol{\alpha}$ levels of confidence. This means we have a way how to find a lower estimate for the highest confidence levels of a game.

Definition 20. *Let G be a game with random payoff and let $\boldsymbol{\lambda}$ denote the highest confidence levels on which G is a Nash-solvable game. If $\forall i \in I_G : \lambda_i > 0.5$ we call G a predictable game. Games with random payoff that are not predictable, we will call unpredictable games.*

In this section we discussed the case when every player uses the same confidence level α_i for each mixed strategy $x \in X_i$. More generally we could consider a separate confidence level for each mixed strategy. That is we would consider $\alpha_i(x)$ the confidence level to be a function of the mixed strategy. In such a case, however, most of the results we have shown in this section may not be possible to be proven.

Now let us also consider the case for best response strategy sets defined using joint probabilistic constrains.

Definition 21. *Let G be a game with random payoff and $\alpha_i \in [0,1]$. We define the set of best response strategies on a confidence level of α_i for the player i, given strategy profile of other players $\boldsymbol{x}_{-i} \in X_{-i}$ as*

$$B_i^{\alpha_i} = \{x_i \in X_i ; \mathbb{P}(\forall y \in X_i : u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(y, \boldsymbol{x}_{-i})) \geq \alpha_i\}$$

Now the condition for best response strategies changes to joint probability of the strategy to be the best response. This probability is given by the intersection of all mixed strategies. But we can weaken it to the intersection of pure strategies.

Lemma 12. *Let P_i be countable set of pure strategies and X_i be the corresponding set of mixed strategies. Then*

$$\mathbb{P}(\forall y \in X_i : u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(y, \boldsymbol{x}_{-i})) = \mathbb{P}(\forall p \in P_i : u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(p, \boldsymbol{x}_{-i})).$$

Proof. Denote $A = \{\omega \in \Omega ; \forall y \in X_i : u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(y, \boldsymbol{x}_{-i})\}$ and $B = \{\omega \in \Omega ; \forall p \in P_i : u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(p, \boldsymbol{x}_{-i})\}$. Clearly, $A \subseteq B$, because $\forall p \in P_i : \delta_p \in X_i$. Let $\omega \in B$ so that $\forall p \in P_i : u_i(x_i, \boldsymbol{x}_{-i}, \omega) \geq u_i(p, \boldsymbol{x}_{-i}, \omega)$. Take $y \in X_i$ then

$$u_i(y, \boldsymbol{x}_{-i}, \omega) = \int u_i(p, \boldsymbol{x}_{-i}, \omega) dy(p) \leq \int u_i(x_i, \boldsymbol{x}_{-i}, \omega) dy(p) = u_i(x_i, \boldsymbol{x}_{-i}, \omega).$$

Therefore, $u(y, \boldsymbol{x}_{-i}, \omega) \leq u_i(x_i, \boldsymbol{x}_{-i}, \omega)$ and so $\omega \in A$. From which we have $A = B$ and consequently $\mathbb{P}(A) = \mathbb{P}(B)$.

□

By this lemma we know that to construct a set of best response strategies it is sufficient to use only the pure strategies. Which in most cases will be a finite set instead of potentionally infinite sets of mixed strategies.

We will now proceed to define the Nash equilibrium as in the previous chapter using the sets of best response strategies.

Definition 22. *Let G be a game with random payoff and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots) \in [0,1]^{I_G}$. If $\forall i \in I_G : x_i \in B_i^{\alpha_i}(\boldsymbol{x}_{-i})$, we say that $\boldsymbol{x} \in X$ is a Nash equilibrium of G on confidence levels $\boldsymbol{\alpha}$ or that it is a $\boldsymbol{\alpha}$-Nash equilibrium of G. We denote the set of $\boldsymbol{\alpha}$-Nash equilibria as $\boldsymbol{\alpha}$-$NE(G)$. Vector $\boldsymbol{\alpha}$ is called the confidence level vector.*

Convexity of those sets is similiarly as in the case of individual constrains derived by the methods of stochastic programming.

Lemma 13. *Let P_i be a finite set and $\forall p \in P_i : u_i(x_i, \boldsymbol{x}_{-i}, \omega) - u_i(p, \boldsymbol{x}_{-i}, \omega) = f_i(x_i, p) - g(\omega)$, where f_i is a concave function of x_i and $g(\omega)$ has a log-concave distribution, then $B_i^{\alpha}(\boldsymbol{x}_{-i})$ is convex $\forall \alpha \in [0,1]$.*

Proof. Just realize that $B_i^{\alpha}(\boldsymbol{x}_{-i}) = \{x_i \in X_i; \mathbb{P}(\forall p \in P_i : f_i(x_i, p) \geq g(\omega)) \geq \alpha\}$. For the proof of convexity of such sets with joint probabilistic constrains we refer to [2].

□

Lemma 14. *Let G be a game with random payoff. Then:*

1. *$\forall i \in I_G : \forall \boldsymbol{x}_{-i} \in X_{-i} : B_i^{\alpha_i}$ is non-increasing in α_i. That is $\forall \alpha_i \geq \beta_i : B_i^{\alpha_i}(\boldsymbol{x}_{-i}) \subseteq B_i^{\beta_i}(\boldsymbol{x}_{-i})$.*
2. *$\boldsymbol{\alpha}$-$NE(G)$ is non-increasing in $\boldsymbol{\alpha}$. That is $\forall \boldsymbol{\alpha} \geq \boldsymbol{\beta}$ we have that $\boldsymbol{\alpha}$-$NE(G) \subseteq \boldsymbol{\beta}$-$NE(G)$.*

Proof. 1. Let $\alpha_i \geq \beta_i, \boldsymbol{x}_{-i} \in X_{-i}$ and $x_i \in B_i^{\alpha_i}(\boldsymbol{x}_{-i})$ then we have that $\mathbb{P}(\forall y \in X_i : u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(y, \boldsymbol{x}_{-i})) \geq \alpha_i \geq \beta_i$ and so $x_i \in B_i^{\beta_i}(\boldsymbol{x}_{-i})$.

2. Let $\boldsymbol{\alpha} \geq \boldsymbol{\beta}$ and $\boldsymbol{x} \in \boldsymbol{\alpha}$-$NE(G)$ then $\forall i \in I_G : x_i \in B_i^{\alpha_i}(\boldsymbol{x}_{-i})$ and by the first part of this proof we have that $x_i \in B_i^{\beta_i}(\boldsymbol{x}_{-i})$, therefore $\forall i \in I_G : x_i \in B_i^{\beta_i}(\boldsymbol{x}_{-i})$ and so $\boldsymbol{x} \in \boldsymbol{\beta}$-$NE(G)$.

□

Similiarly to the previous subsection we have that $B_i^0(\boldsymbol{x}_{-i}) = X_i$ and $\mathbf{0}$-$NE(G) = X$ and by this lemma, if $B_i^1(\boldsymbol{x}_{-i}) \neq \emptyset$ then $B_i^{\alpha}(\boldsymbol{x}_{-i}) \neq \emptyset, \forall \alpha \in (0,1)$ and as a consequence, if $\mathbf{1}$-$NE(G) \neq \emptyset$ then $\boldsymbol{\alpha}$-$NE(G) \neq \emptyset, \forall \boldsymbol{\alpha} \in (0,1)^{I_G}$. So far there are no results for existence of Nash equilibrium with joint probabilistic constrains in general games with random payoff. There are only some results for the zero-sum games of two players.

Definition 23. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. We say that a family of events $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A}$ satisfies the condition of positivity, if $\forall n \in \mathbb{N} : \mathbb{P}(A_n) > 0 \Rightarrow \mathbb{P}(\bigcap_{n \in \mathbb{N}} A_n) > 0$.*

Theorem 15. *Let G be a finite game with random payoff and $\forall i \in I_G$ let $A_p(\boldsymbol{x}) = \{\omega \in \Omega; u_i(x_i, \boldsymbol{x}_{-i}, \omega) \geq u_i(p, \boldsymbol{x}_{-i}, \omega)\}$ satisfy the condition of positivity $\forall \boldsymbol{x} \in X$ given, then there exists $\boldsymbol{\alpha} \in (0,1]^{I_G}$, such that $\boldsymbol{\alpha}$-$NE(G) \neq \emptyset$.*

Proof. Let G be a finite game with random payoff. By the Theorem 11 there exists an $\boldsymbol{\beta} \in (0,1]^{I_G}$ such that $\boldsymbol{\beta}$-$NE(G)$ with individual constrains is not empty. This means that there exists $\boldsymbol{x} \in X$ such that $\forall i \in I_G : x_i \in B_i^{\beta_i}(\boldsymbol{x}_{-i})$ with individual constrains for some $\beta_i > 0$ and so $\forall p \in P_i : \mathbb{P}(u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(p, \boldsymbol{x}_{-i})) \geq \beta_i > 0$. Since $A_p(\boldsymbol{x})$ satisfy the condition of positivity it holds that

$$\mathbb{P}(\forall p \in P_i : u_i(x_i, \boldsymbol{x}_{-i}) \geq u_i(p, \boldsymbol{x}_{-i})) \geq \alpha_i > 0.$$

Therefore $x_i \in B_i^{\alpha_i}(\boldsymbol{x}_{-i})$ with joint constraints, from which $\boldsymbol{x} \in \boldsymbol{\alpha}$-$NE(G)$ with joint constraints.

□

Now the question is when $A_p(\boldsymbol{x})$ satisfy the condition of positivity? Clearly when $\forall \boldsymbol{p} \in P : u_i(\boldsymbol{p}, \omega)$ are independent random variables, then also $A_p(\boldsymbol{x})$ are indenpendent random events. In such a trivial case the joint and individual probabilistic constrains are equivalent.

2.6.2 Deterministic equivalent games

Different approach to optimality was published in [14]. In this model players want to play a so-called deterministic equivalent game on a $\boldsymbol{\alpha}$ levels of confidenece.

Definition 24 (Deterministic equivalent game)**.** *Let G be a game with random payoff and let $\boldsymbol{\alpha}$ be a confidence level vector. We define $v_i^{\alpha_i} : X \to \mathbb{R}$ as $v_i^{\alpha_i}(\boldsymbol{x}) = \sup\{\delta; \mathbb{P}(u_i(\boldsymbol{x}) \geq \delta) \geq \alpha_i\}$ and say that $\Gamma_{\boldsymbol{\alpha}} = (I, X, \{v_i^{\alpha_i}\}_{i\in I})$ is a deterministic equivalent of G on a confidence levels $\boldsymbol{\alpha}$.*

Theorem 16 ([14])**.** *Let G be a game with random payoff in mixed strategies with finite set of players. Denote $\hat{u}_i$ the payoff function of player i in a corresponding game in pure strategies. If $P = \bigtimes_{i\in I} P_i$ is a finite set of pure strategy profiles and $\forall i \in I$ vector $(\hat{u}_i(\boldsymbol{p}))_{\boldsymbol{p}\in P}$ has an elliptically symmetric distribution with location parameter $(\mu_i(\boldsymbol{p}))_{\boldsymbol{p}\in P}$ and possitivelly definite scale matrix Σ_i, then there exists a Nash equilibrium of the deterministic equivalent of G for all $\boldsymbol{\alpha} \in (0.5, 1]^{|I|}$.*

Proof. For the proof of this theorem we refer to [14].

□

Family of elliptically symmetric distributions is quite large and contains for example multivariate Normal distribution, multivariate Cauchy distribution, multivariate Student's t-distribution or logistic distribution.

However, problem with deterministic equivalent games is that even, if G is a game with random payoff and finite number of pure strategies the corresponding deterministic equivalent game $\Gamma_{\boldsymbol{\alpha}}$ on a confidence levels $\boldsymbol{\alpha}$ generally is a game with infinite number of pure strategies and therefore we do not know whether it has a Nash equilibrium. This means that after transition from the game with random payoff to its deterministic equivalent form we lose the von Neuman - Morgenstern property that the mixed strategy payoff is given as a expected payoff from pure strategies given the distribution of the mixed strategy. With the fact, that quantiles of distribution are generally hard to express and the resulting payoff function is not convex it makes it hard to prove the existence of an optimal solution in this model.

2.6.3 Generalizing the minimax program

We have seen that one possible approach to optimality in the case of deterministic games is by considering the minimax problem where players want to defend themselves against the worst possible case scenario. In this model the player $i \in I$ wants to solve program

$$\max_{x_i \in X_i} \min_{\boldsymbol{x}_{-i} \in X_{-i}} u_i(x_i, \boldsymbol{x}_{-i}). \tag{2.10}$$

We have two possible options how to rewrite this program for the stochastic case. First is to let player $i \in I$ choose a confidence level $\alpha_i \in (0, 1)$ and let them defend against the worst outcome of the game on this confidence level. That is

$$\begin{aligned} &\max_{x_i \in X_i, \delta \in \mathbb{R}} \delta \\ &\text{s.t. } \mathbb{P}[\min_{\boldsymbol{x}_{-i} \in X_{-i}} u_i(x_i, \boldsymbol{x}_{-i}) \geq \delta] \geq \alpha_i. \end{aligned} \tag{2.11}$$

We will refer to this approach the worst payoff method.

Second possible approach is by maximizing the minimal payoff subject to the minimal probability of it happening being at least α_i. This corresponds to a program

$$\begin{aligned} &\max_{x_i \in X_i, \delta \in \mathbb{R}} \delta \\ &\text{s.t. } \min_{\boldsymbol{x}_{-i} \in X_{-i}} \mathbb{P}[u_i(x_i, \boldsymbol{x}_{-i}) \geq \delta] \geq \alpha_i. \end{aligned} \tag{2.12}$$

We will refer to this approach as the least likely payoff method.

Since we have no definite proof of the existence of $\boldsymbol{\alpha}$-Nash equilibria in general case game, it is better suited for us to consider these programs as our optimality criteria for the further work. Those generalizations were originally considered for the case of a matrix games with random payoff in [5], [4],[3] and [6].

2.7 Stochastic dominance in games with random payoff

In previous sections we have seen an approach to finding optimal solutions of the game with random payoff based on stochastic programs, hereby we distinguished two models based on the common type of a stochastic program they relate to. We have seen that our assumptions may be further generelized, especially in the case of the model with individual constrains. Now let us take a look on a different approach to finding equilibria of the game. In Subection 2.5 of this chapter we have seen a notion of dominance for deterministic games. In the context of games with random payoff it is natural to generelize this idea using methods of stochastic dominance, which serves to compare two different probability distributions. Let us briefly discuss this idea.

Definition 25 (Stochastic dominance)**.** *Let $\mathcal{G}$ be a set of (utility) functions. We say that the distribution X weakly stochasticaly dominates distribution Y with respect to the generator $\mathcal{G}$, if $\mathsf{E}\, f(X) \geq \mathsf{E}\, f(Y), \forall f \in \mathcal{G}$, we will write $X \succeq_{\mathcal{G}} Y$. We say that the distribution X strictly stochasticaly dominates distribution Y with respect to the generator $\mathcal{G}$, if $X \succeq_{\mathcal{G}} Y$ and $\exists f \in \mathcal{G} : \mathsf{E}\, f(X) > \mathsf{E}\, f(Y)$, we will write $X \succ_{\mathcal{G}} Y$.*

This basic concept of stochastic dominance is based on the utility theory under uncertainty, where we want to compare two random possibilities based on some class of utility functions. This is so usefull in practice because it is generaly hard to specify a utility function of a person, however we can quite easily find a sufficient class of utility functions in which at least approximatelly lies the utility function of our agent. The most commonly used is the N-th order stochastic dominance or NSD. This class dominance is generated by following class of utility functions.

Definition 26. *We define the set of N-order utility functions as $\mathcal{U}_N = \{f : \mathbb{R} \to \mathbb{R}; (-1)^{n-1} f' \geq 0, n = 1, \ldots, N\}$.*

Definition 27 (NSD)**.** *We say that the random variable X weakly stochastically dominates Y under NSD, if $\mathsf{E}\, f(X) \geq \mathsf{E}\, f(Y), \forall f \in \mathcal{U}_N$. We write $X \succeq_{NSD} Y$. We say that X strictly dominates Y under NSD, if $X \succeq_{NSD} Y$ and $\exists f \in \mathcal{U}_N$: $\mathsf{E}\, f(X) > \mathsf{E}\, f(Y)$. We will write $X \succ_{NSD} Y$.*

Since $\forall K \geq N, \mathcal{U}_K \subseteq \mathcal{U}_N$ it clearly follows that

$$X \succeq_{NSD} Y \implies X \succeq_{KSD} Y.$$

The most common is to use first or second order stochastic dominance, which have following properties.

Theorem 17 (1SD characterization)**.** *Let X and Y be real random variables. The following statements are equivalent:*

1. $X \succeq_{1SD} Y$,
2. $F_X(x) \leq F_Y(x), \forall x \in \mathbb{R}$,
3. $F_X^{-1}(\alpha) \geq F_Y^{-1}(\alpha), \forall \alpha \in [0, 1]$.

Similiarly for the strict stochastic dominance following statements are equivalent:

1. $X \succ_{1SD} Y$,
2. $F_X(x) \leq F_Y(x), \forall x \in \mathbb{R}$ *and* $\exists x_0 \in \mathbb{R} : F_X(x_0) < F_Y(x_0)$,
3. $F_X^{-1}(\alpha) \geq F_Y^{-1}(\alpha), \forall \alpha \in [0, 1]$ *and* $\exists \alpha_0 \in [0, 1] : F_X^{-1}(\alpha_0) > F_Y^{-1}(\alpha_0)$.

Where F_X and F_X^{-1} denote the cummulative distribution function and quantile function of X respectively.

If we further define the integrated cummulative distribution function and integrated quantile function as

$$F_X^{(2)}(x) = \int_{-\infty}^{x} F_X(t)dt, x \in \mathbb{R},$$
$$F_X^{(-2)}(\alpha) = \int_{0}^{\alpha} F_X^{-1}(\theta)d\theta, \alpha \in [0,1],$$

we may obtain similiar results for the second order stochastic dominance.

Theorem 18 (2SD characterization)**.** *Let X and Y be real random variables. The following statements are equivalent:*

1. $X \succeq_{2SD} Y$,
2. $F_X^{(2)}(x) \leq F_Y^{(2)}(x), \forall x \in \mathbb{R}$,
3. $F_X^{(-2)}(\alpha) \geq F_Y^{(-2)}(\alpha), \forall \alpha \in [0,1]$.

Similarly for the strict stochastic dominance following statements are equivalent:

1. $X \succ_{2SD} Y$,
2. $F_X^{(2)}(x) \leq F_Y^{(2)}(x), \forall x \in \mathbb{R}$ *and* $\exists x_0 \in \mathbb{R} : F_X^{(2)}(x_0) < F_Y^{(2)}(x_0)$,
3. $F_X^{(-2)}(\alpha) \geq F_Y^{(-2)}(\alpha), \forall \alpha \in [0,1]$ *and* $\exists \alpha_0 \in [0,1] : F_X^{(-2)}(\alpha_0) > F_Y^{(-2)}(\alpha_0)$.

In the context of game with random payoff we may use the stochastic dominace to compare two strategy profiles as follows.

Definition 28. *Let $\boldsymbol{x}, \boldsymbol{y} \in X$ be strategy profiles. We say that $\boldsymbol{x}$ weakly dominates $\boldsymbol{y}$ under NSD, if*

$$\forall i \in I : u_i(\boldsymbol{x}) \succeq_{NSD} u_i(\boldsymbol{y}).$$

We say that $\boldsymbol{x}$ strictly dominates $\boldsymbol{y}$ under NSD, if

$$\forall i \in I : u_i(\boldsymbol{x}) \succ_{NSD} u_i(\boldsymbol{y}).$$

As the strategy profiles determine the distribution of the payoff function we will for simplicity write that $\boldsymbol{x} \succeq_{NSD} \boldsymbol{y}$ or $\boldsymbol{x} \succ_{NSD} \boldsymbol{y}$ and mean that $\boldsymbol{x}$ weakly dominates $\boldsymbol{y}$ under NSD or $\boldsymbol{x}$ strictly dominates $\boldsymbol{y}$ under NSD respectivelly.

Definition 29. *Let $\boldsymbol{x} \in X$ if $\nexists \boldsymbol{y} \in X$ such that $\boldsymbol{y} \succ_{NSD} \boldsymbol{x}$, we say that strategy profile $\boldsymbol{x}$ is NSD efficient. Let us denote the set of NSD efficient strategy profiles of the game with random payoff G as*

$$EF_{NSD}(G) = \{\boldsymbol{x} \in X; \boldsymbol{x} \text{ is } NSD \text{ efficient}\}.$$

By this we may then restrict ourselves on the set of NSD efficient strategy profiles. However, there are to the best of our knowlege no results on properties of these strategy sets.

3. Sequential games

In this chapter we will take a look on deterministic sequential games in more details. In the second chapter of this thesis we mostly focused on games in 'normal' or 'strategic' form. This means that for games with multiple stages we considered a pure strategy of the player to be a combination of possible actions in every stage of the game. To ilustrate this lets consider a repeated game G with 3 stages. Let player $i \in I$ have at each stage two possible actions $A_t = \{A, B\}, t = 1, 2, 3$ then in the normal form of game G the set of pure strategies is in the form $P_i = \{a_1 \times a_2 \times a_3; a_1 \in A_1, a_2 \in A_2, a_3 \in A_3\}$. Now we will also take a look on an extensive form of the game, which is given by an oriented game tree $T_G = (V_G, E_G)$, where each vertex $v \in V_G$ defines a stage of the game. Based on the actions played in the corresponding stage game moves to the following vertex of the game tree. This form of the game is more suitable for study of a case when players are able to play their strategies in current stage based on the outcome of the previous stage of the game.

Let us formalize this idea in a following definitions. First we will define several graph theory terms that we will use in this chapter. For this we use definitions based on [12].

Definition 30 ([12]). *Let V be the set of verticies and $E \subseteq \binom{V}{2}$ (the set of 2 element subsets of the set of verticies) be the set of edges. The touple $G = (V, E)$ is called an undirected graph. If $E \subseteq \{(u, v); u, v \in V\}$ (the set of touples of verticies), we call $G = (V, E)$ a directed graph. Directed graph with the property that if $(u, v) \in E$ then $(v, u) \notin E$ is called an oriented graph.*

Definition 31 ([12]). *Let G be a graph we will define function $d : V \to \mathbb{N}$ such that $d(v) = |\{e \in E; v \in e\}|$ and call it a degree of the vertex.*

In more general for directed graphs there are considered two types of degrees of a vertex. There is an outcoming and incoming degree of the vertex, where the outcoming degree is defined as number of edges starting in the vertex and incoming is defined as number of edges ending in the vertex. However, for the purposes of our thesis we will consider only the total number of edges that either start or end in the vertex as the vertex's degree.

Definition 32 ([12]). *Sequence of verticies $\{v_n\}_{n=1}^{K}$ such that $(v_j, v_{j+1}) \in E, j = 1, \ldots, K-1$ is called a path in the graph G. Verticies $u, v \in V$ are connected via path from u to v, if there exists a path $\{v_n\}_{n=1}^{K}$ such that $v_1 = u$ and $v_K = v$, we will denote this $u \to v$.*

Definition 33 ([12]). *Let G be a graph, if $\forall u, v \in V$ there exists a path from u to v or v to u, G is called a connected graph.*

Definition 34 ([12]). *Connected graph $T = (V, E)$ in which for any two verticies u and v there exists at most single path from u to v is called a tree. If T is an oriented graph, we call it an oriented tree and vertex $u \in V$ such that $\forall v \in V$ there exists a path from u to v is called a root of the tree T. We define $L = \{v \in V; d(v) = 1\}$ and call it the set of leafs of the tree T. Elements of L are called leafs of the tree.*

Definition 35 ([12])**.** *Let $T(v) = (V(v), E(v))$ be a oriented tree with root in v and let u be an arbitrary vertex in T. We define $T(u) = (V(u), E(u))$ such that $V(u) = \{x \in V(v); u \to x\}$ and $E(u) = E(v) \cap \{(u, v); u, v \in V(v)\}$ and call it a subtree (or a branch) of tree $T(v)$ starting at u.*

Definition 36 ([11])**.** *Let I be a set of players, $T(v_0) = (V(v_0), E(v_0))$ be a directed tree with the root v_0. Let $\forall v \in V(v_0) : G_v = (I, \{X_i^v\}_{i\in I}, \{u_i^v\}_{i\in I})$ be a game and $s_v : X^v \to V(v_0)$ be an associated successor function, that fullfills for a subtree $T(v)$ with the root in the vertix v, $\forall \boldsymbol{x} \in X^v : s_v(\boldsymbol{x}) \in V(v)$ and $(v, s_v(\boldsymbol{x})) \in E(v)$. Then the triple $\Gamma = (T(v_0), \{G_v\}_{v\in V(v_0)}, \{s_v\}_{v\in V(v_0)})$ is called a game in the extensive form.*

Further, notice that, if we consider only some branch of the game tree as defined above it is once again a game in the extensive form. This motivates us to define a subgame of a extensive form game.

Definition 37 ([11])**.** *Let Γ be a game in the extensive form and let u be a stage of Γ. Extensive form game defined as $\Gamma(v) = (T(v), \{G_u\}_{u\in V(v)}, \{s_u\}_{u\in V(v)})$ is called a subgame of Γ starting at stage v.*

For repeated games this tree is a linear with trivial successor function. If the number of repeatings of such a game is some finite $T \in \mathbb{N}$ then it may be simplified to a form $V = \{1, \ldots, T\}$ and $E = \{(n, n+1) \in \mathbb{N} \times \mathbb{N}; n+1 \leq T\}$ with

$$\forall \boldsymbol{x} \in X : s_t(\boldsymbol{x}) = \begin{cases} t+1, & \text{if } t < T, \\ \text{END}, & \text{if } t = T. \end{cases}$$

In a case when $T = \infty$, then $\forall t \in V : s_t(\boldsymbol{x}) = t + 1$. Whatsmore, we can think of two cases of representation of the payoff in cases when the number of stages is finite. Either it may occur at each stage of the game as it is assumed by our definition or it may be allocated at the leaf of the game tree. In the later case we will just consider all previous stage's payoff functions as trivial 0. Another commonly used representation in practice is that players can only strategize during some stages and take turns when they can influence the outcome. That is in every stage exactly one player may influence the payoff and the successor function. This is also a special case of our definition, when we will consider stages in which each player makes a move and let both the payoff and successor functions to be constant with respect to the given player's strategy outside of those stages. We decided to use our more general definition so that we have a more easily interpretable form of the game in a case when there are actuall payoffs allocated for each player in each stage and as it is more straigt-forward to generalize the case of a game with finite time into a game with infinite time.

Definition 38. *Lenght of a path from the root of the game tree to a given stage $v \in V$ is called a time of the stage. If time of all stages of the game is finite, we call it a game with finite number of stages or finite time. If game is not a game with finite time, we call it a game with infinite time or infinite number of stages.*

3.1 Games with finite number of stages

Let us first discuss the case when the set of stages V is finite. This means that every stage has a finite time. This game may be easily transformed into its strategic form. In this form consequently every player has a finite number of pure strategies and therefore by the Fundamental theorem of game theory there exists a mixed strategy Nash equilibrium of such a game.

We will start by showing how to transform an extensive form game into its strategic form and how to interpret mixed strategies from the strategic form of the game in the sense of extensive form.

Definition 39. *Let Γ be a game in the extensive form. Let $P_i(v)$ denote the set of pure strategies of the player i in the stage v and its elements as p_i^v and similiarly $P(v) = \prod_{i \in I} P_i(v)$ the set of pure strategy profiles in the stage v and its elements as $\boldsymbol{p}^v$. We define a transformation S from the extensive form into the strategic form of the game as follows:*

- *The set of players I of the game $S(\Gamma)$ is defined as the set of players of any stage v of Γ.*

- *Let L be the set of leafs of the game tree T.For $\forall l \in L$ denote $p(l) = (v_0, v_1, \ldots, v_K, l)$ the path from the root v_0 of T to l and consider $s_u^{-1}(v) = \{\boldsymbol{p} \in P(u); s_u(\boldsymbol{p}) = v\}$. Now we may consider the set of pure strategy profiles generating outcome $l \in L$ as $P^l = s_{v_0}^{-1}(v_1) \times s_{v_1}^{-1}(v_2) \times \cdots \times s_{v_K}^{-1}(l)$. We define the pure strategy of the player i in $S(\Gamma)$ as the unique sequence of the i-th coordinate of elements of P^l for some $l \in L$ and P_i the set of such sequences is called the set of pure strategies of the player i in $S(\Gamma)$. The set of pure strategy profiles of $S(\Gamma)$ is defined as $P = \bigcup_{l \in L} P^l$.*

- *u_i the payoff function of the player i in the game $S(\Gamma)$ is defined as $u_i(\boldsymbol{p}) = \sum_{v \in p(l)} u_i^v(\boldsymbol{p}^v)$ for a $\boldsymbol{p} \in P^l$.*

To obtain an equivalent of the Definition 5 with this transformation into strategic form, we require for all the paths in the extensive form game to have the same lenghts of some $T \in \mathbb{N}$. This is done by considering the maximum length of all paths in the game tree and defining trivial (idle) stages along the shorter paths. This formally means that we select $T = \max_{l \in L} |p(l)|$ and for $l \in L$ such that $T_l = |p(l)| < T$ define new stages $l_t, t = 1, \ldots, T - T_l$ with $X^{l_t} = \{\mathcal{I}\}, t = 1, \ldots, T - T_l$, $u_i^{l_t}(\mathcal{I}) = 0$ and the successor function $\forall \boldsymbol{x} \in X^l : s_l(\boldsymbol{x}) = l_1$, $s_{l_t}(\mathcal{I}) = l_{t+1}, t = 1, \ldots, T - T_l - 1$. By $\mathcal{I}$ we denote the single idle (or trivial) strategy profile, which formally has no effect on the game's outcome.

It is easy to see that if T is a finite tree then also L is a finite set and if in every stage the set of pure strategy profiles is finite then consequently the set of pure strategy profiles of the transformed game $S(\Gamma)$ must be also finite, therefore it is a finite game in the strategic form as was considered in the Chapter 1 of this thesis and from the Fundamental theorem of game theory (Theorem 2), there exists a Nash equilibrium of such a game.

Interpretation of mixed strategies of the strategic form of the game in the sense of the extensive form is that they measure probability of choosing a given path in such a game. This is a viable approach to strategizing in the sense of games with perfect information. That is in games when all players know exact possible strategies of all the other players and their respective outcomes and consider all of them to be perfectly rational and strategize accordingly.

As we mentioned before, in many applications of finite games in the extensive form it is considered for the payoff of each player in every non-leaf stage to be trivial 0 and in the leaf that there is only a single possible outcome, which defines the overall payoff for the player and so the players are trying to control the successor function in such a way the game will end up in the leaf that is most profitable for them. This idea was considered by John von Neuman for the case of two players, where he considered a game where the only possible outcomes of the game are $O = \{\text{Player 1 wins}, \text{Draw}, \text{Player 2 wins}\}$. In our representation it would correspond to stating that our game in extensive form Γ has $\forall i \in I = \{1, 2\}, \forall v \notin L, \forall \boldsymbol{x} \in X^v : u_i^v(\boldsymbol{x}) = 0$ and $\forall i \in I, \forall l \in L : X^l = \{\mathcal{I}\}, u_i^l(\mathcal{I}) \in \{-1, 0, 1\}$ and that $S(\Gamma)$ the game in its strategic form is a zero-sum game. Here we assume that the payoff from the game is given at the leafs of the game tree and that the leaf is a trivial game where there is outcome always given by the results of the previous stages. So players want to optimize their path allong the game tree to end in a leaf with highest possible payoff. In such a game we will call a pure strategy $\boldsymbol{p}_1 \in P_1$ that yields $u_1(\boldsymbol{p}_1, \boldsymbol{p}_2) = 1, \forall \boldsymbol{p}_2 \in P_2$ a winning strategy of the Player 1. Pure strategy $\boldsymbol{p}_1 \in P_1$ that yields $u_1(\boldsymbol{p}_1, \boldsymbol{p}_2) \geq 0, \forall \boldsymbol{p}_2 \in P_2$ will be called a strategy guarateeing at least draw. Similiarly for the Player 2. In such a case, von Neumann managed to prove a following theorem.

Theorem 19 (Theorem 3.13, [11])**.** *In every two-player game with the perfect information where the set of outcomes is defined as above, one and only one of the following alternatives holds:*

1. *Player 1 has a winning strategy.*
2. *Player 2 has a winning strategy.*
3. *Each of the two players has a strategy guaranteeing at least a draw.*

3.2 Games with infinite number of stages

The theory for games with infinite time is somewhat different from the theory of games with finite time. One important distinction comes from the type of players that may be involved in a game with infinite time.

Definition 40. *Let Γ be a game in extensive form with infinite time. Player $i \in I$ is called a finite player, if for every path $\{v_t\}_{t=1}^{\infty}$ in the game tree and any $\boldsymbol{x} \in \prod_{t=1}^{\infty} X^{v_t} : \sum_{t=1}^{\infty} |u_i^{v_t}(\boldsymbol{x}^{v_t})| < \infty$. Player that is not finite is called an infinite player.*

The main difference between behaviour of finite and infinite players is given by the fact that for finite players there exists a time $T < \infty$ such that they can gain (or lose) at most some arbitrary small $\varepsilon > 0$ after T. This means that for such players it is satisfactory to strategize only until the time T based on their sensitivity to optimality. That is in the case of a finite player with some $\varepsilon > 0$ sensitivity to the payoff, they want to find $T < \infty$ such that their payoff will differ by at most ε after T and solve the corresponding T-stage game. On the other hand infinite players may opt to lose any finite payoff as long as they are able to get higher payoffs in the future. Interesting game dynamics occurs in a situation when a finite player faces infinite player. In such a game a viable strategy for the infinite player is to bankrupt the finite player by deliberatelly decreasing their own payoff for some finite time in return for the game continuing in a path where the finite player may not control further evolution of the game (they are bankrupt). Example of such a behaviour may be seen, if a new company tries to enter an unregulated market where there exists a monopoly. The monopoly may deliberately decrease the price of goods in the market such that it is not profitable to produce them until the smaller company bankrupts.

Therefore, for a finite player it is important to know whether they play against finite or infinite players and strategize acordingly.

Games with infinite number of stages may be similiarly as the games with finite number of stages represented in their strategic form. Now however the definition of a viable transformation is a little bit more tricky, as we can no longer identify all the possible paths in a infinite tree with leafs of the tree. Now instead we would consider a set of all the possible paths from the root of the tree and as a pure strategy of the player in the strategic form of the game we would consider a unique sequence of pure strategies in each stage of the game corresponding to this path.

This representation of game would however result in a game which is not finite, therefore there are no general results on existence of a Nash equilibrium in such a game.

3.3 Games with imperfect information

So far we have discussed the case when all the players knew exactly what stage they are in. However in most of the real world scenarios this is not true. In fact, it is more common for the players to know only that they are in certain class of stages. In this class player knows what are the actions he can take, but has no knowledge of what will be the successor vertex uppon playing them. So that the player recognizes that he is in some information set, but has no prior knowledge in which exact stage. For example the player may know that in certain stages he can influence the outcome, but does not know where exactly in the game tree those stages are located, e.g. he does not know how many stages are there.

Let us formalize this idea in the following definition.

Definition 41 ([11]). *Let Γ be a game in the extensive form a subset U_i of V is called the information set of the player i, if $\forall v, u \in U_i : X_i^v = X_i^u$ and $u_i^v = u_i^u$. Partition $(U_i^k)_{k=1}^{K_i}$ of the set V, where each U_i^k is an information set is called the player i's information about the Γ. Player i is called a player with the perfect information, if each of his information sets contains only a single vertex. Γ is called a game with the perfect information, if every player in Γ is a player with the perfect information. If Γ is not a game with perfect information we call it a game with imperfect information.*

3.3.1 Behavior strategies

So far we considered strategies as of Definition 6 and 8, those in the extensive form of the game correspond to paths and probability distributions over the set of those paths. In the sense of an extensive form of the game with imperfect information one may ask what would change for our results, if we would consider more natural definition of strategies for the extensive form games with imperfect information, where we would consider a probability distributions over pure strategies of a certain information set? Such strategies are called the behavior strategies.

Definition 42 (Definition 6.2., [11])**.** *A behavior strategy of a player in a game in the extensive form is a function mapping each information set of the game to a probabillity distribution over the set of pure strategies of that player in that information set.*

To ilustrate this consider a two-stage game of two players I and II. In the first stage player I has two possible pure strategies A and B and in the second stage he has two pure strategies C and D. The second player in the first stage has the single idle strategy $\mathcal{I}$ and in the second stage he has pure strategies E and F. This means that he makes move only in the second stage of the game and in the first stage he observes the outcome based on the action of the first player. If in the first stage player I plays strategy A both players get payoff of 0 and the game proceeds to second stage. If the player I plays in the first stage B the game ends and $u_I^1(B, \mathcal{I}) = 2$, $u_{II}^1(B, \mathcal{I}) = 1$. In the second stage of the game the outcome is given by the following payoff table.

	E	F
C	3,2	1,1
D	0,2	4,2

In this example the pure strategies of the first player in the strategic form of the game are $P_I = \{(A, C), (A, D), (B, \mathcal{I})\}$ and of the second player $P_{II} = \{(\mathcal{I}, E), (\mathcal{I}, F)\}$ and so the mixed strategies are given as probability distributions over those sets. For example mixed strategy x_I of the player I is gives $x((A, C)) = 1/4$, $x((A, D)) = 1/2$ and $x((B, \mathcal{I})) = 1/4$. On the other hand the behavior strategy maps stage of the game to a given probability distribution over the pure strategies available in that given stage so for example $\tau_I(1)(A) = 1/2$, $\tau_I(1)(B) = 1/2$, $\tau_I(2)(E) = 1/3$ and $\tau_I(2)(F) = 2/3$ is a behavior strategy of the first player. From this example we may deduce that in the case when the information sets are trivially just single verticies of the game tree (it is a game with perfect information) we may to every behavior strategy find an equivalent mixed strategy. This idea is formalized in the following discussion.

Definition 43 (Definition 6.3., [11])**.** *A mixed/behavior strategy profile is a vector of strategies $\sigma = (\sigma_i)_{i \in I}$ where σ_i is either behavior or mixed strategy of the player i. Denote the set of all mixed/behavior strategy profiles as Σ. Let $\rho(v, \sigma)$ denote the probability that vertex $v \in V$ will be visited given mixed/behavior strategy profile σ.*

Definition 44 (Definition 6.5., [11])**.** *A mixed strategy x_i and behavior strategy σ_i are called equivalent, if $\forall v \in V, \forall \sigma_{-i} \in \Sigma_{-i} : \rho(v, x_i, \sigma_{-i}) = \rho(v, \sigma_i, \sigma_{-i})$.*

Theorem 20 (Theorem 6.6., [11])**.** *If a mixed strategy x_i of the player i is equivalent to a behavior strategy b_i, then for every mixed/behavior strategy profile of the other players σ_{-i} and every player $j \in I$ it holds,*

$$u_j(x_i, \sigma_{-i}) = u_j(b_i, \sigma_{-i})$$

where u_j denotes the payoff function of the player j in the strategic form of the game.

This theorem says that, if strategies are equivalent in the sense of the Definition 44 then they are also equivalent in the allocated payoff. This is something we would expect to happen as the payoff in the strategic form of the game is uniquely determined by the payoffs in each stage and the probability of visiting that stage.

Notice that for a mixed/behavior strategy profile σ and a vertex $v \in V$ it holds that $\rho(v, \sigma) = \int \mathbb{I}(\exists u \in V : v = s_u(\boldsymbol{p}^u))d\sigma(\boldsymbol{p})$. From which it is simple to deduce that behavior and mixed strategies are in the context of games with perfect information equivalent. This is formalized in the following theorem.

Theorem 21 (Theorem 6.11.,[11])**.** *Let Γ be a game in the extensive form that satisfies that at every vertex there are at least two pure strategies. Every behavior strategy of player i has an equivalent mixed strategy if and only if each information set of player i intersects every path from the root of the game tree at most once.*

As a consequence of this theorem, in every game with perfect information the behavior strategies are equivalent to the mixed strategies and therefore the existence of a Nash equilibrium in such games is equivalent to the existence of a Nash equilibrium in the strategic form of the game.

Theory of behavior strategies may be further developed to answer the question when do Nash equilibria in games with imperfect information exist. However in this thesis we will further consider only games with perfect information. This result also means that in the case of games with perfect information, there is no formal difference between players strategizing at the begining of the game for its whole duration and by strategizing at each stage of the game, a perfectly rational player with perfect information about the game will always end up with the same strategy. This may not be the best model for many real-world situations, but due to its simplicity it is better aplicable in practice.

3.4 Repeated games

Now lets examine an important model for sequential games called a repeated game. The repeated game is based upon a game in the strategic form, that is repeated for some finite or infinite number of stages. In this type of game players have a full knowlege about what happened in past stages and may strategize according to it. This sometime yield different optimal strategies as in the single occurence of the base game. We will ilustrate this on the famous example of the prisoner's dilema.

Definition 45 ([11])**.** *Let G be a game in the strategic form. The repeated version of G (or repeated G) is an extensive form game $\Gamma = (T, \{G_v\}_{v \in V}, \{s_v\}_{v \in V})$, where $V = \{1, \dots, N\}$ for some $N \in \mathbb{N} \cup \{\infty\}$, $\forall v \in V : G_v = G$ and $\forall v \in V, v < N, \forall \boldsymbol{x} \in X^v : s_v(\boldsymbol{x}) = v + 1$.*

In repeated games we can think of multiple ways to represent the overall payoff u_i that the player would optimize against, that is his payoff in the strategic form of the game. First way is to consider the overall payoff from the game as a sum of payoffs from each stage of the game, that is

$$u_i = \sum_{v=1}^{N} u_i^v$$

in the case of a finite number of stages N this will always result in a standard game with finite players, however in the case when N is infinite every player whose payoff function is not a trivial 0 would be considered as a infinite player. This would be problematic, since infinite players would consider any strategy, that yields infinite payoff as optimal and will not distinguish between them. This model is suited for the case, if the main problem that players want to solve is how to survive. To model different situation it is better to consider the contribution of the payoff in a certain stage to the overall payoff of the player to be dependent on the time in which it will be gained. The most common way to do this is by considering a discounting for each player.

Definition 46. *A sequence of numbers $\{\beta_i^v\}_{v\in V}$ is called a discounting of the player $i \in I$. If discounting of the player i satisfies $\beta_i^v \leq \beta_i^{v+1}$ and $\sum_{v\in V} \beta_i^v < \infty$ we call it a consistent discounting. If $u_i = \sum_{v\in V} \beta_i^v u_i^v$ we call i a player with discounting or a discounting player.*

Disciunting is a measure of how player values future game. Consistent discounting says that he considers later payoff less than the earlier. The idea of discounting comes from the financial mathematics, where discouned payment represents current value of future payments. That is, if player may ensure having return $r_{s,t}$ between time s and the future time t, the value of payoff in time t as represented in the time s is $\beta_i^{s,t} u_i^t = \frac{1}{1+r_{s,t}} u_i^t$. In many applications it is usefull to think that the return, that the player may ensure without playing the game between two separate stages remains constant, that is $r_{t,t+1} = r$ in such a case the corresponding discounting is given as $\beta^t = (\beta)^t = (\frac{1}{1+r})^t$.

However in some non-financially related situations it may be hard to know exactly how players value their future payoffs. It may be uknown for the players themselves. In those cases we may consider the overall payoff as na average payoff from each individual stage. That is, if N is finite we would consider

$$u_i = \frac{1}{N} \sum_{v=1}^{N} u_i^v.$$

In the case of infinitelly repeated game a natural extension is to consider a limit of the former expression

$$u_i = \lim_{N\to\infty} \frac{1}{N} \sum_{v=1}^{N} u_i^v.$$

In the case of a finite repeated game it is clear that the overall mixed strategies are still equivalent with the product of mixed strategies in each occurence of the game because

$$u_i(\boldsymbol{x}) = \mathsf{E}_{\boldsymbol{x}} \frac{1}{N} \sum_{v=1}^{N} u_i^v = \frac{1}{N} \sum_{v=1}^{N} \mathsf{E}_{\boldsymbol{x}} u_i^v.$$

In the case of a infinitelly repeated game we have that

$$u_i(\boldsymbol{x}) = \mathsf{E}_{\boldsymbol{x}} \lim_{N\to\infty} \frac{1}{N} \sum_{v=1}^{N} u_i^v$$

Futhermore from the Lebesgue theorem it holds that, if there exists $g \in \mathcal{L}(\boldsymbol{x})$ such that $\forall N \in \mathbb{N} : |\frac{1}{N}\sum_{v=1}^{N} u_i^v| \leq g, \boldsymbol{x} - a.s.$ then

$$u_i(\boldsymbol{x}) = \mathsf{E}_{\boldsymbol{x}} \lim_{N\to\infty} \frac{1}{N} \sum_{v=1}^{N} u_i^v = \lim_{N\to\infty} \frac{1}{N} \sum_{v=1}^{N} \mathsf{E}_{\boldsymbol{x}}\, u_i^v.$$

The Lebesgue condition in the context of Game theory means that, if the average payoff until the stage N for all such N is bounded by some dominant function for every pure strategy profile that is played with non-zero probability in the mixed strategy profile $\boldsymbol{x}$, then this strategy profile may be represented as a product of mixed strategies from each stage of the game. And the corresponding payoff fullfils the von Neuman - Morgensterns axioms, that it is a linear functional. This holds true, if the payoff functions of the base game are bounded, which means that if the base game is finite, then this holds true. In such a case from the proof of the existence of the Nash Equilibrium, (see [11]) it follows that the payoff function in such a game is always bounded and continuous in mixed strategies.

Let us first consider the case for repeated games with average payoff. Interesting question is to ask how does an equilibrium of the base game relate to the equilibria of its repeated version? This is stated in the following theorem.

Theorem 22 (Theorem 13.6, [11])**.** *Let Γ be the repeated version of G with $N < \infty$ stages. Let $\boldsymbol{x}^1, \dots, \boldsymbol{x}^N$ be Nash equilibria of G. Then the strategy profile $\boldsymbol{x} = \boldsymbol{x}^1 \times \cdots \times \boldsymbol{x}^N$ is a Nash equilibrium of Γ.*

Proof. For the proof of this theorem we refer to [11].

□

Lets now recall the definition of upper value v_i^U of the player i as

$$v_i^U = \min_{\boldsymbol{x}_{-i}\in X_{-i}} \max_{x_i\in X_i} u_i(x_i, \boldsymbol{x}_{-i}).$$

Theorem 23 (Theorem 13.8, [11])**.** *Let $\boldsymbol{x}^*$ be an equilibrium of the finitelly repeated version of G. Let v_i^U be the upper value of the player i in the base game G, then $u_i(\boldsymbol{x}^*) \geq v_i^U$, where u_i is the average payoff of the player i.*

Proof. For the proof we refer to [11].

□

Definition 47 ([11])**.** *Set $V = \{\boldsymbol{u} \in \mathbb{R}^I; \forall i \in I : u_i \geq v_i^U\}$ is called the set of individually rational payoffs. Let P be the set of pure strategy profiles in the strategic form of the game G we define $F = conv(\{\boldsymbol{u}(\boldsymbol{p}); \boldsymbol{p} \in P\})$ the set of feasible payoffs.*

Theorem 24 (The Folk theorem, [11])**.** *The set of equilibrium payoffs (payoffs that correspond to some Nash equilibrium $\boldsymbol{x}^*$) in a infinitelly repeated game G with average payoffs is the set $F \cap V$.*

3.4.1 Repeated prisoner's dilema

So far we have thought about non-cooperative games as games in which players have no incentive for cooperation. However a better way of understending them would be to think of them as games in which players have no means of coordination of their actions. In some cases it is possible for players to coordinate just trough their self-interest.

Let us examine this idea in the famous example of Prisoner's dilema. This is a game of two players with two possible pure strategies 'cooperate' (denoted C) or 'defect' (denoted D) and a following payoff matrix

	D	C
D	a_1, a_2	c_1, b_2
C	b_1, c_2	d_1, d_2

where $c_1 > d_1 > a_1 > b_1$ and $c_2 > d_2 > a_2 > b_2$. In such a game it is apparent that the most beneficial strategy profile for both players would be to cooperate, however, if it is a single-stage game the Nash equilibrium of the game is for both players to defect. Now suppose this same payoff matrix describes the proportional payoff at each stage of the repeated game with infinite time, where the first player discounts their payoff by a consistent valuation of future gains $\boldsymbol{\alpha} = \{\alpha_t\}_{t=1}^{\infty}$ and the second player by a consistent valuation of future gains $\boldsymbol{\beta} = \{\beta_t\}_{t=1}^{\infty}$. So that both players are finite. At each time t both players may decide whether to continue cooperating or to defect. Now each player from the perspective of strategic form game has infinite number of pure strategies in the form $\{p_t; p_t \in \{C, D\}\}_{t=1}^{\infty}$, this means that the existence of the Nash equilibrium of such game is not guaranteed by the Fundamental theorem of game theory and computing the possible Nash equilibrium in such a form is problematic. If we look at the game in its extensive form, we can much clearly see the dynamics that will follow, if both players are rational. Both players may observe the strategy of the other player at each given stage of the game and strategize according to the past. Both players know, that they have higher payoff, if they both decide to cooperate.

If neither of the players defects at any given time their payoff is given as

$$u_1(\boldsymbol{x}, \boldsymbol{y}) = \sum_{t=1}^{\infty} \alpha_t d_1$$

and

$$u_2(\boldsymbol{x}, \boldsymbol{y}) = \sum_{t=1}^{\infty} \beta_t d_2.$$

If we suppose that $\alpha_t = \alpha^t$ for $\alpha \in (0, 1)$ and $\beta_t = \beta^t$ for $\beta \in (0, 1)$ this simlifies to

$$u_1(\boldsymbol{x}, \boldsymbol{y}) = \frac{d_1 \alpha}{1 - \alpha}$$

and

$$u_2(\boldsymbol{x}, \boldsymbol{y}) = \frac{d_2 \beta}{1 - \beta}.$$

If the first player decides to defect from time $\tau \in \mathbb{N}$, the second player decidedes to react by always defecting from time $\tau + 1$ by this he ensures he will get at least $a_2 > b_2$ at each following round of the game. This will result in payoffs

$$u_1(\boldsymbol{x}, \boldsymbol{y}) = \frac{d_1 \alpha (1 - \alpha^{\tau - 1})}{1 - \alpha} + \alpha^\tau c_1 + \frac{a_1 \alpha^{\tau + 1}}{1 - \alpha}$$

and

$$u_2(\boldsymbol{x}, \boldsymbol{y}) = \frac{d_2 \beta (1 - \beta^{\tau - 1})}{1 - \beta} + \beta^\tau b_2 + \frac{a_2 \beta^{\tau + 1}}{1 - \beta}.$$

This is only beneficial for the first player only if

$$\begin{aligned} \frac{d_1 \alpha (1 - \alpha^{\tau - 1})}{1 - \alpha} + \alpha^\tau c_1 + \frac{a_1 \alpha^{\tau + 1}}{1 - \alpha} &> \frac{d_1 \alpha}{1 - \alpha}, \\ \alpha^\tau c_1 + \frac{a_1 \alpha^{\tau + 1}}{1 - \alpha} &> \frac{d_1 \alpha^\tau}{1 - \alpha}. \end{aligned}$$

Which implies that

$$\begin{aligned} -\alpha^{\tau + 1} c_1 + a_1 \alpha^{\tau + 1} &> d_1 \alpha^\tau - c_1 \alpha^\tau \\ (a_1 - c_1) \alpha^{\tau + 1} &> (d_1 - c_1) \alpha^\tau \\ \alpha &< \frac{d_1 - c_1}{a_1 - c_1}. \end{aligned}$$

Similiarly, the second player will decide to defect at some time τ only if

$$\beta < \frac{d_2 - c_2}{a_2 - c_2}.$$

In this example the players decide to defect only if their discounting factor is smaller than the reward from both cooperating adjusted for the reward from betraying the other player proportional to the reward from both defecting adjusted for the reward from betraying the other player.Strategy, that generates this outcome is called the Grim trigger strategy. This strategy is based on enforcing the cooperation by the prospect of "mutual destruction". It is however discutable how aplicable in reality such strategies are? In a sense real-world agents outside of finance has no real reason to apply the discounting factor way of thinking as in every iteration of the game they would recalculate the future strategy based on the current prospect of future payoff and generarly in situations as is foreign relations it is not so reasonable to assume, that they will for some reason also account for past benefits of the cooperation. Because for the most part those are hard to quantify in such a way. For such situation the model without discounting may be better suited and the question is to compare finite and infinite players instead of their potential payoffs and future valuations.

3.5 Subgame perfect, perfect and uniform equilibria

In the context of extensive form of a game, authors tried to refine the concept of Nash equilibrium. In the context of a sequential game it is only natural to assume that players strategize in each stage only with respect to the possition they are currently in. As we discussed in the previous section for a perfectly rational player with perfect information there should be no difference in this perspective, but let us formalize this idea. For this reason Reinhard Selten in 1965 introduced the idea of a subgame perfect Nash equilibrium.

Definition 48 ([11])**.** *Let Γ be a game in extensive form. A strategy profile $\boldsymbol{x}^*$ is called a subgame perfect equilibrium, if for every subgame $\Gamma(v)$ of Γ is a Nash equilibrium of $\Gamma(v)$.*

As Γ is a trivial subgame of itself this immidiatelly yields, that $\boldsymbol{x}^*$ must be a Nash equilibrium of Γ.

Theorem 25 (Theorem 7.4., [11])**.** *Let Γ be a game in the extensive form without nontrivial subgames, that is subgames that not end in leafs, then every Nash equilibrium of Γ is a subgame perfect equilibrium.*

Theorem 26 (Theorem 7.5, [11])**.** *Let Γ be an extensive form game and let $\boldsymbol{x}^*$ be its Nash equilibrium. If $\rho(v, \boldsymbol{x}^*) > 0$ for some $v \in V$, then $\boldsymbol{x}^*$ restricted to the subgame $\Gamma(v)$ is a Nash equilibrium of $\Gamma(v)$.*

Proof. For the proof of those theorems we refer to [11].

□

Theorem 27 (Theorem 7.9., [11])**.** *Every finite extensive-form game with perfect information has a subgame perfect equilibrium in pure strategies.*

Another possible refinement of the Nash equilibrium proposed by Selten in 1975 is related to the strategic form representation of the game, however motivation for it came from a extensive form of a game. This concept is based on the idea that players should consider also the possibility, that the other players can make mistakes in their strategizing with some small probability and defend against them. Therefore they may consider a perturbed version of the game where even dominated strategies may be played with some small probability.

Definition 49 (Definition 7.9., [11])**.** *Let G be a finite game in the strategic form. A perturbation vector of player $i \in I$ is a vector $\varepsilon_i = \{\varepsilon_i(p_i)\}_{p_i \in P_i}$ satisfying $\varepsilon_i(p_i) > 0$, $p_i \in P_i$ and*

$$\sum_{p_i \in P_i} \varepsilon_i(p_i) \leq 1.$$

$\boldsymbol{\varepsilon} = \{\varepsilon_i\}_{i \in I}$ is called a perturbation vector, where ε_i is a perturbation vector of the player $i \in I$. For a perturbation vector $\boldsymbol{\varepsilon}$ we define a $\boldsymbol{\varepsilon}$-perturbed game $G(\boldsymbol{\varepsilon})$ as a game in constrained strategies with constrained sets of strategies

$$X_i(\varepsilon_i) = \{x_i \in X_i; x_i(p_i) \geq \varepsilon_i(p_i), p_i \in P_i\},\ i \in I.$$

Theorem 28 (Theorem 7.20, [11]). *Every finite $G(\boldsymbol{\varepsilon})$ perturbed game has an equilibrium.*

Denote $M(\boldsymbol{\varepsilon}) = \max_{i \in I, p_i \in P_i} \varepsilon_i(p_i)$ and $m(\boldsymbol{\varepsilon}) = \min_{i \in I, p_i \in P_i} \varepsilon_i(p_i)$.

Theorem 29 (Theorem 7.22, [11]). *If p_i is a weakly dominated strategy, then every equilibrium of the $\boldsymbol{\varepsilon}$-perturbed game fullfils*

$$x_i(p_i) = \varepsilon_i(p_i).$$

Theorem 30 (Theorem 7.23, [11]). *Let $\{\boldsymbol{\varepsilon}^k\}_{k \in \mathbb{N}}$ be a sequence of perturbation vectors satisfying $\lim_{k \to \infty} M(\boldsymbol{\varepsilon}^k) = 0$. For every mixed strategy $x_i \in X_i$, there exists a sequence $\{x_i^k\}_{k \in \mathbb{N}}$ of mixed strategies of the player i satisfying:*

- *$x_i^k \in X_i(\varepsilon_i^k)$ for each $k \in \mathbb{N}$.*
- *$\lim_{k \to \infty} x_i^k = x_i$.*

Theorem 31 (Theorem 7.24, [11]). *Let G be a game in the strategic form. For $k \in \mathbb{N}$ let $\boldsymbol{\varepsilon}^k$ be a perturbation vector, and let $\boldsymbol{x}^k$ be an equilibrium of $G(\boldsymbol{\varepsilon})$, If*

1. *$\lim_{k \to \infty} M(\boldsymbol{\varepsilon}^k) = 0$,*
2. *$\lim_{k \to \infty} \boldsymbol{x}^k = \boldsymbol{x} \in X$,*

then $\boldsymbol{x}$ is a Nash equilibrium of the game G.

Definition 50 (Definition 7.25, [11]). *A mixed strategy profile $\boldsymbol{x}$ is called a perfect equilibrium of game G in the strategic form, if there exists a sequence of perturbation vectors $\{\boldsymbol{\varepsilon}^k\}_{k \in \mathbb{N}}$ with $\lim_{k \to \infty} M(\boldsymbol{\varepsilon}^k) = 0$ and for each $k \in \mathbb{N}$ there exists an equilibrium $\boldsymbol{x}^k$ of $G(\boldsymbol{\varepsilon}^k)$ such that*

$$\lim_{k \to \infty} \boldsymbol{x}^k = \boldsymbol{x}.$$

As a direct consquence of the previous discussion every perfect equilibrium is a Nash equilibrium.

Theorem 32 (Theorem 7.27, [11]). *Every finite game G in the strategic form has at least one perfect equilibrium.*

Lastly in the repeated game with discounting we may consider a so-called uniform equilibrium that has the property, that the strategy profile remains an equilibrium as the valuation of the future by the players increases to the point, where they value future gains same as the current.

Definition 51 ([11]). *Strategy profile $\boldsymbol{x}^* \in X$ is called a uniform equilibrium for games with discounting, if $\lim_{\beta \to 1} \boldsymbol{u}_\beta(\boldsymbol{x}^*)$ exists and there exists $\beta_0 \in [0,1)$, such that $\boldsymbol{x}^*$ is a Nash equilibrium for every game with $\boldsymbol{u}_\beta(\boldsymbol{x}) = \sum_{t=1}^{\infty} \beta^{t-1} \boldsymbol{u}_t(\boldsymbol{x})$ with $\beta \in [\beta_0, 1)$.*

Theorem 33 (The Folk theorem for discounted games, [11]). *Let G be a base game in which there exists $\hat{\boldsymbol{u}} \in F \cap V$, such that $\hat{u}_i > v_i^U, \forall i \in I$. Then $\forall \varepsilon > 0$ there exists $\beta_0 \in [0,1)$ such that $\forall \beta \in [\beta_0, 1)$ and $\forall \boldsymbol{u} \in V \cap F$, there exists an equilibrium $\boldsymbol{x}^*$ of the repeated version of G with discounting β, such that*

$$\max_{i \in I} |u_i^\beta(\boldsymbol{x}^*) - \boldsymbol{u}_i| < \varepsilon.$$

4. Sequential games with random payoff

In this chapter we will more closely examine the case of sequential games with random payoff. We will first start with the special case of a Markov decision process and its generalization in the form of a stochastic game as a motivation for the further generalization. As we will show, those two models are a special case of what we call a sequential game with random payoff. Then we will take a look at special forms of sequential game with random payoff with Markov and martingale payoffs and possible solution models for those types of games, lastly we will propose a solution model for the sequential game with random payoff with general ergodic payoff.

4.1 Markov decision process

The basic idea of the Markov decision process theory comes from the theory of Markov chains. A Markov decision process is a single player multi-stage game in which the player (called the decision maker) wants to maximize the expected payoff (reward) he will get based on his strategies (decisions). This reward is influenced by a current state of the system for which there exists non-deterministic transitions influenced by decisions taken in the preceeding stage of the game by the player. This may be represented as a collection of Markov chains, that are controled by the decision maker, this means that the decision maker chooses which of the Markov chains' transition matricies is used in the current stage to transit to a different state in the following stage. Decision maker in such a model wants to find an optimal control, which is a series of decisions that maximize the expected reward from the process.

Definition 52 ([13]). *Let $T = \{1, 2, \dots\}$ be a finite or countably infinite set of stages. Let for $t \in T : S_t$ be a finite or countably infinite set of possible states of the system in the time t and let $A_{t,s}$ be the set of possible actions at the time t if the system is in the state $s \in S_t$. Denote $r_t(s, a, j)$ the immediate reward at the stage t if the system is in the state s and the decision maker choses action a and the system will be in the state j in the stage $t+1$ and denote $p_t(j|s, a)$ the probability that the system will be in the state j in the time $t + 1$ given it is in the state s at the time t and the decision maker choses an action a. A markov decision process is the collection $(T, \{S_t\}_{t\in T}, \{A_{t,s}\}_{t\in T, s\in S_t}, \{r_t(s, a, j)\}_{t\in T, s\in S_t, a\in A_{t,s}, j\in S_{t+1}}$, $\{p_t(j|s, a)\}_{t\in T, s\in S_t, a\in A_{t,s}, j\in S_{t+1}})$.*

For simplicity in further discussion we will omit index sets. We will further assume that $\sum_{j\in S_{t+1}} p_t(j|s, a) = 1$ and that $A_{s,t} = A_t$ the set of actions does not depend on the current state of the system.

The expected reward in the stage t may be computed as

$$r_t(s, a) = \sum_{j\in S_{t+1}} r_t(s, a, j)p_t(j|s, a).$$

Further notice that for a given $t \in T$ and a given $a \in A_t$

$$(p(j|s,a))_{s\in S_t, j\in S_{t+1}}$$

is a transition matrix of some Markov chain.

Definition 53 ([13]). *A function $d_t : S_t \to A_t$ is called a (Markovian) decision rule. The set of allowed decision rules is denoted as D_t. Decision rule is called a randomized decision rule if for each state it is a probability distribution over the set of actions. A collection of decision rules for each $t \in T$, i.e. $\pi = (d_1, d_2, \dots)$ is called a policy. The set of policies is denoted by Π. If $\pi = (d_t)_{t\in T}$, where $\forall t \in T : d_t = d$, then policy π is called stationary.*

As stated in [13], each policy π specifies a stochastic process $\{X_t^\pi\}_{t\in T}$ with state space $\{S_t\}_{t\in T}$. Markov policy specifies a Markov chain and stationary policy specifies stationary Markov chain. This Markov process further induces a stochastic process of rewards $\{r(X_t^\pi, d_t(X_t^\pi))\}_{t\in T}$ which is called the Markov reward process.

For each policy we have a corresponding reward process and now the question is how to optimize this process in a way that we can consider it optimal.

4.1.1 Finite horizont

One way is to consider a finite horizont N and maximize the expected inflow of rewards until N. This is described as

$$v_N^\pi(s) = \mathsf{E}\,[\sum_{t=1}^{N} r_t(X_t^\pi, d_t(X_t^\pi)) + r_{N+1}(X_{N+1}^\pi)|X_1^\pi = s],$$

given that the initial state of the system was s and the decision maker chose policy π. This value exists and it is bounded, if $r_t(s,a)$ are bounded functions. Policy that maximizes the expected inflow is called the optimal policy. If the set of stages is finite, then the time horizont is naturally selected as the number of stages. If the number of stages is infinite, there may be multiple ways of selecting appropriate time horizont. If

$$v^\pi(s) = \mathsf{E}\sum_{t=1}^{\infty} r_t(X_t^\pi, d_t(X_t^\pi)) < \infty$$

for all possible choises of policy $\pi \in \Pi$, then the decision-maker may select an $\varepsilon > 0$ and consider some N such that

$$\forall \pi \in \Pi : \mathsf{E}\sum_{t=N+1}^{\infty} r_t(X_t^\pi, d_t(X_t^\pi)) < \varepsilon.$$

The benefit of choosing a finite time horizont in such a situation is that, if both the state and action spaces are finite, then it is easy to recurently compute the expected value until the time N and the corresponding optimal policy. Optimal policy π_N found in this way for the finite horizont N would correspond to a ε-optimal policy $\pi = (\pi_N, d_{N+1}, \dots)$, where $d_t, t \geq N+1$ may be chosen arbitrarly. In the case that for some of the policy choises the expected reward income is infinite, the choise of N is not so straight-forward, but the decision maker may still consider a finite time horizont he values the most and solve program

$$\max_{\pi \in \Pi} v_N^\pi(s),$$
$$\text{s.t. } v^\pi(s) = \infty.$$

That is they consider only policies with infinite inflow and select such that it maximizes the rewards until the time N.

Now lets consider the exact algorithm for solving the finite time horizont problem. First let us define a history of the decision process.

Definition 54 ([13]). *History of the Markov decision process until stage t is defined as $H_t = \{S_1, A_1, \dots, A_{t-1}, S_t\}$ for $t = 2, \dots, N+1$. Elements of the history set may be inductivelly constructed as $h_t = (h_{t-1}, a_{t-1}, s_t)$ for $h_{t-1} \in H_{t-1}$.*

We will define the expected value from time n onward, given history h_n as

$$u_n^\pi(h_n) = \mathsf{E} \sum_{t=n}^{N+1} r_t(X_t^\pi, d_t(X_t^\pi))$$

The finite horizont policy evaluation algorithm as proposed in [13] for a given policy $h_{N+1} = (h_N, a_N, s_{N+1}) \in H_{N+1}$ is of the form

1. Set $t = N+1$ and compute $u_{N+1}^\pi(h_{N+1}) = r(s_{N+1})$.
2. Lower t by 1 and compute $u_t^\pi(h_t)$ for every $h_t \in H_t$ as

$$u_t^\pi(h_t) = r_t(s_t, d_t(s_t)) + \sum_{j \in S_{t+1}} p(j|s_t, d(s_t)) u_{t+1}^\pi(h_t, d(s_t), j)$$

3. If $t = 1$ stop, otherwise repeat 2.

In [13] Theorem 4.2. shows that policy that maximizes valuation by this algorithm is an optimal policy for the finite horizont problem and that to find such a policy we can use the backward induction algorithm defined as follows

1. Set $t = N+1$ and compute $u_t(s_t) = r_t(s_t)$ for each $s_t \in S_t$.
2. Lower t by 1 and compute $u_t(s_t)$ for each $s_t \in S_t$ as

$$u_t(s_t) = \max_{a \in A_t} [r_t(a, s_t) + \sum_{s_{t+1} \in S_{t+1}} p(s_{t+1}|s_t, a) u_{t+1}(s_{t+1})].$$

3. If $t = 1$ stop, otherwise repeat 2.

Where the optimal policy's decision rule d_t at the time t takes only values of the argument maxima of the equation from the step 2 of the algorithm. This algorithm allows us to immediately compute both the optimal policies and the corresponding expected inflow of rewards.

4.1.2 Infinite horizont

In the case when T is infinite we can consider similiar approaches as in the case of repeated games in the Chapter 3 of this thesis. First case is when

$$\forall \pi \in \Pi, \forall s \in S_1 : v^{\pi}(s) = \mathsf{E} \sum_{t=1}^{\infty} r_t(X_t^{\pi}, d_t(X_t^{\pi})) < \infty.$$

In this case we want to find $\pi^* \in \Pi$ such that

$$\forall \pi \in \Pi, \forall s \in S_1 : v^*(s) := v^{\pi^*}(s) \geq v^{\pi}(s)$$

The second possible approach is to consider a discounting factor $0 < \lambda < 1$ and compute

$$v_{\lambda}^{\pi}(s) = \mathsf{E} \sum_{t=1}^{\infty} \lambda^{t-1} r_t(X_t^{\pi}, d_t(X_t^{\pi})).$$

In the case when $\forall t \in T : \sup_{a \in A_t} \sup_{s \in S_t} |r_t(s,a)| \leq M < \infty$ it holds that $\forall s \in S_1, \forall \pi \in \Pi : |v_{\lambda}^{\pi}(s)| \leq \frac{M}{1-\lambda} < \infty$.

The last approach is to consider the mean average reward given as

$$v^{\pi}(s) = \lim_{N \to \infty} \frac{1}{N} v_N^{\pi}(s)$$

where $v_N^{\pi}(s)$ is the finite horizont valuation as defined above.

4.2 Stochastic game

The idea of stochastic games is based on the generalization of the Markov decision process for multiple players, who have their own payoff functions and their strategies influence the transitions between possible states.

Definition 55 ([11]). *Let $T = \{1, \ldots, N\}$ be a set of stages for $N \leq \infty$ and I be a set of players. Let for each stage t S_t denote the set of states in the time t and $A_{i,s,t}$ be the set of pure strategies available to the player i, if the game is in the stage t at the state s, denote $A_{s,t} = \bigtimes_{i \in I} A_{i,s,t}$. Let $P_t(j|\boldsymbol{a}_t, s_t, \boldsymbol{a}_{t-1}, s_{t-1}, \ldots, \boldsymbol{a}_1, s_1)$ denote the probability, that the game will be in a state j in the stage $t+1$ given the history of the game at the time t was $h_t = (s_1, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{t-1}, s_t)$ for $s_k \in S_k$ and $\boldsymbol{a}_k \in A_{s_k,k}$, $k = 1, \ldots, t$. Denote the set of histories until the time t as H_t. Let $u_i^t : S_t \times \bigtimes_{s \in S_t} A_{s,t} \to \mathbb{R}$ be the payoff function of the player i in the stage t. We define a stochastic game as a collection $\Sigma = (T, \{S_t\}_{t \in T}, \{P_t(\cdot|h_t)\}_{t \in T, h_t \in H_t}, I,$ $\{A_{i,s,t}\}_{t \in T, s \in S_t, i \in I}, \{u_i^t\}_{t \in T, i \in I})$.*

In the stochastic game model it is very natural to use behavior strategies as were first described in the Chapter 3 of this thesis. However the definition of behavior strategies is in this case somewhat different and takes a more natural form of assigning a distribution over pure strategies to every history of the game.

Definition 56 ([11]). *Let H_t be a set of histories until time t and $H = \bigcup_{t \in T} H_t$ be the set of histories. Function τ_i assigning a history $h = (s_1, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{t-1}, s_t)$ a probability distribution over pure strategies available to the player i in the time t is called a behavior strategy. Vector $\tau = \{\tau_i\}_{i \in I}$ is called a behavior strategy profile.*

It is possible to show that in the case of a stochastic game the behavior strategies are equivalent to mixed strategies as was described in Chapter 3. We just have to realize that the game may be rewritten in an extensive form as a tree determined by unique combinations of stages and states. In this tree the information sets are single verticies and therefore every behavior strategy has an equivalent mixed strategy.

Behavior strategy profiles together with probabilities of transition between the states of the game induce a probability distribution over the sets of histories for a history $h = (s_1, \boldsymbol{a}_1, \dots)$ given as

$$P_{s_1,\tau}(h) = (\prod_{t=1}^{N} \prod_{i \in I} \tau_i(a_i^t | s_1, \boldsymbol{a}_1, \dots, s_{t-1}, \boldsymbol{a}_{t-1}, s_t))(\prod_{t-1}^{N-1} P_t(s_{t+1} | \boldsymbol{a}_t, s_t, \dots, \boldsymbol{a}_1, s_1)),$$

where a_i^t denotes the action of the player i in the time t.

4.2.1 Finite horizont

Let us again begin by considering the case for stochastic games with finite horizont, that is $N < \infty$.

In this case commonly considered approach is to consider the game in the strategic form with the payoff defined as an expected average payoff with respect to $P_{s_1,\tau}$, that is

$$\gamma_i^N(s_1, \tau) = \mathsf{E}_{s_1,\tau}[\frac{1}{N} \sum_{t=1}^{N} u_i^t].$$

This is an analogue of the approach seen in the Markov decision processes, where the player (decision maker) wants to maximize his expected reward.

If we denote X_i^N the set of behavior strategies of the player i in the stochastic game Σ we can define the strategic form of the game Σ with initial state s_1 as $G(s_1) = (I, \{X_i\}_{i \in I}, \{\gamma_i(s_1)\}_{i \in I})$

Definition 57 ([11])**.** *Let $s_1 \in S_1$ be the initial state of the game. Behavior strategy profile τ^* is called a Nash equilibrium of the stochastic game Σ with the initial state s_1, if it is a Nash equilibrium of $G(s_1)$.*

Definition 58 ([11])**.** *Behavior strategy profile τ^* is called a Nash equilibrium of the stochastic game Σ, if it is a Nash equilibrium for every initial state $s_1 \in S_1$.*

Theorem 34 ([11])**.** *Every finite stochastic game has a Nash equilibrium.*

The stochastic game the game is described using a (controled) stochastic process with values in the sets of states of the game. If this process reaches the same state by different paths the question is whether the player would or should react differently. If the state process is homogenous Markov, that is

$$\forall t \in T : P_t(.|\boldsymbol{a}_t, s_t, \dots, \boldsymbol{a}_1, s_1) = P(.|\boldsymbol{a}_t, s_t)$$

we may consider the following definition.

Definition 59 ([11])**.** *Behavior strategy τ_i of the player $i \in I$ is called Markov, if $\forall t \in T, \forall h, \hat{h} \in H_t$ such that $h = (s_1, \boldsymbol{a}_1, \dots, s_t)$ and $\hat{h} = (\hat{s}_1, \hat{\boldsymbol{a}}_1, \dots, s_t)$ it holds $\tau_i(h) = \tau_i(\hat{h})$.*

Theorem 35 ([11]). *Every N-stage stochastic game has a Nash equilibrium in Markov strategies.*

This means that the natural approach to react in the same situation in the same manner does not constrain us from reaching an optimal solution for the game.

4.2.2 Infinite horizont

In the case of a infinite horizont we need to make sure that the game has a finite expected payoff. For this reason discounting is used for each stage of the game in the standard way as it was used in the case of a Markov decision processes.

Definition 60 ([11]). *Let $\{\beta_t\}_{t\in\mathbb{N}}$ be a discounting such that $\beta_t = \beta^{t-1}$ for $\beta \in [0, 1)$. We define the β-discounted payoff for a behavioral strategy profile $\boldsymbol{\tau} \in X$ as*

$$\gamma_i^\beta(s_1, \boldsymbol{\tau}) = E_{s_1,\boldsymbol{\tau}} \sum_{t=1}^{\infty} \beta_t u_i^t.$$

Definition 61 ([11]). *The β-discounted infinite stochastic game with initial state s_1 is the strategic form game $G(s_1) = (I, X, \{\gamma_i^\beta\}_{i\in I})$.*

Similarly as previous we define the Nash equilibria.

Definition 62 ([11]). *Strategy profile $\boldsymbol{\tau} \in X$ is a β-discounted equilibrium for the initial state s_1, if it is a Nash equilibrium of $G(s_1)$. Strategy profile $\boldsymbol{\tau} \in X$ is a β-discounted equilibrium, if it is a β-discounted equilibrium for every initial state s_1.*

Here we define a stationary strategy in a similiar way than Markovian strategies from the previous section.

Definition 63 ([11]). *A behavior strategy τ_i of player i in an infinite stochastic game is a stationary strategy if for every finite history $h = (s_1, \boldsymbol{a}_1, \dots, s_{t-1}, \boldsymbol{a}_{t-1}, s_t)$ the mixed action $\tau_i(h)$ depends only on the state s_t.*

That is stationary strategies react only to the current state of the game and do not consider past of the game relevant.

Theorem 36 (Fink, [11]). *For every discount factor $\beta \in [0, 1)$, every β - discounted stochastic game has an equilibrium in stationary strategies.*

For the proof of this theorem we refer to [11].

4.3 Sequential game with random payoff

Now let us define a more general version of game with randomness.

Definition 64. *Let $\Gamma = (T, \{G_v\}_{v\in V}, \{s_v\}_{v\in V})$ be a game in the extensive form, where G_v is a game with random payoff for some $v \in V$, then Γ is called a sequential game with random payoff.*

By definition the information sets of this game are equal to individual verticies of the game tree and therefore we will not distinguish between notation for mixed and behavioral strategies. As the behavior strategies are better imagined we will mostly use them when speaking about strategies in the sequential game with random payoff.

Definition 65. *We will say that two games in extensive form G and H are equivalent in the payoff, if $(V_G, E_G) = T_G = T_H = (V_H, E_H)$, where T_G and T_H are the game trees of G and H respectivelly and for every mixed or behavior strategy profile $\boldsymbol{x}$ it holds that $u_G^v(\boldsymbol{x}) = u_H^v(\boldsymbol{x})$, a.s. for every $v \in V_G = V_H$.*

Theorem 37. *Let Σ be a stochastic game, then there exists a payoff equivalent sequential game with random payoff.*

Proof. Let Σ be a stochastic game. This means that the game tree of Σ is given as $V = T = \{1, \dots, N\}$ for some $N \leq \infty$ and $E = \{(n, n+1); n \in T, n+1 \in T\}$. Denote $\mathcal{S}_1$ the distribution of the initial states of the game. As we have discussed earlier in the case of a stochastic game the behavior and mixed strategies are equivalent, therefore we just need to show that there exists a game with random payoff such that for every behavior strategy profile $\tau \in X$ the distribution of payoff in the stochastic game is the same as in the game with random payoff. In the stochastic game the distribution of the payoff is given by the transformation of the distribution of the state of the game. We can express this distribution in the stage $v \in V$ given behavior strategy τ as $\mathcal{S}_v^\tau(h_v) = \int P_{s_1,\tau}(h_v)\mathcal{S}_1(ds_1)$, where h_v is the history of the game Σ until the stage $v \in V$. Let $(\Omega_v, \mathcal{A}_v, \mathbb{P}_v)$ be a probability space and S_v^τ be a random variable with values in the set of states of the stochastic game in the stage v with the distribution $\mathcal{S}_v^\tau$. If we denote $u_i^v(\tau) = u_i^v(S_v^\tau, \tau)$ the random payoff of the stage $v \in V$, then clearly $u_i^v(\tau) \sim u_i^v(S_v^\tau, \tau), \forall \tau \in X$. Denote $G_v = (I, X^v, \{u_i^v\}_{i \in I})$ the game with random payoff $u_i^v(\tau)$, then $\Gamma = (T, \{G_v\}_{v \in V}, \{s_v\}_{v \in V})$ is payoff-equivalent game with random payoff to the stochastic game Σ, where $s_v(\tau) = v + 1$ is the successor function in the extensive form of the game.

$\square$

This theorem states that there is no difference in considering the random transitions between states of the game and the randomness in the payoff itself. The advantage of random payoff model is that in reality we may not be able to observe the states of the game but only the resulting payoffs. As we have seen in previous section in the model of stochastic game the main goal is to optimize the expected payoff in the game. In the model of a game with random payoff we want to use optimality criteria as described in the second chapter of this thesis.

4.3.1 Optimality in sequential games with random payoff

In further discussion we will consider the case of sequential games with random payoff, which have only a single branch. That is we will assume that $V = \{1, \dots, N\}$ and $E = \{(n, n+1); n \in V, n+1 \in V\}$ and $s_t(\boldsymbol{x}) = t+1$ if $t+1 \in V$. In this situation given mixed or behavior strategy profile $\boldsymbol{x} = (\boldsymbol{x}^1, \boldsymbol{x}^2, \dots) \in X$ yields for the player i a payoff process $\{u_i^t(\boldsymbol{x})\}_{t \in \mathbb{N}}$, where we use the notation $u_i^t(\boldsymbol{x}) = u_i^t(\boldsymbol{x}^t | u_i^{t-1}(\boldsymbol{x}^{t-1}), \dots, u_i^1(\boldsymbol{x}^1))$, *a.s.*.

Definition 66. *Let Γ be a sequential game with random payoff. We will say that the payoff of the player i in the Γ has property the $\mathcal{P}$ if $\forall \boldsymbol{x} \in X$ mixed or behavior strategy profile the associated payoff process $\{u_i^t(\boldsymbol{x})\}_{t\in T}$ has the property $\mathcal{P}$. We will say, that the payoff in the game Γ has the property $\mathcal{P}$ if for every player $i \in I$ the payoff of the player i has the property $\mathcal{P}$.*

First let us consider the stepwise approach, where the player i wants to construct a behavioral strategy by finding the optimal strategy at the time t given the history of the game until time t was h_t. This means the players want to optimize their payoff at the time t given as $u_i^t(\boldsymbol{x}^t|h_t) = u_i^t(\boldsymbol{x}^t|u_i^{t-1}(\boldsymbol{x}^{t-1}) = u_i^{t-1}, \dots, u_i^1(\boldsymbol{x}^1) = u_i^1)$ where $h_t = (\boldsymbol{x}^1, u_i^1, \boldsymbol{x}^2, u_i^2, \dots, \boldsymbol{x}^{t-1}, u_i^{t-1})$. Here $u_i^t(\boldsymbol{x}^t|h_t)$ is a conditional payoff under the condition that the history of the game was h_t. Notice here the history h_t is no longer described by the states and actions but by strategies and their respective realizations of payoff. Let us start by considering that the players want to maximize their immediate reward in a given time t. In this case we may consider the generalizations of the minimax program to the stochastic case as was presented in the Chapter 2. First we will write the least likely payoff program.

$$\begin{aligned} &\max_{x_i^t \in X_i^t, \delta \in \mathbb{R}} \delta \\ &\text{s.t. } \mathbb{P}[u_i^t(x_i^t, \boldsymbol{x}_{-i}^t|h_t) \geq \delta] \geq \alpha_i^t, \boldsymbol{x}_{-i}^t \in X_{-i}^t, \end{aligned} \tag{4.1}$$

where $\alpha_i^t \in (0,1)$ is a given confidence level of the player i for the time t. X_{-i}^t is usually a infinite set of strategies of the other players, which means that (4.1) is a program with infinite number of bounds each corresponding to a given mixed strategy of the other players. This can be overcome if we are able to compute $\inf_{\boldsymbol{x}_{-i} \in X_{-i}} \mathbb{P}[u_i^t(x_i^t, \boldsymbol{x}_{-i}^t|h_t) \geq \delta]$. This however, may be complicated and so in practice we would use a relaxed form of this program, where the player i wants to defend himself only against particular actions (pure strategies) of the other players this yields a program

$$\begin{aligned} &\max_{x_i^t \in X_i^t, \delta \in \mathbb{R}} \delta \\ &\text{s.t. } \mathbb{P}[u_i^t(x_i^t, \boldsymbol{p}_{-i}^t|h_t) \geq \delta] \geq \alpha_i, \boldsymbol{p}_{-i}^t \in P_{-i}^t. \end{aligned} \tag{4.2}$$

Where $\boldsymbol{p}_{-i}^t$ denotes a profile of pure strategies the other players at the time t and P_{-i}^t denotes the set of pure strategies available to the other players at the time t.

The second generalization of the minimax program as was discussed in the Chapter 2 is the worst payoff method, where players want to solve a program

$$\begin{aligned} &\max_{x_i^t \in X_i^t, \delta \in \mathbb{R}} \delta \\ &\text{s.t. } \mathbb{P}[\forall \boldsymbol{x}_{-i}^t \in X_{-i}^t : u_i^t(x_i^t, \boldsymbol{x}_i^t|h_t) \geq \delta] \geq \alpha_i^t. \end{aligned} \tag{4.3}$$

With this the player choses strategy $\boldsymbol{x}_i^* \in X_i^t$ for the time t so that he is guaranted to gain at least δ^* with probability higher than α_i^t given the history of the game was h_t.

In a sequential game with random payoff we may also consider using prediction of the payoff and optimize with respect to the predicted payoff function at the time t. This means that we construct $\hat{u}_i^t|h_t$ the predicted payoff function, given the history of the game until time t was h_t. In a standard manner we want to minimize the prediction error given as

$$e_{\hat{u}_i}^2(\boldsymbol{x}) = \mathsf{E}\ |u_i^t(\boldsymbol{x}) - \hat{u}_i^t(\boldsymbol{x})|^2.$$

Where $\hat{u}_i^t(\boldsymbol{x})$ we want to be $\sigma(u_i^{t-1}(\boldsymbol{x}), \dots, u_i^1(\boldsymbol{x}))$-measurable random variable. We know that for a given $\boldsymbol{x} \in X$ mixed or behavioral strategy this is minimized by the conditional mean value $\mathsf{E}\,[u_i^t(\boldsymbol{x})|u_i^{t-1}(\boldsymbol{x}), \dots, u_i^1(\boldsymbol{x})]$. As $u_i^{t-k}(\boldsymbol{x}), k = 1, \dots, t-1$ are known at the time t this yields a deterministic game $\hat{G}_t = (I, X^t, \{\hat{u}_i\}_{i \in I})$ with the predicted payoff functions. So that the player wants to construct his strategy with as follows

- At $t = 1$ compute optimal strategy x_i^1 with respect to the game with deterministic payoff $\mathsf{E}_{\mathbb{P}}\,\boldsymbol{u}$.
- At $t \geq 2$ given the history until time t was $h_t = (\boldsymbol{x}^1, \boldsymbol{u}^1, \dots, \boldsymbol{x}^{t-1}, \boldsymbol{u}^{t-1})$ compute optimal strategy x_i^t with respect to the deterministic game $\hat{\boldsymbol{u}}(h_t) = \mathsf{E}\,[\boldsymbol{u}|h_t]$.

Another approach is based on considering an aggregated payoff from the game. Either we may consider some finite horizont T of the game and optimize with respect to

$$u_i(\boldsymbol{x}) = \sum_{t=1}^{T} u_i^t(\boldsymbol{x}), a.s.$$

or infinite horizont with

$$u_i(\boldsymbol{x}) = \sum_{t=1}^{\infty} u_i^t(\boldsymbol{x}), a.s.$$

where we define $\sum_{t=1}^{\infty} u_i^t(\boldsymbol{x}) = \lim_{T\to\infty} \sum_{t=1}^{T} u_i^t(\boldsymbol{x})$ by the convergence almost surely. If $u_i(\boldsymbol{x}) < \infty, a.s.$ we may consider the optimality conditions for strategic form of the game as were presented in the second chapter. Disadvantage of this approach is that especially in the case with infinite horizont it may be difficult to find a closed form representation of the payoff and of the sets of strategies of the strategic form of the game. For that reason we may prefer to use finite horizont instead. To do this we can use a similiar approach as was presented in the Chapter 3, where we choose an $\varepsilon > 0$ and find $T = T(\varepsilon) \in \mathbb{N}$ such that

$$\forall i \in I, \forall \boldsymbol{x} \in X : |\sum_{t=1}^{T} u_i^t(\boldsymbol{x}) - \sum_{t=1}^{\infty} u_i^t(\boldsymbol{x})| < \varepsilon, a.s..$$

With this we may now consider the ε-version of the original sequential game with random payoff in the strategic form as was described in the second chapter.

In some cases it may happen that $\mathbb{P}[\sum_{t=1}^{\infty} u_i^t(\boldsymbol{x}) = \infty] > 0$. In such a case we need to regularize the game so that it is possible to consider it as a strategic form game. To do this we may use the standard techniques as were presented in the Chapter 3. First way is to consider a discounting $\{\beta_t\}_{t\in T}$ and optimize with respect to the discounted overall payoff given as

$$u_i(\boldsymbol{x}) = \sum_{t=1}^{\infty} \beta_t u_i^t(\boldsymbol{x}), a.s..$$

Here we assume that $\forall i \in I, \forall \boldsymbol{x} \in X : u_i(\boldsymbol{x}) < \infty, a.s.$ and we have the same possible optimality conditions as for a standard game in the strategic form. If the discounting factors are not suitable for the scenario (for example if there is no clear discounting factor to use), we may consider the average payoff model, where players want to optimize the average payoff function given as

$$\bar{u}_i(\boldsymbol{x}) = \liminf_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u_i^t(\boldsymbol{x}), a.s..$$

Let us now take a closer look on the case when the associated payoff processes have some additional properties.

4.3.2 Sequential markov game

We will start with a repeated sequential game with random payoff where the payoff process has the Markov property. This means that given mixed or behavior strategy profile $\boldsymbol{x} \in X$ it holds that $\forall i \in I : \mathbb{P}[u_i^t(\boldsymbol{x}) \in A | u_i^{t-1}(\boldsymbol{x}) = u_{t-1}, \ldots, u_i^0(\boldsymbol{x}) = u_0] = \mathbb{P}[u_i^t(\boldsymbol{x}) \in A | u_i^{t-1}(\boldsymbol{x}) = u_{t-1}]$ for $A \in \mathcal{B}(\mathbb{R})$.

The corresponding Markov transition kernell of the player i is given as

$$P_i^t(\boldsymbol{x}_t, u, A) = \mathbb{P}[u_i^t(\boldsymbol{x}) \in A | u_i^{t-1}(\boldsymbol{x}) = u]$$

Further we will assume that $\forall t \in T : P_i^t = P_i$, or that the payoff process has homogenous transitions.

For a given mixed strategies profile $\boldsymbol{x} \in X$ and given distribution of the payoff of the player i in the time $t-1 : U_i^{t-1}(\boldsymbol{x})$ we can express the distribution of the payoff at the time t as

$$U_i^t(\boldsymbol{x})(A) = \int P_i(\boldsymbol{x}_t, u_{t-1}, A) U_i^{t-1}(\boldsymbol{x})(du_{t-1}).$$

In the case of the Markov game the distribution U_i^t may be described by u_i^{t-1}, which is known in the time t and P_i. Here u_i^{t-1} is a realization of the payoff function at the time $t-1$. And so, if we denote the strategy profile played at time $t-1$ as $\boldsymbol{x}_{t-1}$ the payoff was $u_i^{t-1}(\boldsymbol{x}_{t-1})$.

And so the program (4.1) may be reformulated in the form

$$\begin{aligned} &\max_{x_i^t \in X_i^t, \delta \in \mathbb{R}} \quad \delta, \\ &\text{s.t. } P_i(x_i^t, \boldsymbol{x}_{-i}^t, u_i^{t-1}(\boldsymbol{x}_{t-1}), [u_i^t(x_i^t, \boldsymbol{x}_{-i}^t) \geq \delta]) \geq \alpha_i^t, \boldsymbol{x}_{-i}^t \in X_{-i}^t. \end{aligned} \tag{4.4}$$

Where P_i denotes the Markov transition kernell of the game.

4.3.3 Sequential martingale game

Another interesting case is when for a given mixed or behavioral strategy profile $\boldsymbol{x} \in X$ the resulting payoff process is a martingale. This means that we assume $\forall i \in I, \forall \boldsymbol{x} \in X, \forall t \in \mathbb{N} : \mathsf{E}\ u_i^t(\boldsymbol{x}) < \infty$ and that $\mathsf{E}\left[u_i^t(\boldsymbol{x}^t) | u_i^{t-1}(\boldsymbol{x}), \ldots, u_i^1(\boldsymbol{x})\right] = u_i^{t-1}(\boldsymbol{x}^t), a.s.$. For this type of a game it is natural to consider the model with predicted payoff. As we have discussed earlier the $\mathcal{L}_2$ optimal predicted payoff function with respect to the payoff process σ-algebra $\mathcal{H}_t = \sigma(u_i^{t-1}(\boldsymbol{x}), \ldots, u_i^1(\boldsymbol{x}))$ is the conditional expected payoff function. In the case of a martingale game this is exactly the realization of the payoff function at the stage $t-1$. In this model the player i would want to construct his strategy so that x_i^1 is optimal with respect to the expected payoff function $\mathsf{E}\ u_i^1$ and then for a $t \geq 2$ the strategy x_i^t is optimal with respect to the realization of the payoff function u_i^{t-1} at the time $t-1$, which is known for the player at the time of the decision.

4.3.4 Sequential game with ergodic payoff

Lastly let us consider the case when the payoff process is ergodic for every possible mixed or behavioral strategy. For the purposes of our thesis we will define an ergodic process in the following way.

Definition 67. *Let $X = \{X_t\}_{t \in T}$ be a random process. We say that X is ergodic if*

- $\forall t \in T : \mathsf{E}\ X_t = \mu \in \mathbb{R}$ *and* $var(X_t) = \sigma^2 > 0$.
- $\bar{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t \xrightarrow{a.s.} \mu$.
- $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{D} N(0, a\sigma^2)$ *for some* $a > 0$

That is we call process ergodic if it follows the strong law of large numbers and satisfies the requirements of some version of the Central limit theorem. For our payoff processes we assume that $\forall i \in I : \forall \boldsymbol{x} \in X$ the associated payoff process $\{u_i^t(\boldsymbol{x})\}_{t \in T}$ is ergodic.

In this case it follows that the average payoff function is of the form

$$\bar{u}_i(\boldsymbol{x}) = \liminf_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} u_i(\boldsymbol{x}) = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} u_i(\boldsymbol{x}) = \mathsf{E}\ u_i(\boldsymbol{x}), a.s..$$

Therefore in this type of a game the players with the average payoff want to maximize the mean (or expected) payoff function $\mathsf{E}\ u_i$. Let us denote $\sigma_i^2(\boldsymbol{x}) = \mathrm{var}(u_i(\boldsymbol{x}))$ and assume that

$$\sqrt{N}(\bar{u}_i(\boldsymbol{x}) - \mathsf{E}\ u_i(\boldsymbol{x})) \xrightarrow{D} N(0, a_i(\boldsymbol{x})\sigma_i^2(\boldsymbol{x}))$$

for some $a_i(\boldsymbol{x}) > 0$. We can also express the approximate $1-\alpha$ confidence interval for the payoff after N rounds of the game as

$$(\mathsf{E}\ u_i(\boldsymbol{x}) - \frac{q_{1-\alpha/2}\sqrt{a_i(\boldsymbol{x})}\sigma_i(\boldsymbol{x})}{\sqrt{N}}, \mathsf{E}\ u_i(\boldsymbol{x}) + \frac{q_{1-\alpha/2}\sqrt{a_i(\boldsymbol{x})}\sigma_i(\boldsymbol{x})}{\sqrt{N}}),$$

where $q_{1-\alpha/2}$ is a $(1-\alpha/2)$-quantile of the normal distribution.

5. Numerical analysis

5.1 Cornout model of duopoly

5.1.1 Static model of demand

The Cornout model was developed by french mathematician Antoine Augustin Cornout in 1838. This model describes behaviour of duopoly situation on a closed market with homogenous good. Example of such situation may be the competition of companies such as Coca-cola and Pepsi, who produce very similar products and have a dominant position on the market. The main assumptions of this model are that the goods are homogenous (or that they are perfect substitutes), demand for the products is linear, there is no cost of production and companies want to maximize their profit independetly of one another.

Let us formalize this mathematically in the following way. Let I and II be companies (players). Let Q_I, Q_{II} be the quantity produced by I and II respectivelly. $Q = Q_I + Q_{II}$ be the total quantity of goods produced on the market.

The demand for the goods is given as $D(Q) = a - bQ$ for some $a > 0, b > 0$ such that $a - bQ \geq 0, \forall Q$, where a denotes the market cap and b is the market saturation rate. Without the loss of generality we can use $D(Q) = a - Q$, where a denotes the relative market cap with respect to the saturation rate. Here D describes the price of the goods as a function of available quantity of goods Q. It is further assumed that $a - Q \geq 0$ for all possible values of Q.

Each company respectivelly wants to maximize its payoff given as

$$u_i(Q_I, Q_{II}) = D(Q)Q_i = aQ_i - QQ_i, i = I, II.$$

This is clearly a concave function of Q_i and so the optimal solution is given by a system of equations

$$\frac{\partial u_I(Q_I, Q_{II})}{\partial Q_I} = 0, \tag{5.1}$$

$$\frac{\partial u_{II}(Q_I, Q_{II})}{\partial Q_{II}} = 0, \tag{5.2}$$

$$a - Q_{II} - 2Q_I = 0, \tag{5.3}$$

$$a - Q_I - 2Q_{II} = 0. \tag{5.4}$$

Which yields $Q_I^* = Q_{II}^* = \frac{a}{3}$.

Another variation of this model describes the demand as a hyperbolic function of quantity given as $D(Q) = \frac{a}{bQ}$, where $a > 0, b > 0$ and $Q > 0$. Again we may without the loss of generality assume that $D(Q) = \frac{a}{Q}$, where $a > 0$ and $Q > 0$.

In this case the payoff function is given as

$$u_i(Q_I, Q_{II}) = D(Q)Q_i = \frac{a}{Q}Q_i, i = I, II.$$

In this case the resulting game is a constant-sum game because

$$u_I(Q_I, Q_{II}) + u_{II}(Q_I, Q_{II}) = \frac{a}{Q}Q_I + \frac{a}{Q}Q_{II} = a.$$

5.1.2 Dynamic model of demand

Let $T = \{1, \dots\}$ be the set of time. Denote Q_I^t, Q_{II}^t the quantity of goods produced by I and II respectivelly at time $t \in T$. Let $Q_t = Q_I^t + Q_{II}^t$ be the total quantity of goods produced on the market at the time $t \in T$.

The demand for the goods at time t is given as $D_t(Q_t) = a_0 - a_1(Q_t - Q_{t-1}) - bQ_t$, where $Q_0 = 0$ (or in more general $Q_0 = q_0 \in \mathbb{R}$) and $a_0 > 0, a_1 > 0, b > 0$ such that $a_0 - a_1(Q_t - Q_{t-1}) - bQ_t \geq 0, \forall Q_t, Q_{t-1}$. Which can again be simplified to just $a_0 - a_1(Q_t - Q_{t-1}) - Q_t$. Here a_0 denotes the relative market cap and a_1 the change of the market cap with respect to the change of supply. With this the payoff function at the time t is given as

$$u_i^t(Q_I^t, Q_{II}^t) = D_t(Q_t)Q_i^t, i = I, II.$$

We can derive the optimal solution at the time t given the quantity produced at time $t-1$ was Q_{t-1} as before by solving

$$\frac{\partial u_I(Q_I^t, Q_{II}^t)}{\partial Q_I^t} = 0, \tag{5.5}$$

$$\frac{\partial u_{II}(Q_I^t, Q_{II}^t)}{\partial Q_{II}^t} = 0, \tag{5.6}$$

$$a_0 - 2a_1 Q_I^t - a_1 Q_{II}^t + a_1 Q_{t-1} - 2Q_I^t - Q_{II}^t = 0, \tag{5.7}$$

$$a_0 - 2a_1 Q_{II}^t - a_1 Q_I^t + a_1 Q_{t-1} - 2Q_{II}^t - Q_I^t = 0. \tag{5.8}$$

This results in the optimal solution

$$Q_I^t = \frac{a_0 + a_1 Q_{t-1}}{3(a_1+1)},$$

$$Q_{II}^t = \frac{a_0 + a_1 Q_{t-1}}{3(a_1+1)}.$$

Which means that the recurent expression for overal quantity produced by duopolists at time t is $Q_t = \frac{2(a_0 + a_1 Q_{t-1})}{3(a_1+1)}$. This may be expressed in a finite geometric sum

$$Q_t = \frac{2}{3}\frac{a_0}{a_1+1}\sum_{n=0}^{t}\left(\frac{2}{3}\frac{a_1}{a_1+1}\right)^n$$

with closed form

$$Q_t = \frac{a_0}{a_1+1}\frac{2}{3}\left(\frac{1-(\frac{2a_1}{3(a_1+1)})^t}{1-(\frac{2}{3}\frac{a_1}{a_1+1})}\right).$$

With $t \to \infty$ this expression converges to

$$Q = \frac{2a_0}{a_1+3}.$$

Notice that as $a_1 \to 0$ we get exatly the solution from the static model. The resulting optimal strategy is a behavioral strategy assigning each time $t \in T$ an action Q_t. If we were to assume that the players are not perfectly rational, that is that they can make an error the optimal strategy of the player i would be to use the recurent formula. However, under our assumptions from the Chapter 2, there is no difference between the recurent and closed form of the strategy. In this approach we considered the reaction of the player in a given round of the game and computed the optimal strategy.

Now let us consider different approaches as presented in the Chapter 3. We will start by assuming that $T = \mathbb{N}$ and that both players have an oportunity to invest their assets with a risk-free interest rate $0 < r$. This means that their time value of payoff is described by a discounting factor $\beta = \frac{1}{1+r}$ and so they want to play a game with discounting $\{\beta_t\}_{t\in T}$, where $\beta_t = \beta^{t-1}$. Since we assume that $0 < r$ it is clear that $\{\beta_t\}_{t\in T}$ is a consistent valuation of time ($\beta_t \geq \beta_{t+1}$ and $\beta_1 = 1$). Denote $Q_I = (Q_I^1, Q_I^2, \dots)$ and similarly $Q_2 = (Q_{II}^1, Q_{II}^2, \dots)$. This means that the players want to maximize the payoff function given as

$$u_i(Q_I, Q_{II}) = \sum_{t\in T} \beta_t u_i^t(Q_I^t, Q_{II}^t), i = I, II.$$

We can get the optimal solution by computing

$$\frac{\partial u_i(Q_I, Q_{II})}{\partial Q_i} = \mathbf{0}, i = I, II$$

which yields two systems of equations (each for a given player).

$$\beta_t \frac{\partial u_I^t(Q_I^t, Q_{II}^t)}{\partial Q_I^t} + \beta_{t+1} \frac{\partial u_I^{t+1}(Q_I^{t+1}, Q_{II}^{t+1})}{\partial Q_I^t} = 0, t = 1, 2, \dots, \tag{5.9}$$

$$\beta_t \frac{\partial u_{II}^t(Q_I^t, Q_{II}^t)}{\partial Q_{II}^t} + \beta_{t+1} \frac{\partial u_{II}^{t+1}(Q_I^{t+1}, Q_{II}^{t+1})}{\partial Q_{II}^t} = 0, t = 1, 2, \dots. \tag{5.10}$$

Notice that now the value Q_{II}^t is present in both the payoff in the round t and $t+1$. We can again get the recurent formula by computing

$$a_0 - 2a_1 Q_I^t - a_1 Q_{II}^t + a_1 Q_{t-1} - 2Q_I^t - Q_{II}^t + \beta a_1 Q_I^t = 0, \tag{5.11}$$

$$a_0 - 2a_1 Q_{II}^t - a_1 Q_I^t + a_1 Q_{t-1} - 2Q_{II}^t - Q_I^t + \beta a_1 Q_{II}^t = 0. \tag{5.12}$$

Which has a solution

$$Q_I^t = \frac{a_0 + a_1 Q_{t-1}}{(3-\beta)a_1 + 3},$$

$$Q_{II}^t = \frac{a_0 + a_1 Q_{t-1}}{(3-\beta)a_1 + 3}.$$

This yields that the overal demand $Q_t = \frac{2(a_0 + a_1 Q_{t-1})}{(3-\beta)a_1 + 3}$. Denote $q = \frac{2}{(3-\beta)a_1+3}$ then we can express the overall quantity on the market at the time t using geometric sum

$$Q_t = a_0 q \sum_{n=0}^{t} (a_1 q)^n.$$

Since we assume that $r > 0$ which implies $\beta < 1$ it holds that $a_1 q < 1$, therefore this is a well defined geometric sum with closed form

$$Q_t = a_0 q \left(\frac{1 - (a_1 q)^t}{1 - a_1 q} \right).$$

With $t \to \infty$ this converges to

$$Q = \frac{2a_0}{(1-\beta)a_1 + 3}.$$

Another possible approach is to evaluate

$$\gamma_i(Q_I, Q_{II}) = \liminf_{t\to\infty} \frac{1}{t} \sum_{n=1}^{t} u_i^n(Q_I^n, Q_{II}^n), i = I, II,$$

this is however possible only numerically.

5.1.3 Stochastic model of demand

Now let us discuss the stochastic version of the Cornout model. In this case we will assume that both the relative market cap A_0^t at the time t is a random variable such that $A_0^t > 0, a.s.$. We also assume that the demand is non-negative, that is $D_t(Q_t) \geq 0, a.s.$. The reaction to the change in quantity produced will be assumed to be known constant a_1^t. We will futher assume that $\{A_0^t\}_{t\in T}$ is independent and identically distributed. This means we will consider $D_t(Q_t) = \max(A_0^t - a_1^t(Q_t - Q_{t-1}) - Q_t, 0)$. This results in a sequential game with random payoff

$$u_i^t(Q_I^t, Q_{II}^t) = D_t(Q_t)Q_i^t, i = I, II.$$

For this payoff we may use techniques as were discussed in the previous chapter. In the following example we will consider, that the players want to maximize the agregate payoff for a given time horizont.

5.2 Competition of internet providers

Now let us discuss this model for an example of a duopoly. Here we will consider the case of two internet providers on a single street. We will assume that both providers I and II are approximatelly similar in size and each day they want to decide the average daily internet network speed limit (capacity) for a given street in a city in Gigabites per second (Gb/s). The providers want to find an optimal amount to supply each day for a given 7-day period.

We will assume that the price of the service is linear with respect to the total daily speed limit provided by both companies and that, if the supplied amount is higher than demand the resulting price of the network is 0. Let the base demand for the network on the street in a given day be random with two scenarios

1. high demand $D + L$ with probability p.
2. Standard demand D with probability $1 - p$.

If the demand is higher than the quantity suplied this means that the individual network speed for the customer will be slower than he requires for a higher cost. On the other hand, if the demand is significantly lower the customers may use a higher speed network than they required for a cheeper cost.

We will assume that the reaction of customers is exact to the change in the supplied capacity and so $a_1^t = 1$. Further assume that the base demand each day is independent from the previous day.

With this the overall price at which providers sell the service at day t is given by

$$D_t(Q_t) = D + X_t L - (Q_t - Q_{t-1}) - Q_t,$$

where $X_t \sim Alt(p)$.

Further we will consider $D = 20\, Gb/s$, $L = 5\, Gb/s$, $p = 0.4$. We will also consider that the providers may provide the capacity up to maximum of $M = 10\, Gb/s$. We can check that with this it follows that $\mathsf{E}\ D_t(Q_t) \geq 0, \forall Q_t, Q_{t-1} \in [0, 2M]$.

This motivates us to consider the model for a finite agregate payoff

$$u_i(Q_I, Q_{II}) = \sum_{t=1}^{7} u_i(Q_I^t, Q_{II}^t), i = I, II.$$

Here $Q_i = (Q_i^1, Q_i^2, Q_i^3, Q_i^4, Q_i^5, Q_i^6, Q_i^7)^T \in [0, M]^7$ are the respective daily quantities for each day supplied by company i. We will assume that the providers want to defend themselves from the worst outcome, that is they want to solve

$$\begin{aligned} &\max_{Q_i \in [0,M]^7, \delta \in \mathbb{R}} \delta, \\ &\text{s.t. } \mathbb{P}[\forall Q_{-i} \in [0, M]^7 : u_i(Q_i, Q_{-i}) \geq \delta] \geq \alpha_i. \end{aligned} \tag{5.13}$$

Here we assume that the two companies use the same confidence of $\alpha_I = \alpha_{II} = 0.95$ for their worst payoff scenario.

In this setting we have $K = 2^7 = 128$ different scenarios for the 7-day period. Denote them ω_k with respective probabilities p_k, $k = 1, \ldots, 128$. We can rewrite this program in the form of a mixed-integer non-linear program (MINLP) as

$$\begin{aligned} &\max_{Q_i^t, \delta, z_k} \delta, \\ &\text{s.t. } \delta - u_i(Q_i, Q_{-i})(\omega_k) \leq B(1 - z_k), \forall Q_{-i}^\tau \in [0, M], \tau = 1, \ldots, 7,\ k = 1, \ldots, K \\ &\sum_{k=1}^{K} p_k z_k \geq \alpha_i, \\ &Q_i^t \in [0, M],\ t = 1, \ldots, 7, \\ &\delta \in \mathbb{R}, z_k \in \{0, 1\},\ k = 1, \ldots, K. \end{aligned} \tag{5.14}$$

Where

$$u_i(Q_i, Q_{-i})(\omega_k) = \sum_{t=1}^{7} \max(D + X_t(\omega_k)L - (Q_i^t + Q_{-i}^t - Q_i^{t-1} - Q_{-i}^{t-1}) - Q_i^t - Q_{-i}^t, 0) Q_i^t$$

and B is a sufficiently big constant such that the bound is always satisfied whenever $z_k = 0$. Notice that the k-th bound of the program is satisfied whenever it is satisfied for a $\delta - \min_{Q_{-i} \in [0,M]^7} u_i(Q_i, Q_{-i})(\omega_k)$.

For a given $k = 1, \ldots, K$ the gradient of $u_i(Q_i, Q_{-i})(\omega_k)$ with respect to Q_{-i} is

$$-Q_i^t, \; t = 1, \ldots, 7.$$

Since $Q_i^t \geq 0$ the resulting gradient is always non-positive and therefore the minimum for each scenario is when the other supplier always supplies the maximal amount M each day.

With this we can simplify the program into the form

$$\begin{aligned} &\max_{Q_i^t, \delta, z_k} \delta, \\ &\text{s.t. } \delta - f(Q_i, \omega_k) \leq B(1 - z_k), \; k = 1, \ldots, K \\ &\sum_{k=1}^{K} p_k z_k \geq \alpha_i, \\ &Q_i^t \in [0, M], \; t = 1, \ldots, 7, \\ &\delta \in \mathbb{R}, z_k \in \{0, 1\}, \; k = 1, \ldots, K. \end{aligned} \tag{5.15}$$

Where

$$f(Q_i, \omega_k) = \sum_{t=1}^{7} \max(D + X_t(\omega_k)L - 2M - (Q_i^t - Q_i^{t-1}) - Q_i^t, 0) Q_i^t.$$

We also have to rewrite the maxima in the bounds using integer variables c_k^t and variables representing the maximum m_k^t in the following form

$$\begin{aligned} &\max_{Q_i^t, \delta, z_k, m_k^t, c_k^t} \delta, \\ &\text{s.t. } \delta - \sum_{t=1}^{7} m_k^t Q_i^t \leq B(1 - z_k), \; k = 1, \ldots, K \\ &m_k^t \leq (1 - c_k^t) g(Q_i, \omega_k), \; k = 1, \ldots, K, \; t = 1, \ldots, 7 \\ &\sum_{k=1}^{K} p_k z_k \geq \alpha_i, \\ &Q_i^t \in [0, M], \; t = 1, \ldots, 7, \\ &\delta \in \mathbb{R}, z_k \in \{0, 1\}, \; k = 1, \ldots, K. \\ &m_k^t \in \mathbb{R}, c_k^t \in \{0, 1\}, \; k = 1, \ldots, K, \; t = 1, \ldots, 7, \end{aligned} \tag{5.16}$$

where $g(Q_i, \omega_k) = D + X_t(\omega_k)L - 2M - (Q_i^t - Q_i^{t-1}) - Q_i^t$. To solve this program we use the APOPT solver of Python GEKKO package for optimization introduced in [1].

In its implementation GEKKO requires string lenght of equations less than 15000 characters, which is not satisfied for our probability bound. Therefore we use the adjusted version of our program in the form

$$
\begin{aligned}
&\max_{Q_i^t,\delta,z_k,P_j} \delta, \qquad (5.17)\\
&\text{s.t. } \delta - \sum_{t=1}^{7} m_k^t Q_i^t \le B(1-z_k),\ k=1,\dots,K\\
&m_k^t \le (1-c_k^t) g(Q_i,\omega_k),\ k=1,\dots,K,\ t=1,\dots,7\\
&P_j = \sum_{k=1+K/4(j-1)}^{jK/4} p_k z_k,\ j=1,\dots,4\\
&P_1+P_2+P_3+P_4 \ge \alpha_i,\\
&Q_i^t \in [0,M],\ t=1,\dots,7,\\
&\delta\in\mathbb{R}, z_k\in\{0,1\},\ k=1,\dots,K,\\
&P_j\in[0,1],\ j=1,\dots,4.\\
&m_k^t\in\mathbb{R}, c_k^t\in\{0,1\},\ k=1,\dots,K,\ t=1,\dots,7,
\end{aligned}
$$

The computed optimal strategy is

$$Q_I^* = Q_{II}^* = (5.28, 5.28, 5.28, 5.28, 5.28, 5.28, 5.28)^T$$

with optimal $\delta^* = 250.69$.

We evaluated this strategy on 100 000 simulations of the demand process parameters and compared it with the optimal strategy as based on the static model, the dynamic model behavior strategy and the expected strategy from the dynamic model which is a limit of the behavior strategy. The deterministic strategies are computed with respect to the expected value of the parameters of the model. That is with $a_0^t = \mathsf{E}\ A_1^t = D + pL$ and $a_1^t = 1$. We will refer to those as

$$
\begin{aligned}
Q_{static} &= (7.33, 7.33, 7.33, 7.33, 7.33, 7.33, 7.33)^T,\\
Q_{exp} &= (4.40, 4.40, 4.40, 4.40, 4.40, 4.40, 4.40)^T
\end{aligned}
$$

and

$$Q_{beh} = (2.44, 3.53, 4.01, 4.23, 4.33, 4.36, 4.38)^T$$

and to the optimal strategy from our program as $Q_{opt} = Q_I^*$ all rounded up to two decimal points. We consider also two model cases. In the first both companies use the same above mentioned strategy and in the second only the first company uses the optimal strategy and the other company uses the most harmful strategy

$$Q_{worst} = (M, M, M, M, M, M, M).$$

In the following table we show the characteristics of the payoff distribution computed from the simulations for the standard case.

Strategy	Minimal payoff	Maximal payoff	Mean payoff
Q_{opt}	299.07	533.64	386.32
Q_{static}	234.67	530.44	338.87
Q_{beh}	300.34	462.93	362.95
Q_{exp}	306.24	498.96	379.50

Table 5.1: Total 7-day payoff characteristics computed from the simulations. In the standard strategy case. Rounded up to two decimal units.

We see that the optimal and the dynamic strategies all three performed similarly well in this situation. The static model strategy was by far the most volatile as would be expected. The optimal strategy also has the highest mean payoff. The number of simulations in which the optimal strategy yielded payoff higher than $\delta^* = 250.69$ was in all 100 000 simulations.

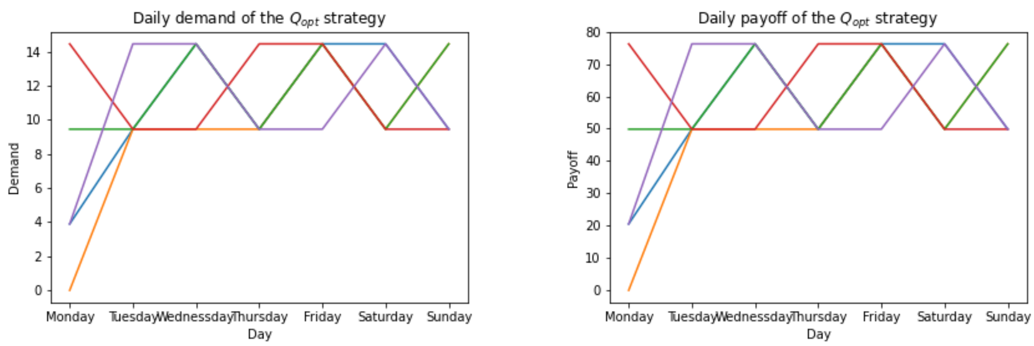
In the following table we show the characteristics of the payoff distribution computed from the simulations for the case when the other player plays the most harmful strategy.

Strategy	Minimal payoff	Maximal payoff	Mean payoff
Q_{opt}	93.83	359.18	191.96
Q_{static}	9.78	393.56	150.65
Q_{beh}	125.72	299.70	191.68
Q_{exp}	109.12	326.48	189.77

Table 5.2: Total 7-day payoff characteristics computed from the simulations. In the most harmful strategy case. Rounded up to two decimal units.

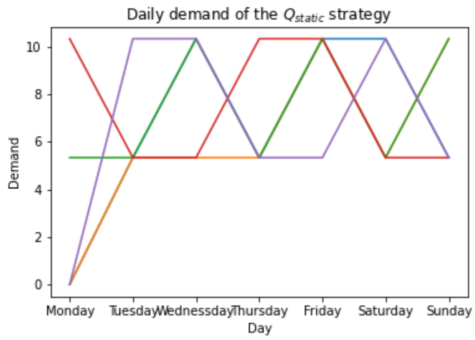
Here we may see that the gap is bit bigger between the three strategies based on the dynamic models. We see that the optimal and dynamic strategies had very similar mean payoffs, but differ in the minimal payoff which is slightly lower for the optimal strategy.

To ilustrate the payoff processes we sample 5 random simulation runs of our strategies and plot the resulting demands and payoffs in the following figures.

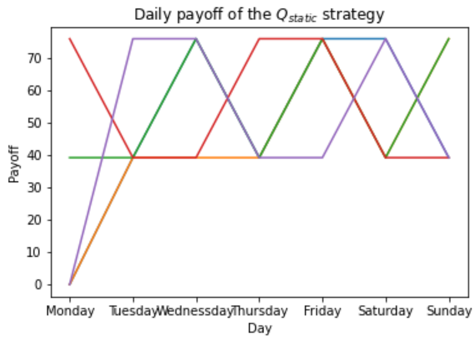


(a) Five simulations of the demand process for the optimal strategy Q_{opt}

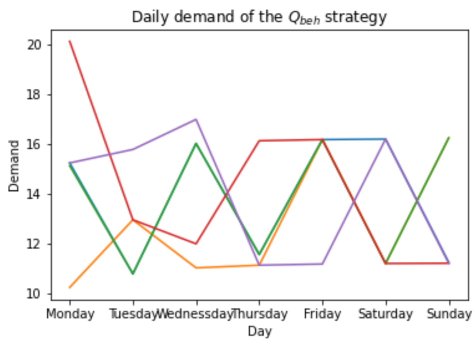
(b) Five simulations of the payoff process for the optimal strategy Q_{opt}



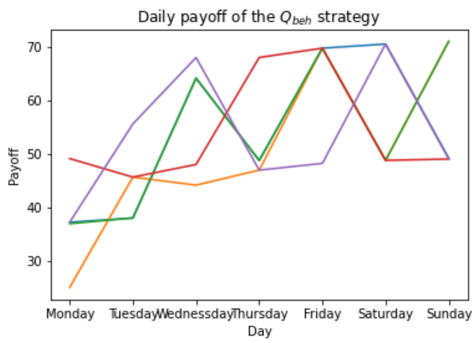
(a) Five simulations of the demand process for the static strategy Q_{static}



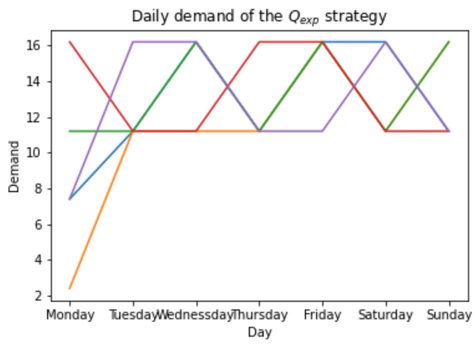
(b) Five simulations of the payoff process for the optimal strategy Q_{static}



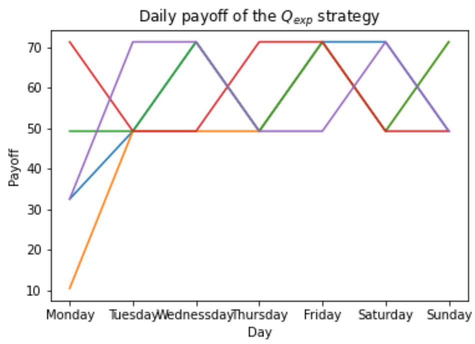
(a) Five simulations of the demand process for the dynamic strategy Q_{beh}



(b) Five simulations of the payoff process for the dynamic strategy Q_{beh}



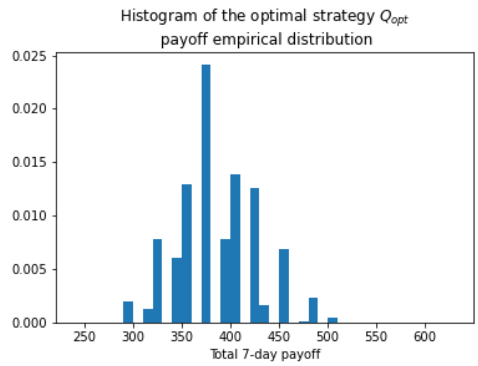
(a) Five simulations of the demand process for the expected strategy Q_{exp}



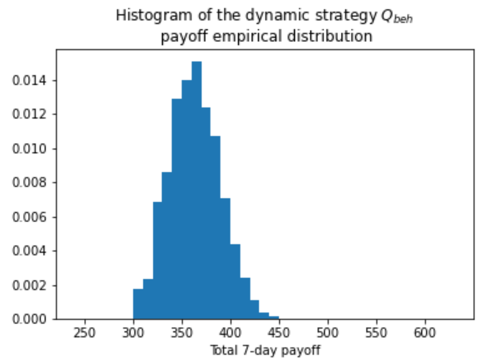
(b) Five simulations of the payoff process for the expected strategy Q_{exp}

From this figure we may see that the reaction to the change is in the case of the dynamic strategy less drastic in the first stage as it gradually grows. This may explain the difference in the minimal payoffs from the simulations. Also in this small time window the effect of change in the first day of the week is more dramatic as it goes from theoretical 0. The possible refinment would be to compute the reaction at the first day as change from the expected supply level of the static model, which would be more realistic.

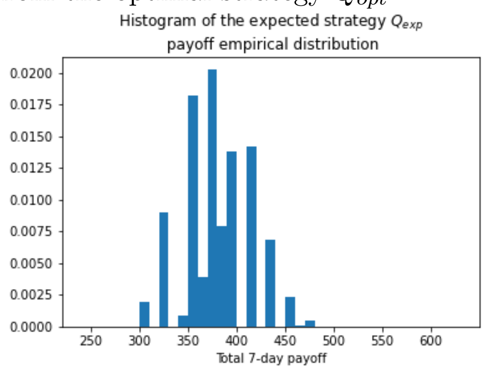
In the following figures we may see the resulting distributions of payoffs using each strategy and their respective probabilities.



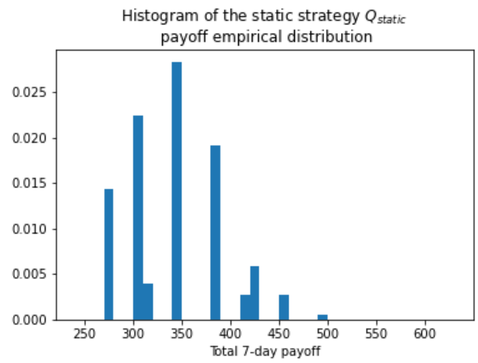
(a) Histogram of the total 7-day payoff from the optimal strategy Q_{opt}



(b) Histogram of the total 7-day payoff from the dynamic strategy Q_{beh}



(c) Histogram of the total 7-day payoff from the expected strategy Q_{exp}



(d) Histogram of the total 7-day payoff from the static strategy Q_{static}

In the three cases when the chosen strategy is constant we see that the payoff distribution takes only discrete steps which correspond to the number of realizations of the higher demand. Therefore the effect of a change in supply is only present in the first day.

5.2.1 Discussion of the results

From our numerical analysis of the modeled scenario we may conclude that the behavior deterministic and expected dynamic strategies performed very well. In the comparison with our base optimal strategy from the worst payoff model they scored similar in some of the metrics. This is probably due to the fact that the dependence of the payoffs in our model is very low and therefore it has ergodic payoff for which we know that the optimal strategy is the strategy with expected parameters. This is well approximated by the behavior deterministic strategy and therefore both strategies are well suited for this model. We also can conclude that the optimal solutions based on the dynamic and stochastic models supply less than the optimal solution from the static model. Possible refinement of this model would be to manage the higher change effect in the first day by assuming that the reaction in the first day is only given by the change from the expected strategy of the static model.

6. Conclusion

In our thesis we studied sequential games or games in the extensive form. We focused primarily on the case when the payoff from the game is a random variable. First we discussed the general results of the game theory for games in the strategic form, which both apply to single-stage games and may be used to represent sequential games (or in other words games in the extensive form). In the second chapter we proposed several criteria for optimality in the games with random payoff in the strategic form. Here we considered the deterministic equivalent game approach discussed by earlier authors as well as new approach based on the generalization of the concept of Nash equilibria. In the third chapter we discussed the known theory for games in the extensive form with deterministic payoff. Based on those results we proposed possible approaches to the games in extensive form with random payoff in the fourth chapter. In the fifth chapter we developed Cornout model of oligopoly for dynamic and stochastic demand and used this model on an example of a game in the telecomunications industry where companies want to allocate their network capacity for a given day in a week. For this example we considered 7-day decision period and used the model with total worst payoff to compute optimal allocation of capacity for each day. We compared this solution with the deterministic strategies based on the expected values of the parameters.

6.1 Main results of our thesis

The main theoretical result of our thesis is the concept of $\boldsymbol{\alpha}$-Nash equilibria that we developed in the second chapter of this thesis. Here we generalized the idea of best response for games with random payoff and by that we were able to define an equivalent of the concept of Nash equilibrium for a game with random payoff. We proved several properties of those equilibria in the following discussion. Mainly we created existence criteria for the non-trivial $\boldsymbol{\alpha}$-Nash equilibria for the case with individual probabilistic constraints in the Theorem 9. Then we considered the case when there is only finitely many realizations of the payoff function and managed to prove that those criteria are always satisfied with which we were able to formulate the stochastic equivalent of the Fundamental theorem of game theory in the Theorem 11. We used this theorem to provide existence criteria for the non-trivial $\boldsymbol{\alpha}$-Nash equilibria with the join constrains. In this chapter we also formulated the generalized versions of minimax problem, which were considered as optimality criteria for matrix games with random payoff by previous authors.

In the third chapter we introduced the concept of a sequential game and showed how it may be transformed from the extensive form into strategic form. In this chapter we summarized results for these types of games and provided an example in the form of a repeated Prisoner's dilema with discounting. On this example we showed how the results change from a single stage to a repeated form of a game.

In the fourth chapter we summarized results on the theory of Markov decision processes and their generalizations in the form of a stochastic game. Mainly in this chapter we introduced the sequential games with random payoff and showed how they relate to the stochastic games. We considered several possible approaches to the optimality that were based on the approaches standardly used for the deterministic games. In more detail we examined cases when the payoff process is either Markov chain, martingale or an ergodic process.

In the fifth chapter we discussed the Cornout model of duopoly and extended standard results to the case when the demand is dynamic and stochastic. We then applied this model on a model example of a competition of two internet network providers that have to optimize allocation of their network capacity over a 7-day period. In this example we used the worst payoff model to find optimal solutions and derived its equivalent in the form of a mixed-integer non-linear program. We then compared the results of our optimal strategy with strategies derived by the deterministic model.

6.2 Potential for further development

Main potential in developing of this thesis in studying the properties of $\boldsymbol{\alpha}$-Nash equilibria and their application in the case of a extensive form of a game. So far we lack the proof of their existence in a game with general payoff. Also we need to develop a good way to approximate the maximal confidence levels on which the game has a $\boldsymbol{\alpha}$-Nash equilibrium and methods to find those equilibria efficently. Our proof provides only very rought estimate of the confidence levels. We could further focus on the case when the payoff has infinite number of realizations. For the case with joint probability constrains a better existence criteria would be required. The conection of optimal strategy profiles in the different concepts of optimality such as the stochastic domination or the generalized minimax programs and of the $\boldsymbol{\alpha}$-Nash equilibria will require further development as well.

In the stochastic Cornout model of duopoly we could use the techniques of probability theory to deduce more theoretical results on the existence of optimal solutions and consider other approaches to the optimality than the worst payoff model.

Bibliography

[1] Logan Beal, Daniel Hill, R Martin, and John Hedengren. Gekko optimization suite. *Processes*, 6(8):106, 2018.

[2] R. J. Birge and F. Louveaux. *Introduction to Stochastic Programming.* Second Edition. Springer, New York, 2011.

[3] R.A. Blau. Random-payoff two-person zero-sum games. *Operations Research*, 22(6):1243–1251, 1974.

[4] R.G. Cassidy, C.A. Field, and M.J.L. Kirby. Solution of a satisfying model for random payoff games. *Management science*, 19(3):266–271, 1972.

[5] A. Charnes, M.J.L. Kirby, and W.M Raike. Zero-zero chance-constrained games. *Theory of probability and its applications*, 8(4):628–646, 1968.

[6] J. Cheng, J. Leung, and A. Lisser. Random-payoff two-person zero-sum games with joint constrains. *European Journal of Operational Research*, 252:213–219, 2016.

[7] A Dhillon and B. Lockwood. When are plurality rule voting games dominance-solvable? *Games and Economic Behavior*, 46:55–75, 2004.

[8] M. Dufwenberg and M. Stegeman. Existence and uniqueness of maximal reductions under iterated strict dominance. *Econometrica*, 70:2007–2023, 2002.

[9] M. Dufwenberg and M. Stegeman. Iterated elimination procedures. *Economic Theory*, 70:437–465, 2020.

[10] J. Dupačová and P. Lachout. *Úvod do optimalizace.* Matfyzpress, Praha, 2011.

[11] M. Maschler, E. Solan, and S. Zamir. *Game Theory.* Second Edition. Cambridge University press, Cambridge, 2020.

[12] J. Matoušek and J. Nešetřil. *Kapitoly z diskrétni matematiky.* 5. vydání. Karolinum, Praha, 2022.

[13] M. L. Puterman. Chapter 8 markov decision processes. In *Stochastic Models*, volume 2 of *Handbooks in Operations Research and Management Science*, pages 331–434. Elsevier, 1990.

[14] V.V. Singh, O. Jouini, and A. Lisser. Existance of nash equilibrium for chance-constrained games. *Operations Research Letters*, 44:640–644, 2016.